

A. d'Onofrio, P. Cerrai, A. Gandolfi
Editors

New Challenges for Cancer Systems Biomedicine

To Prof. Franco Giannessi, with gratitude

SIMAI Springer Series

Series Editors: Nicola Bellomo • Luca Formaggia (*Editors-in-Chief*)

**Wolfgang Bangerth • Fabio Nobile • Lorenzo Pareschi • Pablo Pedregal Tercero •
Andrea Tosin • Jorge P. Zubelli**

Alberto d'Onofrio • Paola Cerrai • Alberto Gandolfi
(Eds.)

New Challenges for Cancer Systems Biomedicine

 Springer

Editors:

Alberto d'Onofrio
Department of Experimental Oncology
European Institute of Oncology
Milan, Italy

Paola Cerrai
Department of Mathematics
University of Pisa
Italy

Alberto Gandolfi
Istituto di Analisi dei Sistemi ed Informatica
"A. Ruberti" – CNR
Rome, Italy

ISSN: 2280-840X

ISSN: 2280-8418 (electronic)

SIMAI Springer Series

ISBN 978-88-470-2570-7

ISBN 978-88-470-2571-4 (eBook)

DOI 10.1007/978-88-470-2571-4

Springer Milan Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012941090

© Springer-Verlag Italia 2012

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

9 8 7 6 5 4 3 2 1

Cover design: Beatrice d, Milano

Typesetting: PTP-Berlin, Protago T_EX-Production GmbH, Germany (www.ptp-berlin.de)

Printing: Grafiche Porpora, Segrate (MI)

Printed in Italy

Springer-Verlag Italia S.r.l., Via Decembrio 28, I-20137 Milano

Springer-Verlag is part of Springer Science+Business Media (www.springer.com)

Preface

The future of oncology has a name: Molecular Medicine (MM). Molecular Medicine is a new science based on three pillars. Two of them are well known and evident in its very name: medical science and molecular biology. However, there is a general unawareness that MM is firmly based on a third but equally important pillar: Systems Biomedicine. Currently this term mainly evokes Bioinformatics and modern Applied Statistics, but increasingly it shall have to include (as in part it already does) the interacting complex of scientific fields such as Mathematical Biology, Systems Biology, Theoretical Biophysics.

The data from MM of tumors are complex and heterogeneous (e.g. clinical data paired with -omics data) but – and this is their most important feature - are unified by their dynamical nature. Indeed, cancers are a family of dynamic diseases, endowed by multiple temporal and spatial scales, and their polymorphic macroscopic instances are emergent properties originating from a wide number of microscopic interplays at intracellular and intercellular level. The complexity of these multiscale data cannot be deciphered by natural language reasoning, or by classical data analysis based on static data mining and model-unrelated time series analysis. These classical tools no longer suffice to cope with MM data in order to understand them and to produce meaningful and useful predictions.

As a consequence, it is mandatory to build mechanistic mathematical models of biomedical phenomena with complex outputs. These models could allow a deeper understanding of the “internal dynamics” of single patients or classes of patients, hopefully opening the road for tailored therapies. This is a huge challenge at the frontier of contemporary mathematical modeling, since dynamic modeling in MM is what allows to bridge the bench to the bedside, and in perspective it will be increasingly instrumental in aiding the cure of patients. By no means this implies that future medical doctors will be like electronic engineers, skilfully using special software to cure patients. Nevertheless, in a realistic perspective, future generation of oncologists will be more similar to cardiologists that rely on basic knowledge of the physics of heart and circulation, and use devices from bioengineering in their everyday clinical work.

Strict collaboration between biomedical researchers and Systems Biomedicine scientists is mandatory to make these hypotheses true in the future. What is the current state of this collaboration? A small number of outstanding experimental groups are seriously collaborating with biomathematicians, physicists and computer scientists, still maintaining separate competences. This is an important phenomenon. However, what is happening mainly in Systems Biology is even more interesting. An increasing number of inter-disciplinary groups are forming thanks to a new generation of group-leaders whose undergraduate background is in biomedical sciences. This trend leads to a far closer contact between two worlds quite separate in the past, and to the use of a common language. Many life scientists then become – in different degrees – confident and aware of the Systems Biomedicine potential. As an example of this potential, we can mention the possible role of mathematical modeling in drug development. Post-genomic drug discovery is indeed revealing serious shortcomings in the current way of performing clinical trials, which appears inadequate to face the age of personalized medicine. Systems Biomedicine, in the future generation of clinical trials, could thus play a fundamental role in shaping cancer treatments for single patients or groups of patients.

Many are the challenges that Systems Biomedicine of cancer must face. We have the responsibility of showing to the Life Sciences community that the potential of this discipline may become reality.

The present book has been inspired by the ideas underlying the Workshop Mathematical Oncology: New Challenges for Systems Biomedicine, held at the Ettore Majorana Centre for Scientific Culture in Erice (Italy), September 26–30, 2011.

The aim of this book is not only to illustrate the state of the art of tumor systems biomedicine, but also (and especially) to explicitly capture and collect results of the above-mentioned collaborative trends. Indeed, this volume is characterized by a well-structured presence of a large number of life scientists working directly in Systems Biomedicine, and a number of mathematical biology researchers working in biomedical institutions. With this book we wish to provide a coherent view of tumor modeling, based on the concept that mathematical modeling is the third pillar of molecular medicine. We hope that these features give to this work an unprecedented tone, providing an original interdisciplinary insight into the biomedical applications. We also hope the book may foster and encourage new fruitful communications and cooperations.

The present volume covers five basic topics of interest in oncology: comprehensive theories of cancer growth, systems biology of cancer, basic mechanisms of tumor progression, tumor-immune system interplay and immunotherapy, computational methods for improving chemotherapies. All the scales are so addressed, from the intracellular molecular networks to the therapy of patients. Moreover, relevance is given to recent mathematical methodologies such as nonlinear analysis, control and optimization theory, cellular automata and cellular-Potts modeling, agent-based modeling, and formal methods of computer science.

We wish to thank Professor Nicola Bellomo for the kind invitation to edit this book as part of the new SIMAI-Springer Series, and for his constant encouragement. We also wish to thank Professor Zvia Agur for her effort as co-organizer in setting the scientific shape of the Erice Workshop, and for the help in defining the project of this book.

Milan, Pisa, Rome, May 2012

Alberto d'Onofrio
Paola Cerrai
Alberto Gandolfi

Contents

Part I Towards a Comprehensive Theory of Cancer Growth

Combining Game Theory and Graph Theory to Model Interactions between Cells in the Tumor Microenvironment	3
Attila Csikász-Nagy, Matteo Cavaliere, and Sean Sedwards	
Growth as the Root of all Evil in Carcinomas: Synergy between pH Buffering and Anti-Angiogenesis Prevents Emergence of Hallmarks of Cancer	19
Ariosto Silva and Robert Gatenby	
Phase Transitions in Cancer	35
Ricard V. Solé	

Part II Cancer Related Signalling Pathways

Spatio-Temporal Modelling of Intracellular Signalling Pathways: Transcription Factors, Negative Feedback Systems and Oscillations	55
Mark A.J. Chaplain, Marc Sturrock, and Alan J. Terry	
Understanding Cell Fate Decisions by Identifying Crucial System Dynamics	83
Dirk Fey, David R. Croucher, Walter Kolch, and Boris N. Kholodenko	
Modelling Biochemical Pathways with the Calculus of Looping Sequences	105
Paolo Milazzo, Antonella Del Corso, Andrea Maggiolo-Schettini, Umberto Mura, and Roberto Barbuti	
Dynamic Simulations of Pathways Downstream of TGFβ, Wnt and EGF-Family Growth Factors, in Colorectal Cancer, including Mutations and Treatments with Onco-Protein Inhibitors	127
Lorenzo Tortolina, Nicoletta Castagnino, Cristina De Ambrosi, Annalisa Barla, Alessandro Verri, Gabriele Zoppoli, Luca Bagnasco, Daniela Piras, Franco Patrone, Alberto Ballestrero, and Silvio Parodi	

Part III Basic Mechanisms of Tumor Progression

Some Results on the Population Behavior of Cancer Stem Cells	145
Edoardo Beretta, Vincenzo Capasso, Annick Harel-Bellan, and Nadya Morozova	
Glucose Metabolism in Multicellular Spheroids, ATP Production and Effects of Acidity	173
Antonio Fasano	
Cell-Cell Interactions in Solid Tumors – the Role of Cancer Stem Cells .	191
Xuefeng Gao, J. Tyson McDonald, Lynn Hlatky, and Heiko Enderling	
Hybrid Cellular Potts Model for Solid Tumor Growth	205
Marco Scianna and Luigi Preziosi	

Part IV Tumor-Immune System Interplay and Immunotherapy

Computational Models as Novel Tools for Cancer Vaccines	227
Filippo Castiglione, Pier Luigi Lollini, Santo Motta, Arianna Paladini, Francesco Pappalardo, and Marzio Pennisi	
On the Dynamics of Tumor-Immune System Interactions and Combined Chemo- and Immunotherapy	249
Alberto d’Onofrio, Urszula Ledzewicz, and Heinz Schättler	
Modeling the Kinetics of the Immune Response	267
Ami Radunskaya and Sarah Hook	

Part V Computational Method for Improving Chemotherapy

Optimizing Cancer Chemotherapy: from Mathematical Theories to Clinical Treatment	285
Zvia Agur and Yuri Kheifetz	
A Systems Biomedicine Approach for Chronotherapeutics Optimization: Focus on the Anticancer Drug Irinotecan	301
Annabelle Ballesta, Jean Clairambault, Sandrine Dulong, and Francis Levi	
Modeling the Dynamics of HCV Infected Cells to Tailor Antiviral Therapy in Clinical Practice: Can This Approach Fit for Neoplastic Cells?	329
Piero Colombatto, Filippo Oliveri, Ferruccio Bonino, and Maurizia R. Brunetto	
Introducing Drug Transport Early in the Design of Hypoxia Selective Anticancer Agents Using a Mathematical Modelling Approach	337
Kevin Hicks	

Top-Down Multiscale Simulation of Tumor Response to Treatment in the Context of In Silico Oncology. The Notion of *Oncosimulator* 355
Georgios Stamatakos

Challenges in the Integration of Flow Cytometry and Time-Lapse Live Cell Imaging Data Using a Cell Proliferation Model 377
Paolo Ubezio, Francesca Falcetta, and Monica Lupi

Combining Game Theory and Graph Theory to Model Interactions between Cells in the Tumor Microenvironment

Attila Csikász-Nagy, Matteo Cavaliere, and Sean Sedwards

Abstract Mathematical concepts of graph theory and game theory both influence models of biological systems. We combine these two approaches to understand how game-like interactions influence the cellular topology of a planar tissue. We review the literature on the role of cell to cell interactions in tumourigenesis and survey the mathematical approaches that have been used to simulate such cell-cell interactions. We present how this game-graph approach can be used to simulate epithelial tissue growth and how it can foster our understanding of the role of cell-cell communication in the early stages of cancer development. We present computational models that we use to test how cooperating and non-cooperating cells build planar tissues and compare the simulated tissue topologies with literature data. We further discuss how such system could be used to model microenvironmental communications between cancer cells and the surrounding tissue.

1 Introduction

Mathematical approaches to investigate cellular behaviour have a long history [10, 11, 20, 25, 86]. Game theory and graph theory have both provided extensive contri-

A. Csikász-Nagy (✉)

The Microsoft Research, University of Trento Centre for Computational and Systems Biology,
38123 Povo (Trento), Italy
e-mail: csikasz@cosbi.eu

M. Cavaliere

Logic of Genomic Systems Laboratory, Spanish National Biotechnology Centre (CNB-CSIC), Darwin 3, Campus de Cantoblanco, 28049 Madrid, Spain
e-mail: mcavaliere@cnb.csic.es

S. Sedwards

INRIA Rennes, Bretagne Atlantique, Campus universitaire de Beaulieu, 35042 Rennes Cedex, France
e-mail: sean.sedwards@inria.fr

bution to our biological knowledge [15, 57, 76, 89, 104, 109]. Game theory investigates the interactions between ‘players’, while graph theory focuses on the large scale topology of an interaction network. Tissues of an organism are made up of highly organized spatial interactions of individual cells. In this chapter we investigate how such interactions can influence the large scale cytoarchitecture (cell to cell interaction network topology) of the tissue. First we review the corresponding biological literature and the published mathematical modelling approaches. We then introduce a method where game and graph theory are combined to treat a tissue as a dynamical graph that is updated based on the local interactions of individual cells.

2 Role of Social Interactions in Complex Biological Systems

Multicellular organisms emerged from individual cells cooperating to deal with the changes in their environment [105], while currently existing single celled slime moulds reunite to form fruiting bodies if food sources are limited [111]. This complex structure helps the survival of some of the cells in the population but requires other cells to sacrifice their proliferation capacity. Biofilm formation is another example where the individual cells pay some cost (produce adhesive factors) for the benefit of the whole population (better reach of resources) [87]. In both of these cases, cheaters (defectors) can emerge in the population which do not pay the cost of sacrificing themselves or producing adhesive factor, but still benefit from the ‘public good’ (i.e., they always take part of the fruiting body or enjoy the elevated resources in a biofilm). As these cheaters save energy by not working towards the public good, they can reproduce faster than their cooperating neighbours. The uncontrolled proliferation of cheaters can cause the entire population to collapse [106]: cooperators are outgrown by cheaters, but they cannot maintain the nurturing environment. Such population collapses can happen without any external influence, since cheaters can arise from genetic mutation of regular cooperators. Indeed, there is an evolutionary benefit to cheat as long as many cooperators provide public good to the population [89]. Individual cells of the human body can also be viewed as cooperators: somatic cells cannot pass on their DNA but help germ cells pass on their genetic material. However, cheaters can also appear in the human body: mutations can lead to the emergence of cancer cells that proliferate as long as nutrients and oxygen are present; their uncontrolled proliferation eventually leading to the collapse of the population of cells when the organism dies. Cooperation between individual cells can happen through cell to cell interactions. Direct membrane-bound signalling molecules can send information between cells or diffusible signalling molecules may be secreted by a cell and picked up by the receptors of neighbouring cells. These receptors induce some response in the receiver cell and may lead to the release of another extracellular signalling molecule. Such complex signalling crosstalk between neighbouring cells is needed to keep a cooperating tissue intact. A mutated cell might lose some of the controls from its environment and can behave differently. The nature of communication between normal and mutated cells (such as those found in cancer) is not

well understood [55]. Cell to cell interactions influence tissue topology and tissue architecture feeds back to influence microenvironmental cell to cell communications. Here we discuss how ideas from systems biology, evolutionary game theory, mathematical biology and computer science can help to uncover the details of cell to cell interactions.

2.1 Cell to Cell Interactions in the Tumour Microenvironment

The fate of a cell is determined by signals received from its environment [61, 66]. This is true for cells in a developing tissue, for quiescent cells and also for proliferating malignant cells. It was observed a long time ago that the neighbouring cells of cancer cells behave differently to the same cells in a normal context, but recently it has also been proposed that they can play an active role in controlling the behaviour of cancer cells [72, 101]. As Cairns stated [21] “Survival of rapidly renewing tissues of long-lived animals like man requires that they be protected against the natural selection of fitter variant cells (that is, the spontaneous appearance of cancer).” He went on to propose that stem cells and tissue architecture might be important for tumour suppression. In the following year Nowell established the evolutionary view of cancer development [92]. Recently, by combining these ideas with newer results on tumour microenvironments, the concept has emerged that cancer is an evolutionary ecological process [51, 80]. Recent reviews discuss what type of interactions might happen between different cells and give examples demonstrating the presence of almost all types of possible ecological interactions [129] between mammalian cells [55, 80, 84, 100]. The question remains as to which of these interactions might be important for mutated cells to spread in an otherwise homogenous tissue. Microenvironmental regulation is a sum of cell to cell and cell-extracellular matrix interactions. The role of microenvironmental signalling during metastasis is crucial, since malignant cells can only colonize tissues where they can find a supporting microenvironment [18, 55, 60]. As distal metastasis is most often established by a single clonal tumour cell, the initial signals it receives when it first reaches a homeostatic tissue might play a crucial role. There is evidence for the role of cell-cell competition during the development of the *Drosophila* wing disc [74], as well as recent research suggesting that a similar process occurs between normal and cancer cells [14, 64, 84]. Cell to cell contact inhibition is another layer, where cells can control their neighbour’s behaviour. How cells sense and signal when they are closely packed is only partially understood [114, 115], while the failure of this cell to cell signalling is a critical step in tumour formation [54]. Over-proliferating cancer cells reach a state where they are limited for oxygen and nutrients, forcing them to compete for resources with other cancer cells and with the neighbouring tissue [3]. The results in *Drosophila* suggest competition for growth factors occurs even in a normal tissue: “Winner” cells survive and divide to fill up the space left over by dead “loser” cells [84]. Other researchers [13, 18, 29] instead suggest that mutated cells might cooperate: they aid the survival and proliferation of each other by producing complementary growth factors. Genetic heterogeneity of tumours [75] is usu-

ally explained as the winner clone developing genetic instability; but at a previous stage, this clone might have interacted with other genetically heterogeneous cancer cells [51]. When the clonal population of the winner clone reaches a critical size, it stops proliferating as it cannot obtain enough oxygen and nutrients. It needs to secrete VEGF that induces angiogenesis [54]. The process of angiogenesis is complex and cancer cells might activate it indirectly by inducing neighbouring normal cells to secrete VEGF by competing with them for oxygen and resources [72]. Thus, there could be both altruistic and competitive interactions between neighbouring cells during the different stages of cancer progression.

3 Tissue Topology Dynamics

One of the early results [69] on tissue topology research was the discovery that in an epithelial tissue, cells are packed in a specific way in a quasi 2D monolayer. Development of the *Drosophila* wing is widely used to understand epithelial tissue patterning [70]. In the wing disc most cells have six direct neighbours, but cells with four or nine neighbours are also observed [48]. Importantly, other epithelial tissues show very similar distribution of sidedness (number of neighbours) [48, 68]. Several theoretical ideas have appeared to explain how this observed distribution emerges, using elasticity theory [41], mechanics [2, 117], or simple mathematical models [48, 98]. Furthermore, as mentioned above, competition also plays a crucial role, suggesting that there might be multiple levels of control working to determine the complex cytoarchitecture of living tissue. It has been shown that during mitosis the number of neighbours changes and the directionality of the cell division plane has also been investigated [49], but the dynamic changes in tissue topology in the presence of competing cells were not followed in detail. Similar investigations have not yet been performed on mammalian tissues, although it is reasonable to assume that epithelial tissue topology (at least in single layer, simple epithelia) looks quite similar to the ones observed in other organisms [40, 88]. Graph rewriting [110], an extension of L-systems [109], is widely used to model plant development [104]. In this generative paradigm, fractal structures are created by the iterative and recursive application of a finite set of elementary rewriting rules. These rules determine how an initial simple graph is transformed into a more complex one. Similar concepts have been used to reproduce the patterning of seashells [77] and to model plant epithelial tissue development [79, 120]. Below we present an approach that combines graph rewriting with game theory to model cell to cell interactions in epithelial tissues.

4 Evolutionary Dynamics of Graphs

The dynamics of complex networks have recently been investigated through evolutionary game theoretic models [97, 118, 122], while the evolution of cooperation has

also been investigated on graphs [63, 93, 112]. These approaches either update the interaction graphs based on decisions of the interacting agents (as in co-evolutionary games [52, 99]) or fix the graph topology and let the game evolve on the fixed graph [71, 94]. Cell to cell interactions induce cell proliferation or cell death, which lead to changes in the tissue topology (aka. cytoarchitecture or cell to cell interaction graph). The dividing cell cannot influence who will be its neighbour after division, thus the process cannot be captured by co-evolutionary networks. The topology also changes after each division, thus fixed graphs are not an adequate model either. In what follows we describe how topological changes in epithelial tissue can be induced by evolutionary game-like interactions between cells.

5 Computational Models of Cancer Formation

Several published models deal with combinations of spatial and evolutionary aspects of tumour formation [10, 17, 20, 27, 30, 37, 102]. Many of the existing spatial models use cellular automata on a fixed grid to represent tissue structure, in order to simplify calculations [4, 31, 83]. Recently, more complex models utilizing biophysical rules have been published [6, 34, 108], challenging previous dogmas [38, 39]. The majority of these models start simulations from a given number of cancer cells and follow how the tumour grows and then progresses through angiogenesis by invading nearby healthy tissue [8, 27]. Other models treat cancer as an evolutionary disease and use evolutionary game theory methods to understand multistep tumour formation [12, 45, 46, 82], but these models do not deal with changes in tissue architecture or, if they do, they follow the tumour shape on a fixed grid [9, 121, 131]. Several multi-scale models of the development of various cancer types have appeared recently [5, 23, 30, 116, 132]. These simulate intra- and extracellular interactions on the molecular scale. Each cell of the simulated tissue has its own intracellular network model and responds to extracellular signals.

Models of epithelial tissue patterning also exist [1, 42, 44, 50, 119, 126]. Some of these models have been carefully matched to experimental observations [123]. Several approaches have been developed to deal with tissue dynamics during cell birth, death and movement and it depends on the specific scientific question under investigation as to which is the most suitable model [96, 108]. The simplest ones use cellular automata and let cells fill neighbouring squares when they move or divide. These do not allow users to follow dynamical changes in the sidedness of the cells. Cellular automata with hexagonal neighbourhoods have also been considered but also keep the shape of the cells fixed [67]. As a further step of complexity, cells can be modelled as dynamically changing convex polygons [36, 58, 81, 113]. Voronoi tessellation of a tissue can give the precise cell to cell contacts and the dual-graph of such Voronoi cells is the Delaunay network, which is itself the direct cell to cell contact network of cells (Fig. 2). Delaunay triangles can be dynamically changed in a way that matches how cells divide, die or move both in two and three dimensions [19, 36, 58, 81, 113]. This approach allows fast computation and can be used to

follow changes in cytoarchitecture, although it does not explicitly consider physical forces between cells. More realistic models that can account for cell shape changes include immersed boundary approaches [107] (taken from fluid dynamics) as well as single cell off-lattice biophysical models [33, 43] – the drawback of these models is that the computational time required to simulate entire tissues can become prohibitive. Cellular Potts models also use a grid, but cells can change their shapes and neighbours as each cell occupies more than one pixel of the grid [62, 78]. Continuous models that deal with diffusing compounds [103] have been combined with agent based models to become hybrid discrete-continuum models [9].

On a complementary research line, several models deal with the evolutionary aspects of tumour formation [12, 45, 82]. Some models use a discrete time branching process, starting from normal cells with driver and passenger mutations, to capture the waves of clonal expansions and understand the emergence of genetic instability [16, 82, 90]. Others consider the microenvironment in their evolutionary mathematical model [47, 124] or focus on interactions of a few cell types [26, 28]. Game theoretical approaches have also been used in agent based computational models without following tissue topology changes [7, 32, 73].

6 A Dynamical Tissue Model of Cell to Cell Interactions

A multicellular organism provides a rich environment for its cells to proliferate, thus requiring complex mechanisms to control cell growth in space and time. Cell to cell signalling plays an important role in controlling both apoptosis and proliferation. Thus cells can send both positive and negative signals to their neighbours depending on what inputs they have received. These signals determine which cell divides or dies, leading to changes in tissue topology. At the same time, the cytoarchitecture determines the neighbourhood of a cell, leading to a crosstalk between topology and signalling. As the authors of [55] state: “growth factor signals controlling cell number and position within tissues are thought to be transmitted in a temporally and spatially regulated fashion from one cell to its neighbors; such paracrine signaling is difficult to access experimentally”. At this time, it still is not well understood when, how and what type of cell to cell interactions happen during cancer formation. The initial steps of tumourigenesis, up to the formation of in situ carcinomas of epithelial tissues [127] can be highly influenced by local cell to cell interactions. Since the majority of cancers are of epithelial origin, understanding the details of local cell to cell interactions and their effects on the epithelial tissue architecture during neoplasm formation could help to design new methods to identify and stop cancer cells before angiogenesis and cancer invasion. The presented method focuses on such early appearance of the first mutant clones in the population. A second step with crucial cell to cell interactions is during metastatic colonization, when a cancer cell arrives to a new microenvironment and it starts to proliferate and interact with the surrounding cells [51, 55]. It could be of further interest to develop methods to investigate cell to cell interactions at this step. We employ computational approaches to

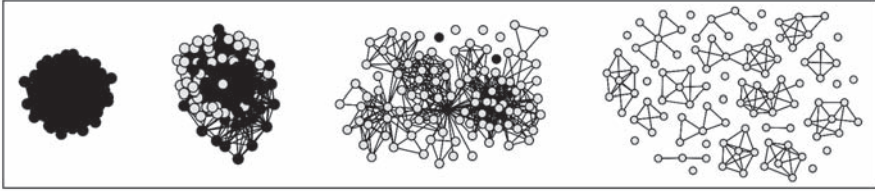


Fig. 1. Network topologies of cooperators and cheaters. Dense network of cooperators (dark gray) collapse after the spreading of defectors (light grey) in a network created by payoff based preferential attachment and copying of parent's links. Modified from [22]: M. Cavaliere et al. *J Theor Biol* **299**, 126–38 (2012) © Elsevier Science

test how the topology of a simulated tissue changes when we assume different types of interactions between different types of cells. We merge concepts from evolutionary game theory and network dynamics to test how the coexistence of cooperating and cheating cells influence tissue topology. The simulated tissue topologies can be compared to in-vivo samples of various stage preangiogenic solid tumours [53]. Combining evolutionary game theory and ideas from graph rewriting led to our earlier results [22] that show how local embedding of an offspring near to its parent can influence the overall topology of the network. In [22], the interaction network is abstract and does not closely model biological tissue, however the evolutionary process that dynamically changes the topology of the network is biologically plausible: the fittest individuals are more likely to reproduce and new members are placed near the fittest individual. The fitness of an individual was based on the payoff it received by playing the prisoner's dilemma game with all its neighbours. In this classical game, cooperators pay a cost to distribute benefit to their neighbours and defectors (aka. cheaters) do not pay this cost but enjoy the benefit of being near to a cooperator. Simulations of the model showed that when a cheater appears in a dense (highly connected) network of cooperators it has a high chance of spreading, eventually leading to a fragmentation of the network (Fig. 1 from left to right). Such fragmentation of a well-structured system resembles the collapse of biofilm- and fruiting body forming populations [130]. This abstract model can also be seen as mimicking the effects of the appearance of 'cheating' tumour cells in an otherwise cooperating normal tissue. This result gives a hint that the structure of a network depends on the type of interactions between the individuals. Here we investigate what tissue structures can be formed by similar rules applied on a 2D tissue-like dynamic network. In this approach, we focus on simulating the dynamics of planar cell to cell contact graphs (Fig. 2) and employ simple rules for the updating of the tissue after each cell division and cell death events [36]. After a cell divides or dies we apply graph rewriting rules [110] to construct the new tissue layout that keeps the graph properly triangulated. In this version of the model we do not deal explicitly with physical forces between neighboring cells, although the cooperating signals we assume in our model could come from molecular as well as from physical signals [41]. We simulate a planar tissue with periodic boundaries to avoid misleading effects of the interactions at the edge of a tissue. In the first analysis we do not allow cell death and only follow

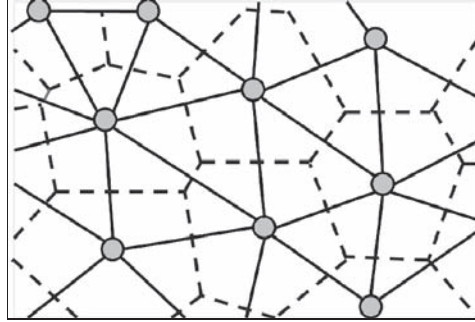


Fig. 2. Voronoi cells (dashed) and Delaunay interaction graph (solid) for a 2D tissue layout

how the tissue evolves as it reaches a size of 10000 cells (from an initial of 6 cells). The employed interaction game is the prisoner's dilemma, where cooperators pay a cost to distribute benefit to neighbouring cells. Defectors do not give benefit and do not need to pay cost, but enjoy the benefits of surrounding cooperators [56, 90]. The probability of a cell dividing is a function of the sum of the payoffs it receives by playing the prisoner's dilemma with all its neighbours (see [22] for details on calculations). Cooperators (C) in a fully cooperating tissue can be imagined as cells that keep communicating with their neighbours. Such communication could be realised by growth factor or other signalling induced proliferation or physical force induced division of cells [95, 125]. More connected cooperators will communicate more and thus have higher payoff, leading to a higher chance to divide. If a tissue is made of only defectors (D), interpreted as cells that do not communicate with each other such that all D-D interactions lead to zero payoff, their selection will be random (i.e., independent of the number of neighbours they have). Samples of simulated tissues of cooperating cells look far more regular (Fig. 3a) than tissues built by cheaters (Fig. 3b). Topology measures of such cooperator and defector- built tissues can be directly compared with literature data on sidedness of cells in epithelial tissues from different organisms (Fig. 4). Fig. 4 shows that cooperators build tissue with a distribution of sidedness that matches normal tissue in a variety of organisms, while tissue built by defectors under equivalent conditions is clearly different. The cooperating tissue results also match the best modelling results of [98], where the authors used a cleavage plane selection model with several assumptions to match the experimental results. In our model we did not use special assumptions on cleavage plane selection (treating it as random) and our results thus show that cell to cell communication is sufficient to explain the observed distribution of sidedness in epithelial tissue topology. The game-like interactions are controlled only by the benefit and cost of cooperation and the selection strength parameters. The sidedness distribution of the simulated cooperator tissue can be matched to the observed values by varying just one of these parameters with respect to the other (since only C-C interaction take place). The resultant strength of C-C interactions determines, probabilistically, which cell will divide. Note that the sidedness distribution would not match the experimental

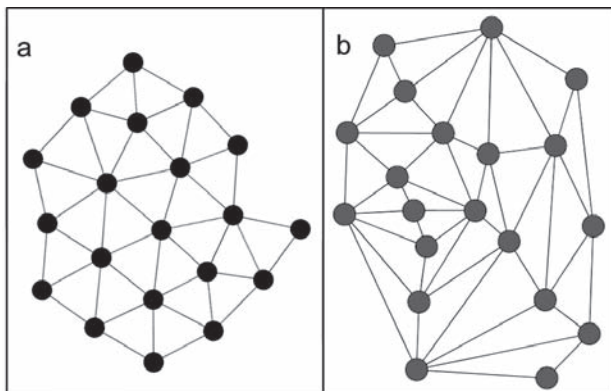


Fig. 3. Tissue topologies of cooperators and defectors. Delaunay networks built by tissue-like embedding of cooperators (a) and defectors (b). Connections inside the communities of 20 cells, sampled from a simulated tissue of 1000 cells

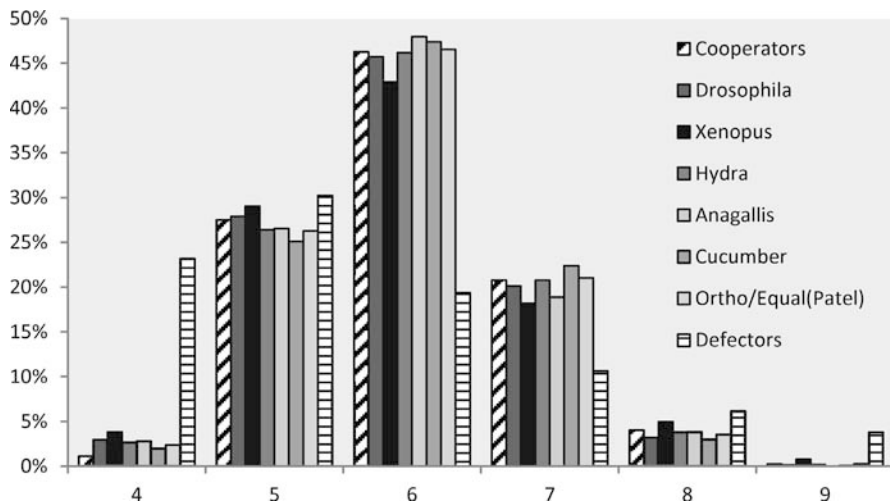


Fig. 4. Distribution of sidedness (number of neighbours) in various epithelial tissues and in simulations of tissues built by cooperators and defectors. Experimental data from [48], the Ortho/Equal (Patel) columns show the best results from the model of [98]

data if the cell with the highest number of neighbours is always selected [35]. When a defector appears in a simulated tissue of all cooperators it has a higher chance of proliferating than cooperators since it does not pay the cost of giving benefit to its neighbours. As more and more defectors appear in the neighbourhood, however, the spreading of this clone slows down and eventually stops when most of the neighbours of new defectors are other defectors. Since the neighbouring cooperators will not divide (they will have low payoff by being connected to defectors), they just pick up more and more connections to defectors as those divide. The final picture

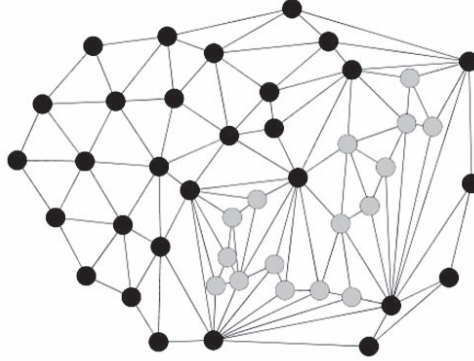


Fig. 5. Spreading of a defector (grey) clone in a tissue of cooperators (black). Simulated Delaunay network of 16 defectors and some surrounding cooperators from a tissue of 1200 interacting cells

looks like a ‘benign tumour’ of defector cells residing in a normal cooperator tissue (Fig. 5). The original neighbours of the first defector gain many connections to new defectors, but the proliferation of defectors soon results in no defector having more than two cooperator neighbours. These defectors will therefore receive benefit (b) from just two cooperators while paying zero cost (c) (payoff = $2b$). Cooperators not connected to defectors will have an average of 6 cooperator neighbours (a property of large planar graphs) and thus an average payoff of $6(b - c)$. Under these conditions, with $b/c \gg 1.5$ defectors are unlikely to divide and cannot spread further. In the simulated case of $b/c = 1.4$ we observe that the defector population grows very slowly. The highly connected cooperators surrounding the defector population are biologically unrealistic. Either physical forces will cut some of the cell to cell contacts we simulate here or these cells might die. In the present model these cooperators constitute a barrier between the populations of defectors and cooperators and prevent the defectors spreading. Since they have the lowest payoff in the whole population, payoff based cell removal would remove them and allow the further spreading of defectors.

7 Possible Extensions of the Model

It would be interesting to see what tissue topology is observed at early stage neoplasms and non-vascularized tumours and see if these have any similarity to the simulated defector population structures. To capture such complex tissue architectures with our model it could be extended to three dimensions. Since models of 3D Delaunay polyhedron dynamics are available [58] the model could be extended this way, but as other 3D tissue dynamics simulation tools are also available [43, 107] one can consider to use other methods to simulate cell to cell interactions in more complex tissues. An obvious extension is the introduction of physical forces in the model [41, 125], which would resolve some issues with highly connected cells. Fur-

ther extension could be to use different game theory models. The public goods game (multi-person prisoner's dilemma) might be more appropriate to capture cell communication with diffusible signalling molecules [24]. Also, the addition of multiple strategies – resembling the behaviour of various cell types – could lead to more realistic models. Non-cancer cells of a tissue can be of different types as well. Considering a fixed ratio of cells in a tissue to be stem cells [64, 85] could have a significant influence in cancer formation [59]. To investigate the role of stem cells, the model could be extended by the possibility of asymmetric division of cells to simulate cell differentiation. This would allow simulations of a topological model of stem cell differentiation to understand the interactions in the stem cell niche [128]. Further extensions would eventually reach towards multi-scale models of tumour formation, where the proposed dynamic graph update idea might bring new concepts to an emerging field [23, 30].

8 Conclusion

Here we introduced a model of cell to cell interaction induced tissue dynamics, based on a combination of evolutionary game theory and graph theory approaches. To achieve this, we extended our earlier model of cooperation induced network formation [22]. We presented simulations of evolutionary game induced cell divisions on a planar graph, updating the graph after each cell division using graph rewriting. The model is deliberately kept as simple as possible, but is nevertheless able to capture some quantitative measures of epithelial tissue topologies. Further extensions of the model will make it more realistic and increase its predictive power. We have proposed some of these extensions here and look forward to see how this dynamic game-graph approach could be used to understand various aspects of tumour formation.

Acknowledgements The authors are grateful to Peter Csermely for comments on the manuscript and acknowledge support from the Italian Research Fund FIRB (RBPR0523C3) and from program JAEDoc15 (“Programa junta para la ampliacion de estudios”).

References

1. Adra, S., Sun, T., MacNeil, S., Holcombe, M., Smallwood, R.: Development of a Three Dimensional Multiscale Computational Model of the Human Epidermis. *PLoS ONE* **5**, e8511 (2010)
2. Aegerter-Wilmsen, T., Smith, A.C., Christen, A.J., Aegerter, C.M., Hafen, E., Basler, K.: Exploring the effects of mechanical feedback on epithelial topology. *Development* **137**, 499–506 (2010)
3. Aguirre-Ghiso, J.A.: Models, mechanisms and clinical evidence for cancer dormancy. *Nat Rev Cancer* **7**, 834–846 (2007)
4. Alarcón, T., Byrne, H.M., Maini, P.K.: A cellular automaton model for tumour growth in inhomogeneous environment. *J Theor Biol* **225**, 257–274 (2003)
5. Alarcón, T., Byrne, H.M., Maini, P.K.: A Multiple Scale Model for Tumor Growth. *Mult Mod Sim* **3**, 440–475 (2005)

6. Anderson, A.R.A., Chaplain, M.A.J., Rejniak, K.A.: *Single-Cell-Based Models in Biology and Medicine*. Birkhauser, Basel, Switzerland (2007)
7. Anderson, A.R.A., Hassanein, M., Branch, K.M., Lu, J., Lobdell, N.A., Maier, J., Basanta, D., Weidow, B., Narasanna, A., Arteaga, C.L., Reynolds, A.B., Quaranta, V., Estrada, L., Weaver, A.M.: Microenvironmental Independence Associated with Tumor Progression. *Cancer Res* **69**, 8797–8806 (2009)
8. Anderson, A.R.A., Rejniak, K., Gerlee, P., Quaranta, V.: Microenvironment driven invasion: a multiscale multimodel investigation. *J Math Biol* **58**, 579–624 (2009)
9. Anderson, A.R.A., Weaver, A.M., Cummings, P.T., Quaranta, V.: Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* **127**, 905–915 (2006)
10. Anderson, A.R.A., Quaranta, V.: Integrative mathematical oncology. *Nat Rev Cancer* **8**, 227–234 (2008)
11. Araujo, R., McElwain, D.: A history of the study of solid tumour growth: The contribution of mathematical modelling. *Bull Math Biol* **66**, 1039–1091 (2004)
12. Attolini, C.S.O., Michor, F.: Evolutionary Theory of Cancer. *Ann NY Acad Sci* **1168**, 23–51 (2009)
13. Axelrod, R., Axelrod, D.E., Pienta, K.J.: Evolution of cooperation among tumor cells. *PNAS* **103**, 13474–13479 (2006) 10.1073/pnas.0606053103
14. Baker, N.E., Li, W.: Cell Competition and Its Possible Relation to Cancer. *Cancer Res* **68**, 5505–5507 (2008)
15. Barabasi, A.L., Albert, R.: Emergence of Scaling in Random Networks. *Science* **286**, 509–512 (1999)
16. Beerenwinkel, N., Antal, T., Dingli, D., Traulsen, A., Kinzler, K.W., Velculescu, V.E., Vogelstein, B., Nowak, M.A.: Genetic Progression and the Waiting Time to Cancer. *PLoS Comput Biol* **3**, e225 (2007)
17. Bellomo, N., Li, N.K., Maini, P.K.: On the foundations of cancer modelling: Selected topics, speculations, and perspectives. *Math Model Meth Appl Sci* **18**, 593–646 (2008)
18. Bidard, F.C., Pierga, J.Y., Vincent-Salomon, A., Poupon, M.-F.: A “class action” against the microenvironment: do cancer cells cooperate in metastasis? *Cancer and Metastasis Rev* **27**, 5–10 (2008)
19. Bock, M., Tyagi, A., Kreft, J.U., Alt, W.: Generalized Voronoi Tessellation as a Model of Two-dimensional Cell Tissue Dynamics. *Bull Math Biol* **72**, 1696–1731 (2010)
20. Byrne, H.M.: Dissecting cancer through mathematics: from the cell to the animal model. *Nat Rev Cancer* **10**, 221–230 (2010)
21. Cairns, J.: Mutation selection and the natural history of cancer. *Nature* **255**, 197–200 (1975)
22. Cavaliere, M., Sedwards, S., Tarnita, C.E., Nowak, M.A., Csikász-Nagy, A.: Prosperity is associated with instability in dynamical networks. *J Theor Biol* **299**, 126–38 (2012)
23. Chaplain, M.A.J.: Multiscale mathematical modelling in biology and medicine. *IMA J Appl Math* **76**, 371–388 (2011)
24. Craig Maclean, R., Brandon, C.: Stable public goods cooperation and dynamic social interactions in yeast. *J Evol Biol* **21**, 1836–1843 (2008)
25. Csikász-Nagy, A.: Computational systems biology of the cell cycle. *Brief Bioinform* **10**, 424–434 (2009)
26. d’Onofrio, A.: A general framework for modeling tumor-immune system competition and immunotherapy: Mathematical analysis and biomedical inferences. *Physica D* **208**, 220–235 (2005)
27. d’Onofrio, A., Gandolfi, A.: A family of models of angiogenesis and anti-angiogenesis anti-cancer therapy. *Math Med Biol* **26**, 63–95 (2009)
28. d’Onofrio, A., Gatti, F., Cerrai, P., Freschi, L.: Delay-induced oscillatory dynamics of tumour–immune system interaction. *Math Comp Modell* **51**, 572–591 (2010)
29. d’Onofrio, A., Tomlinson, I.P.M.: A nonlinear mathematical model of cell turnover, differentiation and tumorigenesis in the intestinal crypt. *J Theor Biol* **244**, 367–374 (2007)

30. Deisboeck, T.S., Stamatakos, G.S.: Multiscale cancer modeling. CRC Press, Boca-Raton (2010)
31. Deutsch, A., Dormann, S.: Cellular automaton modeling of biological pattern formation: characterization, applications, and analysis. Birkhauser, Basel (2005)
32. Dingli, D., Chalub, F.A.C.C., Santos, F.C., Van Segbroeck, S., Pacheco, J.M.: Cancer phenotype as the outcome of an evolutionary game between normal and malignant cells. *Br J Cancer* **101**, 1130–1136 (2009)
33. Drasdo, D., Hoehme, S.: A single-cell-based model of tumor growth in vitro : monolayers and spheroids. *Phys Biol* **2**, 133 (2005)
34. Drasdo, D., Hoehme, S., Block, M.: On the Role of Physics in the Growth and Pattern Formation of Multi-Cellular Systems: What can we Learn from Individual-Cell Based Models? *J Stat Phys* **128**, 287–345 (2007)
35. Dubertret, B., Aste, T., Ohlenbusch, H.M., Rivier, N.: Two-dimensional froths and the dynamics of biological tissues. *Phys Rev E* **58**, 6368 (1998)
36. Duvdevani-Bar, S., Segel, L.: On Topological Simulations in Developmental Biology. *J Theor Biol* **166**, 33–50 (1994)
37. Edelman, L.B., Eddy, J.A., Price, N.D.: In silico models of cancer. *WIRE Syst Biol Med* **2**, 438–459 (2010)
38. Enderling, H., Anderson, A.R.A., Chaplain, M.A.J., Beheshti H., Hlatky, L., Hahnfeldt, P.: Paradoxical Dependencies of Tumor Dormancy and Progression on Basic Cell Kinetics. *Cancer Res* **69**, 8814–8821 (2009)
39. Enderling, H., Hahnfeldt, P.: Cancer stem cells in solid tumors: Is ‘evading apoptosis’ a hallmark of cancer? *Prog Biophys Mol Biol* **106**, 391–399 (2011)
40. Escudero, L.M., da F. Costa, L., Kicheva, A., Briscoe, J., Freeman, M., Babu, M.M.: Epithelial organisation revealed by a network of cellular contacts. *Nat Commun* **2**, 526 (2011)
41. Farhadifar, R., Röper, J.C., Aigouy, B., Eaton, S., Jülicher, F.: The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. *Curr Biol* **17**, 2095–2104 (2007)
42. Fozard, J., Kirkham, G., Buttery, L., King, J., Jensen, O., Byrne, H.: Techniques for analysing pattern formation in populations of stem cells and their progeny. *BMC Bioinf* **12**, 396 (2011)
43. Galle, J., Aust, G., Schaller, G., Beyer, T., Drasdo, D.: Individual cell-based models of the spatial-temporal organization of multicellular systems—Achievements and limitations. *Cytometry A* **69**, 704–710 (2006)
44. Galle, J., Loeffler, M., Drasdo, D.: Modeling the Effect of Deregulated Proliferation and Apoptosis on the Growth Dynamics of Epithelial Cell Populations In Vitro. *Biophys J* **88**, 62–75 (2005)
45. Gatenby, R.A., Gillies, R.J.: A microenvironmental model of carcinogenesis. *Nat Rev Cancer* **8**, 56–61 (2008)
46. Gatenby, R.A., Gillies, R.J., Brown, J.S.: The evolutionary dynamics of cancer prevention. *Nat Rev Cancer* **10**, 526–527 (2010)
47. Gatenby, R.A., Vincent, T.L.: An Evolutionary Model Of Carcinogenesis. *Cancer Res* **63**, 6212–6220 (2003)
48. Gibson, M.C., Patel, A.B., Nagpal, R., Perrimon, N.: The emergence of geometric order in proliferating metazoan epithelia. *Nature* **442**, 1038–1041 (2006)
49. Gibson, W.T., Veldhuis, J.H., Rubinstein, B., Cartwright, H.N., Perrimon, N., Brodland, G.W., Nagpal, R., Gibson, M.C.: Control of the Mitotic Cleavage Plane by Local Epithelial Topology. *Cell* **144**, 427–438 (2011)
50. Grant, M.R., Mostov, K.E., Tlsty, T.D., Hunt, C.A.: Simulating Properties of In Vitro Epithelial Cell Morphogenesis. *PLoS Comput Biol* **2**, e129 (2006)
51. Greaves, M., Maley, C.C.: Clonal evolution in cancer. *Nature* **481**, 306–313 (2012)
52. Gross, T., Blasius, B.: Adaptive coevolutionary networks: a review. *J R Soc Interface* **5**, 259–271 (2008)
53. Guillaud, M., Clem, C., MacAulay, C.: An in silico platform for the study of epithelial pre-invasive neoplastic development. *Biosystems* **102**, 22–31 (2010)
54. Hanahan, D., Weinberg, R.A.: The hallmarks of cancer. *Cell* **100**, 57–70 (2000)

55. Hanahan, D., Weinberg, R.A.: Hallmarks of Cancer: The Next Generation. *Cell* **144**, 646–674 (2011)
56. Hofbauer, J., Sigmund, K.: Evolutionary games and population dynamics. London Mathematical Society Student Texts. Cambridge University Press, Cambridge (1988)
57. Hofbauer, J., Sigmund, K.: Evolutionary game dynamics. *Bull Am Math Soc* **40**, 479 (2003)
58. Honda, H., Tanemura, M., Nagai, T.: A three-dimensional vertex dynamics cell model of space-filling polyhedra simulating cell behavior in a cell aggregate. *J Theor Biol* **226**, 439–453 (2004)
59. Hsu, Y.C., Fuchs, E.: A family business: stem cell progeny join the niche to regulate homeostasis. *Nat Rev Mol Cell Biol* **13**, 103–114 (2012)
60. Hu, M., Polyak, K.: Microenvironmental regulation of cancer development. *Curr Opin Gen Dev* **18**, 27–34 (2008)
61. Hunter, T.: Signaling—2000 and Beyond. *Cell* **100**, 113–127 (2000)
62. Izaguirre, J.A., Chaturvedi, R., Huang, C., Cickovski, T., Coffland, J., Thomas, G., Forgacs, G., Alber, M., Hentschel, G., Newman, S.A., Glazier, J.A.: CompuCell, a multi-model framework for simulation of morphogenesis. *Bioinformatics* **20**, 1129–1137 (2004)
63. Jackson, M.O.: Social and Economic Networks. Princeton University Press, Princeton (2008)
64. Johnston, L.A.: Competitive Interactions Between Cells: Death, Growth, and Geography. *Science* **324**, 1679–1682 (2009)
65. Johnston, M.D., Maini, P.K., Jonathan Chapman, S., Edwards, C.M., Bodmer, W.F.: On the proportion of cancer stem cells in a tumour. *J Theor Biol* **266**, 708–711 (2010)
66. Kholodenko, B.N.: Cell-signalling dynamics in time and space. *Nat Rev Mol Cell Biol* **7**, 165–176 (2006)
67. Kim, S., Debnath, J., Mostov, K., Park, S., Hunt, C.A.: A computational approach to resolve cell level contributions to early glandular epithelial cancer progression. *BMC Syst Biol* **3**, 122 (2009)
68. Lecuit, T., Lenne, P.-F.: Cell surface mechanics and the control of cell shape, tissue patterns and morphogenesis. *Nat Rev Mol Cell Biol* **8**, 633–644 (2007)
69. Lewis, F.T.: The correlation between cell division and the shapes and sizes of prismatic cells in the epidermis of cucumis. *Anat Rec* **38**, 341–376 (1928)
70. Lewis, J.: From Signals to Patterns: Space, Time, and Mathematics in Developmental Biology. *Science* **322**, 399–403 (2008)
71. Lieberman, E., Hauert, C., Nowak, M.A.: Evolutionary dynamics on graphs. *Nature* **433**, 312–316 (2005)
72. Liotta, L.A., Kohn, E.C.: The microenvironment of the tumour-host interface. *Nature* **411**, 375–379 (2001)
73. Mansury, Y., Diggory, M., Deisboeck, T.S.: Evolutionary game theory in an agent-based brain tumor model: Exploring the ‘Genotype-Phenotype’ link. *J Theor Biol* **238**, 146–156 (2006)
74. Martin, F.A., Herrera, S.C., Morata, G.S.: Cell competition, growth and size control in the *Drosophila* wing imaginal disc. *Development* **136**, 3747–3756 (2009)
75. Marusyk, A., Polyak, K.: Tumor heterogeneity: Causes and consequences. *BBA-Rev Cancer* **1805**, 105–117 (2010)
76. Maynard Smith, J.: The theory of games and the evolution of animal conflicts. *J Theor Biol* **47**, 209–221 (1974)
77. Meinhardt, H.: The algorithmic beauty of sea shells. Springer-Verlag New York, Inc., (1995)
78. Merks, R.M.H., Glazier, J.A.: A cell-centered approach to developmental biology. *Physica A* **352**, 113–130 (2005)
79. Merks, R.M.H., Guravage, M., Inzé, D., Beemster, G.T.S.: VirtualLeaf: an open-source framework for cell-based modeling of plant tissue growth and development. *Plant Physiol* **155**, 656–666 (2011)
80. Merlo, L.M.F., Pepper, J.W., Reid, B.J., Maley, C.C.: Cancer as an evolutionary and ecological process. *Nat Rev Cancer* **6**, 924–935 (2006)
81. Meyer-Hermann, M.: Delaunay-Object-Dynamics: Cell Mechanics with a 3D Kinetic and Dynamic Weighted Delaunay-Triangulation. *Curr Top in Dev Biol* **81**, 373–399 (2008)

82. Michor, F., Iwasa, Y., Nowak, M.A.: Dynamics of cancer progression. *Nat Rev Cancer* **4**, 197–205 (2004)
83. Moreira, J., Deutsch, A.: Cellular Automation Models of Tumor Development: A Critical Review. *Adv Compl Sys* **5**, 247–268 (2002)
84. Moreno, E.: Is cell competition relevant to cancer? *Nat Rev Cancer* **8**, 141–147 (2008)
85. Morrison, S.J., Spradling, A.C.: Stem Cells and Niches: Mechanisms That Promote Stem Cell Maintenance throughout Life. *Cell* **132**, 598–611 (2008)
86. Murray, J.D.: *Mathematical biology*, vol. 2. Springer, (2002)
87. Nadell, C.D., Xavier, J.B., Foster, K.R.: The sociobiology of biofilms. *FEMS Microbiol Rev* **33**, 206–224 (2009)
88. Nagpal, R., Patel, A., Gibson, M.C.: Epithelial topology. *Bioessays* **30**, 260–266 (2008)
89. Nowak, M.A.: *Evolutionary Dynamics: Exploring the Equations of Life*. Harvard University Press, Cambridge, MA (2006)
90. Nowak, M.A., Michor, F., Komarova, N.L., Iwasa, Y.: Evolutionary dynamics of tumor suppressor gene inactivation. *PNAS* **101**, 10635–10638 (2004)
91. Nowak, M.A., Sigmund, K.: *Evolutionary Dynamics of Biological Games*. *Science* **303**, 793–799 (2004)
92. Nowell, P.C.: The clonal evolution of tumor cell populations. *Science* **194**, 23–28 (1976)
93. Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A.: A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502–505 (2006)
94. Ohtsuki, H., Nowak, M.A.: Evolutionary stability on graphs. *J Theor Biol* **251**, 698–707 (2008)
95. Orr, A.W., Helmke, B.P., Blackman, B.R., Schwartz, M.A.: Mechanisms of Mechanotransduction. *Develop Cell* **10**, 11–20 (2006)
96. Osborne, J.M., Walter, A., Kershaw, S.K., Mirams, G.R., Fletcher, A.G., Pathmanathan, P., Gavaghan, D., Jensen, O.E., Maini, P.K., Byrne, H.M.: A hybrid approach to multi-scale modelling of cancer. *Phil Trans Roy Soc A* **368**, 5013–5028 (2010)
97. Pacheco, J.M., Traulsen, A., Nowak, M.A.: Coevolution of Strategy and Structure in Complex Networks with Dynamical Linking. *Phys Rev Lett* **97**, 258103–258104 (2006)
98. Patel, A.B., Gibson, W.T., Gibson, M.C., Nagpal, R.: Modeling and Inferring Cleavage Patterns in Proliferating Epithelia. *PLoS Comput Biol* **5**, e1000412 (2009)
99. Perc, M., Szolnoki, A.: Coevolutionary games—A mini review. *Biosystems* **99**, 109–125 (2010)
100. Pienta, K.J., McGregor, N., Axelrod, R., Axelrod, D.E.: Ecological therapy for cancer: defining tumors using an ecosystem paradigm suggests new opportunities for novel cancer treatments. *Transl Oncol* **1**, 158–164 (2008)
101. Polyak, K., Haviv, I., Campbell, I.G.: Co-evolution of tumor cells and their microenvironment. *Trends Genet* **25**, 30–38 (2009)
102. Preziosi, L.: *Cancer modelling and simulation*, vol. 3. CRC Press, (2003)
103. Preziosi, L., Tosin, A.: Multiphase modelling of tumour growth and extracellular matrix interaction: mathematical tools and applications. *J Math Biol* **58**, 625–656 (2009)
104. Prusinkiewicz, P., Lindenmayer, A., Hanan, J.: *The algorithmic beauty of plants*. Springer Verlag, (1990)
105. Rainey, P.B., Kerr, B.: Cheats as first propagules: A new hypothesis for the evolution of individuality during the transition from single cells to multicellularity. *BioEssays* **32**, 872–880 (2010)
106. Rankin, D.J., Bargum, K., Kokko, H.: The tragedy of the commons in evolutionary biology. *Trends Ecol Evol* **22**, 643–651 (2007)
107. Rejniak, K.A.: An immersed boundary framework for modelling the growth of individual cells: An application to the early tumour development. *J Theor Biol* **247**, 186–204 (2007)
108. Rejniak, K.A., Anderson, A.R.A.: Hybrid models of tumor growth. *WIRE Syst Biol Med* **3**, 115–125 (2011)
109. Rosen, R.: The theory of graphs and its applications. *Bull Math Biol* **24**, 441–443 (1962)

110. Rozenberg, G.: Handbook of graph grammars and computing by graph transformation: volume I: Foundations. World Scientific, River Edge, NJ, USA (1997)
111. Santorelli, L.A., Thompson, C.R.L., Villegas, E., Svetz, J., Dinh, C., Parikh, A., Sucgang, R., Kuspa, A., Strassmann, J.E., Queller, D.C., Shaulsky, G.: Facultative cheater mutants reveal the genetic complexity of cooperation in social amoebae. *Nature* **451**, 1107–1110 (2008)
112. Santos, F.C., Pacheco, J.M.: Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation. *Phys Rev Letters* **95**, 098104 (2005)
113. Schaller, G., Meyer-Hermann, M.: Kinetic and dynamic Delaunay tetrahedralizations in three dimensions. *Comp Phys Comm* **162**, 9–23 (2004)
114. Schrader, J., Deuster, O., Rinn, B., Schulz, M., Kautz, A., Dodel, R., Meyer, B., Al-Abed, Y., Balakrishnan, K., Reese, J., Bacher, M.: Restoration of contact inhibition in human glioblastoma cell lines after MIF knockdown. *BMC Cancer* **9**, 464 (2009)
115. Seluanov, A., Hine, C., Azpurua, J., Feigenson, M., Bozzella, M., Mao, Z., Catania, K.C., Gorbunova, V.: Hypersensitivity to contact inhibition provides a clue to cancer resistance of naked mole-rat. *PNAS* **106**, 19352–19357 (2009)
116. Shirinifard, A., Gens, J.S., Zaitlen, B.L., Popławski, N.J., Swat, M., Glazier, J.A.: 3D Multi-Cell Simulation of Tumor Growth and Angiogenesis. *PLoS ONE* **4**, e7190 (2009)
117. Shraiman, B.I.: Mechanical feedback as a possible regulator of tissue growth. *PNAS* **102**, 3318–3323 (2005)
118. Skyrms, B., Pemantle, R.: A dynamic model of social network formation. *PNAS* **97**, 9340–9346 (2000)
119. Smallwood, R.: Computational modeling of epithelial tissues. *WIRE Syst Biol Med* **1**, 191–201 (2009)
120. Smith, R.S., Guyomarc’h, S., Mandel, T., Reinhardt, D., Kuhlemeier, C., Prusinkiewicz, P.: A plausible model of phyllotaxis. *PNAS* **103**, 1301–1306 (2006)
121. Spencer, S.L., Gerety, R.A., Pienta, K.J., Forrest, S.: Modeling Somatic Evolution in Tumorigenesis. *PLoS Comput Biol* **2**, e108 (2006)
122. Szabó, G., Fáth, G.: Evolutionary games on graphs. *Phys Rep* **446**, 97–216 (2007)
123. Thorne, B.C., Bailey, A.M., Peirce, S.M.: Combining experiments with multi-cell agent-based modeling to study biological tissue patterning. *Brief Bioinform* **8**, 245–257 (2007)
124. Tomlinson, I.P.M.: Game-theory models of interactions between tumour cells. *Eur J Cancer* **33**, 1495–1500 (1997)
125. Vogel, V., Sheetz, M.: Local force and geometry sensing regulate cell functions. *Nat Rev Mol Cell Biol* **7**, 265–275 (2006)
126. Walker, D., Georgopoulos, N., Southgate, J.: From pathway to population – a multiscale model of juxtacrine EGFR-MAPK signalling. *BMC Sys Biol* **2**, 102 (2008)
127. Wang, C.C., Jamal, L., Janes, K.A.: Normal morphogenesis of epithelial tissues and progression of epithelial tumors. *WIRE Sys Biol Med* **4**, 51–78 (2011)
128. Wang, L.D., Wagers, A.J.: Dynamic niches in the origination and differentiation of haematopoietic stem cells. *Nat Rev Mol Cell Biol* **12**, 643–655 (2011)
129. West, S.A., Griffin, A.S., Gardner, A.: Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *J Evol Biol* **20**, 415–432 (2007)
130. Xavier, J.B.: Social interaction in synthetic and natural microbial communities. *Mol Syst Biol* **7**, 483 (2011)
131. Zhang, L., Athale, C.A., Deisboeck, T.S.: Development of a three-dimensional multiscale agent-based tumor model: Simulating gene-protein interaction profiles, cell phenotypes and multicellular patterns in brain cancer. *J Theor Biol* **244**, 96–107 (2007)
132. Zhang, L., Wang, Z., Sagotsky, J.A., Deisboeck, T.S.: Multiscale agent-based cancer modeling. *J Math Biol* **58**, 545–559 (2009)

Growth as the Root of all Evil in Carcinomas: Synergy between pH Buffering and Anti-Angiogenesis Prevents Emergence of Hallmarks of Cancer

Ariosto Silva and Robert Gatenby

Abstract The term “cancer” is applied to a wide variety of illnesses that stem from uncontrolled growth of cells originating within the host. Cancer leads to death when the tumor burden is sufficiently large to interfere with the normal physiological functions. Each tumor is itself, a heterogeneous collection of habitats, predominantly due to variability in blood flow and cellular populations that exploit multiple different survival mechanisms and adaptive strategies. The resulting spatial and temporal heterogeneity typically foils cancer treatment, so that complete eradication of the cancer cells is only rarely observed and, indeed, may simply be impossible. All cancers, however, retain several conserved traits, or hallmarks, which represent convergent evolution of cellular strategies that are indispensable for malignancy. In this article we study the evolution of different strategies during solid tumor growth and invasion of the stroma. Using spatial computational models, we simulated the growth of phenotypically heterogeneous tumor populations within multiple microenvironmental habitats formed by variations in blood flow. The simulations demonstrate the dynamics by which hypoxia, pH and nutrient gradients formed during tumor growth, affect the choice among different cellular adaptive strategies: invasion/migration, proliferation, quiescence, and angiogenesis. We propose that the canonical hallmarks of cancer stem from the natural environmental pressures imposed by the tumor-host microenvironment during tumor growth, and that the use of treatments that remove or relieve these selective pressures, as opposed to exacerbating them or killing as many cancer cells as possible, may induce tumor dormancy, and lead to more benefits than aggressive therapy in cases of currently incurable cancers.

A. Silva · R. Gatenby (✉)

H. Lee Moffitt Cancer Center and Research Institute, 12902 Magnolia Dr., Tampa, FL 33612, USA
e-mail: ariosto.silva@moffitt.org, robert.gatenby@moffitt.org

1 Introduction

Hanahan and Weinberg [13] proposed six basic hallmarks of cancer: (1) self-sufficiency in growth signals, (2) insensitivity to anti-growth signals, (3) evasion of apoptosis, (4) limitless replicative potential, (5) sustained angiogenesis, and (6) tissue invasion and metastasis. Later, four more were added, including notably the characteristic increased glycolytic metabolism in cancers, also known as the Warburg effect [14].

One purpose of this work is identification of general phenotypic cancer traits that can serve as possible therapeutic targets. Here we focus on understanding how these traits are fixed in the tumor population during the carcinogenic process, and how possible manipulations of the tumor microenvironment could nullify this selective advantage. As Theodosius Dobzhansky famously said, “nothing makes sense in biology except in the light of evolution” [5], we propose that, similarly, it is essential to understand the evolutionary forces behind the hallmarks of cancer.

Gatenby and collaborators [10, 12, 24] have proposed a model for emergence of such hallmarks during evolution of carcinomas due to proliferation constraints that emerge during cellular proliferation on epithelial surfaces. In the early, *in situ* stages of the development of these tumors, nutrient and oxygen delivery are limited by reaction-diffusion through the basement membrane, and increasing super-imposed layers of cancer cells. As a result of the reaction-diffusion dynamics, regions of the *in situ* tumor that are most distant from the basement membrane become hypoxic. This requires a shift to anaerobic (glycolytic metabolism) which both increases acid production and decreases oxygen utilization. The former results in regional acidosis, which is cytotoxic, and selects for acid-resistant phenotypes. The latter may result in local cycles with varying oxygen concentrations. We propose that eventually cells that had increased capacity for anaerobic metabolism would be selected by this microenvironment, and the Warburg Effect would thus become fixed in the tumor.

We here propose that all adaptations in cancer, prior to therapy, should arise from normal tissue proliferation constraints and the negative consequences of uncontrolled growth, which inevitably lead to death due to low extracellular pH, reduced availability of oxygen, nutrients and growth factors. Growth and competition under these harsh conditions results in Darwinian natural selection, and combined with cancer’s inherent genetic instability, of evolution.

Thus, insensitivity to anti-growth signals from the host emerge early in carcinogenesis to allow cells to proliferate while not in contact with the basement membrane (defeating anoikis governed by cell-basement membrane signaling) and exploiting the large potential, unoccupied space above the normal epithelial layer. Proliferation might also be increased by permitting greater crowding through elimination of cell-cell inhibitor signaling.

Two other hallmarks, resistance to apoptosis and independence of host growth factors, would be next selected in tumor cells growing in regions further away from the host tissue and vasculature. That is, proliferating cells require growth promoting signals often present as ligands (i.e. EGF, estrogen, etc.) in the blood. As the cells

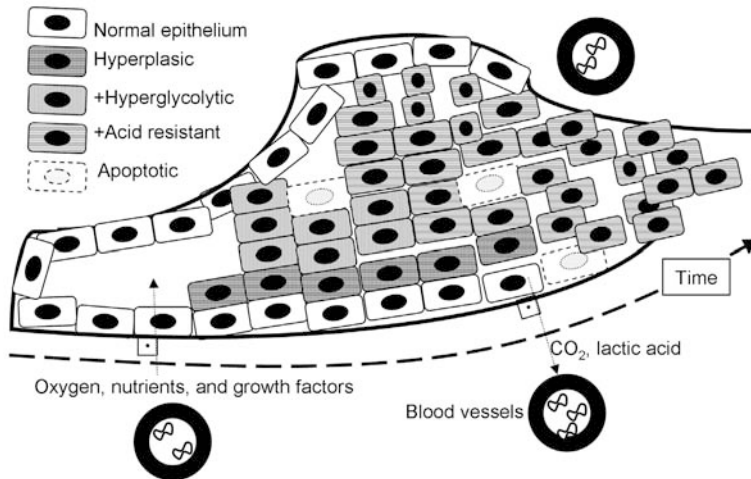


Fig. 1. Environmental barriers mediate the transition from in situ to invasive carcinoma. Early carcinogenesis proceeds from normal tissues through initiation to a hyperplastic state to interstitial neoplasia, progressing to carcinoma in situ. Until this stage, epithelial cancers are avascular. Following breakdown of the basement membrane, cells gain access to existing and newly formed blood and lymphatic vascular routes for metastasis. The stages of tumor growth and their associated physiological states are diagrammed, showing that progression from one stage to the next is governed by state processes. Normal epithelial cells (white) become hyper proliferative (cross lines) following induction. As they reach the oxygen diffusion limit, they become hypoxic, which can either lead to cell death (dashed contours) or adaptation of a glycolytic phenotype (vertical lines), which allows cells to survive. As a consequence of glycolysis, lesions become acidic, which selects for acid resistant cells (horizontal lines) that eventually breach the basement membrane [10]

move further from the membrane they also move further from the blood vessels, which remain on the opposite side of the membrane (Fig. 1). As a result, growth factor concentrations will decline in those cells furthest from the membrane. This may be sufficient to induce apoptosis. However, these cells, because of their location, have the greatest opportunity to proliferate. Thus, these populations will experience strong selection pressure to develop strategies that both defeat apoptosis (permitting survival) and allow independence from host growth factors (permitting proliferation).

The unlimited replicative potential hallmark is perhaps the most straightforward strategy. Cancer cells that are not capable of replicating due to senescence will be overcome by competitors. Thus, there will be strong selection pressures to upregulate telomerase or for other adaptive strategies to overcome the Hayflick limit.

The angiogenic hallmark is considered a requirement for tumors to grow beyond a volume limited by the reaction-diffusion of oxygen, nutrients and growth factors from the well vascularized surrounding tissue. While most authors propose an angiogenic switch, where cancer cells constitutively secrete angiogenic cytokines, we have previously proposed the increase in angiogenic potential of tumors to be a result of an increased number of hypoxia-adapted cells in the hypoxic regions of

tumors [22]. The main difference between the two mechanisms is that the first is an altruistic one, where all cells in a neighborhood (angiogenic and non-angiogenic alike) would benefit from the activity of angiogenic cells, while the second results from the natural selection for survival in a hypoxic environment. Our simulations in a non-spatial model pointed in the direction of hypoxia adaptation, and in this work we revisited the emergence of this hallmark in a bi-dimensional spatial model.

The final of the original six cancer hallmarks is invasive and metastatic potential. While it is not generally appreciated, the invasive hallmark adds an explicitly spatial component to the hallmarks. That is, when viewed from an evolutionary perspective, invasion can arise as an adaptive trait only if it promotes proliferation of those cells that invade. In other words, the invasive phenotype must be rewarded by increased proliferation in its new location. This is most obvious in the invasion of cancer cells into normal adjacent tissue where resources are abundant and space is available. Thus, the invasive phenotype can only evolve and be maintained as a viable population at the tumor-host interface. This will tend to enforce a specific geographic heterogeneity in a malignant tumor - the edge should contain invasive populations while the core will contain only non-invasive cells that can proliferate only as a result of angiogenesis (above). Interestingly, such spatial phenotype variability is observed in invasive species. For example, the cane toads at the edge of their invading population in Australia possess larger legs and greater leaping range when compared to their counterparts in the core of the population [20].

In a previous work we used spatial computational models, and *in vitro* experiments with human pancreatic cell lines, to determine if tumor growth by itself was a sufficient evolutionary force, selecting for more metastatic cells [18]. The original hypothesis was that the edge of a growing tumor is the ideal place for proliferation, and that even cells with increased proliferative advantage, but low invasive potential, would eventually be left behind the invasive edge, and be trapped in the tumor's interior, in a possibly increasingly hypoxic region.

Both the computational simulations and the *in vitro* experiments reached the same conclusion, the cancer cells in the invading edge of a tumor are both more motile (higher migration directionality) and proliferative (shorter cell cycle) than the original population. This result was an indication that the invasive hallmark of cancer is also the byproduct of evolutionary forces acting during the uncontrolled tumor growth.

The Warburg effect is a well-known characteristic that consists in an increased glycolytic metabolism of cancer cells when compared to surrounding healthy tissue. When Otto Warburg first identified this increased production of lactic acid by cancer cells in aerobic conditions, he believed that cancer would be consequence of mitochondrial defects [10]. This metabolic feature has been proposed to be another adaptation that would allow cancer cells to maintain their energetic metabolism even in hypoxic tumor regions. The lactic acid produced by this glycolytic metabolism would also select for low pH resistant cancer cells, facilitate the degradation of the extracellular matrix and death of stromal cells in the tumor-host interface, consisting in the Acid Mediated Tumor Invasion hypothesis [9].

In this article we used spatial computational models to simulate how the tumor invasion and angiogenic hallmarks naturally emerge from tumor growth. These two hallmarks can also be translated into two diverging strategies used by human societies: some groups maximize the local carrying capacity by expending available resources to build infrastructure such as villages, roads, and irrigation. Other groups eschew permanent dwelling and crops and are instead nomadic - investing resources energy in mobility as they move from place to place using up resources in each location prior to moving. At its most extreme this strategy manifests as invasive populations (such as the army of Genghis Kahn) which thrives by invading and plundering adjacent civilizations. We also study how to use anti-angiogenic and pH buffering therapies to prevent both strategies, and induce tumor dormancy.

2 Materials and Methods

The computational model used in this work is a bi-dimensional agent-based model, also referred as a hybrid cellular automaton. This is an adaptation of our previous work with modeling the competition of replicative and motile pancreatic cell lines [18]. The main difference between the two models is that, in the current work, the tumor is surrounded by extracellular matrix, which in turn is permeated by blood vessels. Thus, tumor growth is not limited only by nutrients, space, and cell cycle, but also by the capacity of the tumor cells to digest and invade extracellular matrix.

In each simulation, a tumor is seeded as a heterogeneous population of cancer cells embedded in extracellular matrix interspersed with blood vessels. Each cell in the seeded tumor differs in terms of the phenotypes described below:

- motility, represented by the variable p_M , it is the probability that a cell will attempt to move in the current time step. Motion will only occur if there is available space in the cell's vicinity;
- proliferation, represented by $pRep$, which is the probability that a cell will attempt a replication in the current time step. Replication will only occur if there is available space in the cell's vicinity, if the cell is not quiescent, and there are no energetic restrictions preventing proliferation;
- minimum energy requirements for replication, atp_{MIN} , the minimum ATP production rate that ensures 100% of chance of cell replication at the end of cell cycle, all other conditions being satisfied;
- hypoxia threshold for VEGF secretion, pO_{2VEGF} , is the maximum partial oxygen pressure that incites cancer cells to secrete angiogenic factors;
- glycolytic metabolism, k_{GLU} , represents the rate at which cancer cells uptake and metabolize glucose. In hypoxia, high glycolytic metabolism leads to acidification of the extracellular environment, by lactic acid production, and accelerated digestion of the extracellular matrix [23];
- acid resistance, which are the low extracellular pH levels that induce cell quiescence and death [23].

Each simulation step was composed of four levels: the first consists in cells dying or replicating at every generation (24h interval); each replication or cell death decision is taken based on extracellular concentration of species, intracellular ATP production rate and the proliferation value. The extracellular concentration of species depends on the equilibrium between the production/consumption of these species (O_2 , glucose, bicarbonate anions) and the diffusion of these species from blood where they originate, to the cells where they are consumed and converted into H^+ and CO_2 .

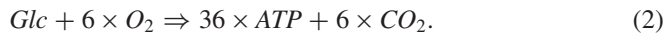
The second level of simulation was computed every two hours, when every cell was tested for motility. Cell motility was implemented as a random walk with each cell being able to randomly move to any neighboring empty slot every two hours with a probability equal to its motility value. The third step consisted in glucose metabolism, and the fourth diffusion of species in media. The diffusion of species was calculated numerically using the heuristic described in [23]. Every replication step (24h-long) consisted in 12 motility steps (2h-long), each containing 50 metabolic steps (50s-long), which finally encompassed 10 diffusion steps (0.1s-long). The steady-state for metabolism was calculated in a transient of 50 steps, each one corresponding to 1 second because our preliminary simulations showed that the transients due to cell motion, death and replication reached steady-state in around 20 seconds.

Cells could proliferate, die or remain quiescent depending on their ATP production rate or extracellular pH (pHe). Should the pHe value be below the quiescence threshold, the tumor cells would become quiescent and would not replicate until the next cell cycle, if energy and pHe restrictions were satisfied. Tumor cells would die when exposed to pHe below the death threshold. Both thresholds were determined by the value of the “acid resistance” phenotype.

At every generation the probability of a cell replication increased if its ATP production was above a minimum level of 0.85uM/s [23] reaching a maximum at 8.5uM/s. ATP production rates lower than the minimum value lead to cell death due to starvation as described in Eq. (1):

$$pRep = \begin{cases} \frac{(ATP - 0.1 \times atp_{MIN})}{0.9 \times atp_{MIN}} & \text{if } 0.1 \times atp_{MIN} \leq ATP \leq atp_{MIN} \\ 0 & \text{if } ATP \leq 0.1 \times atp_{MIN} \\ 1 & \text{if } atp_{MIN} \leq ATP. \end{cases} \quad (1)$$

Glucose aerobic metabolism in cytoplasm and mitochondria:



Glucose anaerobic metabolism in cytoplasm:



In hypoxic conditions (excess of glucose):

$$ATP_{ANAER} = 2 \times \left(k_G \times [Glu]_e - k_{O2} \times \frac{[O2]_e}{6} \right), \quad (4)$$

$$ATP_{AER} = 36 \times \left(k_{O2} \times \frac{[O2]_e}{6} \right). \quad (5)$$

In hyperoxic or hypoglycemic conditions (excess of O2):

$$ATP_{AER} = 36 \times k_G \times [Glu]_e, \quad (6)$$

$$ATP_{ANAER} = 0. \quad (7)$$

$[Glu]_e$ and $[O2]_e$ are the extra-cellular concentrations of glucose and oxygen, and k_G and k_{O2} are their cellular uptake rates, respectively. While k_{O2} is the same for all cells in the tumor, as it reflects the passive diffusion of oxygen across the cell membrane, k_G is higher in hyper-glycolytic cells, and represents the Warburg effect. ATPANAER and ATPAER correspond to the production rate of ATP anaerobically and aerobically, respectively, and their sum is the total ATP production rate. The production rate, and not the intracellular ATP concentration, is used because the model assumes equilibrium is reached long before one generation is complete. Thus, the total ATP in a cell, given a fixed time between generations in steady state, is a function of the production rate. The probability of division is thus based on the ATP production rate at equilibrium while the probability of cell death is decided at every metabolic step: if cell ATP production at a given moment is below the minimum necessary for survival, the cell dies. The entire algorithm of the simulation as well as the numeric values of all parameters, are listed in the Supplemental Table 1 (http://www.i-genics.com/art_1/SuppTable1.pdf).

3 Results

In this work we studied how the natural growth of a phenotypically heterogeneous tumor, within the well-vascularized host tissue, may select for particular traits, such as the well-known hallmarks of cancer. We also hypothesized that by interfering in the evolutionary forces that select for these hallmarks, rather than by attempting to kill as many tumor cells as possible, it would be possible to prevent their emergence and freeze tumors in a less malignant state. This approach would be particularly interesting in patients with surgically removed primary tumors, but at risk of still carrying undetectable metastasis.

3.1 *Growth as the Root of all Evil*

Our simulations of tumor growth showed a consistent pattern: as the tumor population grew, there was an increase in the intra- and peri-tumoral hypoxia and acidification, due to the unbalance between oxygen consumption and lactic acid produced by

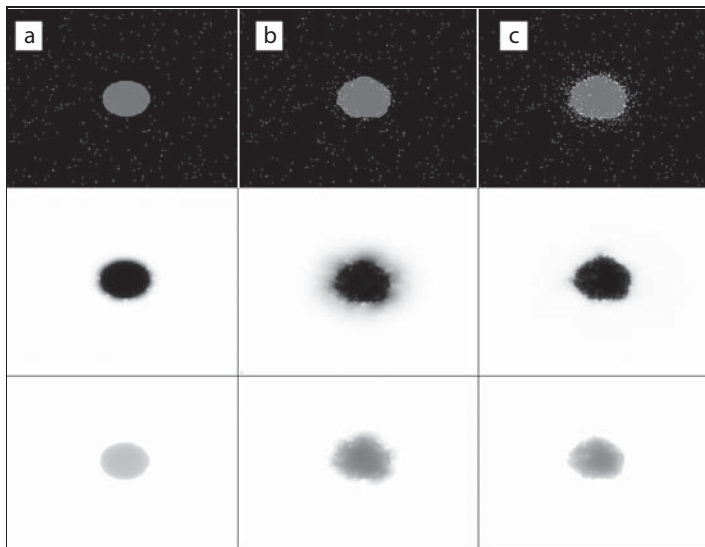


Fig. 2. Evolution driven by the growth of an untreated tumor. On the top, blood vessels and tumor cells are represented by white and grey dots, respectively. On the center and the bottom rows, the oxygen and pH levels are represented by a gray scale ranging from normal physiological in white, to anoxia and low pH in black, respectively. The original tumor (a) grew alternating intervals of hypoxia and invasion (b) with angiogenesis (c)

the growing tumor, and the capacity of the tumor-host interface to provide oxygen and mobile pH buffers (Fig. 2a). The chronic low pH and hypoxic conditions then led to cell death and quiescence within the tumor, and in the tumor-host interface, selecting for acid resistant tumor cells, and facilitating the invasion of the tumor cells into the host tissue (Fig. 2b).

In parallel, the angiogenic factors produced by the surviving acid resistant cancer cells induced an increase in the microvessel density around the tumor edge, which in turn reduce the tumor acidity and hypoxia. This tumor continued its growth at a steady rate alternating destruction of blood vessels and extracellular matrix, with the recruitment of new vasculature (Fig. 2c).

As observed in Fig. 5a, acid resistance was the only trait selected in this population, while motility and the hypoxia angiogenic threshold remained essentially unchanged.

3.2 Anti-Angiogenesis and Invasion

Emergence of angiogenesis is considered one of the most important barriers a solid tumor must overcome, before it is able to grow beyond a volume of a few cubic millimeters. Anti-angiogenic therapies, however, have faced the dilemma of actually increasing tumor invasion despite the evidence of reduced peri-tumoral vascularization and edema [4, 25]. While tumor vessels have been proposed to be leaky

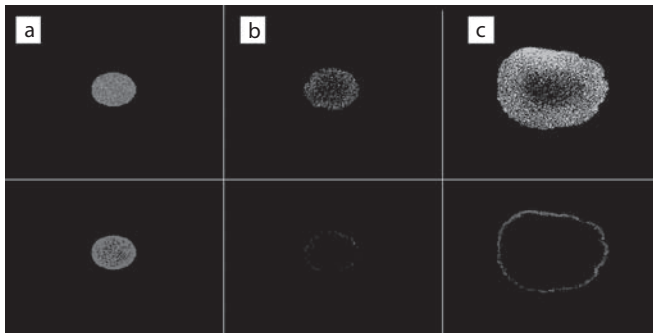


Fig. 3. Anti-angiogenic therapy exacerbates acid mediated tumor invasion. On top, the acid resistance of tumor cells are depicted in a gray scale ranging from white (quiescence at pH lower than 6.6 and death at pH lower than 6.4) to black (quiescence at pH lower than 7.1 and death at pH lower than 6.8). On the bottom row, the proliferative cells (non quiescent) are represented in gray, the quiescent cells are not visible. Tumors treated with anti-angiogenic therapy evolve towards acid resistance, grow mostly by invasion rather than replication, and have a major percentage of quiescent cells

and chaotic, and that anti-angiogenesis might indeed increase blood flow and oxygenation of the tumor [3], here we propose that anti-angiogenic therapy actually exacerbates tumor hypoxia, acidification, and increases the selection for motile acid-resistant cells, resulting in a more aggressive tumor population.

To test this hypothesis, we repeated the original simulations, but with cancer cells incapable of inducing angiogenesis. The simulations showed that these tumors evolve much like the ones without treatment, but in absence of angiogenesis, there is a stronger selection for acid resistance and motility (Fig. 3). These tumors also showed increased necrotic core and larger number of quiescent cells.

This was also the scenario where there was the largest variation of proliferative potential (Fig. 7), what can be explained by the small number of non-quiescent cells in the tumor, what makes the duration of the cell cycles less important. These results could be plainly explained by the interpretation that the harsher the environment, the bigger the payoff of the strategy of moving out of the environment, in contrast of the strategy of settling down in it (replication).

3.3 pH Buffering

In previous works, Gatenby and Gillies suggested that increases in the blood pH buffering capacity could be a strategy used to reduce tumor growth. While animal experiments have not shown changes in the growth of primary tumors, the chronic administration of sodium bicarbonate had a significant effect in reducing the formation of metastases [21], indicating that acid mediated tumor invasion may have an important role in the early stages of the growth of smaller tumors.

We simulated the effect of an increase of 40% in the concentration of sodium bicarbonate in blood, which is similar to the increase obtained in animal models [23].

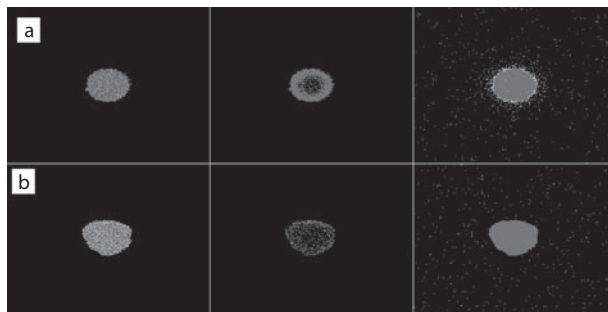


Fig. 4. pH buffering reduces invasion of micro-metastases. On the top (a), a tumor treated with pH buffering therapy shows no selection of acid resistance (left), most of its cells are not quiescent (in gray, center), and show increased microvessel density due to hypoxia-induced angiogenesis. Treating the same tumor with combined anti-angiogenic and pH buffering therapies (b) reduces the effect of the pH buffering in inhibiting invasion and selection of acid resistance, but prevents the increase in the vascularization around and within the tumor, thus reducing the risk of metastases

The original tumor was again simulated and the results showed that there was no cell death induced by low pH, and that the hypoxic cells were capable of partially reducing hypoxia by inducing angiogenesis. The cells in these tumors showed low motility, low acid resistance, and the tumor-host interface was preserved (Fig. 4).

3.4 pH Buffering and Anti-Angiogenesis

The small but well-vascularized tumor treated with sodium bicarbonate would however continue to grow, although more slowly than if left untreated, and the new vascularization would be a risk for metastases. Thus, we simulated the combination of an anti-angiogenic treatment with the pH buffering: the first aiming to reduce the blood supply and nutrients that would otherwise fuel tumor growth, and the second restoring intra-tumoral pH aiming to avoid the selection of acid-resistant cells, and thus inducing tumor cell quiescence and preservation of the tumor host interface.

The simulation of the combined treatments caused tumors to grow more slowly than anti-angiogenic treatment alone, and there was no increase in the motility or acid resistance of the tumor cells, indicating that the combined treatment might be an alternative for failed anti-angiogenic therapy alone (Fig. 5d).

4 Discussion

Figueredo and collaborators [8] have proposed that cancers adopt general adaptive strategies similar to animal species confronted with unpredictable environments: either dedicate more time and resources to raise a small number of offspring, or to breed a large number of offspring and dedicate a minimum of resources to each. The first strategy leads to populations that grow slowly but steadily, with a small

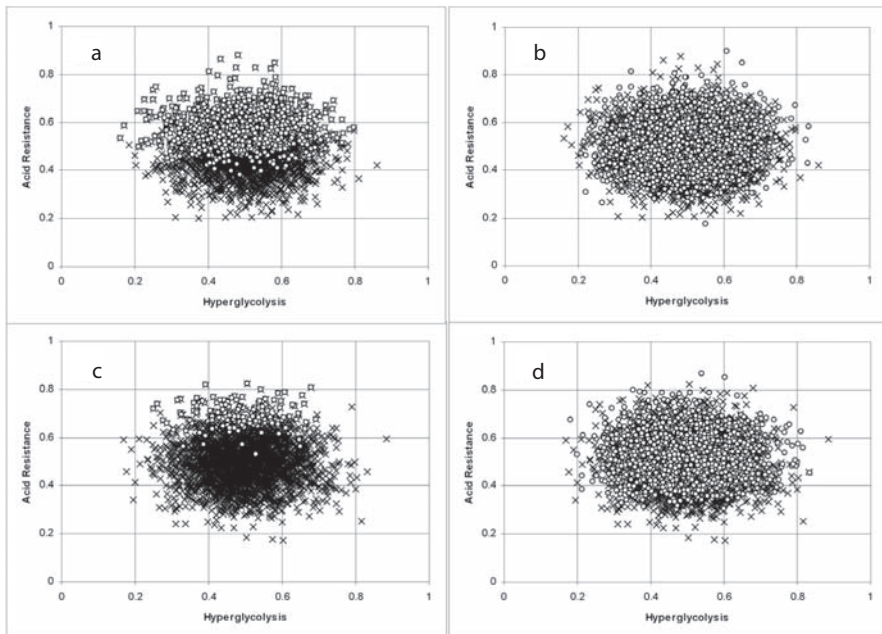


Fig. 5. Evolution of acid resistance and hyperglycolysis in growing tumors. The original tumor used for the simulation of the four scenarios untreated (a), pH buffered (b), anti-angiogenic therapy (c), and combination of pH buffering and anti-angiogenic therapy (d) already had high values of hyperglycolysis, so there was no selection towards hyperglycolysis, and death was only induced by low pH. In both scenarios where pH buffering was used (b and d), there was no selection for acid resistance, however the untreated tumor (a), and more evidently in the anti-angiogenic treatment scenario (c) there was a significant increase in the acid resistance of the tumor population

mortality in infancy, while the second consists of fast growing populations in periods of abundance, and by massive death rates in harsher conditions. In other words, while slow life history species control their number by a small birth rate and a long delay between birth and maturity, the fast life history populations reach equilibrium by a high death rate.

All the tumors simulated in this work started from the same heterogeneous population, composed of cells with random phenotypic values for energy metabolism, motility, acid resistance, and angiogenesis. The evolutionary forces imposed by the tumor microenvironment, however, drove these tumors into different directions by rewarding more aggressive or more conservative strategies. When compared with an untreated tumor, for instance, tumors treated with anti-angiogenic agents adopt a fast life history form of growth, with a higher number of dead and quiescent cells, but a more invasive behavior.

The use of pH buffering therapy, which reduces the stress on the cancer cells, led the tumors to slower life history approaches, with these cells growing more slowly,

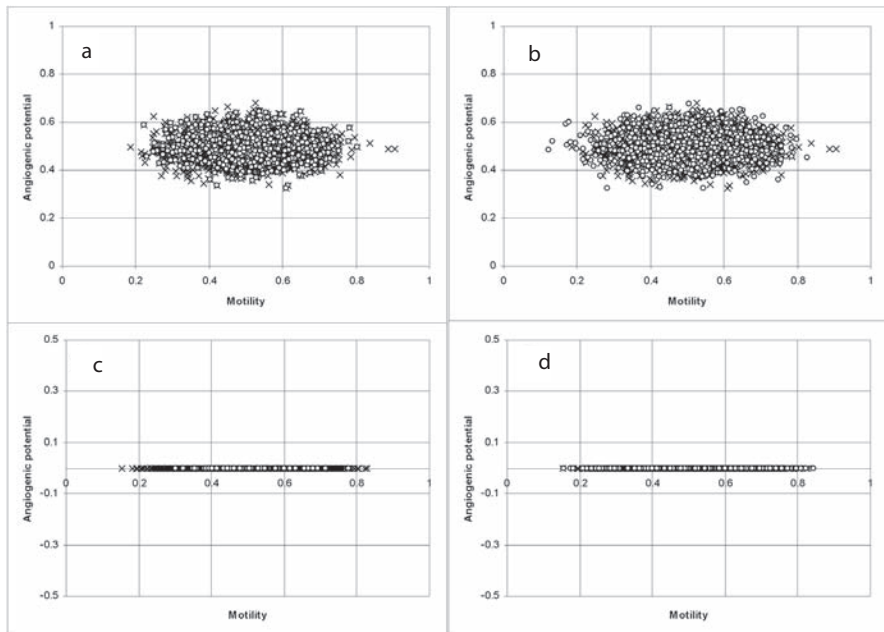


Fig. 6. Motility versus angiogenesis: invest or migrate? None of the four scenarios showed increase in the angiogenic potential of cancer cells, indicating that the hypoxic threshold for their production of angiogenic factors was unchanged. There was a significant increase in the motility of cancer cells of tumors treated with anti-angiogenic therapy (c), as is observed in brain tumor patients treated with anti-angiogenic drugs

invading less, and investing in further improving their microenvironment through angiogenesis.

It is important to notice, however, that in none of the simulations here studied were cells with constitutive angiogenic potential selected by the microenvironment (Fig. 6), which reinforces the previous results, that adaptation to survival in hypoxia and acidic conditions is sufficient for the explanation of the angiogenic switch by evolutionary forces alone.

The simulations of untreated tumors clearly depicted that there are at least two strategies adopted by the tumor cells as a response to their micro-environment, the first consists in the secretion of angiogenic factors by the tumor cells in hypoxic regions of the tumor, and the second is the invasion of extracellular matrix and replication. It is important to notice that these strategies are not “chosen” by the tumor cells, but they are rather a response to the intra-tumoral harsh conditions initiated by the tumor growth itself, leading to a vicious cycle of invasion-angiogenesis-growth.

Tumors are known to have narrow rims of normoxia in the tumor-host interface [15], and thus most of the cancers cells in the tumor are both hypoxic and quiescent, which minimizes any advantages of constitutive angiogenesis and high proliferative potential.

These results suggest that the treatment of cancer, a heterogeneous population of cells armed with genetic strategies representing billions of years of ingenious survival and adaptation mechanisms entrenched among trillions of other cells from the host, can be a futile exercise, if the only strategy used is the total eradication using aggressive treatments, even with very specific targeted drugs. The realization that cancer progression is the result of the interaction of the tumor and its microenvironment, and that all the hallmarks observed are consequence of the adaptation to these environmental barriers, makes it possible to imagine that malignancy could be avoided simply by removing these barriers (thus removing natural selection) while physically confining the tumor population to avoid uncontrolled growth and metastases.

This conservative strategy amounts to the conversion of the tumor population from a fast life history to a slow life history strategy, much similar to the Adaptive Therapy approach proposed by Gatenby and colleagues [11], which similarly recommended the use of the minimum amount of therapy necessary to prevent tumor growth instead of the maximum tolerated dose. The rationale being that the preservation of the drug sensitive cells would be a break that would slow down the growth of the drug resistance population.

In both cases the idea is to remove, or at least reduce as much as possible, the interference of evolution and thus transform tumors in benign quiescent masses.

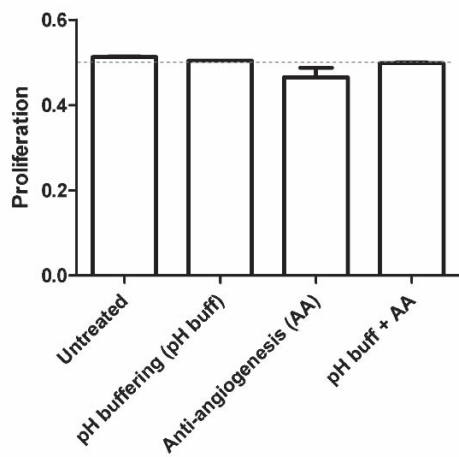


Fig. 7. Evolutionary fitness: replication rate or less death? The proliferative potential depicted above is the probability that a cell will replicate at the end of a 24h interval, given that energy and pHe requirements are met, and thus is inversely correlated to the duration of the cell's cycle. In the untreated tumor, there is a small increase in this value when compared to the original tumor (horizontal dashed line) but no changes are observable in the pH buffered, slow growing tumors (pH Buffering, pH buff + AA). There is an increased variation in the proliferative potential in the tumor treated with anti-angiogenic therapy, mainly due to the narrow rim of non-quiescent cells (Fig. 2), which puts replication in a lower priority when compared to motility and acid resistance

The conclusions obtainable from the computational model used in this work are obviously limited by the simplifications inherent to it. For instance, in this model cell replication was limited only by availability of metabolic substrates and pH, as well as “free space” for growth. In reality, cell-cell interactions inhibit replication in cell lines at different degrees [19], which is probably also the case in cancer *in vivo*, where interstitial fluid pressure leads to compression, and thus reduced volume of tumor cells. Cell growth is also dependent on growth factors produced by the stroma, which must diffuse through layers of avascular tumors before reaching the inner layers of the tumor. We believe that these two limitations, however, do not compromise the conclusions drawn from these simulations because: (A) while actual cells are elastic and may deform under pressure, allowing for regions of higher cell density and prolonged replication, these cells eventually will reach a maximum density that will induce quiescence. Computational models that have modeled the mechanic process of cell division and its importance for solid tumors have shown that these tumors have an initial exponential growth phase that later is replaced by a more linear growth [6, 7]. This is what is observed in this work’s model, with only the cells in the tumor-host interface being able to replicate, thus reinforcing the natural selection for not only proliferation but also for being in the tumor-host interface. Thus highly proliferative but poorly motile cells will become trapped in both models, even though ours underestimates the density at which this happens. (B) This model implementation does not account for stroma-produced growth factors, however, the diffusion of these species is ruled by the same dynamics as the diffusion of oxygen and glucose, and thus the cells in the inner core of the tumor would still become quiescent or die with a probability that increases with their distance to the tumor-host interface.

Our simulation results also agree with prior studies that have used evolutionary game theory to determine the outcome of a game played by cancer cells with different degrees of proliferative and motile strategies. Basanta et al. [1, 2] have shown that motility is rewarded when cells are in a harsh environment (closer to the tumor inner core) while proliferation is the best strategy when cells are closer to the tumor-host interface, and that cells using either strategy alternate during tumor growth. In the current model, cells are randomly seeded with different values for both phenotypes/strategies, and we observe that our model predicts the selection of cells capable of implementing both.

The results from this work thus propose that the combination of anti-angiogenic therapy with pH buffers, such as sodium bicarbonate, IEPA [16], or Lysine free base [17], would be synergistic and prevent development of metastases in cancer patients.

References

1. Basanta, D., Hatzikirou, H., Deutsch, A.: Studying the emergence of invasiveness in tumours using game theory. *Eur. Phys. J. B* **63**, 393–397 (2008)
2. Basanta, D., Simon, M., Hatzikirou, H., Deutsch, A.: Evolutionary game theory elucidates the role of glycolysis in glioma progression and invasion. *Cell Prolif.* **41**, 980–987 (2008)
3. Carmeliet, P., Jain, R.K. Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases. *Nat. Rev. Drug Discov.* **10**, 417–427 (2011)
4. de Groot, J.F., Fuller, G., Kumar, A.J., Piao, Y., Eterovic, K., Ji, Y., Conrad, C.A.: Tumor invasion after treatment of glioblastoma with bevacizumab: radiographic and pathologic correlation in humans and mice. *Neuro-Oncol.* **12**, 233–242 (2010)
5. Dobzhansky, T.: Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* **35**, 125–129 (1973)
6. Drasdo, D., Hoehme, S.: Individual-based approaches to birth and death in avascular tumors. *Math. Comput. Modelling* **37**, 1163–1175 (2003)
7. Drasdo, D., Hoehme, S.: A single-cell-based model of tumor growth in vitro: monolayers and spheroids. *Phys. Biol.* **2**, 133–147 (2005)
8. Ellis, B.J., Figueredo, A.J., Brumbach, B.H., Schlomer, G.L.: The impact of harsh versus unpredictable environments on the evolution and development of life history strategies. *Hum. Nat.* **20**, 204–268 (2009)
9. Gatenby, R.A., Gawlinski, E.T., Gmitro, A.F., Kaylor, B., Gillies, R.J.: Acid-mediated tumor invasion: a multidisciplinary study. *Cancer Res.* **66**, 5216–5223 (2006)
10. Gatenby, R.A., Gillies, R.J. Why do cancers have high aerobic glycolysis? *Nat. Rev. Cancer* **4**, 891–899 (2004)
11. Gatenby, R.A., Silva, A.S., Gillies, R.J., Frieden, B.R. Adaptive therapy. *Cancer Res.* **69**, 4894–4903 (2009)
12. Gatenby, R.A., Smallbone, K., Maini, P.K., Rose, F., Averill, J., Nagle, R.B., Worrall, L., Gillies, R.J.: Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer. *Br. J. Cancer* **97**, 646–653 (2007)
13. Hanahan, D., Weinberg, R.A.: The hallmarks of cancer. *Cell* **100**, 57–70 (2000)
14. Hanahan, D., Weinberg, R.A.: Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011)
15. Helminger, G., Yuan, F., Dellian, M., Jain, R.K.: Interstitial pH and pO₂ gradients in solid tumors in vivo: high-resolution measurements reveal a lack of correlation. *Nat. Med.* **3**, 177–182 (1997)
16. Ibrahim-Hashim, A., Cornnell, H.H., Coelho Ribeiro, M.L., Abrahams, D., Cunningham, J., Lloyd, M., Martinez, G.V., Gatenby, R.A., Gillies, R.J.: Reduction of metastasis using a non-volatile buffer. *Clin. Exp. Metastasis* **28**, 841–849 (2011)
17. Ibrahim-Hashim, A., Wojtkowiak, J., Ribeiro, M., Estrella, V., Bailey, K., Cornnell, H., Gatenby, R., Gillies, R.: Free base lysine increases survival and reduces metastasis in prostate cancer model. *J. Cancer Sci. Ther.* **S1-004** (2011)
18. Lee, H.O., Silva, A.S., Concilio, S., Li, Y.S., Slifker, M., Gatenby, R.A., Cheng, J.D.: Evolution of tumor invasiveness: the adaptive tumor microenvironment landscape model. *Cancer Res.* **71**, 6327–6337 (2011)
19. Liu, L., Sun, B., Pedersen, J.N., Aw Yong, K.M., Getzenberg, R.H., Stone, H.A., Austin, R.H.: Probing the invasiveness of prostate cancer cells in a 3D microfabricated landscape. *Proc. Natl. Acad. Sci. USA* **108**, 6853–6856 (2011)
20. Phillips, B.L., Brown, G.P., Webb, J.K., Shine, R.: Invasion and the evolution of speed in toads. *Nature* **439**, 803 (2006)

21. Robey, I.F., Baggett, B.K., Kirkpatrick, N.D., Roe, D.J., Dosescu, J., Sloane, B.F., Hashim, A.I., Morse, D.L., Raghunand, N., Gatenby, R.A., Gillies, R.J.: Bicarbonate increases tumor pH and inhibits spontaneous metastases. *Cancer Res.* **69**, 2260–2268 (2009)
22. Silva, A.S., Gatenby, R.A.: Adaptation to survival in germinal center is the initial step in onset of indolent stage of multiple myeloma. *Mol. Pharm.* **8** 2012–2020 (2011)
23. Silva, A.S., Yunes, J.A., Gillies, R.J., Gatenby, R.A.: The potential role of systemic buffers in reducing intratumoral extracellular pH and acid-mediated invasion. *Cancer Res.* **69**, 2677–2684 (2009)
24. Smallbone, K., Gatenby, R.A., Gillies, R.J., Maini, P.K., Gavaghan, D.J.: Metabolic changes during carcinogenesis: potential impact on invasiveness. *J. Theor. Biol.* **244** 703–713 (2007)
25. Zhang, W., Fulci, G., Buhrman, J.S., Stemmer-Rachamimov, A.O., Chen, J.W., Wojtkiewicz, G.R., Weissleder, R., Rabkin, S.D., Martuza, R.L.: Bevacizumab with angiostatin-armed oHSV increases antiangiogenesis and decreases bevacizumab-induced invasion in U87 glioma. *Mol. Ther.* **20**, 37–45 (2011)

Phase Transitions in Cancer

Ricard V. Solé

Abstract Cancer dynamics is sometimes associated to well-defined transitions between qualitative properties of neoplasms, or even to the shift from presence to regression. Cancer is largely understood as a Darwinian evolution experiment within organisms, characterized by a break of cooperation between cells. Some patterns of sudden change have been identified as phase transitions, similar to those known from the physics of phase changes. Two of them are analyzed here: (a) the shift from cancer to cancer-free tissues under the interaction with the immune system, and (b) the predicted existence of a threshold of instability associated to unstable tumorigenesis. The nature, evidence and implications of these transitions are discussed.

1 Cancer as a Complex System

Cancer development involves a number of dynamical events which are context-dependent (what tissue and group of cells is involved), age-related (time is relevant) but is also an evolutionary process [13, 18, 19, 29]. This is precisely the major difference between normal and cancerous tissues: within a normal cell population, cooperation is the rule, and appropriate negative control feedbacks guarantee a stable tissue architecture and homeostasis. Once some of these regulatory feedbacks breaks down, clonal expansion of malignant cells becomes possible and a neoplasm can develop. Sometimes, growth develops in a monotonous way, generating a mass of cells which keeps growing until some selection barrier stops it. If enough mutations allow it, selection barriers can be overcome and progression towards invasion and metastasis unfold. Each step towards tumorigenesis involves some changes. Each

R.V. Solé (✉)

ICREA-Complex Systems Lab, Universitat Pompeu Fabra, Dr Aiguader 88, 08003 Barcelona, Spain

Institut de Biologia Evolutiva, CSIC-UPF, Psg Barceloneta 37–49, 08003 Barcelona, Spain

Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA

e-mail: ricard.sole@upf.edu

change marks some qualitative transition affecting the future outcome of the progression process.

The other side of the coin here is the potential pathways from tumor presence to tumor absence. Tumors can be reduced or even eliminated if appropriate therapies are applied. In the most basic scheme, these therapies deal with cytotoxic drugs acting on cancer (but also normal) cells inhibiting cell growth or simply triggering cell death. Radiation and chemotherapy must be effective enough to remove all potentially proliferative anomalies, and success is typically marked by a wide removal of both normal and cancer cells. But tumor disappearance can also occur driven by endogenous causes, without treatment. This is the case, immune responses are able to properly identify abnormal cells and eliminate them. A dramatic illustration of this scenario is provided by so called *spontaneous regression*, a rare and enormously effective process able to fully eradicate advanced malignancies. Here we study two examples of such cancer disappearance process using the approach of the physics of complex systems. By looking at transitions in cancer as phase transitions between different phases we hope to gain insight into the nature of tumorigenesis and its potential Achille's heels.

2 Phase Transitions in Physics and Biology

Phase changes are well known in physics: boiling or freezing are just two examples of changes of phase where the nature of the basic components is not changed. In physics, phases and their transitions are usually associated to qualitative transitions between order and disorder (see [23] and references cited) associated to continuous changes in external parameters, such as temperature. As temperature grows, a piece of ice experiences a slow modification of its structural properties; atoms dance around the regular locations in the crystal lattice. But such a dance is abruptly interrupted as soon as we reach the critical melting temperature, when ice suddenly transforms into water. In this case, the change involves a shift from one phase (ice) to the other (liquid water). If we consider the water-steam transition, its nature is clearly different, since the two phases (water and steam) coexist through all the process.

Transitions can also occur, sometimes dramatically, between molecular patterns of gene activity within cells, behavioral patterns of collective exploration in ants or the success or failure of epidemics to propagate (see [23, 25]). Complex systems theory is deeply tied to the search for general principles associated to the presence of different qualitative behaviors and how can we shift from one pervading self-organized, emergent phenomena and defining its potential patterns.

In order to illustrate these ideas in our context, let us consider a classic problem in evolutionary dynamics, the so called *error catastrophe* problem [6–8, 21]. It was originally defined within the context of an abstract population of replicating molecules (or entities) generically called *replicators* competing for limiting resources and displaying mutations. More precisely, we consider a (large) population of strings (genomes or polymers) where each sequence can replicate at some rate. Replication rate will be sequence-dependent and the relation between sequence and

growth rate should be expected to be complex. Additionally, we assume that every time a chain replicates, mutations can occur at a given rate μ . Eigen predicted in 1971 that mutation imposes a limit to the amount of information (in terms of genome length, L) that is consistent with stable information. Specifically, it was shown that there is a critical mutation rate $\mu_c \approx 1/L$ beyond which no Darwinian evolution can occur. Instead, for $\mu > \mu_c$, a random drift would be observed. Below the threshold, selection operates and information can be maintained in stable ways.

The general problem is very difficult to tackle, but a simple scenario can be fully analyzed and provides an illustration to the phase change associated to the error catastrophe problem. It considers a further simplification of the quasispecies model. The master sequence (say 11111111) replicates at a given rate, but mutations also occur leading to new sequences, like 11101111 including one or several flips. Assuming that mutation rate is low enough, we want to know how large mutation can be in order move from a stable quasispecies to a random phase. The so called Swetina–Schuster model considers that the master sequence has a large replication rate, whereas all other chains, no matter how many zeros they have or where they are located, have all a smaller and identical replication rate. Moreover, it is assumed that mutations occur from the master to the rest, but not the other way around. This might seem a rather strong assumption, but we must notice that any of the nearest (one-mutant) neighbors of it will mutate back to the master sequence with an exponentially small probability¹.

If all non-master sequences can be considered equal, and we label as M and N the total population of sequences in the master and non-master classes, respectively, the equations describing their evolution are:

$$\frac{dM}{dt} = r_M(1 - \mu)M - M\phi(M, N), \quad (1)$$

$$\frac{dN}{dt} = \mu r_M M + r_N N - N\phi(M, N) \quad (2)$$

where $r_M(1 - \mu)M$ indicate the growth of M due to correct replication whereas $\mu r_M M$ is the “flow” from M to N associated to incorrect replication. The term $r_N N$ corresponds to the “error free” copy of N (under our hypotheses). Finally, the right-hand side of both equations introduces competition between both populations due to limited resources. In order to find $\phi(M, N)$ we will introduce a further simplification: let us assume that the total population is constant and, for convenience, normalized to $N + M = 1$. This defines the so called constant population constraint, and because of the normalization N and M become relative frequencies. We have then:

$$\frac{d(M + N)}{dt} = \frac{dM}{dt} + \frac{dN}{dt} = 0, \quad (3)$$

¹ Specifically, if we have a chain of length L involving a sequence with only one zero, and assuming that μ_b is the probability of mutation per unit and replication round, the probability of recovering the master is given by $P(N \rightarrow M) = \mu_b(1 - \mu_b)^{L-1}$. If μ_b is very small, we can approximate $(1 - \mu_b)^{L-1} \approx \exp(-L + 1) \approx \exp(-L)$. Since L is typically large, the exponential term makes the probability of back mutation extremely small.

and gives as a result:

$$\phi(M, N) = r_M M + r_N N \quad (4)$$

which is nothing but the average replication rate. Since $M = 1 - N$, there is only one effective variable. Using M as our key variable, we have now:

$$\frac{dM}{dt} = r_M(1 - \mu)M - M(r_M M + r_N(1 - M)) \quad (5)$$

$$= (r_M - r_N)M(1 - M) - \mu r_M M. \quad (6)$$

Two equilibrium points are consistent with this system $dM/dt = 0$, namely

$$M^* = 0, \quad (7)$$

$$M^* = 1 - \frac{\mu r_M}{r_M - r_N}. \quad (8)$$

The first point corresponds to the absence of the master sequence, which cannot be maintained anymore. These are actually our two phases. The second is associated to the presence of master sequences together with a “cloud” of non-master ones. It is easy to check that the second point exists if the mutation rate is below a critical value

$$\mu_c(r_N, r_M) = 1 - \frac{r_N}{r_M}. \quad (9)$$

The two phases can be displayed using a two-dimensional representation of our parameter space. Let us use the plane (μ, r_M) for our purposes, and fix $r_N = 1/4$. In Fig. 1a we display the phase space where the two possible domains are indicated as white and gray areas separated by a critical curve μ_c . The white domain is associated to the drift phase (M_0^*) whereas the second (M_1^*) defines the quasispecies phase.

In spite that the model is an oversimplification, it provides a clever picture of the error catastrophe. In Fig. 1b we display (for a fixed r_M) the frequency of master chains M against μ . Close to criticality, we can see a decrease in M , which drops to zero at μ_c . An important point to be made here is the irreversible character of the transition: once the extinction (drift) phase is reached, there is no way out from it². This is a so called *absorbing* phase transition, since one of the states indefinitely captures the dynamics once reached.

We can gather some additional insight into the nature of this transition by considering a discrete representation of the dynamics using a two-dimensional $L \times L$ lattice Ω . Imagine that we place our strings on the sites $(i, j) \in \Omega$ of our lattice (here $1 \leq i, j \leq L$). If we indicate as M and N the state of each site (here no empty sites are considered) the discrete dynamics of our system, defining a stochastic cellular automaton, allows to recover the previous mathematical results as well as some new

² This is in contrast with standard, so called equilibrium transitions, where we move from one phase to another by increasing the temperature, but the previous phase can be recovered by decreasing the temperature again, which cannot occur in our example by lowering mutation rates.

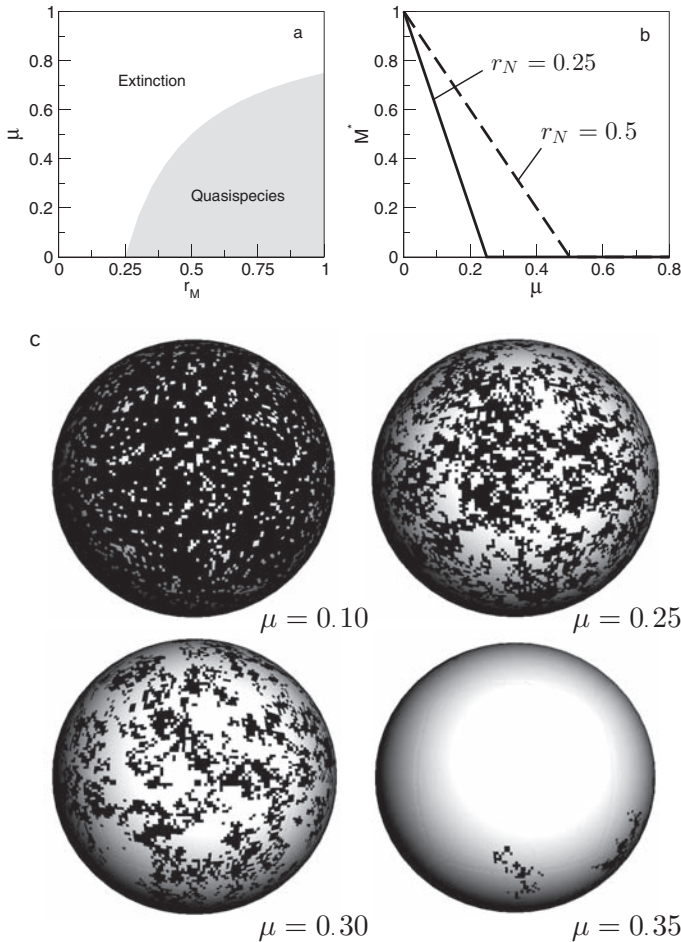


Fig. 1. The error catastrophe phase transition, illustrated by its simplest version, the so called Swetina–Schuster model (see text). In (a) we plot the two phases associated to this model, using mutation rate and r_M as key parameters (here we fix $r_N = 0.25$). These phases are separated by a critical line $\mu_c(r_M)$. The gray domain includes all pairs consistent with a stable master sequence population, whereas the white one indicates the parameter combinations where no master chain can survive. In (b) the abundance of master chains at steady state (M^*) is displayed against varying mutation rates. A transition is observed for $\mu_c = 0.25$. Several snapshots of the cellular automaton model are displayed in (c) for increasing levels of mutation, using $p_M = 1$ and $p_N = 0.5$. Although the mathematical model predicts extinction at $\mu = 0.5$ it occurs earlier due to spatial effects

properties. Let us assume that we choose at random one site on the lattice. We consider its state and decide whether or not it replicates. Let us assume that a M site, if chosen, will replicate. It does so by choosing one of its eight nearest neighbors as target and replace it by an M with probability $1 - \mu$ and by an N with probability μ . Similarly, if we find an N state, the same operation is made with probability p , with

no mutations. This can be summarized by means of a set of transitions including all the possible events introducing changes:



It is not difficult to see that the corresponding critical boundary for mutations is now $\mu_c = 1 - p/p_M$ and $M^* = 1 - \mu p_M/(p_M - p)$. In Fig. 1c we show several examples of the spatial dynamics of our cellular automaton for different μ values, after fixing $p_M = 1$ and $p_N = 0.5$. As we increase mutation rates, the amount of master sites is reduced and larger fluctuations in their spatial organization develop. At some threshold value, the black clusters of master sites no longer persist and all the lattice points are occupied by non-master sites.

We can appreciate that the model is very simple and yet very rich in dynamical complexity and – what is more important – makes a well defined prediction: There is an upper bound to the rate of disorder that an adaptive system can tolerate. A key discovery made much later revealed that RNA viruses replicate with very high mutation rates, as high as the error threshold allows. In other words, RNA viruses live at the error threshold [6, 21]. This observation tells us that, in order to adapt to a changing environment (the plastic immune system) high mutation rates compatible with preserving genetic information are reached. Such spontaneous evolution towards high instability levels will play a key role in our discussion below on unstable cancer.

3 Phase Changes in Unstable Tumors

The error threshold problem has been analyzed in some detail because it is our theoretical background for our next example. One widespread feature of most tumors is the presence of genetic instability, which is observable at different levels, from chromosomal aberrations to pathological tissue organization. The later is well illustrated in Fig. 2, where histological views of different stages of colon cancer are shown [12]. As we can clearly appreciate, a systematic loss of tissue coherence occurs as progression develops. From a clearly defined structure, we shift towards a more and more featureless organization. In all these cases, there is a failure in the maintenance of genome integrity associated to mutations, chromosomal rearrangements and gene losses [14, 29]. These changes are known to affect those mechanisms dealing with detecting and controlling potential mistakes in the process of DNA replication, associated to the emergence of the so called mutator phenotype [9, 15]. This is a self-propelling process, since increased levels of failure will favor further mutations in other genes associated to genome stability. As instability grows, so it does cellular heterogeneity, tissue disorder, plasticity and diversity. Clearly, there must be some

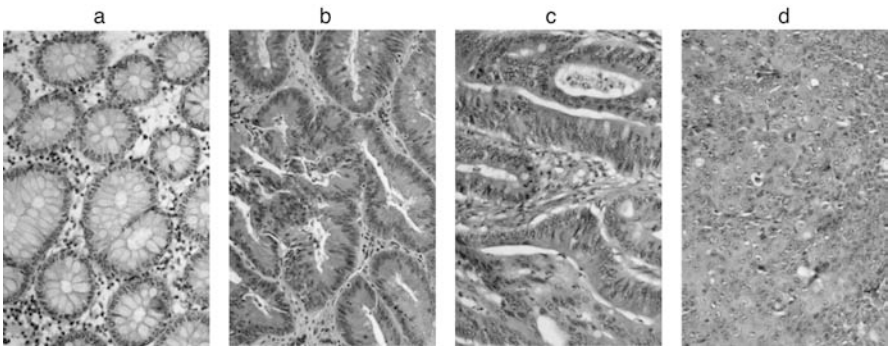


Fig. 2. Degradation of tissue organization at different stages of the evolution of a colon cancer. Here we have: (a) normal tissue (b) colon adenomatous polyps with low-grade, (c) dysplasia and (d) undifferentiated, invasive colon cancer with highly disordered morphology and loss of normal colon tissue organization. Modified from [12]; R.A. Gatenby and B.R. Frieden, *Cancer Res.* **62**, 3675–3684 (2002) © American Association for Cancer Research

advantage in having some degree of instability, which allows cancer populations to adapt to changing conditions and, if required, develop mechanisms to overcome selection barriers. But of course genetic instability has a dark side: increasing levels of mutation jeopardize cell survival, making cell death also more likely to occur.

The previous description reminds us some of the features we just described for RNA viruses. Some of these similarities are summarized in Table 1. In both cases, the system evolves towards increasing instability levels [21, 24]. As a consequence, they can rapidly change as they adapt, but in both systems most individuals (either viruses or cells) will be low-fit or even unviable. Most RNA particles are defective and unable to infect, and most cells from unstable cancers will not grow once inoculated in a culture medium. By taking these analogies a step further, we proposed the idea that cancer might also exhibit survival thresholds similar to those associated to the quasispecies model for RNA viruses. Following these similarities, we suggested that basic models of tumor growth exhibiting increasing levels of instability should also display an analog error threshold to genomic instability [2, 22, 24] (see also [12]). If true, then such thresholds could be exploited for cancer treatment and provide some insight into the complex organization of unstable tumors. Let us see how such model can be constructed.

Table 1. Comparison between RNA virus populations and unstable tumors

	<i>RNA viruses</i>	<i>Unstable tumors</i>
Mutation rate	Critical $\mu \sim 1/L$	Very high
Population structure	Quasispecies	Heterogeneous
Average individual viability	very low infectivity	very low proliferation
Genome complexity	minimal, close to critical	reduced, unstable

Assuming that two populations of cells are at play, namely normal (n) and cancer (x), we will introduce a simple model of cancer instability. The model assumes that both populations are homogeneous, thus being all cells in each subset identical in their replication capacities. This is of course a strong oversimplification, particularly when dealing with unstable cancer, where by definition a high diversity is expected to be observed. Under the assumption that these populations compete for available space and resources [11] the simplest model reads:

$$\frac{dn}{dt} = rn - n\phi(n, x), \quad (13)$$

$$\frac{dx}{dt} = r\Gamma(\mu)x - x\phi(n, x), \quad (14)$$

where r is the replication rate of normal cells in the original tissue and μ indicates the rate of genetic instability. The function $\Gamma(\mu)$ includes the (non-linear) effect of instability on the growth of unstable cancer cells.

For simplicity, we assume that there is a maximum population size and that the total population is constant, i.e. $n + x = C$ which for simplicity we fix to $C = 1$. This is the constant population constraint (see chapter 6 of [23]), which implies that $dx/dt + dn/dt = 0$. Using this constraint, we obtain:

$$\phi(n, x) = rn + r\Gamma(\mu)x. \quad (15)$$

And since $n = 1 - x$ we can reduce our previous model to the following one-dimensional system:

$$\frac{dx}{dt} = r(\Gamma(\mu) - 1)x(1 - x). \quad (16)$$

Two fixed points are present: the zero-population one $x^* = 0$ and the maximum population state, here $x^* = 1$. It is easy to see that the first is stable if $\Gamma(\mu) < 1$ and unstable otherwise. By properly defining the function $\Gamma(\mu)$ we might be able to define the conditions under which genetic instability allows cancer growth to occur and overcompete the host tissue.

In order to propose a reasonable form for $\Gamma(\mu)$, two features of instability need to be considered. The first is that, at low levels, instability favors the emergence of fast growing clones able to overcome selection barriers. In other words, some amount of instability is desirable for the tumor: it provides a source for adaptability and selective advantages. The second deals with the negative part of the story: at high levels, mutations tend to be harmful and cells are non-viable. The function should be such that $\Gamma(0) = 1$ (i. e. no effect is at work) whereas we should expect $\Gamma(\mu) \rightarrow 0$ as instability grows (i. e. $\mu \rightarrow \infty$). One possible choice [26] is:

$$\Gamma(\mu) = (1 + \alpha\mu)e^{-\mu/\mu^*} \quad (17)$$

which is consistent with all the previous requirements. Here $\alpha \geq 0$ is a parameter that weights the selective (carcinogenic) advantage of genetic instability: the higher its value, the faster the increase in growth rate. The parameter μ^* is some characteristic

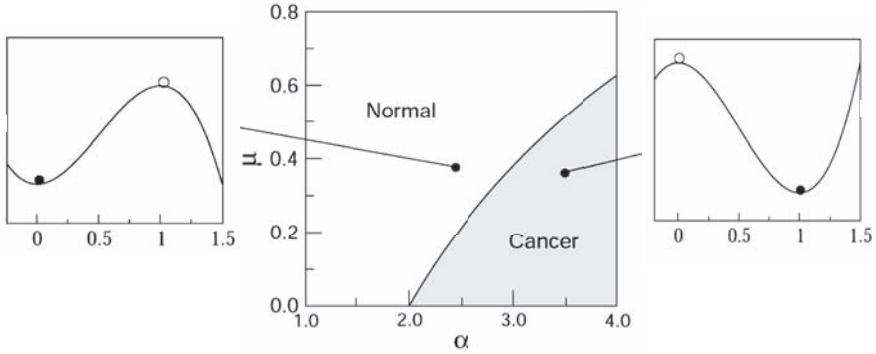


Fig. 3. Phases in unstable cancer dynamics. The gray area indicates the domain where unstable cancer overcomes the host tissue. When close to the separation boundary, small changes of instability levels can generate a shift in the behavior (provided that normal cells are still present). Two examples of the associated potentials are shown, with $\mu = 0.4$ and $\alpha = 2.5$ (left) and $\alpha = 3.5$ (right), respectively

rate which we fix here to $\mu^* = 0.5$. By using this functional form for $\Gamma(\mu)$ we find two phases, as indicated in Fig. 3. Since the two alternative possibilities ($x^* = 0$ and $x^* = 1$) are separated by a critical line, a sharp change is expected to take place near this boundary.

Another way of representing the transition is based on the so called potential $\Phi(x)$, defined as the function such that it is possible to write the dynamics of the system as derived from

$$\frac{dx}{dt} = -\frac{d\Phi(x)}{dx} \quad (18)$$

which leads to a definition for the potential as:

$$\Phi(x) = -\int r(\Gamma(\mu) - 1)x(1-x)dx, \quad (19)$$

Integrating the previous expression, we obtain:

$$\Phi(x) = -r(\Gamma(\mu) - 1) \left(\frac{x^2}{2} - \frac{x^3}{3} \right). \quad (20)$$

The shape of the potential is shown in Fig. 3 for two different parameter sets associated to the two different phases. As we can see, the potential displays a minimum at $x^* = 1$ for the cancer phase (the tumor wins and takes all) and $x^* = 0$ (tumor losses and disappears). This all-or-none behavior corresponds to a first-order transition.

This model (and related ones, see [22, 24]) predicts that unstable tumors close to the critical boundary could be highly adaptive but also fragile with respect to small increases of instability levels. If further instability were introduced (as might be the case while using radiotherapy, chemotherapy or cytotoxic drugs) the tumor could eventually cross the line and decay. The potential of applying this approach to ther-

apy is being seriously considered by several authors, particularly in connection with experiments showing that lethal mutagenesis can be achieved on RNA viruses [9, 16].

4 Tumor Decay under Immune Attack

As mentioned in the introduction, a remarkable nonlinear phenomenon that appears from time to time in cancer is the spontaneous regression (SR) of advanced tumors. The problem has been elusive because it often appears in patients displaying extensive invasion of malignant tissues and are thus sent home. Life expectancy in most of these cases is short and patients are not followed for further examinations until they go back once improvement seems obvious and typically when the cancer is already gone. As a consequence, little or no information concerning the cellular and molecular processes involved is available and the origins of SR have remained elusive.

An exception to this situation is given by some particular systems, such as particular strains of mice studied by Zheng Cui and co-workers, displaying intrinsic resistance to tumor transplantation and display age-dependent spontaneous regression [4]. Transgenic mice are exposed to cancer cells from cell cultures able to generate cancer. Cancers develop and expand, but then regression occurs in a rather predictable way and its paths can be properly traced, revealing that cancer cells are able to trigger a massive immune response, in which immune cells (leucocytes; see Fig. 4) are able to infiltrate cancer tissues and effectively destroy the tumor population. This occurs with no damage to normal cells. How does this work in dynamical terms? Mathematical models have been developed in order to address the interactions between cancer and immune cells [1, 17] (and references therein). Let us consider a very simple model, which was early proposed by two physicists in 1978.

An important part of the immune response is based on special types of cells able to detect anomalous molecular markers at the surface of infected cells [20]. These include the cytotoxic T cells (CTLs). Each class of CTL is able to recognize one specific antigen among all the huge set of possible molecules. Once recognition occurs, an attack is launched and the CTL injects toxic granules inside the target cell eventually killing it.

In principle, the previous mechanisms could be a source of cancer cell removal. However, tumor cells can evade the IS recognition in different ways [29] and sometimes the IS becomes ineffective against cancer growth. The interaction between cancer and immune system actually involves multiple layers of complexity [1]. Below we present the Garay-Lefever model of cancer-immune system interactions (1977) which is shown to display both first and second order transition phenomena [10].

Consider a population of tumor cells, to be indicated as X , which can be attacked by some class of cytotoxic killer cells (T -cells). Here cancer cells replicate at a given rate η . In this model, T cells can recognize and kill cancer cells by detecting special markers on the tumor cell surface. The identification occurs at a rate k_1 and triggers the formation of a cell-cell complex C (the basic rules used here are summarized in

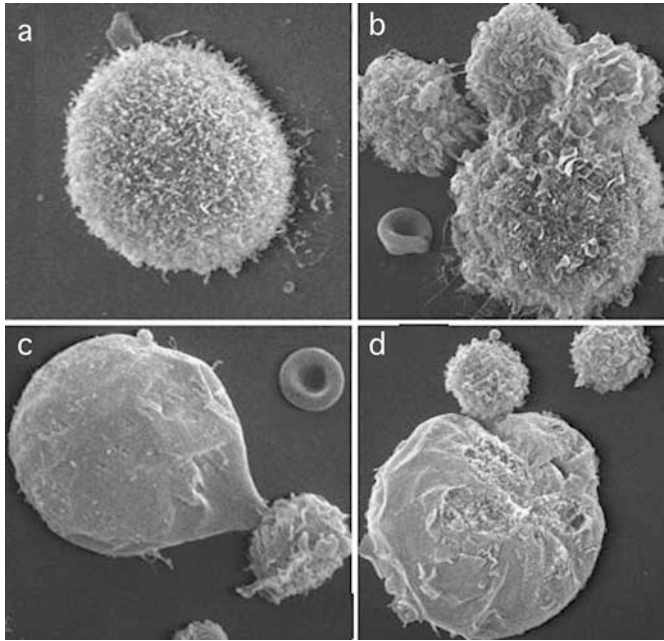


Fig. 4. Cancer regression in model organisms. Using transgenic mice, it was shown that it is possible to replicate the spontaneous regression of advanced tumors and follow the cellular responses in detail. The scanning microscope images show (a) cancer cells, (b) cancer cells being surrounded by immune cells, (c) aballooning effect and membrane collapse (d). Modified from [4] Cui et al. Proc. Natl. Acad. Sci. USA **100**, 6682–6687 (2003) © National Academy of Sciences of USA

Fig. 5). Once the cancer cell has been killed, the complex dissociates and the T cells is free again.

The previous components define a set of reactions, namely:

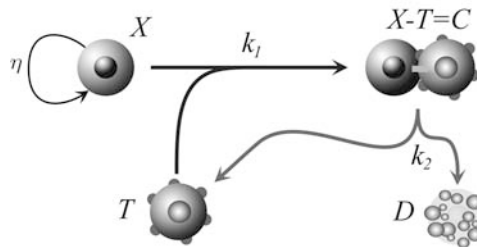


Fig. 5. Basic rules considered in the Garay-Lefever model. Two types of cells, namely lymphocytes and cancer cells are considered, as well as their interacting pairs. Here we summarize the basic interactions, indicating the rates at which some events occur

for the cancer growth kinetics and



for the cancer-CTL interaction. Here D indicates a dead (cancer) cell. The rates k_1 and k_2 measure the speed at which the complex forms and disintegrates, respectively. The corresponding equations describing this dynamics are:

$$\frac{dX}{dt} = \eta X \left(1 - \frac{X}{K}\right) - k_1 TX, \quad (23)$$

$$\frac{dT}{dt} = k_2 C - k_1 TX. \quad (24)$$

here K is the maximum number of cancer cells assumed to exist. The first term in the right hand side for the cancer growth equation is thus consistent with a logistic growth model.

As defined here, cytotoxic cells act as specific predators, reducing the number of cancer cells at a rate $k_1 TX$ i. e. depending on the encounter rates, efficiency of recognition and other phenomena captured by k_1 . Moreover, killer cells forming complexes get back into the free T cell pool as soon as the cancer cell is killed at a rate k_2 .

This two-dimensional system can be further reduced by considering the approximation the lifetime associated to the C complex is very short and thus it takes a comparatively short time to form and decay [10] i.e.:

$$\frac{dT}{dt} \approx 0. \quad (25)$$

This equilibrium condition just tells us that cytotoxic cells form complexes that rapidly split (compared to other phenomena involved in the model). The total number of cytotoxic cells, E , can be now computed as follows

$$E = T + C, \quad (26)$$

since E includes both complex-associated and free killer cells. In the following, we assume that E is a conserved (constant) value.

Using the steady state assumption for T cell behaviors, we have now $T(X) = k_2 C / k_1 X$ and using $C = E - T$, we obtain:

$$T(X) = \frac{E}{1 + \frac{k_1}{k_2}}, \quad (27)$$

which allows us to write down the whole mean field equation for tumor growth dynamics in terms of a one-dimensional system, namely:

$$\frac{dX}{dt} = \eta X \left(1 - \frac{X}{K}\right) - \frac{k_1 X E}{1 + \frac{k_1}{k_2}}, \quad (28)$$

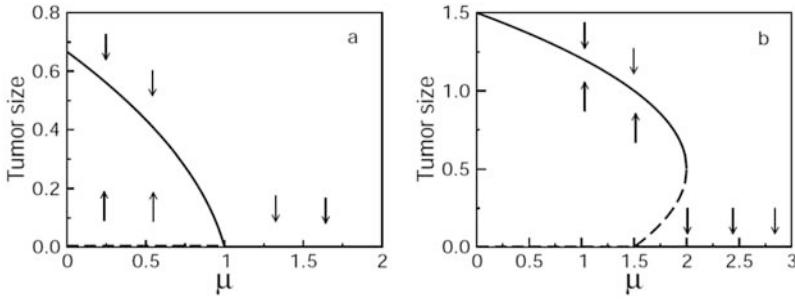


Fig. 6. Bifurcation diagrams in the Garay-Lefever model. Here we plot the stable solutions (thick, continuous lines) for (a) $\theta = 1.5$ and (b) $\theta = 0.5$ corresponding to the monostable and bistable domains (see text). These curves predict sharp transitions between cancer and tumor-free phases

following Garay and Lefever, we will reduce the number of parameters in this model by using the following rescaled quantities

$$\mu = \frac{k_1 E}{\eta} \quad \theta = \frac{k_2}{k_1 N} \quad x = \frac{k_1 X}{k_2} \quad (29)$$

(time t would be rescaled accordingly as $t \rightarrow \eta t$, but we keep the same notation). The model now reads:

$$\frac{dx}{dt} = x(1 - \theta x) - \frac{\mu x}{1 + x}. \quad (30)$$

Using this reduced one-dimensional model, three possible fixed points are obtained (Fig. 7) namely $x^* = 0$ (cancer extinction) and the two nontrivial solutions:

$$x_{\pm} = \frac{1}{2\theta} \left[1 - \theta \pm \sqrt{(1 + \theta)^2 - 4\theta\mu} \right]. \quad (31)$$

The tumor-free fixed point will be stable provided that

$$\lambda_{\mu}(0) = \left(\frac{\dot{x}}{x} \right)_{x=0} < 0, \quad (32)$$

which in our case is simply $\mu > \mu_c = 1$. Since the parameter μ captures the conflict between the rates η (tumor cell formation) and k_1 (tumor cell removal) the critical condition $\mu_c = 1$ naturally separates cancer survival from extinction.

Looking at the nontrivial solutions, two possible types of solutions can be identified:

1. For $\theta > 1$, the fixed point x_- will be negative (and thus biologically irrelevant) whereas x_+ remains positive provided that $\mu < 1$. In this case, a cancer cell population will exist below the critical point μ_c .
2. For $\theta > 1$, when the inequality $(1 + \theta)^2 \geq 4\theta(\mu - 1)$ holds, the two nontrivial solutions exist. It is not difficult to see that, given the previous results, the interval

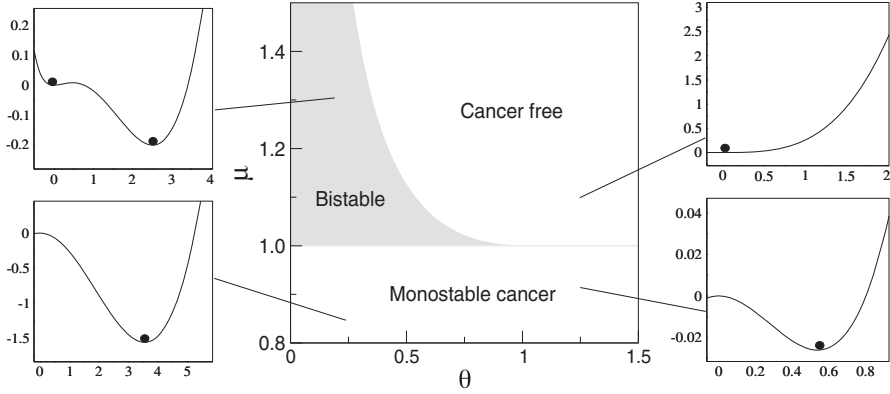


Fig. 7. The three phases exhibited by the Garay-Lefever model (1977) [10]. Here the upper white domain indicates that no cancer cells are stable (tumor extinction) whereas two other areas correspond to the presence of monostable cancer (white low domain) and two alternative states (gray area). Four examples of the associated potentials are also shown, with filled circles indicating the presence of stable states

of existence of these two simultaneous solutions is:

$$1 < \mu < \frac{(1 + \theta)^2}{4\theta} = \bar{\mu}, \quad (33)$$

defining a domain of bistable behavior. For $\mu < 1$ only the first solution x_+ is present. For the bistable domain, we can write down the two solutions as follows:

$$x_{\pm} = \frac{1}{2\theta} \left[1 - \theta \pm \sqrt{\theta(\bar{\mu} - \mu)} \right]. \quad (34)$$

In Fig. 6 we show two examples of the bifurcation diagrams associated to the model. In (a) an example of the continuous transition scenario is obtained for $\theta > 1$ whereas the bistable case, for $\theta < 1$, is shown in (b). In Fig. 7 the basic results are summarized by representing the three phases on the (θ, μ) parameter space. We also plot four examples of the associated potential function, here given by

$$\Phi_{\mu}(x) = -\frac{x^2}{2} + \frac{\theta x^3}{3} + \mu(x - \ln(1 + x)), \quad (35)$$

which also illustrate the exchange of stability between different attractors.

The most interesting result obtained from this model is the presence of a sharp boundary separating viable from extinct cancer. The result is important because the same rules that define stable states are responsible, when the right parameters are tuned, for the regression of cancer from a cancer-advanced state. This type of phenomenon, as we already mentioned above, is well known from statistical physics, but only in the last decade the importance of these processes has been fully appreciated outside the physics community [23]. Under this picture, we can provide a rationale for the singular character of reported spontaneous regressions [3]. The key lesson

from our model is that slight variations in the effectivity of immune responses can allow us to cross critical points and shift the macroscopic state from malign invasion to cure.

5 Cancer as an Evolving System

As pointed out by [18] cancer is a disease of clonal evolution within the body. Although such idea was early formulated both modeling, clinical and molecular data reveal that evolution and genomic instability in particular, are key to carcinogenesis. Appropriate theoretical approaches to the origins and development of cancer need to take such considerations into account. On the other hand, it seems clear that tumors involve large-scale, collective phenomena through their evolutionary dynamics [5]. The observation of spontaneous regression clearly suggests that such phenomena can eventually trigger global changes including a shift from one phase to another. In this case, the nonlinear effects associated to the predator-prey interactions can be strong enough to fully remove the cancer population. Together with the appropriate set of interactions, well-defined thresholds need to be overcome.

Similarly, a transition to extinction can take place when instability makes reliable growth impossible. The process associated to this might be spontaneous too: the failure of some given genetic elements associated to preserving genomic integrity inevitably leads to further failures. Some of them will affect other stability genes or indirectly promote the loss of chromosomes (or parts of them). Instability, no matter how it is causally related to other components of the cellular logic, is expected to increase over time. The existence of a threshold means that there is a boundary to viability beyond which unstable cells will fail to maintain their integrity. If such boundary is approached under normal conditions, our model predicts that small additional increases lead inevitably to tumor death. If such type of instability therapy can be implemented, it would have some important advantages in treating aggressive neoplasms. The path to finding the appropriate molecular targets might be long, since it requires a specificity for cancer cells that should avoid affecting the stability of normal ones. But potential avenues are already being considered [9] and promise new way of thinking about cancer and its evolution. Future models of these processes should consider more realistic approximations, incorporating the basic biology, genomic interactions and spatial degrees of freedom (see [27]).

References

1. Adam, J.A., Bellomo, N.: A survey of models for tumor-immune system dynamics. Birkhauser, Boston (1996)
2. Cahill, D.P., Kinzler, K.W., Vogelstein, B., Lengauer C.: Genetic instability and darwinian selection in tumors. *Trends Genet.* **15**, M57–M61 (1999)

3. Cole, W.H.: Spontaneous regression of cancer and the importance of finding its cause. *Natl. Cancer Inst. Monogr.* **44**, 5–9 (1976)
4. Cui, Z., Willingham, M.C., Hicks, A.M., Alexander-Miller, M.A., Howard, T.D., Hawkins, G.A., Miller, M.S., Weir, H.M., Du, W., DeLong, C.J.: Spontaneous regression of advanced cancer: Identification of a unique genetically determined, age-dependent trait in mice. *Proc. Natl. Acad. Sci. USA* **100**, 6682–6687 (2003)
5. Deisboeck, T.S., Couzin, I.D.: Collective behavior in cancer cell populations. *Bioessays* **31**, 190–197 (2009)
6. Domingo, E., Holland, J.J.: Mutation rates and rapid evolution of RNA viruses. In: Morse, S. (ed.) *The evolutionary biology of RNA viruses*, pp. 161–183, Raven Press, New York (1994)
7. Eigen, M.: Self-organization of matter and the evolution of biological macromolecules. *Naturwiss.* **58**, 465–523 (1971)
8. Eigen, M., McCaskill, Schuster, P.: (1987) The molecular quasispecies. *Adv. Chem. Phys.* **75**, 149–263 (1987)
9. Fox, J.F., Loeb, L.A.: Lethal mutagenesis: targeting the mutator phenotype in cancer. *Seminars Cancer Biol.* **20**, 353–359 (2010)
10. Garay, R.P., Lefever, R.: A kinetic approach to the immunology of cancer: stationary state properties of effector-target cell reactions. *J. Theor. Biol.* **73**, 417–438 (1978)
11. Gatenby, R.A.: Application of competition theory to tumour growth: implications for tumour biology and treatment. *Europ. J. Cancer* **32A**, 722–726 (1996)
12. Gatenby, R.A., Frieden, B.R.: Application of information theory and extreme physical information to carcinogenesis. *Cancer Res.* **62**, 3675–3684 (2002)
13. Greaves, M., Maley, C.C.: Clonal evolution in cancer. *Nature* **481**, 306–313 (2012)
14. Lengauer, C., Kinzler, K.W., Vogelstein, B.: Genetic instabilities in human cancers. *Nature* **396**, 643–649 (1998)
15. Loeb, L.A.: Mutator phenotype may be required for multistage carcinogenesis. *Cancer Res.* **51**, 3075–3079 (1991)
16. Loeb, L.A., Essigmann, J.M., Kazazi, F., Zhang, J., Rose, K.D., Mullins, J.I.: Lethal mutagenesis of HIV with mutagenic nucleoside analogs. *Proc. Natl. Acad. Sci. USA* **96**, 1492–1497 (1999)
17. Mallet, D.G., de Pillis, L.G.: A cellular automata model of tumor-immune system interactions. *J. Theor. Biol.* **239**, 334–350 (2006)
18. Merlo, L.M.F., Pepper, J.W., Reid, B.J., Maley, C.C.: Cancer as an evolutionary and ecological process. *Nature Rev. Cancer* **6**, 924–935 (2006)
19. Nowell, P.C.: The clonal evolution of tumor cell populations. *Science* **194**, 23–32 (1976)
20. Rosenberg, S.A.: Progress in human tumor immunology and immunotherapy. *Nature* **411**, 380–384 (2001)
21. Schuster, P.: How do RNA molecules and viruses explore their worlds? In: Cowan, G.A., Pines, D., Meltzer, D. (eds.) *Complexity: metaphors, models and reality*, pp. 383–418. Addison-Wesley, Reading, MA (1994)
22. Solé, R.V.: Phase transitions in unstable cancer cell populations. *Eur. Phys. J. B* **35**, 117–124 (2003)
23. Solé, R.V.: *Phase transitions*. Princeton University Press, Princeton (2011)
24. Solé, R.V., Deisboeck, T.S.: An error catastrophe in cancer? *J. Theor. Biol.* **228**, 47–54 (2004)
25. Solé, R.V., Goodwin B.: *Signs of Life: How complexity pervades biology*. Basic Books, New York (2001)
26. Solé, R.V., Rodríguez-Caso, C., Deisboeck, T.S., Saldanya, J.: Cancer stem cells as the engine of unstable tumor progression. *Journal of Theoretical Biology* **253**, 629–637 (2008)

27. Spencer, S., Gerety, R., Pienta, K., Forrest, S.: Modeling somatic evolution in tumorigenesis. *PLoS Comput. Biol.* **2**, e108 (2006)
28. Weaver, B.A.A., Silk, A.D., Montagna, C., Verdier-Pinard, P., Cleveland, D.W.: Aneuploidy acts both oncogenically and as a tumor suppressor. *Cancer cell* **11**, 25–36 (2007)
29. Weinberg, R.A.: *The Biology of Cancer*. Garland, New York (2007)

Spatio-Temporal Modelling of Intracellular Signalling Pathways: Transcription Factors, Negative Feedback Systems and Oscillations

Mark A.J. Chaplain, Marc Sturrock, and Alan J. Terry

Abstract There are many intracellular signalling pathways where the spatial distribution of the molecular species cannot be neglected. One such class of pathways is those involving transcription factors (e.g. Hes1, p53-Mdm2, NF- κ B, heat-shock proteins) which often exhibit oscillations in both space and time. In this chapter we present a partial differential equation model of the transcription factor, Hes1. Our model considers the dynamics of Hes1 in a 2-dimensional cellular domain including a nucleus, cytoplasm and microtubule-organising centre (MTOC). Spatial movement of the molecules (protein, mRNA) is assumed to be by diffusion, and also convection along microtubules. Through numerical simulations we find ranges of values for the model parameters such that sustained oscillatory dynamics occur, consistent with available experimental measurements. In order to bridge the gap between *in vivo* and *in silico* experiments we investigate more realistic cell geometries by using an imported image of a real cell as our computational domain.

1 Introduction

All intracellular signal transduction processes are inherently spatial. The eukaryotic cell hosts a variety of spatial compartments (e.g., the nucleus, cytoplasm), with different metabolic activities taking place in each compartment. The appropriate sub-cellular localisation of molecular species is critical for a cell to remain healthy [17, 22, 37]. Signalling molecules reach the appropriate spatial compartments through molecular movement, such as diffusion and active transport. Cell shape and size change over time and are known to influence intracellular signal transduction [31, 36]. The key process of *transcription* occurs at specific sites – genes – in the nucleus,

M.A.J. Chaplain (✉) · M. Sturrock · A.J. Terry
Division of Mathematics, University of Dundee, Nethergate, Dundee DD1 4HN, Scotland
e-mail: chaplain@maths.dundee.ac.uk, msturrock@maths.dundee.ac.uk,
aterry@maths.dundee.ac.uk

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series
DOI 10.1007/978-88-470-2571-4_4, © Springer-Verlag Italia 2012

and some genes are located closer to the nuclear membrane than others, increasing their sensitivity to transcription factors [12]. Nuclear pore complexes determine where species move in and out of the nucleus and how quickly they do so [8]. Within the cytoplasm, another key process – *translation* – occurs in the ribosomes. It is apparent from all these observations that mathematical models of intracellular signalling pathways will be more realistic the more they seek to account for spatial features inherent to these pathways.

Many intracellular pathways exhibit oscillatory dynamics in space and time in response to a range of external stimuli [15, 16, 35, 45]. A negative feedback loop often lies at the core of such pathways, controlling the levels of mRNA and proteins which are frequently transcription factors. These molecules initiate or regulate transcription in eukaryotic cells, and in order for them to function they must bind to specific DNA sequences in the nucleus. Three examples of negative feedback loops are the Hes1 pathway (which plays a role in somitogenesis), the NF- κ B pathway (which is central to many stressful, inflammatory, and innate immune responses) and the p53-Mdm2 pathway (which is critical for regulating the cell-cycle). All three can be deregulated in human cancer [20, 24, 42]. In this chapter, we consider a partial differential equation (PDE) model for the Hes1 system and consider the explicit spatial interactions of the Hes1 protein and mRNA within the cell.

2 Hes1: The Canonical Transcription Factor

Hes1 is a member of the family of basic helix-loop-helix (bHLH) transcription factors. It has been found to be deregulated in some forms of cancer [42] and is known to play a role in somitogenesis, the developmental process responsible for segmentation of the vertebrate embryo. During somitogenesis, a “segmentation clock” controls the timing of the assignment of mesodermal cells to discrete blocks. The segmentation clock depends on the oscillatory expression of a complex network of signalling pathways, including the Hes1 pathway which contains a negative feedback loop (see Fig. 1). This feedback loop is formed through interactions of the Hes1 protein with its own gene, where the Hes1 protein binds to regulatory sequences on the *hes1* promoter and represses the transcription of *hes1* mRNA. Experiments have measured the concentration levels of *hes1* mRNA and Hes1 protein in mouse cell lines [16, 28]. In response to serum treatment, it was found that levels of *hes1* mRNA and Hes1 protein exhibited oscillations with a regular period of approximately two hours. This coincides with the period observed for the mouse segmentation clock.

Mathematical models of oscillatory dynamics in the Hes1 pathway have taken a variety of forms. The first attempt to model this pathway was presented in [16] where an ODE model was used, and in order to reproduce the observed oscillations, a third unknown species was introduced. [33] found that invoking an unknown species could be avoided by introducing a time delay (attempting to account for the processes of transcription and translation), producing a delay differential equation (DDE) system. The effect of low particle numbers in this system was explored in [5]. The details of the Hes1 pathway were scrutinized in greater depth in [32], again using a

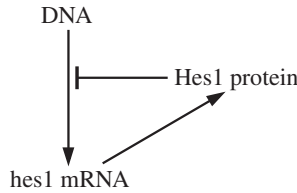


Fig. 1. Schematic diagram of the Hes1 pathway. Hes1 protein is produced from hes1 mRNA via the process of translation. Hes1 protein then inhibits transcription of hes1 mRNA, providing the negative feedback loop

DDE system. In particular, an investigation into the effects of dimerisation and compartmentalisation was presented. The role of Gro/TLE1 was considered in [7]. Other models have examined the role of the Hes1 pathway in somitogenesis [1], while [47, 48] have explored spatio-temporal oscillations in a PDE model. The remainder of this chapter focusses on the recent results of these latter two papers.

2.1 The Mathematical Model of the Hes1 System

We begin by introducing the PDE model first developed in [47] describing the intracellular interactions between hes1 mRNA (concentration denoted by $[m]$) and the Hes1 protein (concentration denoted by $[p]$). The model is considered on a two-dimensional spatial domain representing a cell, with a separate nucleus and cytoplasm. In the equations below, a subscript n denotes a nuclear concentration and a subscript c denotes a cytoplasmic concentration.

We assume both protein and mRNA are subject to diffusion. Diffusion coefficients are denoted depending on the type of species (either a subscript m for mRNA or p for protein) and location (a subsubscript n for nuclear or c for cytoplasmic). For example, the diffusion coefficient for hes1 mRNA in the nucleus is D_{m_n} . Both protein and mRNA are assumed to undergo linear decay, with parameter μ_m denoting mRNA decay and μ_p denoting protein decay. Production of mRNA takes place by the process of transcription in the nucleus. Our production term for nuclear hes1 mRNA is a Hill-like function which decreases as protein levels in the nucleus increase. In this function, the parameters α_m , $\hat{p}$, and h represent, respectively, the basal rate of transcription in the absence of nuclear protein, the concentration of Hes1 protein that reduces the transcription rate to half its basal value, and a Hill coefficient. Proteins are translated from mRNA by ribosomes in the cytoplasm, a process that is likely to occur at least some minimal distance from the nuclear membrane, and hence we assume protein production occurs a small distance outside the nucleus with production rate proportional to the amount of cytoplasmic hes1 mRNA, the constant of proportionality being denoted α_p . The full system of equations is therefore given by:

$$\frac{\partial [m_n]}{\partial t} = D_{m_n} \nabla^2 [m_n] + \underbrace{\frac{\alpha_m}{1 + ([p_n]/\hat{p})^h}}_{\text{transcription}} - \mu_m [m_n], \quad (1)$$

$$\frac{\partial[m_c]}{\partial t} = D_{m_c} \nabla^2[m_c] - \mu_m[m_c], \quad (2)$$

$$\frac{\partial[p_c]}{\partial t} = D_{p_c} \nabla^2[p_c] + H_1(x, y) \underbrace{\alpha_p[m_c]}_{\text{translation}} - \mu_p[p_c], \quad (3)$$

$$\frac{\partial[p_n]}{\partial t} = D_{p_n} \nabla^2[p_n] - \mu_p[p_n], \quad (4)$$

where $H_1(x, y)$ is a function accounting for the localisation of protein production in the ribosomes a distance l from the centre of the nucleus and is defined as follows:

$$H_1(x, y) = \begin{cases} 0, & \text{if } x^2 + y^2 \leq l^2, \\ 1, & \text{if } x^2 + y^2 > l^2. \end{cases} \quad (5)$$

We apply continuity of flux boundary conditions across the (internal) nuclear membrane and zero-flux boundary conditions at the outer cell membrane:

$$D_{m_n} \frac{\partial[m_n]}{\partial \mathbf{n}} = D_{m_c} \frac{\partial[m_c]}{\partial \mathbf{n}} \quad \text{and} \quad [m_n] = [m_c] \quad \text{at the nuclear membrane}, \quad (6)$$

$$D_{p_n} \frac{\partial[p_n]}{\partial \mathbf{n}} = D_{p_c} \frac{\partial[p_c]}{\partial \mathbf{n}} \quad \text{and} \quad [p_n] = [p_c] \quad \text{at the nuclear membrane}, \quad (7)$$

$$\frac{\partial[m_c]}{\partial \mathbf{n}} = 0 \quad \text{at the cell membrane}, \quad (8)$$

$$\frac{\partial[p_c]}{\partial \mathbf{n}} = 0 \quad \text{at the cell membrane}, \quad (9)$$

where $\mathbf{n}$ is a unit normal. We also apply zero-initial conditions, i.e.,

$$[m_n] = [m_c] = [p_n] = [p_c] = 0 \quad \text{at } t = 0. \quad (10)$$

To solve the Hes1 system (1)–(10) numerically, we first non-dimensionalised the model (details of which are provided in the appendix) and then we chose non-dimensional parameter values which were obtained from Eq. (25) in [47]. The model was solved numerically using the finite element method as implemented in the software package COMSOL 3.5a, using triangular basis elements and Lagrange quadratic basis functions along with a backward Euler time-stepping method of integration. This numerical method of solution is used in all the simulations in this paper. The equations were solved on the domain shown in Fig. 2, representing a cell with cytoplasmic and nuclear subdomains. Finally we calculated dimensional parameter values – these are shown in the third column in Table 1. For simplicity, all nuclear and cytoplasmic diffusion coefficients were set equal to the same constant; we denote the dimensional diffusion coefficient by D_{ij} , which indicates diffusion of species i (mRNA or protein) in location j (nucleus or cytoplasm). For this simulation study, we have found ranges of values for all of the parameters such that the system exhibits

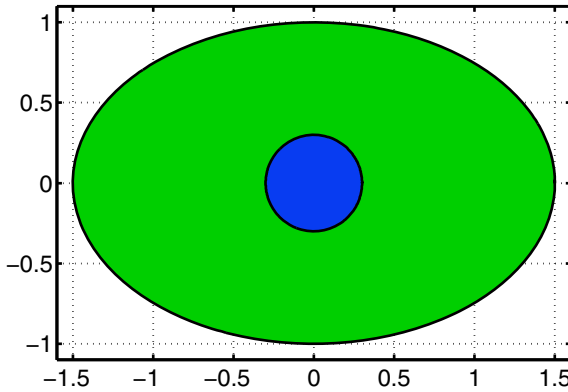


Fig. 2. Diagram showing the domain used in numerical simulations of the initial Hes1 model. Spatial units here are non-dimensional, with one non-dimensional spatial unit corresponding to $10\ \mu\text{m}$. The cell is an ellipse, centre $(0,0)$, with major and minor axes of 3 and 2, respectively. The nucleus is shown here as a black circle, centre $(0,0)$, radius 0.3. The cytoplasm (shown in grey) is the part of the cell that is outside the nucleus

sustained oscillatory dynamics, where we define such dynamics as being able to observe at least five distinct peaks in the total concentration of the transcription factor (Hes1 protein) in the nucleus. These ranges are given in the fourth column in Table 1. In order to find the range for any particular parameter, we varied this parameter whilst holding all the other parameters fixed at their “default” values, the dimensional versions of which are stated in the third column of Table 1.

2.1.1 Hes1 Model Parameter Values

Our range for the diffusion coefficient in Table 1 is consistent with two recent spatio-temporal modelling studies of intracellular signalling pathways similar in scope to the present study [49, 50]. Experimentalists have found the diffusion coefficient of soluble proteins in the cytoplasm to be in the range $10^{-9}\text{cm}^2\text{s}^{-1}$ to $10^{-8}\text{cm}^2\text{s}^{-1}$ [29, 43], which is in agreement with the upper bound of our range. There is also a growing body of evidence suggesting that proteins and mRNA molecules are subject to macromolecular crowding, which generates an environment where diffusion is hindered by obstacles and traps [30]. Taking this into account would likely increase our lower bound to be consistent with experimental measurements. We may consider macromolecular crowding in future work.

The degradation rate μ_m for hes1 mRNA and the degradation rate μ_p for Hes1 protein have both been estimated from experiments. Hirata et al. estimated μ_m to be $4.83 \times 10^{-4}\text{s}^{-1}$ and μ_p to be $5.16 \times 10^{-4}\text{s}^{-1}$ [16]. Our parameter ranges for μ_m and μ_p in Table 1 contain these experimental estimates.

We mentioned above in Sect. 2 that a DDE model of the Hes1 pathway had been explored in [33]. Our range for the Hill coefficient, namely $h \geq 4$, is very similar to the range ($h > 4$) producing sustained oscillatory dynamics in the DDE model in [33]. Note that a larger Hill coefficient corresponds to greater nonlinearity, or

Table 1. Description of parameters in the initial Hes1 model (defined in Sect. 2.1), values used in simulations, and ranges over which sustained oscillatory dynamics are observed

<i>Parameter</i>	<i>Description</i>	<i>Value in simulations</i>	<i>Range over which oscillations are observed</i>
D_j	Diffusion coefficient of species i in compartment j	$3.13 \times 10^{-11} \text{cm}^2 \text{s}^{-1}$	$2.67 \times 10^{-11} \text{cm}^2 \text{s}^{-1}$ to $1.25 \times 10^{-9} \text{cm}^2 \text{s}^{-1}$
α_m	Basal transcription rate of hes1 mRNA	$6.25 \times 10^{-11} \text{Ms}^{-1}$	$\geq 3.87 \times 10^{-12} \text{Ms}^{-1}$
$\hat{p}$	Critical concentration of Hes1 protein	$1.00 \times 10^{-9} \text{M}$	$6.89 \times 10^{-10} \text{M}$ to $1.00 \times 10^{-6} \text{M}$
h	Hill coefficient	5	≥ 4
μ_m	Degradation rate of hes1 mRNA	$1.25 \times 10^{-3} \text{s}^{-1}$	$1.25 \times 10^{-4} \text{s}^{-1}$ to $1.50 \times 10^{-3} \text{s}^{-1}$
α_p	Translation rate of Hes1 protein	0.0555s^{-1}	$\geq 0.0350 \text{s}^{-1}$
μ_p	Degradation rate of Hes1 protein	$1.25 \times 10^{-3} \text{s}^{-1}$	$1.25 \times 10^{-4} \text{s}^{-1}$ to $1.50 \times 10^{-3} \text{s}^{-1}$
l	Minimum radial distance of translation from centre of nucleus	$6.32 \mu\text{m}$	nuclear membrane ($3 \mu\text{m}$) to $6.63 \mu\text{m}$

co-operativity, in the regulation of hes1 transcription by Hes1 protein. According to [33], Hes1 acts as a dimer, which alone would suggest that $h = 2$, so the requirement that $h > 2$ implies that there may be additional interactions involving the three binding sites for Hes1 at the hes1 promoter.

Our value for the critical concentration of Hes1, namely $\hat{p} = 10^{-9} \text{M}$, is the same as the critical concentration for the zebrafish Her1 protein mentioned in [25]. Her1 is similar to Hes1 in that both are believed to belong to simple negative feedback loops. The DDE model of the Hes1 pathway in [33] contains a parameter (namely, p_0) analogous to $\hat{p}$ but representing a number of molecules rather than a concentration. A sensible range is suggested as 10 to 100. By an elementary calculation converting concentration into number of molecules, we find that our value for $\hat{p}$ corresponds to approximately 16 molecules, which clearly lies within the postulated range in [33].

Our estimate for the translation rate α_p of 0.0555s^{-1} is similar to the rate mentioned in [7] of 1min^{-1} or 0.0167s^{-1} . The basal transcription rate α_m has not been measured experimentally, so we leave our estimate of $6.25 \times 10^{-11} \text{Ms}^{-1}$ as a prediction. The distance l of translation from the centre of the nucleus has been studied in Figs. 9–11 in [47] but a range of values such that sustained oscillatory dynamics

occur was not stated. The range for l presented in Table 1 reveals that oscillatory dynamics can be obtained when the minimum distance of protein translation coincides with where the cytoplasm meets the nucleus. However, increasing l too much results in a loss of oscillatory dynamics, implying that the precise spatial location of the ribosome in the cytoplasm is important.

2.2 Extended Hes1 Model

In this section we extend the original Hes1 model of [47] (described in Sect. 2.1), first by including explicitly the structure of the nuclear membrane. Encapsulating the nucleus, the nuclear membrane divides the cell into two compartments, between which there is a constant exchange of molecular material. This physical separation of the nucleus and cytoplasm provides a level of spatial regulation in signal transduction. Nucleocytoplasmic transport occurs through the nuclear pore complex (NPC). The NPCs perforate the two lipid bilayers which form the nuclear membrane and allow for bidirectional transport of a large number of RNA and protein cargoes which vary in size from 1 kDa to nearly 50 MDa (almost 40nm in diameter) [52]. The number of functional NPCs varies depending on the growth state of the cell, which in turn affects the overall permeability of the nuclear membrane [14].

In order to model the nuclear membrane explicitly, we need to account for its thickness d (which is also the depth of the NPC) and the diffusion of molecules across it. This effectively allows us to model its permeability. The nuclear membrane thickness has been estimated to be approximately 100nm [6]. Regarding diffusion across the nuclear membrane, note that since the NPCs are not located everywhere within it, there exist some areas of it that cannot be traversed, and this slows down the average rate at which particles diffuse across it. Molecular crowding may also slow down this average rate. In the restricted space of an NPC, larger molecules, such as proteins, will diffuse more slowly than smaller molecules, such as mRNA [27, 41]. A second explicit step in modelling the nuclear membrane is therefore to assume that diffusion across it is slower than in the cytoplasm or nucleus, with protein diffusion slower than mRNA diffusion across the membrane. Although diffusion coefficients for cytoplasmic, nuclear, and nuclear-embedded proteins have been estimated experimentally [23], we are not aware of experimental estimates for diffusion rates across the NPCs for hes1 mRNA and Hes1 protein. Therefore, still assuming as at the end of Sect. 2.1 that the nuclear and cytoplasmic diffusion coefficients are the same constant D_{ij} , we shall simply choose $D_m = D_{ij}/5$ and $D_p = D_{ij}/15$ for the nuclear membrane diffusion coefficients for hes1 mRNA and Hes1 protein, respectively. In summary, we can take into account nuclear membrane thickness and slower diffusion across it by replacing the boundary conditions in (6) and (7) by those for a thin boundary layer of width d , defined by:

$$D_{m_n} \frac{\partial [m_n]}{\partial \mathbf{n}_1} = \frac{D_m ([m_n] - [m_c])}{d}, \quad (11)$$

$$D_{m_c} \frac{\partial [m_c]}{\partial \mathbf{n}_2} = \frac{D_m([m_c] - [m_n])}{d}, \quad (12)$$

$$D_{p_c} \frac{\partial [p_c]}{\partial \mathbf{n}_1} = \frac{D_p([p_c] - [p_n])}{d}, \quad (13)$$

$$D_{p_n} \frac{\partial [p_n]}{\partial \mathbf{n}_2} = \frac{D_p([p_n] - [p_c])}{d}, \quad (14)$$

where $\mathbf{n}_1$ and $\mathbf{n}_2$ are appropriate outward normals.

The boundary conditions, (11)–(14), describe the flux across the nuclear membrane. This flux can be thought of as a permeability coefficient (defined as the diffusion coefficient of the species in the nuclear membrane divided by the membrane thickness) multiplied by the concentration difference of the species across the nucleocytoplasmic boundary.

Our second extension to the original Hes1 model of [47] is to consider active transport. It is important for transcription factors to be able to move quickly from the cytoplasm to the nucleus, which can be achieved by active transport along microtubules. The microtubules are fibrous, hollow rods that function primarily to help support and shape the cell. For the majority of the cell cycle (i.e., the interphase period), the microtubules are arranged in the cytoplasm as an aster originating from the *microtubule-organising centre* (MTOC) located close to the nucleus (see figure 1, [26]). The microtubules also play a major role in the intracellular trafficking of macromolecules and organelles [8, 11]. This trafficking of cargo molecules occurs as follows: motor proteins bind to the cargoes and then actively transport them along microtubules. Motor proteins can be split into two families – dyneins (which move molecules from the cytoplasm towards the nuclear membrane) and kinesins (which move molecules towards the cell membrane). Motor proteins interact with microtubules via their ATPase domain, while their opposite terminus interacts with the cargo being transported. The movement of proteins along microtubules towards the nucleus can be viewed as a biased random walk. For example, although cargoes bound to dynein mainly move in the direction of the nucleus, there is evidence for detachment and reattachment of cargoes to motor proteins, pauses, and simultaneous attachment to both dynein and kinesin which can change the direction of movement through a “tug-of-war” [34, 46]. For simplicity, we shall model active transport of the transcription factor Hes1 as always being directed towards the nucleus. We do this by adding a convection term to the cytoplasmic Hes1 equation, namely Eq. (3), which becomes:

$$\frac{\partial [p_c]}{\partial t} = D_{p_c} \nabla^2 [p_c] - \underbrace{\nabla \cdot (\mathbf{a} [p_c])}_{\text{active transport}} + \alpha_p [m_c] - \mu_p [p_c], \quad (15)$$

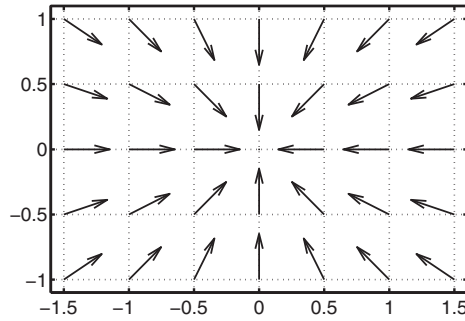


Fig. 3. Plot showing the vector field $\mathbf{a}$ (defined in Eq. (16)) modelling the convective effect of the microtubules

where $\mathbf{a}$ is the convective velocity given by

$$\mathbf{a} = \left[\frac{-ax}{\sqrt{x^2 + y^2}}, \frac{-ay}{\sqrt{x^2 + y^2}} \right], \quad (16)$$

and the parameter a is the convection speed. The vector field $\mathbf{a}$ is depicted in Fig. 3.

Finally, in order to take into account the location of the MTOC, we modify the domain on which our equations are solved. To be specific, we solve on the domain shown in Fig. 4. In this domain, the MTOC is located around the circumference of a circle a small distance away from the nucleus. Since microtubules originate from the MTOC and not from the nucleus, we assume that active transport may occur only in the green region outside the MTOC. Hence, in the outer green region, we assume cytoplasmic Hes1 protein satisfies Eq. (15) but in the orange region between the MTOC and the nuclear membrane we assume it satisfies Eq. (3).

2.2.1 Simulation Results for the Extended Hes1 Model

We solved the extended Hes1 model numerically (given by Eqs. (1)–(4) and (15)), subject to conditions (8)–(14), on the domain shown in Fig. 4. We retained the parameter values used to simulate the original Hes1 model in Sect. 2.1. The diffusion coefficients across the nuclear membrane have already been defined in terms of the diffusion coefficient in the nucleus and cytoplasm, so did not need to be estimated. The nuclear membrane thickness was chosen to be the same as the experimentally measured value of 100nm [6]. The rate of active transport was chosen to produce numerically stable sustained oscillations. We summarise the parameter values used for the extended Hes1 model in the second column of Table 2. Ranges of values such that the model exhibits sustained oscillatory dynamics were also found and are stated in the third column of Table 2.

By comparing Tables 1 and 2, we see that the parameter ranges yielding sustained oscillatory dynamics are widened by the addition of an explicit nuclear membrane and active transport. Our extended model is therefore both more realistic and a more robust oscillator. Note in particular that sustained oscillations may occur in

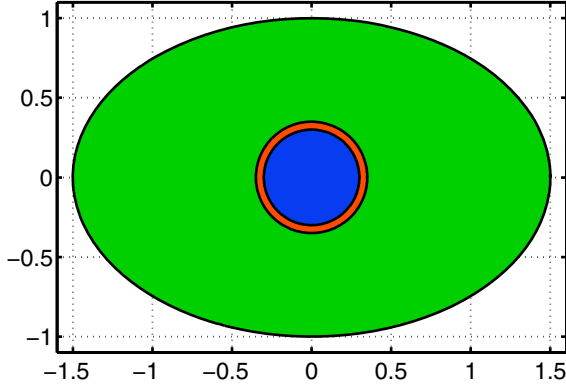


Fig. 4. Diagram showing the domain used in numerical simulations of the extended Hes1 model. Spatial units here are non-dimensional, with one non-dimensional spatial unit corresponding to $10\mu\text{m}$. The cell is an ellipse, centre $(0,0)$, with major and minor axes of 3 and 2, respectively. The nucleus is shown as a black circle, centre $(0,0)$, radius 0.3. The microtubule-organising centre (MTOC) is located around the circumference of the circle, centre $(0,0)$, radius 0.35, which surrounds the nucleus and is close to it. The cytoplasm is the part of the cell that is outside the nucleus (the grey and white regions) and active transport occurs only in the grey region. It does not occur in the white region because microtubules originate from the MTOC and not from the nucleus

the extended model even when the Hill coefficient h is as low as two. As we mentioned above in Sect. 2.1.1, Hes1 acts as a dimer, which suggests that oscillations should be possible with $h = 2$ [33]. We have now found that this is indeed possible, and so it may not be necessary to seek evidence for binding site interactions or other nonlinearities to faithfully model the Hes1 pathway. Furthermore, by fixing $h = 2$ we studied parameter sensitivity and found ranges of the nuclear membrane parameters (i.e. permeability) and active transport speed which produce oscillatory dynamics. As can be seen in Table 3, these ranges are quite broad and suggest that allowing for a Hill coefficient of 2 could be a generic feature of systems including a nuclear membrane and active transport. Oscillatory dynamics are observed over a wider range of the parameter l . This is due to the fact that active transport moves proteins created close to the cell membrane towards the nucleus, ensuring enough protein accumulates in the nucleus to shut down hes1 mRNA production.

It has been estimated that motor proteins transport cargo along microtubules at a speed of approximately $5.00 \times 10^{-5} \text{cms}^{-1}$ [46]. Our value for the rate of active transport in the second column of Table 2, namely $a = 1.25 \times 10^{-6} \text{cms}^{-1}$, is lower than this estimate, but it should be kept in mind that our value incorporates not only transport but also implicitly incorporates reactions required for such transport, such as binding to and dissociation from microtubules, as well as competition between newly synthesised molecules of Hes1 protein to attach to microtubules. It should also be kept in mind that molecules can become temporarily stuck on microtubules, slowing down the average rate of active transport [46]. A more advanced study of active transport would require consideration of stochastic effects, in which context

Table 2. Parameter values used in the extended Hes1 model and ranges over which sustained oscillatory dynamics are observed

<i>Parameter</i>	<i>Value in simulations</i>	<i>Range over which oscillations are observed</i>
D_{ij}	$3.13 \times 10^{-11} \text{cm}^2 \text{s}^{-1}$	$6.67 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$ to $1.13 \times 10^{-9} \text{cm}^2 \text{s}^{-1}$
α_m	$6.25 \times 10^{-11} \text{Ms}^{-1}$	$\geq 2.50 \times 10^{-12} \text{Ms}^{-1}$
$\hat{p}$	$1.00 \times 10^{-9} M$	$3.17 \times 10^{-11} M$ to $7.69 \times 10^{-7} M$
h	5	≥ 2
μ_m	$1.25 \times 10^{-3} \text{s}^{-1}$	$2.08 \times 10^{-4} \text{s}^{-1}$ to $4.00 \times 10^{-3} \text{s}^{-1}$
α_p	0.0555s^{-1}	$\geq 2.50 \times 10^{-3} \text{s}^{-1}$
μ_p	$1.25 \times 10^{-3} \text{s}^{-1}$	$2.08 \times 10^{-4} \text{s}^{-1}$ to $3.79 \times 10^{-3} \text{s}^{-1}$
D_m	$6.25 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$	$\geq 2.50 \times 10^{-14} \text{cm}^2 \text{s}^{-1}$
D_p	$2.08 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$	$\geq 1.67 \times 10^{-14} \text{cm}^2 \text{s}^{-1}$
d	$1 \times 10^{-5} \text{cm}$	$\leq 4.70 \times 10^{-4} \text{cm}$
a	$1.25 \times 10^{-6} \text{cms}^{-1}$	$7.50 \times 10^{-9} \text{cms}^{-1}$ to $1.08 \times 10^{-4} \text{cms}^{-1}$
l	$6.32 \mu\text{m}$	nuclear membrane ($3 \mu\text{m}$) to $10.7 \mu\text{m}$

Table 3. Nuclear membrane (permeability) and active transport parameter values used in the extended Hes1 model with fixed Hill coefficient $h = 2$ and ranges over which sustained oscillatory dynamics are observed

<i>Parameter</i>	<i>Value in simulations</i>	<i>Range over which oscillations are observed</i>
D_m	$6.25 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$	$\geq 1.38 \times 10^{-13} \text{cm}^2 \text{s}^{-1}$
D_p	$2.08 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$	$\geq 7.92 \times 10^{-14} \text{cm}^2 \text{s}^{-1}$
d	$1.00 \times 10^{-5} \text{cm}$	$\leq 3.70 \times 10^{-4} \text{cm}$
a	$1.25 \times 10^{-6} \text{cms}^{-1}$	$1.92 \times 10^{-9} \text{cms}^{-1}$ to $8.33 \times 10^{-5} \text{cms}^{-1}$

it might be fruitful to apply the Gillespie algorithm [3, 4, 19]. In any case, our range of values for the active transport rate such that sustained oscillatory dynamics occur does include the estimate of $5.00 \times 10^{-5} \text{cms}^{-1}$. Moreover our range of values for the nuclear membrane thickness indicate that the numerical solution is robust to variation in this parameter, which is reassuring as this value is likely to vary between cells.

Fig. 5a shows how the total nuclear concentrations of hes1 mRNA and Hes1 protein vary over time, and Fig. 5b shows how the total cytoplasmic concentrations of hes1 mRNA and Hes1 protein vary over time. By comparing Figs. 5a and 5b with, respectively, Figs. 4 and 5 in [47], we see that oscillatory dynamics are retained in the model when a nuclear membrane and active transport are added to it. Yet there are some quantitative differences between our new plots and those for the original model. For example, a greater proportion of Hes1 enters the nucleus in our new plots, for whereas in [47] the height of the peaks in nuclear Hes1 were only approximately

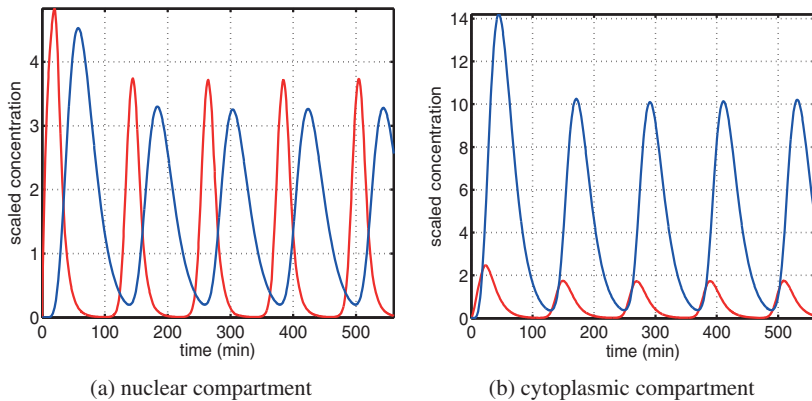


Fig. 5. Plots of the total concentrations (in non-dimensional units) of hes1 mRNA (dashed line) and Hes1 protein (solid line) in (a) the nucleus and (b) the cytoplasm for the extended Hes1 model. The period of oscillation is approximately 120 minutes. Parameter values as per column 2, Table 2

2.5% of those in the cytoplasm, Fig. 5 shows that they are now approximately 33% of those in the cytoplasm. Thus, although our new assumption of slow diffusion across the nuclear membrane hinders the entry of Hes1 into the nucleus, our other new assumption of cytoplasmic active transport of Hes1 is more than enough to overcome this. The increased proportion of Hes1 protein in the nucleus influences the production of hes1 mRNA. To be specific, since Hes1 is a transcription factor which inhibits its own gene, hes1 mRNA production is reduced by the increased proportion of nuclear Hes1 protein. In particular, hes1 mRNA levels in the nucleus drop to zero between consecutive peaks in Figs. 5a, a result not encountered in Fig. 4 in [47].

We have examined the dependence of the nuclear to cytoplasmic ratio of Hes1 on the speed of active transport. Fig. 6 shows 100 different values of a , plotted in increments of $2.08 \times 10^{-7} \text{cms}^{-1}$ (the sixth value, $1.25 \times 10^{-6} \text{cms}^{-1}$ is the default value used in simulations). All other parameter values are found in column 2, Table 2. The nuclear to cytoplasmic ratio is obtained by taking the mean value of the total concentration of Hes1 protein in the nucleus over a 1000 minute time period and dividing it by the mean total value attained in the cytoplasm over the same time period. The plot shows that the nuclear to cytoplasmic ratio of Hes1 protein increases monotonically as a is increased and tends to a limiting value.

Figs. 7a and 7b show respectively how hes1 mRNA and Hes1 protein concentrations vary spatially within the cell from times $t = 150$ minutes to $t = 300$ minutes. At time $t = 150$ minutes, we see that nuclear hes1 mRNA levels are high and that there is also hes1 mRNA concentrated outside the nucleus. The presence of hes1 mRNA in the cytoplasm causes the production by translation of Hes1 protein, which is actively transported towards the nucleus (times 150, 180). When Hes1 reaches the MTOC directly outside the nucleus, it is no longer actively transported but moves by diffusion alone. Hence Hes1 levels build up directly outside the nucleus. This build up is exacerbated by the nuclear membrane, across which Hes1 moves by slow

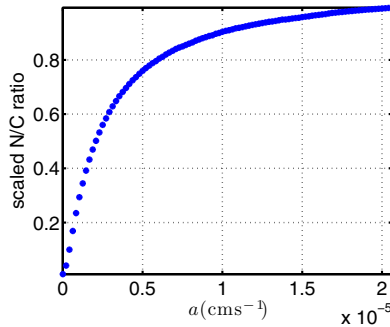


Fig. 6. Graph showing nuclear to cytoplasmic (N/C) ratio of Hes1 protein plotted against a , the active transport speed. The plot shows that the nuclear to cytoplasmic ratio of Hes1 protein increases monotonically as a is increased and tends to a limiting value

diffusion. As levels of Hes1 rise in the nucleus, the transcription of *hes1* mRNA is inhibited (times 180, 210). Without mRNA transcription, no new Hes1 protein can be created by translation. Hence levels of Hes1 fall throughout the cell by natural degradation (times 210, 240). In the absence of Hes1, mRNA transcription is no longer inhibited and this process resumes (time 240 minutes). The cycle just described now repeats, and indeed the oscillatory period of two hours (120 minutes) is clear from comparing times 150 and 180 with times 270 and 300 respectively.

Spatial profiles for the original Hes1 model without an explicit nuclear membrane or active transport can be found in Figs. 6 and 7 in [47]. Although the time information in these figures in [47] does not completely match the times in our new plots (Figs. 7a and 7b), a comparison is possible. The spatial profiles for *hes1* mRNA are not changed qualitatively by our new extensions to the model – the local concentration in the nucleus still reaches a much higher peak than in the cytoplasm. However, the behaviour of Hes1 protein is changed. Instead of building up outside the nucleus as in Fig. 7b, it spreads out across the cytoplasm in the absence of active transport and an explicit nuclear membrane (Fig. 7 in [47]).

2.2.2 Modelling Spatial Effects in the Nucleus

Until now we have assumed that the diffusion coefficients for all species in each compartment are equal. While this assumption helps reduce the number of parameters in the model, it may not be the most accurate approach. For example it is known that proteins experience macromolecular crowding in the nucleus [2], so a different nuclear protein diffusion coefficient may be more appropriate. To this end, in Fig. 8 we present the results of simulations exploring the effect of varying the diffusion coefficients of the molecules in the nucleus. As shown in the plots, changing the diffusion coefficients causes a change in the amplitude and period of the oscillations.

We have also assumed that transcription of *hes1* mRNA occurs throughout the nucleus. However, a more accurate way to model transcription would be to localise

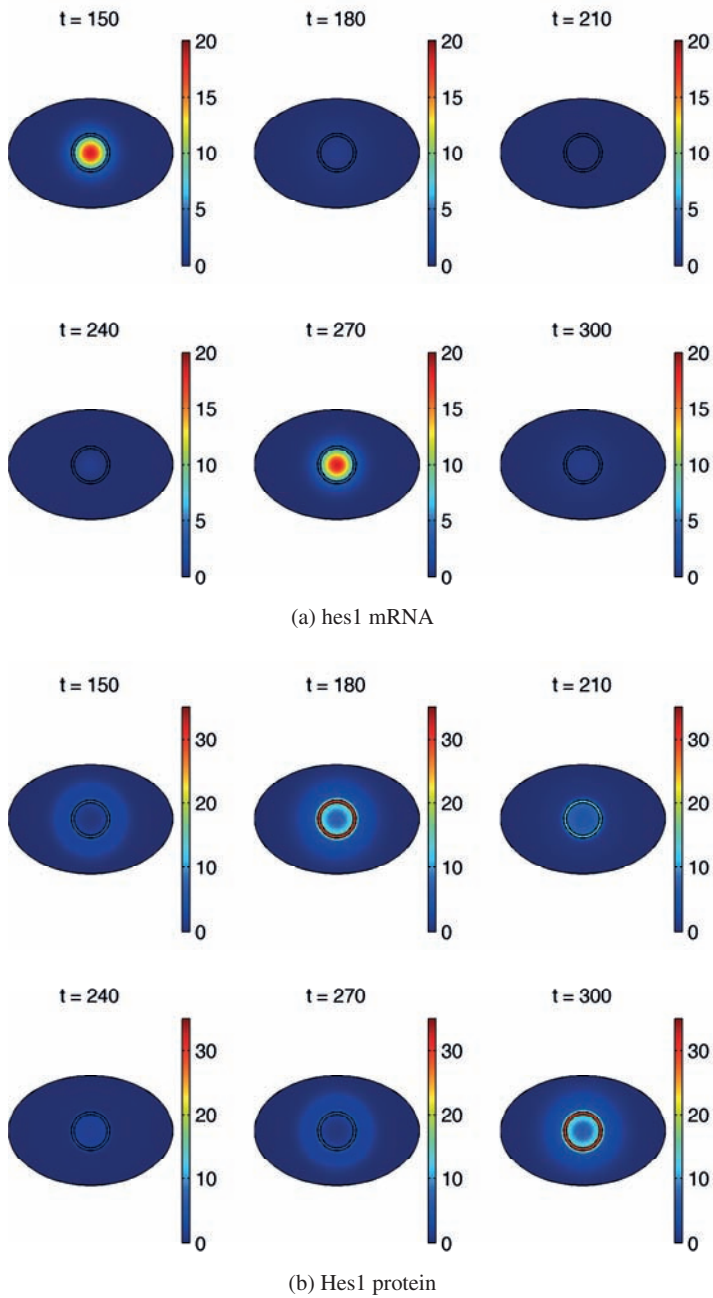


Fig. 7. Plots showing the spatio-temporal evolution of (a) hes1 mRNA and (b) Hes1 protein from times $t = 150$ to $t = 300$ minutes at 30 minute intervals for the extended Hes1 model. The concentrations exhibit oscillatory dynamics in both time and space. Parameter values as per column 2, Table 2

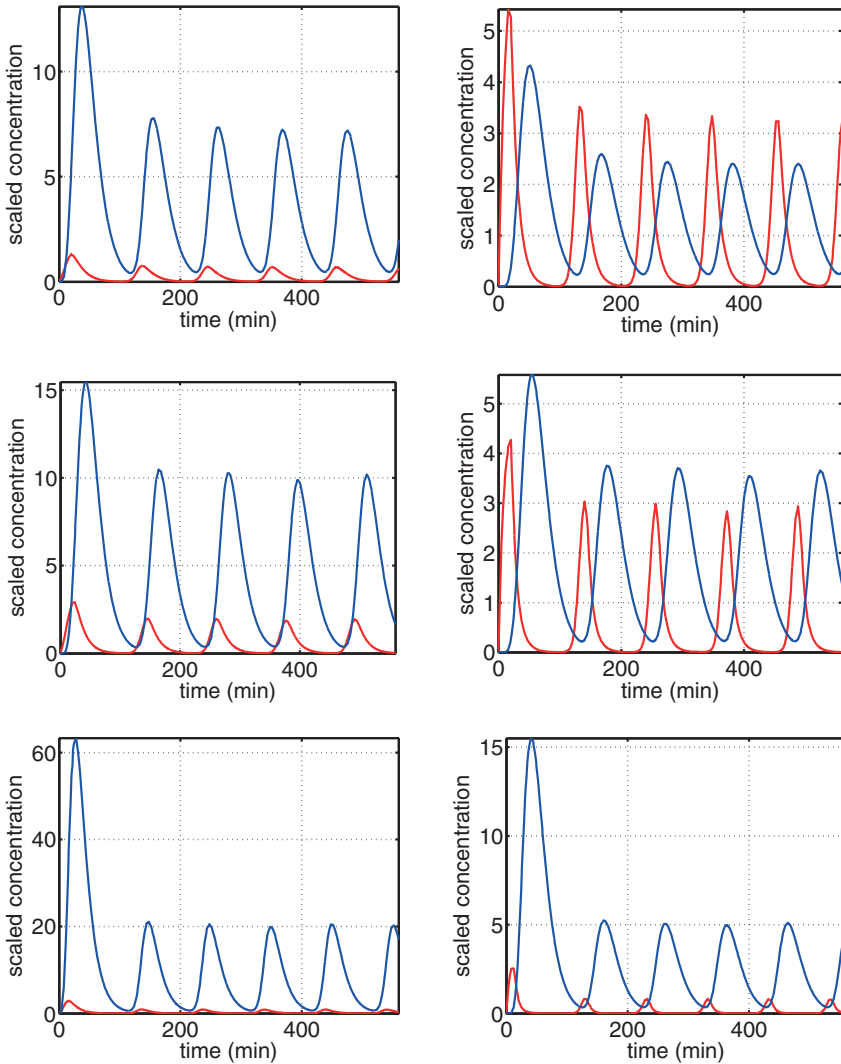


Fig. 8. Plots showing the effect on the concentration profiles of varying the mRNA and protein diffusion coefficients. In each row, the left plot shows the total concentrations in the cytoplasm and the right plot shows the total concentrations in the nucleus (Hes1 protein as solid line, hes1 mRNA as dashed line). Plots in the first row correspond to the case where all four diffusion coefficients are different i.e. $D_{pc} = 3.13 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$, $D_{mc} = 2D_{pc}$, $D_{pn} = 5D_{pc}$, $D_{mn} = D_{pc}/10$. Plots in the second row correspond to the case where the nuclear diffusion coefficients are increased i.e. $D_{pc} = D_{mc} = 3.13 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$ and $D_{pn} = D_{mn} = 10D_{pc}$. Plots in the third row show the result of increasing the diffusion coefficients of mRNA compared with protein diffusion coefficients, i.e., $D_{pc} = D_{pn} = 3.13 \times 10^{-11} \text{ cm}^2 \text{ s}^{-1}$ and $D_{mc} = D_{mn} = 10D_{pc}$. All other parameter values are found in column 2, Table 2

mRNA production to a smaller sub-region of the nucleus. This can be achieved in the model by modifying Eq. (1) as follows:

$$\frac{\partial[m_n]}{\partial t} = D_{m_n} \nabla^2[m_n] + H_2(x, y) \left[\frac{\alpha_m}{1 + ([p_n]/\hat{p})^h} \right] - \mu_m[m_n], \quad (17)$$

where

$$H_2(x, y) = \begin{cases} 1, & \text{if } x^2 + y^2 \leq r^2, \\ 0, & \text{if } x^2 + y^2 > r^2, \end{cases} \quad (18)$$

and where r is the production zone radius. Simulation results from this modified model are presented in Fig. 9. The plots in Fig. 9a show the mRNA concentration in the nucleus over time as we reduce the production zone radius (r) and keep the mRNA production rate (α_m) constant. Oscillatory dynamics are maintained until a critical value of the radius r is reached between $0.90\mu\text{m}$ and $0.49\mu\text{m}$. The lower three plots show that there is a loss of oscillatory dynamics when the production zone is too small. The plots in Fig. 9b show the mRNA concentration in the nucleus over time as we decrease the production zone radius but increase the mRNA production rate (dividing the default value of α_m by the area of the production zone). It is revealed in these plots that oscillatory dynamics can be maintained for smaller values of r .

2.2.3 Convection as the Sole Transport Mechanism of Cytoplasmic Hes1 Protein

Our spatio-temporal modelling approach allows us to address questions which cannot be answered using ODE or DDE models. For example, we can investigate different ratios of active transport and diffusion such that sustained oscillatory dynamics occur in the extended Hes1 model. We begin to investigate this by decreasing the importance of cytoplasmic protein diffusion relative to its active transport. Setting the diffusion coefficient D_{p_c} to zero we were still able to find sustained oscillatory dynamics for a range of active transport rates a . Representative results are shown in Fig. 10 for two different values of a . Consistent with intuition, the plots shown in Fig. 10 show that a greater proportion of protein accumulates in the nucleus as the active transport rate a is increased. Our results suggest that sustained oscillatory dynamics will occur as long as sufficient quantities of Hes1 protein reach the nucleus, regardless of the precise transport mechanism they use to reach it. We leave this result as a prediction of the model since we are not aware of any experiments which can demonstrate this. In the next section, we consider the opposite situation to that considered here, decreasing the importance of active transport relative to diffusion.

2.2.4 Microtubule Disruption

Microtubules are important in a diverse array of cellular functions, ranging from cell division to intracellular trafficking. Consequently microtubule-disrupting drugs are used in cancer therapy and are studied experimentally [9, 18, 21]. Although we

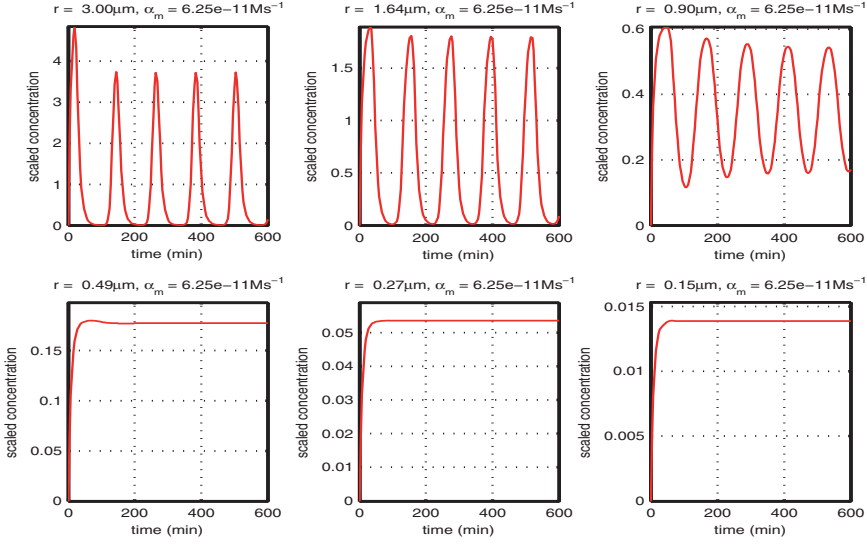
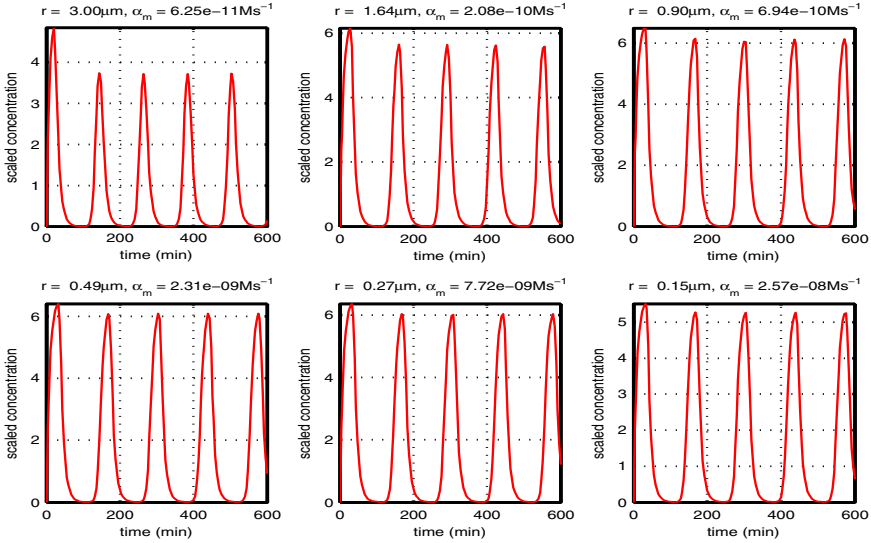
(a) Parameter r varying, constant value of parameter α_m (b) Both parameter r and parameter α_m varying

Fig. 9. Plots showing the effect on mRNA concentration in the nucleus of localising transcription. In these simulations transcription (i.e., mRNA production) is localised to a region in the nucleus defined by Eq. (18). (a) The radius r of the production zone is decreased while the mRNA production rate α_m is kept constant. As can be seen, oscillatory dynamics are present until r becomes too small. (b) The radius r of the production zone is decreased but the mRNA production rate α_m is increased (dividing the baseline value of α_m by the area of the production zone). As can be observed, oscillatory dynamics are present for all values of r . Parameter values as per column 2, Table 2

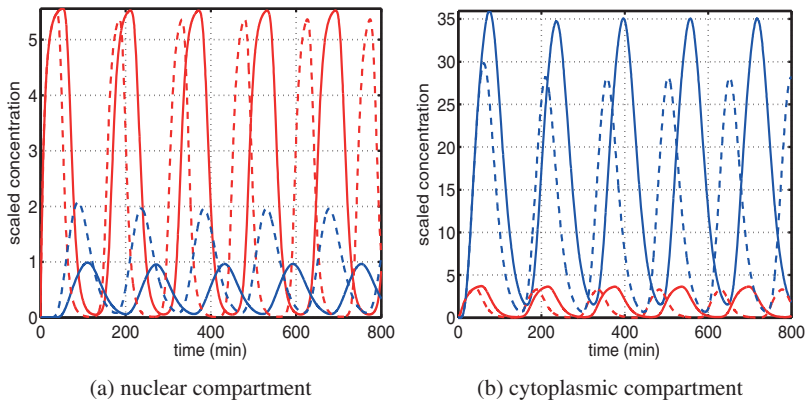


Fig. 10. Plots of the total concentrations of hes1 mRNA and Hes1 protein in (a) the nucleus and (b) the cytoplasm for the extended Hes1 model in the absence of cytoplasmic protein diffusion. Parameter values are as in the second column of Table 2, except that $D_{pc} = 0$. The solid and dashed lines represent Hes1 protein and hes1 mRNA respectively for the case where $a = 1.67 \times 10^{-7} \text{ cm s}^{-1}$ and the dash-dot and dotted lines represent Hes1 protein and hes1 mRNA respectively for the case where $a = 2.50 \times 10^{-7} \text{ cm s}^{-1}$

are not aware of microtubule-disrupting drugs being used on the Hes1 pathway, we shall consider the effect of such drugs in our extended Hes1 model and leave our observations as predictions. Clearly microtubule-disrupting drugs will disrupt active transport along microtubules, so we set the active transport rate a equal to zero in our extended model and otherwise retain the parameter values in the second column of Table 2 (for convenience, the complete set of parameters is stated also in the second column of Table 4). Fig. 11 shows the total concentrations for Hes1 protein and hes1 mRNA over time. The system no longer satisfies the predefined criteria for sustained oscillatory dynamics (at least five distinct peaks in the total concentration of the transcription factor in the nucleus); rather the oscillations are damped. This marks a qualitative change in the dynamics. In a general sense, this is an encouraging result – a qualitative change in dynamics is the type of response we might seek from drug therapy.

Fig. 11 also shows quantitative changes in the dynamics. For example, a smaller proportion of Hes1 now enters the nucleus (relative to Fig. 5) – the total concentration of Hes1 in the nucleus is only roughly 1% of that in the cytoplasm in Fig. 11. This reduction is to be expected since Hes1 is no longer actively transported towards the nucleus.

The damped nature of the oscillations can be seen in Fig. 12, which, like Fig. 7, shows the spatio-temporal evolution of hes1 mRNA and Hes1 protein from times $t = 150$ minutes to $t = 300$ minutes at 30 minute intervals. For the entire 150 minute time interval mRNA levels are high in the nucleus and protein levels are high in the cytoplasm. The nuclear membrane effectively restricts the location of each species to the compartment in which it is produced. Notice that some of the protein

Table 4. Parameter values used in the extended Hes1 model in the case where the active transport rate is set to zero, and ranges over which sustained oscillatory dynamics are observed

Parameter	Value in simulations	Range over which oscillations are observed
D_{ij}	$3.13 \times 10^{-11} \text{cm}^2 \text{s}^{-1}$	$3.33 \times 10^{-11} \text{cm}^2 \text{s}^{-1}$ to $1.46 \times 10^{-9} \text{cm}^2 \text{s}^{-1}$
α_m	$6.25 \times 10^{-11} \text{Ms}^{-1}$	$\geq 6.87 \times 10^{-11} \text{Ms}^{-1}$
$\hat{p}$	$1.00 \times 10^{-9} \text{M}$	$1.05 \times 10^{-9} \text{M}$ to $1.00 \times 10^{-5} \text{M}$
h	5	≥ 6
μ_m	$1.25 \times 10^{-3} \text{s}^{-1}$	$1.25 \times 10^{-4} \text{s}^{-1}$ to $1.21 \times 10^{-3} \text{s}^{-1}$
α_p	0.0555s^{-1}	$\geq 0.0583 \text{s}^{-1}$
μ_p	$1.25 \times 10^{-3} \text{s}^{-1}$	$1.25 \times 10^{-4} \text{s}^{-1}$ to $1.21 \times 10^{-3} \text{s}^{-1}$
D_m	$6.25 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$	$\geq 9.58 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$
D_p	$2.08 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$	$\geq 6.25 \times 10^{-12} \text{cm}^2 \text{s}^{-1}$
d	$1 \times 10^{-5} \text{cm}$	$\leq 8.00 \times 10^{-6} \text{cm}$
a	0	—
l	$6.32 \mu\text{m}$	nuclear membrane ($3 \mu\text{m}$) to $6.24 \mu\text{m}$

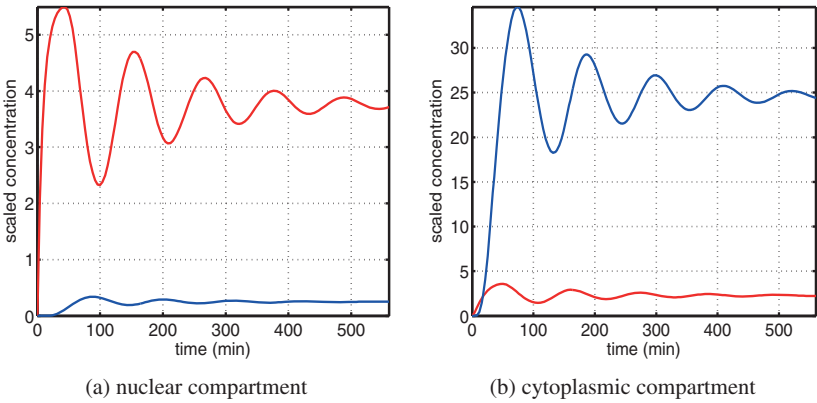


Fig. 11. Plots of the total concentrations of hes1 mRNA (dashed line) and Hes1 protein (solid line) in (a) the nucleus and (b) the cytoplasm for the extended Hes1 model in the absence of active transport. The concentrations exhibit damped oscillations. Parameter values as per column 2, Table 4

has reached the cell membrane, something that was not observed in the model with active transport (see Fig. 7).

Although our set of parameter values (second column, Table 4) in the extended model without active transport did not yield sustained oscillatory dynamics, we found that by varying each parameter individually then such dynamics could occur. Ranges of values for each parameter such that sustained oscillations occur are stated in the third column of Table 4. Note that these ranges are narrower than those presented in Tables 1 and 2 and that, unlike in Tables 1 and 2, they do not contain the

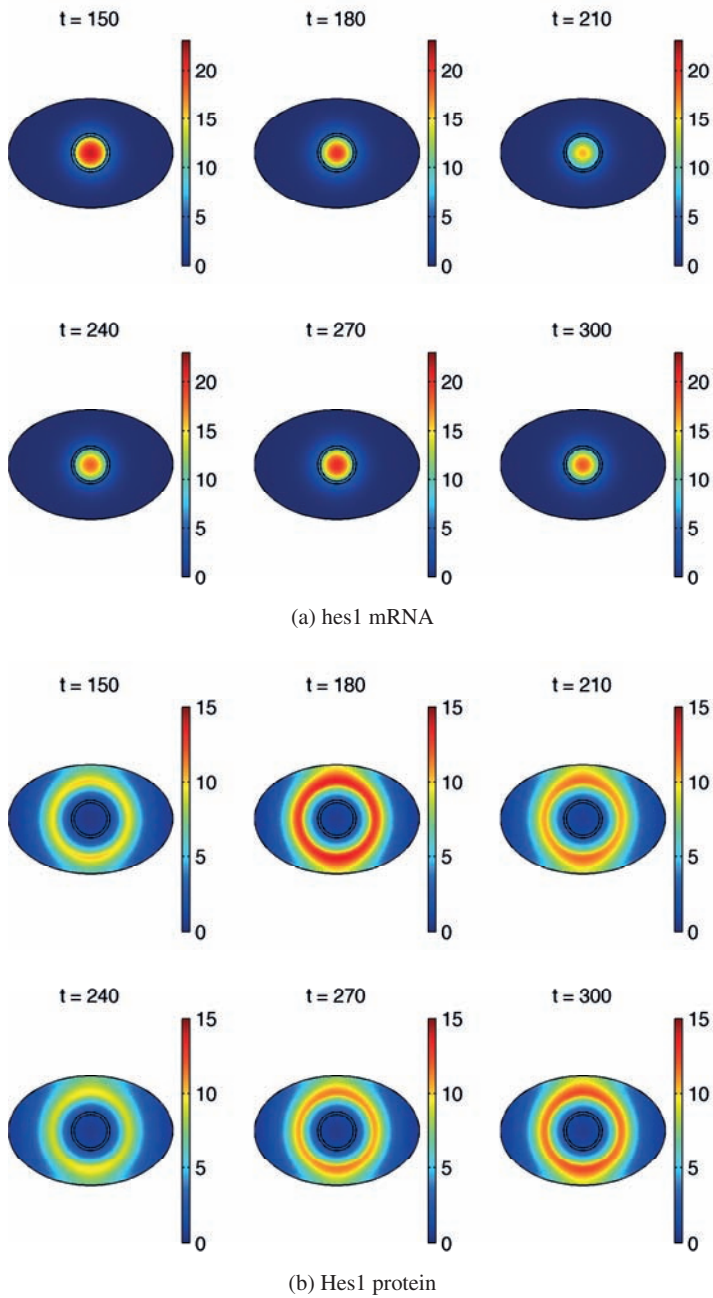


Fig. 12. Plots showing the spatio-temporal evolution of (a) hes1 mRNA and (b) Hes1 protein from times $t = 150$ to $t = 300$ minutes at 30 minute intervals for the extended Hes1 model in the absence of active transport. The concentrations exhibit damped oscillations in time and space. Parameter values as per column 2, Table 4

experimental measurements for the parameters μ_m and μ_p . Furthermore, unlike in Table 2, the experimental measurement for d is not contained in the range for d in Table 4.

2.2.5 Influence of Cell Shape

As mentioned in the Introduction, cell shape can influence intracellular signal transduction [31, 36]. The influence of cell geometry on the Notch-Delta and NF- κ B pathways has recently been investigated by Terry and co-workers, who found through spatio-temporal modelling that oscillatory behaviour in these pathways is to some extent robust to changes in the shapes and relative sizes of the nucleus and cytoplasm [49, 50]. We have performed numerous simulations to study the influence of cell shape on the extended Hes1 model, with parameters as in the second column of Table 2. In general, we found that the qualitative behaviour of the extended Hes1 model is much more robust to variety in cell shape than the quantitative behaviour – oscillatory dynamics can be retained when the domain is altered whilst, for example, the proportion of Hes1 that enters the nucleus will change. Hence, to obtain quantitatively accurate results, we should use a domain that exactly matches a living cell.

We explore the effect of using a realistic cell shape in Figs. 14–15. For this, we have used an image of an osteosarcoma cell since the Hes1 pathway is known to play a critical role in the development of osteosarcomas [54]. The image of the osteosarcoma cell, taken from [13], is shown in Fig. 13a, and the imported domain used for simulations is shown in Fig. 13b. An additional region was added to account for the MTOC.

The realistic cell domain does not change the solution qualitatively – oscillations are evident from the total concentration plots of Fig. 14. However, there are quantitative differences between the total concentration plots in Fig. 14 and those in Fig. 5 (where the only difference in the system being solved is the domain used). For instance, the total concentration of Hes1 protein in both the nucleus and cytoplasm is reduced in Fig. 14 relative to Fig. 5, and there is also a notable reduction in Hes1 protein total concentration compared to hes1 mRNA total concentration in both the nucleus and cytoplasm. Interestingly, the ratio of nuclear to cytoplasmic protein is retained.

Spatial profiles with the osteosarcoma cell domain are presented in Fig. 15. Comparing Figs. 15 and 7 allows us to see why the total protein concentration in Fig. 14 is significantly less than that in Fig. 5. Whilst protein is produced uniformly around the nucleus in Fig. 7, this is not the case in the osteosarcoma cell in Fig. 15 because the nucleus is much nearer to the cell membrane and we have made the assumption that protein is produced a small distance from the nucleus. Protein is mainly produced in the osteosarcoma cell in the areas above and to the left of the nucleus, where there is space for this to occur.

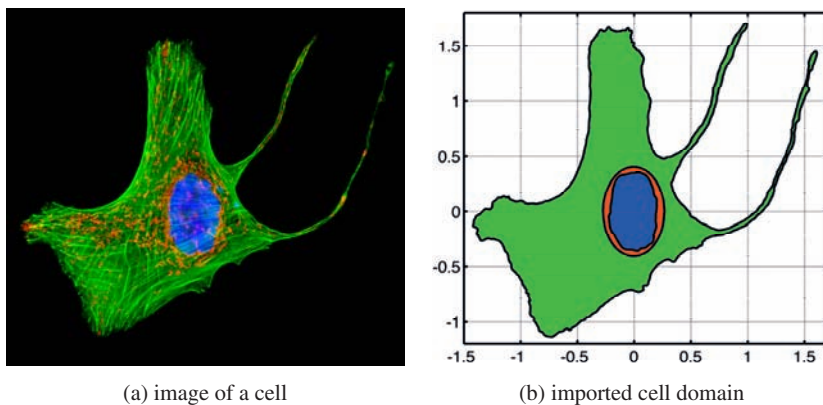


Fig. 13. Images of (a) an osteosarcoma cell (U-2 OS) (reproduced with permission from [13]) and (b) an imported image of this cell with axes displayed in non-dimensional spatial units. (One non-dimensional spatial unit corresponds to $10\ \mu\text{m}$.) A third region directly outside the nucleus was added to the imported domain (shown as white), the outer boundary of which represents the MTOC cf. Fig. 4

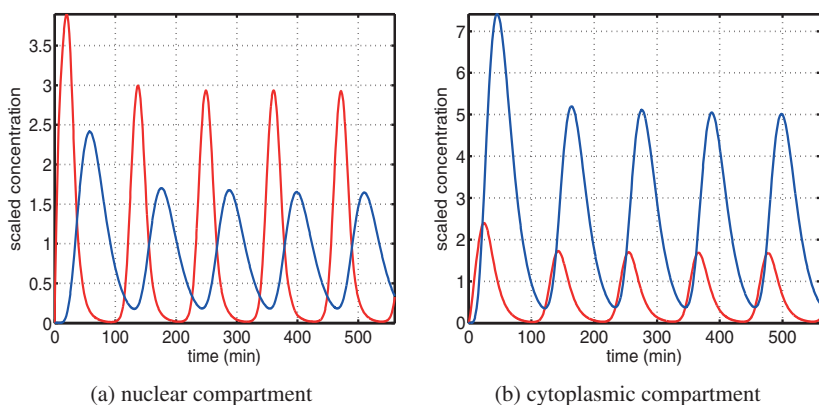


Fig. 14. Plots of the total concentrations of hes1 mRNA (dashed line) and Hes1 protein (dotted line) in (a) the nucleus and (b) the cytoplasm for the extended Hes1 model solved over an osteosarcoma cell domain as shown in Fig. 13b. The period of oscillation is approximately 112.5 minutes. Parameter values as per column 2, Table 2

3 Discussion

In this chapter we have presented two recent spatio-temporal models of the Hes1 intracellular signalling pathway that were originally developed by [47, 48]. Experiments have shown that stimulation of the Hes1 pathway can cause it to exhibit oscillatory dynamics, driven by a negative feedback loop. Therefore we explored numerically our models for this pathway in the context of sustained oscillatory dynamics. We found ranges of values for the model parameters such that sustained oscillatory

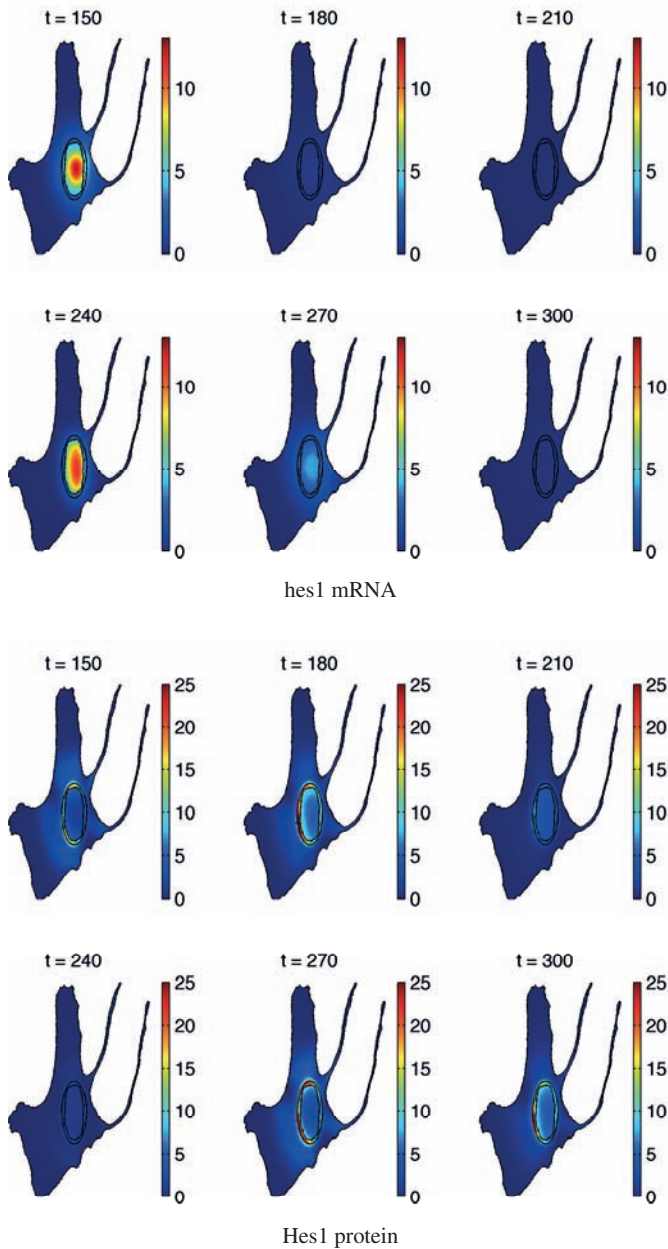


Fig. 15. Plots showing the spatio-temporal evolution of (a) *hes1* mRNA and (b) Hes1 protein within the cell from times $t = 150$ to $t = 300$ minutes at 30 minute intervals for the extended Hes1 model solved over an osteosarcoma cell domain as shown in Fig. 13b. The concentrations exhibit oscillatory dynamics in both time and space. Parameter values as per column 2, Table 2

dynamics occurred, noting that these ranges were consistent with available experimental measurements. This approach to modelling intracellular pathways (i.e. an explicitly spatial approach using PDEs), in particular negative feedback systems, can also be applied to other important systems such as p53-Mdm2 [47, 48] and NF- κ B [49].

Future work will consider extending the models further in several ways. Cells can change shape on the same timescale as oscillatory nuclear-cytoplasmic translocation of Hes1 or p53, and so we may develop a model with a moving boundary on an evolving domain. Based on cell imagery, we will consider more realistic support functions for our translation and active transport terms. We may also study the interactions between different signalling pathways. It is known, for example, that the p53-Mdm2 pathway can co-operate with and antagonise the NF- κ B pathway, which is central to many stressful, inflammatory, and innate immune responses [38, 39]. We are not aware of any spatio-temporal modelling studies of interacting pathways, though there have been temporal studies [40]. There is experimental evidence that molecular movement within a cell can be “subdiffusive” or “superdiffusive” [10, 51, 53], and these are ideas we may investigate from a modelling perspective. Finally, low copy numbers of mRNA and transcription factors can cause stochastic fluctuations in the dynamics of intracellular pathways [44, 55], and we are currently studying this.

Acknowledgements The authors gratefully acknowledge the support of the ERC Advanced Investigator Grant 227619, “M5CGS – From Mutations to Metastases: Multiscale Mathematical Modelling of Cancer Growth and Spread”.

Appendix

Non-dimensionalisation of Hes1 model

We summarise our non-dimensionalisation of the extended Hes1 model (described in Sect. 2.2). The original Hes1 model (described in Sect. 2.1 above) is non-dimensionalised in a similar way – for details, see [47].

To non-dimensionalise the extended Hes1 model given by Eqs. (1)–(4) and (15), subject to the conditions in Eqs. (8)–(14), we first define re-scaled variables by dividing each variable by a reference value. Re-scaled variables are given overlines to distinguish them from variables that are not re-scaled. Thus we can write:

$$[\overline{m}_n] = \frac{[m_n]}{[m_0]}, [\overline{m}_c] = \frac{[m_c]}{[m_0]}, [\overline{p}_n] = \frac{[p_n]}{[p_0]}, [\overline{p}_c] = \frac{[p_c]}{[p_0]}, \overline{t} = \frac{t}{\tau}, \overline{x} = \frac{x}{L}, \overline{y} = \frac{y}{L}, \quad (19)$$

where the right hand side of each equation is a dimensional variable divided by its reference value. From Eq. (19), we can write variables in terms of re-scaled variables and then substitute these expressions into Eqs. (1)–(4) and (15), and into the conditions in Eqs. (8)–(14). This gives a model defined in terms of re-scaled variables which has the same form as the dimensional model but now the parameters

are all non-dimensional. Denoting the non-dimensional parameters with an asterisk, they are related to dimensional parameters as follows:

$$\begin{aligned} D_{ij}^* &= \frac{\tau D_{ij}}{L^2}, \alpha_m^* = \frac{\tau \alpha_m}{[m_0]}, p^* = \frac{[p_0]}{\hat{p}}, \mu_m^* = \tau \mu_m, \alpha_p^* = \frac{\tau [m_0] \alpha_p}{[p_0]}, \\ \mu_p^* &= \tau \mu_p, D_m^* = \frac{\tau D_m}{L^2}, D_p^* = \frac{\tau D_p}{L^2}, d^* = \frac{d}{L}, a^* = \frac{\tau a}{L}, l^* = \frac{l}{L}. \end{aligned} \quad (20)$$

We solve the non-dimensional model using the method described in Sect. 2.1. We simulate the model in COMSOL 3.5a, finding non-dimensional parameter values that yield oscillatory dynamics. We chose the same values as in Eq. (25) in [47] except for those parameters which were new because of our extension to the model. These latter values were chosen as follows: $D_m^* = D_{ij}^*/5$, $D_p^* = D_{ij}^*/15$, $d^* = 0.01$, $a^* = 0.03$, $l^* = 0.63$.

Finally, we calculated the dimensional parameter values. To do this, we needed to estimate the reference values. Since Her1 in zebrafish and Hes1 in mice are both pathways connected with somitogenesis, we used the reference concentrations for Her1 protein and her1 mRNA in [50] as our reference concentrations for Hes1 protein and hes1 mRNA. Thus, we chose $[m_0] = 1.5 \times 10^{-9}\text{M}$ and $[p_0] = 10^{-9}\text{M}$. We assumed a cell to be of width $30\mu\text{m}$. But from Fig. 2 and Fig. 4, the cell width is equal to 3 non-dimensional spatial units or $3L$ dimensional units (using (19)). Hence we set $3L = 30\mu\text{m}$, so that $L = 10\mu\text{m}$. The experimentally observed period of oscillations of Hes1 is approximately 2 hours [16]. Our simulations of the non-dimensionalised model gave oscillations with a period of approximately 300 non-dimensional time units or 300τ dimensional units (using (19)). Hence we set $300\tau = 2\text{hrs} = 7200\text{s}$, so that $\tau = 24\text{s}$. Using our references values and non-dimensional parameter values, we found dimensional parameter values from (20).

Note that we chose our reference time $\tau = 24\text{s}$ based on simulations of the extended Hes1 model since this was our most realistic Hes1 model. For the original Hes1 model and for all special cases of the Hes1 model (for example, setting active transport rates to zero), we retained the reference time $\tau = 24\text{s}$.

References

1. Agrawal, S., Archer, C., Schaffer, D.: Computational models of the notch network elucidate mechanisms of context-dependent signaling. *PLoS Comput. Biol.* **5**, e1000, 390 (2009)
2. Bancaud, A., Huet, S., Daigle, N., Mozziconacci, J., Beaudouin, J., Ellenberg, J.: Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *The EMBO Journal* **28**, 3785–3798 (2009)
3. Barik, D., Baumann, W.T., Paul, M.R., Novak, B., Tyson, J.J.: A model of yeast cell-cycle regulation based on multisite phosphorylation. *Mol. Sys. Biol.* **6**, 405 (2010)
4. Barik, D., Paul, M.R., Baumann, W.T., Cao, Y., Tyson, J.J.: Stochastic simulation of enzyme-catalyzed reactions with disparate timescales. *Biophys. J.* **95**, 3563–3574 (2008)
5. Barrio, M., Burrage, K., Leier, A., Tian, T.: Oscillatory regulation of Hes1: Discrete stochastic delay modelling and simulation. *PLoS ONE* **2**, e117 (2006)

6. Beck, M., Forster, F., Ecke, M., Plitzko, J.M., Melchoir, F., Gerisch, G., Baumeister, W., Medalia, O.: Nuclear pore complex structure and dynamics revealed by cryoelectron tomography. *Science* **306**, 1387–1390 (2004)
7. Bernard, S., Cajavec, B., Pujo-Menjouet, L., Mackey, M., Herzog, H.: Modelling transcriptional feedback loops: the role of Gro/TLE1 in Hes1 oscillations. *Phil. Trans. Roy. Soc. A* **364**, 1155–1170 (2006)
8. Cangiani, A., Natalini, R.: A spatial model of cellular molecular trafficking including active transport along microtubules. *J. Theor. Biol.* **267**, 614–625 (2010)
9. Carbonaro, M., O’Brate, A., Giannakakou, P.: Microtubule disruption targets HIF-1 α mRNA to cytoplasmic P-bodies for translation repression. *J. Cell. Biol.* **192**, 83–99 (2011)
10. Caspi, A., Granek, R., Elbaum, M.: Enhanced diffusion in active intracellular transport. *Phys. Rev. Lett.* **85**, 5655–5658 (2000)
11. Cole, N., Lippincott-Schwartz, J.: Organization of organelles and membrane traffic by microtubules. *Curr. Opin. Cell Biol.* **7**, 55–64 (1995)
12. Cole, C., Scarcelli, J.: Transport of messenger RNA from the nucleus to the cytoplasm. *Curr. Opin. Cell Biol.* **18**, 299–306 (2006)
13. Davidson, M.W.: Micromagnet website. URL <http://micro.magnet.fsu.edu/primer/techniques/fluorescence/gallery/cells/u2/u2cellslarge8.html>. Accessed 26th April 2012
14. Feldherr, C., Akin, D.: Signal-mediated nuclear transport in proliferating and growth-arrested BALB/c 3T3 cells. *J. Cell. Biol.* **115**, 933–939 (1991)
15. Geva-Zatorsky, N., Rosenfeld, N., Itzkovitz, S., Milo, R., Sigal, A., Dekel, E., Yarnitzky, T., Liron, Y., Polak, P., Lahav, G., Alon, U.: Oscillations and variability in the p53 system. *Mol. Syst. Biol.* **2**, E1–E13 (2006)
16. Hirata, H., Yoshiura, S., Ohtsuka, T., Bessho, Y., Harada, T., Yoshikawa, K., Kageyama, R.: Oscillatory expression of the bHLH factor Hes1 regulated by a negative feedback loop. *Science* **298**, 840–843 (2002)
17. Johansson, T., Lejonklou, M., Ekeblad, S., Stålberg, P., Skogseid, B.: Lack of nuclear expression of hairy and enhancer of split-1 (HES1) in pancreatic endocrine tumors. *Horm. Metab. Res.* **40**, 354–359 (2008)
18. Jordan, M.A., Wilson, L.: Microtubules as a target for anticancer drugs. *Nature Rev. Cancer* **4**, 253–265 (2004)
19. Kar, S., Baumann, W.T., Paul, M.R., Tyson, J.J.: Exploring the roles of noise in the eukaryotic cell cycle. *Proc. Natl. Acad. Sci. USA* **106**, 6471–6476 (2009)
20. Karin, M., Cao, Y., Greten, F., Li, Z.: Nf-kappab in cancer: from innocent bystander to major culprit. *Nat. Rev. Cancer* **2**(4), 301–310 (2002)
21. Kavallaris, M.: Microtubules and resistance to tubulin-binding agents. *Nature Rev. Cancer* **10**, 194–204 (2010)
22. Kim, I., Kim, D., Han, S., Chin, M., Nam, H., Cho, H., Choi, S., Song, B., Kim, E., Bae, Y., Moon, Y.: Truncated form of importin alpha identified in breast cancer cells inhibits nuclear import of p53. *J. Biol. Chem.* **275**, 23139–23145 (2000)
23. Klonis, N., Rug, M., Harper, I., Wickham, M., Cowman, A., Tilley, L.: Fluorescence photobleaching analysis for the study of cellular dynamics. *Eur. Biophys. J.* **31**, 36–51 (2002)
24. Lane, D.: p53, guardian of the genome. *Nature* **358**, 15–16 (1992)
25. Lewis, J.: Autoinhibition with transcriptional delay: a simple mechanism for the zebrafish somitogenesis oscillator. *Curr. Biol.* **13**, 1398–1408 (2003)
26. Lomakin, A., Nadezhkina, E.: Dynamics of nonmembranous cell components: Role of active transport along microtubules. *Biochemistry (Moscow)* **75**(1), 7–18 (2010)
27. Marfori, M., Mynott, A., Ellis, J.J., Mehdi, A.M., Saunders, N.F.W., Curmi, P.M., Forwood, J.K., Boden, M., Kobe, B.: Molecular basis for specificity of nuclear import and prediction of nuclear localization. *Biochim. Biophys. Acta Mol. Cell Res.* **1813**(9), 1562–1577 (2011)

28. Masamizu, Y., Ohtsuka, T., Takashima, Y., Nagahara, H., Takenaka, Y., Yoshikawa, K., Okamura, H., Kageyama, R.: Real-time imaging of the somite segmentation clock: revelation of unstable oscillators in the individual presomitic mesoderm cells. *Proc. Natl. Acad. Sci. USA* **103**, 1313–1318 (2006)
29. Matsuda, T., Miyawaki, A., Nagai, T.: Direct measurement of protein dynamics inside cells using a rationally designed photoconvertible protein. *Nature Meth.* **5**, 339–345 (2008)
30. Mendez, V., Fedotov, S., Horsthemke, W.: *Reaction-Transport Systems*. Springer (2010)
31. Meyers, J., Craig, J., Odde, D.: Potential for control of signaling pathways via cell size and shape. *Curr. Biol.* **16**, 1685–1693 (2006)
32. Momiji, H., Monk, N.: Dissecting the dynamics of the Hes1 genetic oscillator. *J. Theor. Biol.* **254**, 784–798 (2008)
33. Monk, N.: Oscillatory expression of Hes1, p53, and NF- κ B driven by transcriptional time delays. *Curr. Biol.* **13**, 1409–1413 (2003)
34. Muller, M., Klumpp, S., Lipowsky, R.: Tug-of-war as a cooperative mechanism for bidirectional cargo transport of molecular motors. *Proc. Natl. Acad. Sci. USA* **105**, 4609–4614 (2008)
35. Nelson, D., Ihekweaba, A., Elliott, M., Johnson, J., Gibney, C., Foreman, B., Nelson, G., See, V., Horton, C., Spiller, D., Edwards, S., McDowell, H., Unitt, J., Sullivan, E., Grimley, R., Benson, N., Broomhead, D., Kell, D., White, M.: Oscillations in NF- κ B signaling control the dynamics of gene expression. *Science* **306**, 704–708 (2004)
36. Neves, S., Tsokas, P., Sarkar, A., Grace, E., Rangamani, P., Taubenfeld, S., Alberini, C., Schaff, J., Blitzer, R., Moraru, I., Iyengar, R.: Cell shape and negative links in regulatory motifs together control spatial information flow in signaling networks. *Cell* **133**, 666–680 (2008)
37. Norvell, A., Debec, A., Finch, D., Gibson, L., Thoma, B.: Squid is required for efficient posterior localization of oskar mRNA during drosophila oogenesis. *Dev. Genes Evol.* **215**, 340–349 (2005)
38. Perkins, N.: Integrating cell-signalling pathways with NF- κ B and IKK function. *Nature Rev. Mol. Cell Biol.* **8**, 49–62 (2007)
39. Pommier, Y., Sordet, O., Antony, S., Hayward, R., Kohn, K.: Apoptosis defects and chemotherapy resistance: molecular interaction maps and networks. *Oncogene* **23**, 2934–2949 (2004)
40. Puszyński, K., Bertolusso, R., Lipniacki, T.: Crosstalk between p53 and NF- κ B systems: pro- and anti-apoptotic functions of NF- κ B. *IET Sys. Biol.* **3**, 356–367 (2009)
41. Rodriguez, M.S., Dargemont, C., Stutz, F.: Nuclear export of RNA. *Biol. Cell* **96**, 639–655 (2004)
42. Sang, L., Collier, H., Roberts, J.: Control of the reversibility of cellular quiescence by the transcriptional repressor HES1. *Science* **321**, 1095–1100 (2008)
43. Seksek, O., Biwersi, J., Verkman, A.: Translational diffusion of macromolecule-sized solutes in cytoplasm and nucleus. *J. Cell. Biol.* **138**, 131–142 (1997)
44. Shahrezaei, V., Swain, P.: The stochastic nature of biochemical networks. *Curr. Opin. Biotechnol.* **19**, 369–374 (2008)
45. Shankaran, H., Ippolito, D., Chrisler, W., Resat, H., Bollinger, N., Opresko, L., Wiley, H.: Rapid and sustained nuclear-cytoplasmic ERK oscillations induced by epidermal growth factor. *Mol. Syst. Biol.* **5**(322) (2009)
46. Smith, D., Simmons, R.: Model of motor-assisted transport of intracellular particules. *Biophys. J.* **80**, 45–68 (2001)
47. Sturrock, M., Terry, A., Xirodimas, D., Thompson, A., Chaplain, M.: Spatio-temporal modelling of the Hes1 and p53-Mdm2 intracellular signalling pathways. *J. Theor. Biol.* **273**, 15–31 (2011)
48. Sturrock, M., Terry, A., Xirodimas, D., Thompson, A., Chaplain, M.: Influence of the nuclear membrane, active transport and cell shape: Insights for spatio-temporal modelling. *Bull. Math. Biol.* pp. DOI 10.1007/s11,538–012–9725–1 (2012)

49. Terry, A., Chaplain, M.: Spatio-temporal modelling of the NF- κ B intracellular signalling pathway: The roles of diffusion, active transport, and cell geometry. *J. Theor. Biol.* **290**, 7–26 (2011)
50. Terry, A., Sturrock, M., Dale, J., Maroto, M., Chaplain, M.: A spatio-temporal model of Notch signalling in the zebrafish segmentation clock: Conditions for synchronised oscillatory dynamics. *PLoS ONE* **6**, e16,980 (2011)
51. Wachsmuth, M., Waldeck, W., Langowski, J.: Anomalous diffusion of fluorescent probes inside living cell nuclei investigated by spatially-resolved fluorescence correlation spectroscopy. *J. Mol. Biol.* **298**, 677–689 (2000)
52. Weis, K.: Regulating access to the genome: nucleocytoplasmic transport throughout the cell cycle. *Cell* **112**, 441–451 (2003)
53. Weiss, M., Hashimoto, H., Nilsson, T.: Anomalous protein diffusion is a measure for cytoplasmic crowding in living cells. *Biophys. J.* **87**, 3518–3524 (2004)
54. Zhang, P., Yang, Y., Zweidler-McKay, P.A., Hughes, D.: Critical role of notch signaling in osteosarcoma invasion and metastasis. *Clin. Cancer Res.* **14**, 2962–2969 (2008)
55. van Zon, J., Morelli, M., Tănase-Nicola, S., ten Wolde, P.: Diffusion of transcription factors can drastically enhance the noise in gene expression. *Biophys. J.* **91**, 4350–4367 (2006)

Understanding Cell Fate Decisions by Identifying Crucial System Dynamics

Dirk Fey, David R. Croucher, Walter Kolch, and Boris N. Kholodenko

Abstract Dysregulation of cell signalling leading to tumourigenic cell fate decisions is a hallmark of cancer. Understanding this dysregulation requires the use of mathematical modelling and analysis. Since mathematical models are, by nature, approximations of the real system, the art of model building is therefore to find the correct level of abstraction. From a systems perspective, approximations result in time varying parameters, which are notoriously difficult to identify. Here, we present an observer based methodology capable of estimating time varying parameters. Importantly, time variant estimates indicate unmodelled dynamics and can be used to identify modelling errors. Strongly time-varying parameters point to modelling errors, and constitute a starting point for model refinements. The methodology is illustrated at the example of receptor tyrosine kinase A mediated activation of protein kinases in a neuroblastoma cell line.

1 Motivation

A hallmark of cancer is dysregulation of pivotal cell fate decisions leading to aberrant proliferation and reduced apoptosis. Healthy cell fate decisions depend on proper sensing of the cell's intra- and extracellular state in a processes called signal transduction. The usual scheme is that receptors sense signals by binding of an extracellular ligand, resulting in conformational changes that trigger the formation of multi-protein complexes and subsequent activation of GTPases and kinases. Hereby, one receptor usually activates several downstream pathways (Fig. 1). A particular cell fate can usually not be attributed to the activity of a single protein in isolation, but rather depends on the context and temporal pattern of the activation and regulating feedback structures within the signalling network [20, 22, 31]. For exam-

D. Fey (✉) · D.R. Croucher · W. Kolch · B.N. Kholodenko (✉)
Systems Biology Ireland, Conway Institute, University College Dublin, Belfield, Dublin 4, Ireland
e-mail: dirk.fey@ucd.ie, david.croucher@ucd.ie, walter.kolch@ucd.ie, boris.kholodenko@ucd.ie

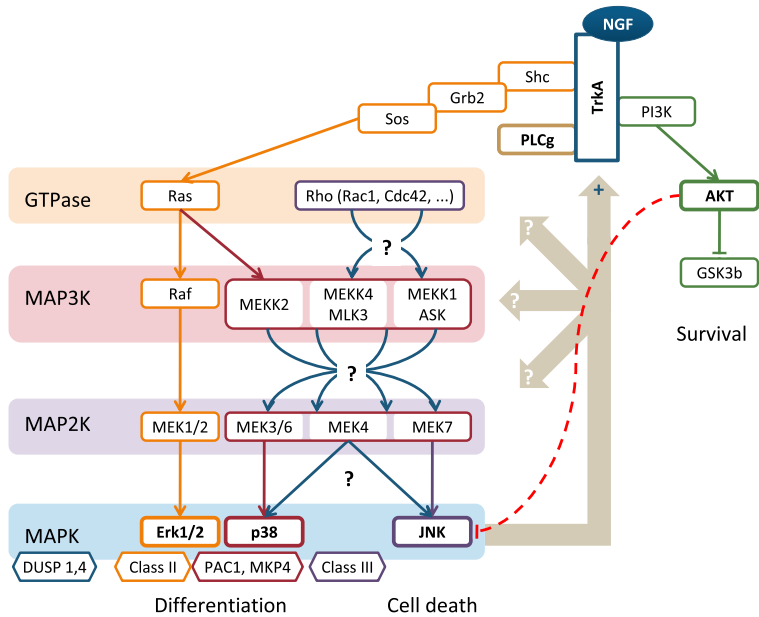


Fig. 1. Signal transduction in a nutshell. After binding to its ligand a receptor (here TrkA) undergoes conformational changes, thus allowing the binding of several adaptor proteins, and triggers the activation of GTPases and kinases. Often the exact cascade of events is poorly understood and includes crosstalk and feedback

ple, treatment of Rat Pheochromocytoma (PC12) cells with epidermal growth factor (EGF) or nerve growth factor (NGF) activates the same signalling cascade (the RAF/MEK/ERK cascade) but has different effects on cell fate. EGF causes transient activation of extracellular regulated kinase (ERK) and proliferation due to negative feedback, whereas NGF causes sustained ERK activation and differentiation due to positive feedback [27, 38]. Because of this complexity, the function of cellular signalling often eludes a naive intuitive understanding, thus calling for the use of mathematical modelling and analysis [17, 18, 24, 25]. Here, we demonstrate how systems modelling and identification can be used to elucidate core processes underlying cellular decision making.

2 Molecular Biology and Systems Theory

Modelling biological systems on the intracellular level has been a research topic for over half a century. All started in 1943, when Erwin Schrödinger gave three talks in Dublin entitled *What is Life* [40, 41]. One of his central thoughts, and at that time a revolutionary idea, was that biological systems follow physical laws. In other words, biological systems can be described by mathematical models. For the control of membrane potential during neuronal excitation, this was achieved in 1952

by Hodgkin and Huxley, who explained and underlined their experimental data with a mathematical model, a key step in understanding how neurons function [15]. A few years later, Denis Noble expanded this model to obtain the first mathematical model of the heart [32]. Nowadays, Hodgkin Huxley models and variants thereof are a vital part of computational neuroscience and widely used in research groups around the world. Part of Hodgkin and Huxley's success relied on the fact that they were able to estimate the model parameters from experimental data. In other areas of molecular biology, such as cell signalling and gene regulation, the identification of mathematical models has proven to be much more challenging.

The advances in biological experimental techniques of the last decades has led to a rapidly growing number of models [28, 30]. In signal transduction, the mitogen activated protein kinase (MAPK) downstream of the epidermal growth factor (EGF) was amongst the first systems to be modelled [16, 21]. From a modelling perspective, cell signalling systems are often composed of a number of similar modules or motifs, such as dimerisation processes or phosphorylation cycles. Biologically these modules are diversely implemented, i.e. composed of different molecular species. As a consequence, one particular module can exhibit very different behaviours, despite having the same interaction pattern. For example, the MAPK pathway consists of phosphorylation dephosphorylation cycles layered in three stages [22]. In nature, this scheme is implemented in several variations, involving different kinases such as ERK, JNK and p38. A model of a MAPK system can exhibit totally different behaviours such as homeostasis, near perfect adaptation and damped or sustained oscillations, depending on the values of the kinetic parameters [20]. This example illustrates the importance of choosing the correct parameters, especially if the model is to be used for predictions.

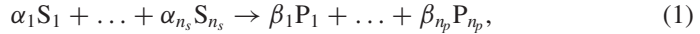
The main bottleneck in obtaining dynamic models of biological systems is the estimation of biological parameters, while structural information like stoichiometry are often known. Unknown parameters can be estimated from time-series data as is common practice in technical applications. Several peculiarities of biological systems hinder a straightforward application of most existing identification methodologies as typically used for technical systems. Biological systems usually have a large number of parameters, though often only a reduced set of experiments are possible, consisting of a few experimental steps and scarce time points. Furthermore, the noise level is usually significant.

Recent years have shown that the control system theoretical viewpoint and approaches are valuable tools for gaining a deeper understanding of biological systems [17, 45]. However, biological systems have particular properties, such as positivity and monotonicity, not often found in technical applications [42, 43]. This requires the design of novel methodologies particularly suited for biological systems. Concerning metabolic pathways, a prominent example is metabolic control analysis for sensitivity analysis developed in the 1970's [9, 14]. Metabolic control theory exploits the fact that enzyme concentrations vary on a much longer timescale than metabolite concentrations and metabolic systems spend most of their time resting in steady state, an assumption not sensible for signalling networks. Therefore, such unified theoretical treatments have not yet been established for signalling net-

works. This is partly due to the fact that signalling systems naturally deal with the temporal integration of ever changing extracellular conditions with the intracellular machinery. As a consequence, the behaviour is often dominated by transient and nonlinear effects. In fact, the nonlinearity of signalling systems often generate the phenomena constituting particular biological functions. Examples are stable limit cycle oscillations in the case of circadian rhythms and cell cycles, or bistable switches in the case of cell fate decisions such as differentiation and apoptosis [6, 29]. From a theoretical perspective, the nonlinearity and behavioural complexity impedes the development of a (unifying) biological theory. As a consequence, the field is diversified with numerous theoretical works treating numerous specialised cases. This diversity is reflected in the recent release of books on the subject, such as *Systems Theory and Systems Biology* edited by Pablo Iglesias and Brian Ingalls [17] or *Systems Modelling in Cellular Biology* edited by Zoltan Szallasi, Jörg Stelling and Vipul Periwal [44] and the dedication of special issues on systems biology in the control community [1, 3]. We can however turn this argument around and exploit the particular form of biological nonlinearities [2, 8, 10–13]. In this spirit, we present an identification methodology that is particularly tailored to biological systems and capable of estimating time-variant parameters.

2.1 Dynamic Modelling

A common framework for the modelling of biochemical reaction networks involves sets of reactions of the following form



where S_i are substrates that are transformed into the products P_i . The factors α_i and β_i are the stoichiometric coefficients of the reactants.

Neglecting spatial and stochastic effects, these reactions are often modelled with systems of ordinary differential equations:

$$\dot{c} = Nv(c, p), \quad (2)$$

where $c \in \mathbb{R}_{\geq 0}^n$ is the vector of concentrations, $p \in \mathbb{R}_{> 0}^{\eta_p}$ the parameter vector and $v \in \mathbb{R}_{\geq 0}^n \times \mathbb{R}_{> 0}^{\eta_p} \mapsto \mathbb{R}_{\geq 0}^m$ the vector of the flows. The stoichiometric matrix $N \in \mathbb{R}^{n \times m}$ depends on the coefficients α_i , β_i and, possibly on factors compensating different units or volumes. For a more detailed introduction, see for example [19, 26, 34].

There is a large variety of possible reaction models [4]. Here, we restrict the reaction models to the most common ones in signalling networks:

- *mass action*: the flow is proportional to each substrate: $v = k \prod_{i \in I} c_i$ where I is a subset of $1, \dots, n$ with possibly repeated entries;
- *power law, S-Systems, generalised mass action*: The flow is polynomial in the substrates $v = k \prod_{i \in [1, n]} c_i^{\alpha_i}$;
- *Michaelis-Menten or Monod*: For low substrates the flow v depends linearly on the substrate s and saturates for large substrate concentrations at $V_{\max}$: $v = V_{\max} \frac{s}{K_M + s}$. At a substrate concentration of K_M , the flow is half the maximum rate;

- *Hill*: the flow is sublinear for low substrate and saturates for large substrate concentrations at $V_{\max}$: $v = V_{\max} \frac{s^h}{K_M^h + s^h}$. The exponent h is larger than one and at a substrate concentration of K_M , the flow is half the maximum rate.

In biochemical reaction modelling, the stoichiometry is usually known, as is the type of reaction kinetics, in contrary to the often quite uncertain parameters. Thus, the identification problem can be formulated as follows:

- *given*: the stoichiometric matrix N and the form of the function $v(c, p)$ describing the reaction rates;
- *unknown*: the kinetic parameters p .

Before we present the identification methodology, we discuss the effect of model approximations as models are by nature approximations of *real* systems.

2.2 Model Approximations and Time-Varying Parameters

Dynamic models are mathematical descriptions of real processes. By nature, formalised descriptions are always approximate. The only model capturing all aspects of a real system is the system itself, or as Rosenblueth and Wiener expressed it “The best material model of a cat is another cat, or preferably the same cat” [37]. Nevertheless, mathematical models are helpful for understanding the behaviour of systems, especially if the underlying system is complex. The art of model building is therefore not to find the most accurate description, but the most helpful one. We have, therefore, to identify the core processes generating the considered behaviour. The process of identifying such a core model, awards us with an (basic) understanding of system that can be deepened by using mathematical analysis.

Complex models can be very accurate in the sense that simulating them replicates experimental data well. In contrast, simple models are usually associated with large quantitative discrepancies. From a theoretical perspective, matching a simplified model to a trajectory generated by a complex system, will result in time-varying parameters. We illustrate this fact using the system in Fig. 2a. The input u regulates the formation and activity of several intermediate signalling molecules and complexes $Z_1, \dots, Z_n$. The output y is the intermediate molecule Z_n , which acts as enzyme for the phosphorylation of a downstream kinase X . Using Michaelis Menten kinetics, the rate of X -phosphorylation is

$$r_{\text{phos}} = k_{\text{phos}} z_n \frac{x}{K_1 + x}. \quad (3)$$

Assuming that the concentration of the intermediate molecules $z_1, \dots, z_n$ are governed by a dynamic system, then z_n is a time varying variable, whose time course depends not only on the initial conditions, but also the input u . We can find an approximate description of the rate r_{phos} in terms of the input by neglecting the dynamics of the intermediate variables. Assume for simplicity that the intermediate variables

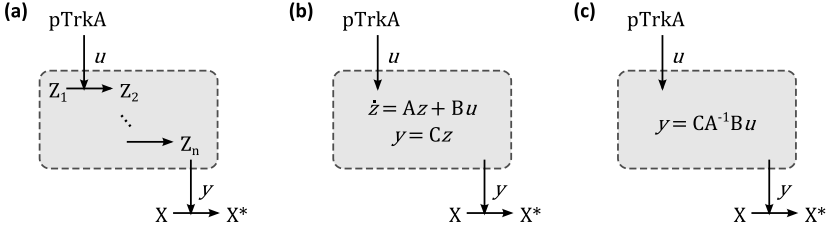


Fig. 2. Conceptual scheme of TrkA signalling approximations. Phosphorylated TrkA (pTrkA) activates downstream kinases X via a system intermediates (grey box). (a) Most accurately, the intermediate system consists of a system of (unspecified) biochemical reactions. (b) Approximating the intermediate dynamics with a system of linear differential equations. (c) Quasi steady state approximation of the scheme in (b)

are described by a linear system of ordinary differential equations (Fig. 2b)

$$\dot{z} = Az + Bu, \quad (4)$$

$$y = Cz. \quad (5)$$

Formally, we obtain a quasi steady state approximation by setting $\dot{z} = 0$ and solving for z (Fig. 2c)

$$z = A^{-1}Bu \Rightarrow y = CA^{-1}Bu. \quad (6)$$

Substituting (6) into (3) gives an approximate expression for the phosphorylation rate

$$r_{\text{phos}} \approx \underbrace{k_{\text{phos}} CA^{-1}Bu}_{=k_1} \frac{x}{K_1 + x}. \quad (7)$$

For constant inputs, (6) describes the steady state of the linear intermediate system. Thus, for a linear intermediate system in steady state, (3) is exact. In the dynamic case, the newly defined parameter k_1 is time varying and approaches the value $k_1 = CA^{-1}B$ for any input trajectory approaching steady state $u(t) \xrightarrow{t \rightarrow \infty} u$.

The above considerations motivate the following (oversimplified) model of kinase phosphorylation

$$\dot{x}^* = k_1 u \frac{x}{K_1 + x} - k_2 \frac{x^*}{K_2 + x^*}, \quad \text{where } x = 1 - x^*, \quad (8)$$

whereby we used the conserved moiety $x_{\text{tot}} = x + x^*$ and normalised concentrations $x_{\text{tot}} = 1$. Because the model neglects the dynamics of the intermediate species, we can imagine that trying to mimic experiments in simulations would perform rather poorly. However, this judgement is based on the assumption of constant parameters.

A consequence of neglecting the intermediate dynamics is that the parameter k_1 is time-varying. Solving the system of linear differential equations (4) gives the time

courses of the intermediates

$$z(t) = e^{At} \int_0^t e^{-A\tau} Bu(\tau) d\tau,$$

where we assumed zero initial conditions $z_i(0) = 0$. By substituting into (3) and comparing to the simplified model (7) we see that the parameter k_1 is a function of time

$$k_1(t) = Ce^{At} \int_0^t e^{-A\tau} Bu(\tau) d\tau \frac{k_{\text{phos}}}{u(t)}.$$

Thus, trying to mimic experiments by simulating (8) would perform rather poorly and advanced identification techniques capable of estimating time-variant parameters are needed. Observers provide such a technique.

3 Parameter Estimation with Observers

The parameter estimation problem is closely related to the state estimation problem as both consider the estimation of unknown quantities. In loose terms, an observer is a mathematical system that estimates internal, non-measured states. Observer based approaches to parameter estimation require a certain system extension. Assuming that the parameters are constant, we can formally extend the state space with the parameters, i.e.

$$\begin{bmatrix} \dot{c} \\ \dot{p} \end{bmatrix} = \begin{bmatrix} f(c, p) \\ 0 \end{bmatrix}. \quad (9)$$

Note that the assumption $\dot{p} = 0$ is a formal construct that enables us to formulate the problem mathematically. The same approach can be taken in order to estimate time varying parameters. Then the assumption $\dot{p} = 0$ corresponds to a time scale separation of system dynamics and time varying parameters.

Given the above system, an observer can achieve a combined state and parameter estimation. However, designing observers for system (9) carries two difficulties: parameter dependency and nonlinearity. The parameter dependency triggers observability issues; for example, linearisations of (9) are generally not observable in steady state [8]. The nonlinearity of the problem means that the observer error depends on unknown states. For linear systems, the so called separation principle states that the dynamics of observer error only depend on the error itself and measured outputs [46]. Unfortunately, there is no separation principle for nonlinear systems. As a consequence, global convergence for all $p \in \mathcal{R}$ can generally not be achieved [5, 7]. An alternative to the state space extension with $\dot{p} = 0$ is the transformation into parameter free coordinates [2, 8, 10, 13]. The transformation exploits particular properties of biochemical reaction systems and facilitates the observer design. The entire identification procedure can be structured into three steps:

1. transformation of the system of ordinary differential equations into a parameter independent form;
2. estimation of all states in the parameter free coordinates using an observer;
3. back transformation to obtain the parameters.

In the following sections, we address each step in detail.

3.1 Transformation into Parameter-Free Coordinates

This section presents the state space extension transforming the system into parameter free coordinates. The approach is first illustrated for a system with a single reaction, before the general extension scheme is presented.

Example 1. Let us consider the following system with $k > 0$, $K > 0$

$$\dot{c} = -v(c), \quad (10a)$$

$$v(c) = k \frac{c^2}{c + K}. \quad (10b)$$

Assuming that c and therefore v are positive, it is possible to derive the differential equation for the relative rate of change of the reaction rate, in essence taking the logarithm and time derivative of (10b). Before doing that, we introduce the new state

$$M = c + K$$

with the derivative

$$\dot{M} = \dot{c}.$$

Now, taking the logarithm and time derivative of (10b) gives

$$\begin{aligned} \frac{\dot{v}}{v} &= \frac{d}{dt} \log v = \frac{d}{dt} (\log k + 2 \log c - \log M) \\ &= 2 \frac{\dot{c}}{c} - \frac{\dot{M}}{M}. \end{aligned}$$

Substituting $\dot{c} = -v$ yields the extended system

$$\begin{aligned} \dot{c} &= -v \\ \dot{M} &= -v \\ \dot{v} &= v \left(-2 \frac{v}{c} + \frac{v}{M} \right) \end{aligned}$$

for which the right-hand-side is parameter free.

As the example illustrates, the states of the parameter free extended system consists of the concentrations c , the denominators of the reaction rates M and the reaction rates v .

In general, the approach considers reaction kinetic systems

$$\dot{c} = Nv(c, p, u), \quad (11a)$$

allowing for fluxes of the form:

$$v_i = k_i \prod_j \frac{c_j^{\alpha_{ij}} u^{\alpha_{u,ij}}}{M_{ij}} \quad \text{where} \quad M_{ij} = c_j^{\eta_{ij}} + K_{ij}, \quad (11b)$$

and where $K_{ij} > 0$, $\eta_{ij} \geq 0$ and $\alpha_{ij}, \alpha_{u,ij} \geq 0$ denote known structural parameters and $u \geq 0$ is a vector of inputs representing concentrations of unmodelled upstream components. If $\eta_{ij} = 0$, then the arbitrary parameter K_{ij} shall be equal to 1. The general formulation of (11b) contains mass action kinetics, generalised mass action kinetics, Michaelis-Menten- and Hill-kinetics as well as their products. For example, setting $\eta_{ij} = 0$ leads to a mass action model, whereas setting $\alpha_{ij} = \eta_{ij} = 1$ gives Michaelis-Menten kinetics.

For $0 < \alpha_{ij} < 1$, the flux v_i is not Lipschitz in $c_j = 0$. To ensure the existence and uniqueness of solutions, we assume that all concentrations are strictly positive.

Assumption 1. The parameters p and the concentrations c are strictly positive along trajectories of (9) and bounded, i.e. $0 < \underline{\delta} \leq c_i(t, c_0) \leq \bar{\delta} < \infty$ holds for all species i and all initial conditions c_0 for some positive constants $\underline{\delta} < \bar{\delta}$.

This condition is satisfied in many biological applications, in particular for models of metabolic pathways.

To simplify the presentation, define the following matrix-valued function $M: \mathbb{R}_{\geq 0}^n \rightarrow \mathbb{R}^{m \times n}$

$$M_{ij} = K_{ij}^{\eta_{ij}} + c_j^{\eta_{ij}}. \quad (12)$$

Then the mapping

$$\Theta : \begin{bmatrix} c \\ p \end{bmatrix} \mapsto \begin{bmatrix} c \\ M(c, p) \\ v(c, p) \end{bmatrix} \quad (13)$$

is diffeomorph if Assumption 1 holds, defining a smooth and bijective state-space transformation of the original system (9) into an equivalent extended system that is parameter free. This means, considering M_{ij} and v_i as additional states, complementing the natural states c_j , results in ordinary differential equations that do not depend on the parameters. The transformed system is

$$\dot{c} = Nv, \quad (14a)$$

$$\dot{M}_{ij} = \eta_{ij} c_j^{\eta_{ij}-1} e_j^T Nv, \quad (14b)$$

$$\dot{v} = \text{diag}(v) \left(\alpha (\text{diag}(c))^{-1} Nv + \alpha_u (\text{diag}(u))^{-1} \dot{u} - \tilde{m} \right), \quad (14c)$$

where

$$\tilde{m}_i = \sum_j \frac{\eta_{ij} c_j^{\eta_{ij}-1} e_j^T N v}{M_{ij}}.$$

Summarising, any biochemical reaction model consisting of flows modelled as in (11b) can be transformed into a system that is free of parameters. In parameter free coordinates, the system is described using an extended state vector and depends only on structural properties of the original system.

The extended state is denoted by $x \in \mathbb{R}_{\geq 0}^{n_x}$

$$x = \begin{bmatrix} c \\ m \\ v \end{bmatrix}, \quad (15)$$

where m is the vector of all non-zero entries of

$$\text{vect } M = [M_{11} \cdots M_{m1} \ M_{12} \cdots M_{mn}]^T.$$

Assuming that the possibly time dependent inputs $u_c = u_c(t)$ are differentiable, the dependence of the extended system on time derivatives of u_c does not pose a problem. We can define $u^T = [u_v^T \ u_c^T \ \dot{u}_c^T]$ and write the system compactly as

$$\dot{x} = f(x, u) = \begin{bmatrix} N v \\ f_M(c, v) \\ f_v(c, M, v, u) \end{bmatrix}. \quad (16a)$$

Remark 1. Step and pulse inputs can be handled either using a differentiable approximation in the form of steep sigmoidal functions e.g. $u_c = A_0(1 + \tanh(t - T_0))$ or by changing the initial condition accordingly and setting $\dot{u}_c = 0$.

To simplify the observer design, we introduce the assumption that the output is a subset of concentrations flows. This is the case in many biological applications.

Assumption 2. The output $y(t) \in \mathbb{R}^{n_y}$ is a subset of concentrations c and flows v :

$$y = h\left(\begin{bmatrix} c \\ m \\ v \end{bmatrix}\right) = \begin{bmatrix} H_c & 0 & 0 \\ 0 & 0 & H_v \end{bmatrix} \begin{bmatrix} c \\ m \\ v \end{bmatrix}, \quad (16b)$$

where the matrices H_c and H_v are composed of columns of the identity matrices.

If some parameters values are already known, the proposed methodology can be adjusted in a straightforward way to not estimate them again. There are basically two cases. First, the parameter is a Hill or Michaelis-Menten constant K_{ij} . Then, there exists a state M_{ij} , which depends on K_{ij} and on some concentrations. This state can therefore be expressed as an algebraic equation of other states and does not require a differential equation. In the second case, the parameter is proportional to a flow, i.e. k_i in a flow v_j . This flow then also contains no unknown, only other states and thus its differential equation can be replaced by an algebraic equation. The reduced

extended system is a differential algebraic system of index one. The algebraic equations can easily be eliminated, thus reducing the state space dimension by the number of known parameters.

3.2 Dissipative Observer

Generally speaking an observer is an algorithm estimating internal states based on output measurements. Usually, observers consists of a copy of the system's equations and a correction term feeding back the estimation error, thus pushing the simulated trajectory towards the true trajectory [39]. In particular, the parameter free system (16a) facilitates the design of a dissipative observer [10, 36]. Let

$$A = \begin{bmatrix} 0 & 0 & N \\ 0 & 0 & 0 \end{bmatrix}, G = \begin{bmatrix} 0 \\ I \end{bmatrix} \text{ and } \Psi(x, u) = \begin{bmatrix} f_M(x, u) \\ f_v(x, u) \end{bmatrix},$$

then the parameter free system 16a writes as

$$\dot{x} = Ax + G\Psi(x, u), \quad (17a)$$

$$y = Cx, \quad (17b)$$

which is Lure's discretion into linear and nonlinear parts. A dissipative observer for (17) is

$$\dot{\xi} = A\xi + G\tilde{\Psi}(\xi + N \cdot (v - y), u) + L \cdot (v - y), \quad (18a)$$

$$v = C\xi, \quad (18b)$$

where the so called observer gain matrices N and L have to satisfy the linear matrix inequality (LMI)

$$\begin{bmatrix} -\frac{1}{\delta^2}I & I + NC & 0 \\ (I + NC)^T & P(A + LC) + (A + LC)^T P + \epsilon I & PG \\ 0 & G^T P & -I \end{bmatrix} \preceq 0, \quad (19)$$

and the observer nonlinearity $\tilde{\Psi}$ has to be constructed such that the mapping $\Phi(z) : z \mapsto \tilde{\Psi}(x, u) - \Psi(x + z, u)$, where $\xi = x + z$, satisfies the Lipschitz condition

$$\|\Phi(z, \sigma)\|_2 \leq \delta \|z\|_2.$$

A simple construction for the observer nonlinearity is

$$\hat{\Psi}_i(x + z) = \begin{cases} \Psi_{\min} & \text{if } \Psi_i(x + z) > \Psi_{\min} \\ \Psi_i(x + z) & \text{if } \Psi_{\min} \leq \Psi_i(x + z) \leq \Psi_{\max} \\ \Psi_{\max} & \text{if } \Psi_i(x + z) < \Psi_{\max} \end{cases},$$

where $\Psi_{\min}$ and $\Psi_{\max}$ respectively are lower and upper bounds on the elements of Ψ on the true trajectory $x(t, x_0)$, i.e. it holds that

$$\Psi_{\min} \leq \Psi(x(t, x_0)) \leq \Psi_{\max}.$$

3.3 Obtaining the Parameter Estimate

The last step is the back transformation into original coordinates, which gives the actual parameter estimate. Inversion of (13) yields an explicit expression. In particular, using (12), the parameters $K_{ij}(t)$ can be estimated via

$$\hat{K}_{ij}(t) = \begin{cases} \left(\hat{M}_{ij}(t) - \hat{c}_j(t) \right)^{1/\eta_{ij}} & \text{for } \eta_{ij} > 0, \\ 1 & \text{for } \eta_{ij} = 0. \end{cases} \quad (20a)$$

Finally, the estimation of the parameters $k_i(t)$ is possible using (11b)

$$\hat{k}_i(t) = \hat{v}_i(t) \prod_j^{n_c} \frac{\hat{M}_{ij}(t)}{\hat{c}_j(t)^{\alpha_{ij}} u(t)^{\alpha_{u,ij}}}. \quad (20b)$$

Because the observer has to be initialised with an unknown initial condition, the parameter estimate is time dependent. It converges to the true, constant values if and only if the observer converges.

3.4 Solutions to Issues Arising in Praxis

Existence of the observer (18) depends on the feasibility of the LMI (19). A necessary condition is that the pair (A, C) is observable [10]. Unfortunately, this condition requires that Hill variables and linearly dependent rates are measured, which can not be achieved in praxis. We can overcome this problem by assuming that some parameters are known and create fake measurements as follows.

- If the K_{ij} are known and $y_c(t) = c(t)$ is measured, we can create fake measurements of the Hill variables:

$$y_M(t) = M(c(t), p) = M(y_c(t), K),$$

where $K = [\dots, K_{ij}, \dots]$ is known and the elements of M are given by

$$M_{ij} = c_j(t)^{\eta_{ij}} + K_{ij}^{\eta_{ij}} = y_{c,j}(t)^{\eta_{ij}} + K_{ij}^{\eta_{ij}}.$$

- If for some i the k_i and K_{ij} are known and $y_c(t) = c(t)$, is measured, we can create fake measurements of the fluxes:

$$y_{v,i}(t) = v_i(c(t), p) = v_i(y_c(t), p_i),$$

where the vector p_i contains the known parameters $p_i = [k_i, K_{i1}, K_{i2}, \dots]$.

The following section demonstrates the power of the observer based approach in praxis.

4 Application to TrkA Induced MAPK Signalling

Signalling via the neurotrophin receptor TrkA is known to be involved in embryonal formation of the neural system through a developmentally controlled expression pattern facilitating a process of neurotrophism and terminal differentiation [35]. Interestingly TrkA is also expressed in neural-derived tumours such as neuroblastoma, an embryonal tumour which arises in the para-vertebral sympathetic ganglion and adrenal medulla. Neuroblastoma tumours expressing high levels of TrkA have an exceptionally good prognosis, with the tumour frequently undergoing spontaneous regression, however the mechanism behind this regression is poorly understood, but should be linked to the regulatory machinery downstream of TrkA. Activation of the TrkA receptor by stimulation with the nerve growth factor (NGF) activates several signalling pathways and downstream kinases involved in cell fate decisions. The ability of TrkA to induce either proliferation, apoptosis or differentiation in vitro depending on the cell type and experimental design suggests a remarkable plasticity within the TrkA signalling network. In order to decipher the regulatory machinery behind the multimodal response, we utilise a combination of experimentation, dynamic modelling and systems identification.

We model the activation of measured TrkA downstream kinases as simple phosphorylation-dephosphorylation cycles of the form

$$\begin{bmatrix} \dot{x} \\ \dot{x}^* \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \text{with} \quad \begin{aligned} v_1 &= k_1 u \frac{x}{K_1 + x}, \\ v_2 &= k_2 \frac{x^*}{K_2 + x^*}, \end{aligned} \quad (21)$$

where u denotes the level of phosphorylated TrkA receptor and x, x^* denote the level of unphosphorylated and phosphorylated kinase, respectively. For each kinase, the parameters k_1, k_2, K_1, K_2 present unknown quantities that have to be identified from experimental data. In particular, we measured timecourses of phosphorylated TrkA, Akt, Erk, JNK, PLC γ and p38 in a TrkA inducible neuroblastoma SY5Y cell line in response to treatment with 1, 10 and 100nM NGF using western blotting (Fig. 3).

The application of the observer based approach requires that several practical issues are overcome. First, the observer requires time-continuous measurements, which we solve by curve fitting the sampled data. Using the Matlab curve fitting toolbox, we fit

$$u(t) = a((1 - \exp^{-t/b}) \exp^{(-t/c)})^2 + d * t^n / (e^n + t^n),$$

to the phosphorylated TrkA data, since it captures the bimodal TrkA response nicely and smoothing splines to all other kinases (Fig. 3). Second, the observer requires that some parameters are known in order to create fake measurements for Hill vari-

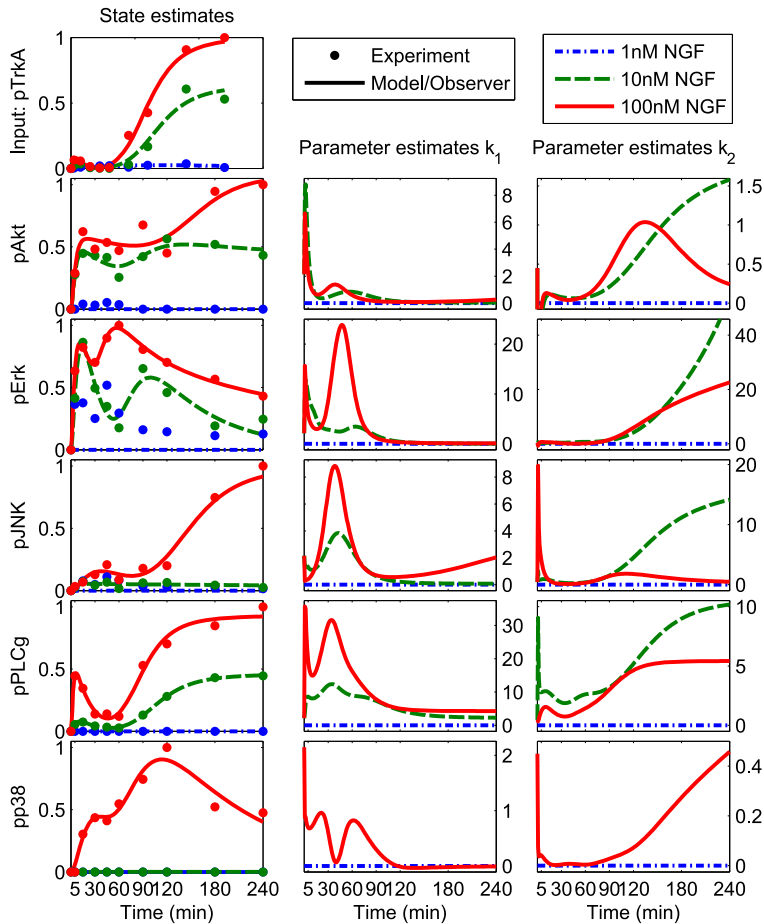


Fig. 3. Time varying state and parameter estimates of the observer. The phosphorylation k_1 and dephosphorylation k_2 estimates were obtained in two different runs of the observer as explained in the main text. Because the 1nM (dash-dotted) dose did not produce detectable phosphorylation levels for most kinases, only the 10nM (dashed) and 100nM (solid) responses were estimated. Further, no estimate was obtained for the p38 10nM response, because phosphorylation levels of p38 in response to 10nM NGF were not detectable in western blots. (a) Dots indicate measurements (quantification of western blots using imageJ), lines represent state estimates of the observer. (b) Estimates of k_1 . (c) Estimates of k_2

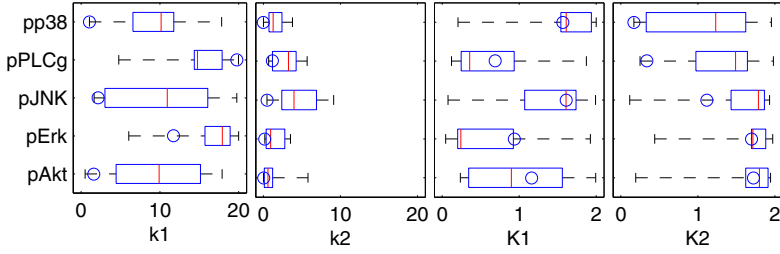


Fig. 4. Parameter estimates of the genetic algorithm. Boxplots represent the distributions of final parameter estimates obtained from several runs of the genetic algorithm. The parameters were initialised from uniform distributions on the interval $[0.1, 20]$ for k_1, k_2 and $[0.01, 2]$ for K_1, K_2 . The red center mark indicates the medians, the edges of the box the 25% and 75% quartiles, the whiskers extend to the most extreme estimates and the circles indicate the best fit

ables and linearly dependent rates (see Sect. 3.4). In our model, the two differential equations (21) are linearly dependent, which means that the observer can only estimate one time varying parameter; either k_1 or k_2 . In order to get estimates of all other parameters (and preliminary estimates of k_1 or k_2), we use a genetic algorithm (Fig. 4). These estimates are then used to create the necessary fake measurements for the observer. Simulating the observer with the output

$$y_i = [1 - x_i^*, \quad x_i^*, \quad K_{1i} + 1 - x_i^*, \quad K_{2i} + x_i^*, \quad v_{2i}(x_i^*)]^T$$

gives a time-varying estimate for k_1 , whereas simulating with

$$y_i = [1 - x_i^*, \quad x_i^*, \quad K_{1i} + 1 - x_i^*, \quad K_{2i} + x_i^*, \quad v_{1i}(x_i^*)]^T$$

gives a time-varying estimate for k_2 . Here, x_i^* is the measured phosphorylation level for each kinase $i = \text{Akt, ERK, JNK, PLC}\gamma, \text{p38}$, whereas all other entries are fake measurements.

4.1 Identification of Time-Variant Parameters Indicate a Sensitive Initial Phase Followed by Desensitising Adaptations

As expected, the parameter estimates of the observer are time varying. The simple model (21) neglects several intermediate signal transducers such as receptor adaptor proteins (SHC, GRB2, ...), GTPases (Ras, Rho, ...) and kinases (MEK, PI3K, ...) as well as pathway crosstalk and feedback structures. Thus, the parameters of the simple model are de facto functions of the neglected components. Because the neglected components change over time, the parameters of the simple model change over time. Indeed, at least two different phases can be distinguished, a highly sensitive initial phase, characterised by high values for the phosphorylation parameter k_1 and low values for the dephosphorylation parameter k_1 , followed by a delayed desensitisation phase, characterised by decreasing k_1 estimates and increasing k_2 estimates, suggesting the presence of desensitising feedback, for example in the form of phosphatase expression (Fig. 3).

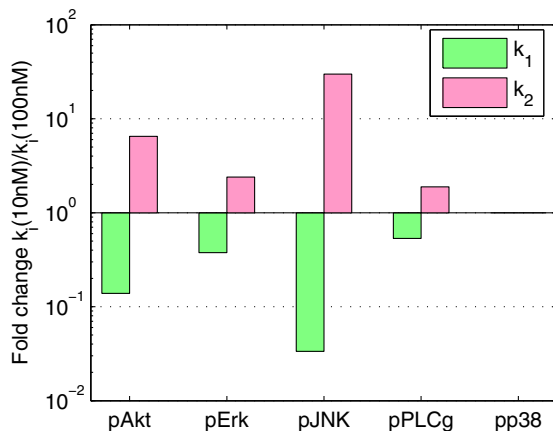


Fig. 5. Fold change of parameter estimates for 10nM versus 100nM NGF at 240min. The JNK estimates markedly differ, showing a 30-fold change, (compared to < 6-fold for Erk and PLC γ). No estimate were obtained for p38, because phosphorylation levels of p38 for 10nM NGF were not detectable by western blotting

4.2 JNK Estimate Stands Out, Indicating that Crucial Dynamics Are Missing

Comparing the estimates for the 10nM and 100nM dose response, we notice remarkable differences for JNK. Generally, the 10nM and 100nM estimates are reasonably alike in the early phase (for $t < 60$ min) and follow similar trends in the late phase, at least for ERK, PLC γ , and p38; however not for Akt and JNK. The Akt and JNK estimates exhibit opposing trends in the late phase. For example, $k_1(t)$ decreases for the 10nM dose, but increases for the 100nM dose (Fig. 3). Comparing the relative differences of the 10nM and 100nM estimates for $t = 240$ min reveals that the JNK difference is notably higher than that of the other kinases (Fig. 5). Together, these observations indicate that the simple model neglects some crucial dynamics affecting the JNK response and that model refinements are required.

4.3 Renement of the JNK Model

Since our analysis hinted at JNK and Akt playing a special role in the systems response, we performed several follow up experiments. In particular, inhibition of JNK blocked the second phase of TrkA activation and NGF induced apoptosis (Fig. 6). In contrast, inhibition of AKT (using a PI3K inhibitor) significantly increased JNK phosphorylation and apoptosis (Fig. 6). The literature suggests that pAkt indirectly inhibits pJNK via phosphorylation and deactivation of JNK upstream kinases ASK1 and MKK4 [23, 33]. Taken together, these data suggest a model in which stimulation of TrkA phosphorylates and activates JNK, which in turn increases the levels of total and phosphorylated TrkA. Additionally, TrkA mediated activation of Akt inhibits phosphorylation of JNK (Fig. 7). Neglecting the exact bio-

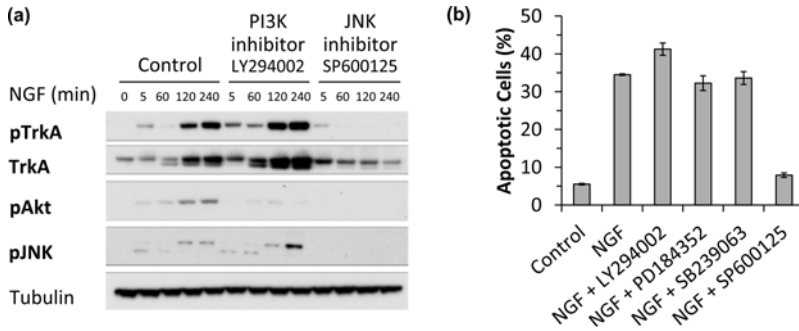


Fig. 6. Results of inhibitor experiments. (a) Western blots of key components phosphorylated TrkA (pTrkA), total TrkA (TrkA), phosphorylated JNK (pJNK) and phosphorylated Akt (pAkt) at indicated time points after NGF treatment. JNK inhibition blocks the induction of total TrkA protein and the second phase of TrkA activation. PI3K inhibition increases TrkA and pTrkA levels and results in notably increased pJNK in the late phase. (b) Percentage of apoptotic cells. Apoptosis was assayed by propidium iodide staining and flow cytometry analysis. NGF induces high levels of apoptosis, which are blocked by inhibition of JNK. Inhibition of PI3K slightly increases the level of NGF induced apoptosis. LY294002 PI3K inhibitor, PD184352 MEK inhibitor, SB239063 p38 inhibitor, SP600125 JNK inhibitor

chemical details, we can model the interaction scheme phenomenologically with two differential equations

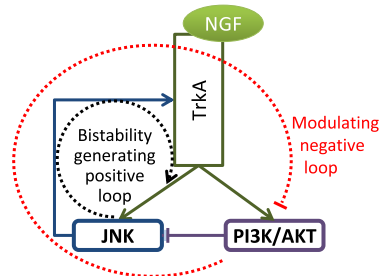
$$\frac{d}{dt}P^* = k_P T^* \frac{P}{K_P + P} - k_{-P} \frac{P^*}{K_{-P} + P^*}, \quad (22a)$$

$$\frac{d}{dt}J^* = k_J T^* (1 - k_I P^*) \frac{J}{K_J + J} - k_{-J} \frac{J^*}{K_{-J} + J^*}, \quad (22b)$$

with $P = 1 - P^*$, $J = 1 - J^*$, where $P^*, J^* \in [0, 1]$ denote the level of active Akt and JNK, respectively. The term $(1 - k_I P^*)$ models the inhibition of JNK phosphorylation by Akt, whereby the $k_I \in [0, 1]$ denotes the inhibition strength with $k_I = 0$ meaning no inhibition and $k_I = 1$ complete inhibition. The level of TrkA activity is given by the algebraic equation

$$T^* = u(0.1 + J^*), \quad (22c)$$

Fig. 7. Refined scheme of TrkA-JNK-Akt interactions. Solid lines represent (indirect) biochemical effects, with arrows indicating activation and bars indicating inhibition. The dotted circles indicate the two feedback loops present in the system



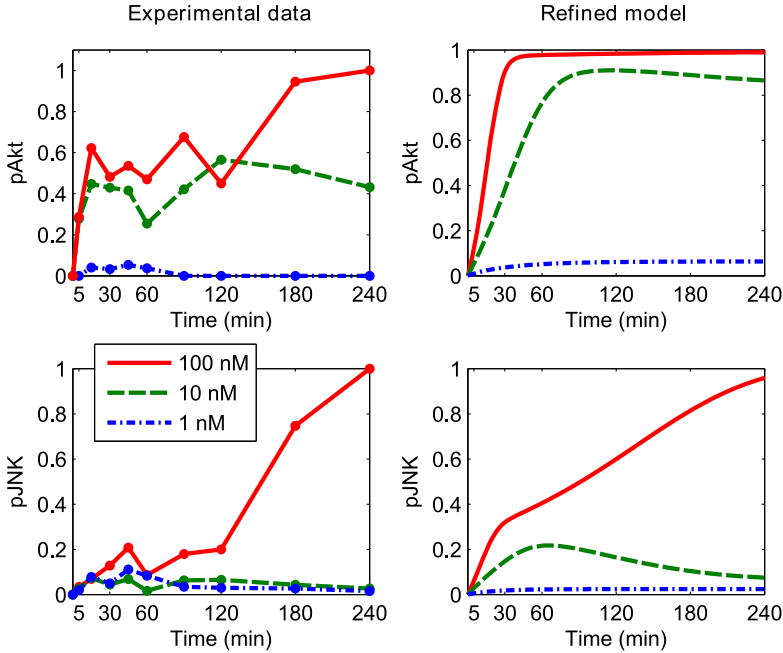


Fig. 8. Qualitative fit of refined model. Left recorded timecourses, right trajectories of the JNK-Akt interaction model

where u denotes the strength of NGF stimulus and the term $(0.1 + J^*) \in [0.1, 1.1]$ models the JNK mediated induction of TrkA with a basal TrkA level of 0.1 and maximal level of 1.1 (since $J^* \in [0, 1]$). The advantage of this phenomenological formulation is that it facilitates the analysis of the systems core structure. (the two interlocked feedback loops) in a few parameters. The advantage of this formulation is that it describes the system's core structure in a few parameters, thus facilitating its identification and mathematical analysis. Indeed, the qualitative behaviour of the system is characterised by two parameters; the input strength u and the inhibition strength k_I and can be summarised in a three dimensional plot (Fig. 9), whereby the remaining parameters were chosen such that the simulated trajectories qualitatively mimic the recorded time courses (Fig. 8).

4.4 Rened JNK Model Exhibits Bistable Behaviour

The refined TrkA-JNK-Akt interaction scheme contains two feedback loops generating complex system behaviour. For certain range of stimuli u , the positive feedback generates bistable behaviour, whereby the range at which bistability occurs depends on the strength of the negative feedback (Fig. 9). Stronger Akt mediated inhibition of JNK (higher values of k_I) shifts the bistable region to higher inputs. Thereby, the size of the JNK switch remains almost unaffected, whereas the size of

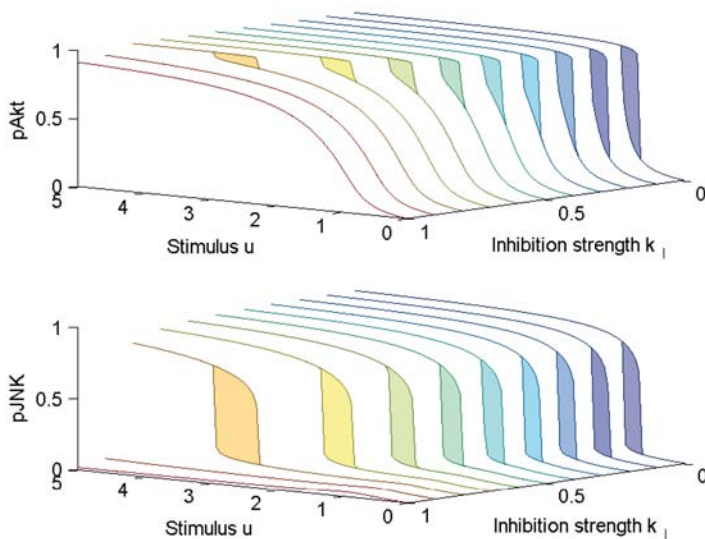


Fig. 9. Complex behaviour of JNK-Akt interaction model summarised in bifurcation diagrams. The system exhibits bistability (patches). Increasing the parameter k_l (describing the strength of Akt mediated JNK inhibition) shift the range of inputs where bistability occurs to higher values and decreases the size of the Akt switch

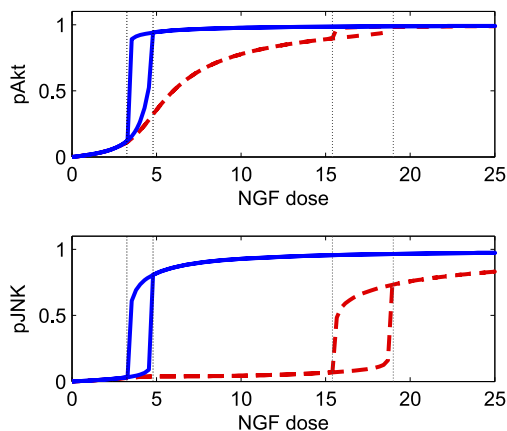


Fig. 10. Bifurcation diagram of the refined JNK-Akt model showing bistability. For increasing NGF stimuli, the JNK response remains almost unaffected (in the off state) until a certain threshold is reached at which JNK switches to the on state. For decreasing stimuli, the system switches back to the off state at a lower threshold, generating a hysteresis effect. For weak negative feedback (solid lines), Akt shows qualitatively behaviour similar to JNK, whereas for strong negative feedback (dashed lines) Akt exhibits a graded response. Blue and red lines correspond to $k_l = 0$ and $k_l = 0.8$, respectively

the Akt switch decreases. For increasing k_I the Akt-off state approaches the Akt-on state until finally, high values of k_I cause the Akt switch to be so small that it can effectively be neglected (Fig. 10). Therewith, negative feedback provides a mechanism by which components normally affected by a bistable switch can effectively be decoupled.

4.5 Concluding Remark

Using the observer based approach, we identified the Akt-JNK interaction network as crucial component for TrkA induced apoptosis in a SY5Y neuroblastoma cell line. The time dependent variations of JNK and AKT associated parameters was markedly higher compared to other measured kinases, hinting at the neglect of crucial dynamics in a simple feedforward model and calling for model refinements. The refined JNK-Akt interactions model contains two feedback loops, and explains the switch to apoptosis with a bistable JNK response. For low stimuli the system is only transiently disturbed and quickly returns to the JNK off state, whereas high stimuli cause JNK to switch into its on state. The threshold at which the system switched depends on Akt mediated negative feedback. Unfortunately, the late phase activation of TrkA seems to be an artefact of the inducible cell line used. We are currently investigating whether the obtained results are transferable to better neuroblastoma models in other cell lines. Nevertheless, the results demonstrate the power of the observer based approach and its applicability to real world data.

5 Conclusion

Dysregulation of cell signalling leading to tumourigenic cell fate decisions is a hallmark of cancer. Understanding this dysregulation requires the use of mathematical modelling and analysis. Since mathematical models are, by nature, approximations of the real system, the art of model building is therefore to find the correct level of abstraction. From a systems perspective, approximations result in time varying parameters, which are notoriously difficult to identify. Here, we presented a methodology capable of estimating time varying parameters. Importantly, time variant estimates indicate unmodelled dynamics and can be used to identify modelling errors. Strongly time-varying parameters point to modelling errors, and are the starting point for model refinements.

References

1. Allgöwer, F., Doyle, F.: Introduction to the special issue on systems biology. *Automatica* **47**, 1095–1096 (2011). Special Issue on Systems Biology
2. Bullinger, E., Fey, D., Farina, M., Findeisen, R.: Identifikation biochemischer Reaktionsnetzwerke: Ein beobachterbasierter Ansatz. *AT-Automatisierungstechnik* **56**, 269–279 (2008)

3. Chesi, G., Chen, L.: Call for papers: Special issue on systems biology. *International Journal of Robust and Nonlinear Control* **20**, 723–724 (2010)
4. Cornish-Bowden, A.: *Fundamentals of Enzyme Kinetics* (3rd edn.). Portland Press, London (2004)
5. Dochain, D.: State and parameter estimation in chemical and biochemical processes: a tutorial. *Journal of Process Control* **13**, 801–818 (2003)
6. Eissing, T., Waldherr, S., Allgöwer, F., Scheurich, P., Bullinger, E.: Steady state and (bi-) stability evaluation of simple protease signalling networks. *Biosystems* **90**, 591–601 (2007)
7. Farina, M., Bullinger, E., Findeisen, R., Bittanti, S.: An observer based strategy for parameter identification in systems biology. In: *Proc. of the 2nd Conference on Foundations of Systems Biology in Engineering*, Stuttgart, Germany, 521–526 (2007)
8. Farina, M., Findeisen, R., Bullinger, E., Bittanti, S., Allgöwer, F., Wellstead, P.: Results towards identifiability properties of biochemical reaction networks. In: *Proc. of the 45th IEEE Conference on Decision and Control*, San Diego, USA, 2104–2109 (2006)
9. Fell, D.: *Understanding the Control of Metabolism*. Portland Press, London (1997)
10. Fey, D., Bullinger, E.: A dissipative approach to the identification of kinetic reaction networks. In: *Proc. of the 15th IFAC Symposium on Systems Identification*, Saint Malo, France, 1259–1264 (2009)
11. Fey, D., Bullinger, E.: Limiting the parameter search space for dynamic models with rational kinetics using semi-definite programming. In: *Proc. of the 15th Symposium on Computer Applications in Biotechnology*, Leuven, Belgium, 150–155 (2010)
12. Fey, D., Findeisen, R., Bullinger, E.: Parameter estimation in kinetic reaction models using nonlinear observers facilitated by model extensions. In: *Proc. of the 17th IFAC World Congress*, Seoul, Korea, 313–318 (2008)
13. Fey, D., Findeisen, R., Bullinger, E.: Identification of biochemical reaction networks using a parameter-free coordinate system. In: Iglesias and Ingalls [17], 297–316
14. Heinrich, R., Schuster, S.: *The Regulation of Cellular Systems*. Chapman and Hall, New York (1996)
15. Hodgkin, A.L., Huxley, A.F.: A quantitative description of membrane current and its application to conduction and excitation in nerve. *J Physiol* **117**, 500–544 (1952)
16. Huang, C.Y., Ferrell, J.E.: Ultrasensitivity in the mitogen-activated protein kinase cascade. *Proc Natl Acad Sci U S A* **93**, 10078–10083 (1996)
17. Iglesias, P.A., Ingalls, B.P. (eds.): *Control Theory and Systems Biology*. MIT Press, Cambridge, MA (2009)
18. Ireton, R., Montgomery, K., Bumgarner, R., Samudrala, R., McDermott, J. (eds.): *Computational Systems Biology*. Springer, New York (2009)
19. Keener, J., Sneyd, J.: *Mathematical Physiology*, second edn. Springer, New York (2001)
20. Kholodenko, B.N.: Cell-signalling dynamics in time and space. *Nat Rev Mol Cell Biol* **7**, 165–176 (2006)
21. Kholodenko, B.N., Demin, O.V., Moehren, G., Hoek, J.B.: Quantification of short term signaling by the epidermal growth factor receptor. *J Biol Chem* **274**, 30,169–30,181 (1999)
22. Kholodenko, B.N., Hancock, J.F., Kolch, W.: Signalling ballet in space and time. *Nat Rev Mol Cell Biol* **11**, 414–426 (2010)
23. Kim, A.H., Khursigara, G., Sun, X., Franke, T.F., Chao, M.V.: Akt phosphorylates and negatively regulates apoptosis signal-regulating kinase 1. *Mol Cell Biol* **21**, 893–901 (2001)
24. Kitano, H.: Systems biology: a brief overview. *Science* **295**, 1662–1664 (2002)
25. Kitano, H.: Grand challenges in systems physiology. *Front Physiol* **1**, 3 (2010)
26. Klipp, E., Herwig, R., Kowald, A., Wierling, C., Lehrach, H.: *Systems Biology in Practice: Concepts, Implementation and Application*. Wiley-VCH, Weinheim (2005)
27. von Kriegsheim, A., Baiocchi, D., Birtwistle, M., Sumpton, D., Bienvenut, W., Morrice, N., Yamada, K., Lamond, A., Kalna, G., Orton, R., Gilbert, D., Kolch, W.: Cell fate decisions are specified by the dynamic erk interactome. *Nat Cell Biol* **11**, 1458–1464 (2009)

28. Le Novère, N., Bornstein, B., Broicher, A., Courtot, M., Donizelli, M., Dharuri, H., Li, L., Sauro, H., Schilstra, M., Shapiro, B., Snoep, J., Hucka, M.: BioModels Database: A free, centralized database of curated, published, quantitative kinetic models of biochemical and cellular systems. *Nucleic Acids Res* **34**, D689–D691 (2006)
29. Leloup, J.C., Goldbeter, A.: Modeling the circadian clock: from molecular mechanism to physiological disorders. *Bioessays* **30**, 590–600 (2008)
30. Li, C., Donizelli, M., Rodriguez, N., Dharuri, H., Endler, L., Chelliah, V., Li, L., He, E., Henry, A., Stefan, M.I., Snoep, J.L., Hucka, M., Le Novère, N., Laibe, C.: BioModels Database: An enhanced, curated and annotated resource for published quantitative kinetic models. *BMC Syst Biol* **4**, 92 (2010)
31. Nakakuki, T., Birtwistle, M.R., Saeki, Y., Yumoto, N., Ide, K., Nagashima, T., Brusch, L., Ogunnaike, B.A., Okada-Hatakeyama, M., Kholodenko, B.N.: Ligand-specific c-fos expression emerges from the spatiotemporal control of erbb network dynamics. *Cell* **141**, 884–896 (2010)
32. Noble, D.: Cardiac action and pacemaker potentials based on the Hodgkin-Huxley equations. *Nature* **188**, 495–497 (1960)
33. Park, H.S., Kim, M.S., Huh, S.H., Park, J., Chung, J., Kang, S.S., Choi, E.J.: Akt (protein kinase b) negatively regulates sek1 by means of protein phosphorylation. *J Biol Chem* **277**, 2573–2578 (2002)
34. Reder, C.: Metabolic control theory: a structural approach. *J Theor Biol* **135**, 175–201 (1988)
35. Reichardt, L.F.: Neurotrophin-regulated signalling pathways. *Philos Trans R Soc Lond B Biol Sci* **361**, 1545–1564 (2006)
36. Rocha-Cózatl, E., Moreno, J.A.: Dissipative design of unknown input observers for systems with sector nonlinearities. *International Journal of Robust and Nonlinear Control* **21**, 1623–1644 (2011)
37. Rosenblueth, A., Wiener, N.: The role of models in science. *Philosophy of Science* **12**, 316–321 (1945)
38. Santos, S.D.M., Verveer, P.J., Bastiaens, P.I.H.: Growth factor-induced mapk network topology shapes erk response determining pc-12 cell fate. *Nat Cell Biol* **9**, 324–330 (2007)
39. Schaffner, J., Zeitz, M.: Variants of nonlinear normal form observer design. In: H. Nijmeijer, T. Fossen (eds.) *New Directions in nonlinear observer design*, 161–180. Springer, Heidelberg (1999)
40. Schrödinger, E.: What is life (1943). Dublin Institute for Advanced Studies, February 1943
41. Schrödinger, E.: What is Life? The physical aspect of the living cell. Cambridge University Press, Cambridge (1944)
42. Sontag, E.: Molecular systems biology and control. *Eur J Control* **11**, 396–435 (2005)
43. Sontag, E., Kiyatkin, A., Kholodenko, B.N.: Inferring dynamic architecture of cellular networks using time series of gene expression, protein and metabolite data. *Bioinformatics* **20**, 1877–1886 (2004)
44. Szallasi, Z., Stelling, J., Periwal, V. (eds.): *Systems Modelling in Cellular Biology*. MIT Press, Cambridge, MA (2006)
45. Wellstead, P., Bullinger, E., Kalamatianos, D., Mason, O., Verwoerd, M.: The rôle of control and system theory in systems biology. *Annu Rev Control* **32**, 33–47 (2008)
46. Xia, X., Zeitz, M.: On nonlinear continuous observers. *International Journal of Control* **66**, 943–954 (1997)

Modelling Biochemical Pathways with the Calculus of Looping Sequences

Paolo Milazzo, Antonella Del Corso, Andrea Maggiolo-Schettini, Umberto Mura, and Roberto Barbuti

Abstract The Calculus of Looping Sequences (CLS) has been proposed as a formalism for representing and simulating biomolecular systems. CLS is based on a set of operators which allow to represent the structure of such systems. Systems evolution is described by means of rewrite rules. A stochastic extension of CLS has been defined in which rewrite rules are equipped with reaction rates. Such rates allow to use CLS for performing both stochastic simulation and deterministic ones. Deterministic simulations are obtained by a standard translation of rewrite rule with reaction rates into differential equations. We use CLS simulations for investigating the behaviour of the (first part of) the EGFR signalling cascade in cells.

1 Introduction

Cell biology, the study of the morphological and functional organization of cells, is now an established field in biochemical research. Computer Science can help the research in cell biology in several ways. For instance, it can provide biologists with models and formalisms capable of describing and analyzing complex systems such as cells. In the last few years many formalisms originally developed by computer scientists to model systems of interacting components have been applied to Biology. Among these, there are Petri Nets [27], Hybrid Systems [1], and the π -calculus [12, 39]. Moreover, new formalisms have been defined for describing biomolecular and membrane interactions [2, 8, 10, 14, 34, 37]. Others, such as P Systems [29, 30],

P. Milazzo (✉) · A. Maggiolo-Schettini · R. Barbuti
Department of Computer Science, University of Pisa, Largo B. Pontecorvo 3, 56127 Pisa, Italy
e-mail: milazzo@di.unipi.it, maggiolo@di.unipi.it, barbuti@di.unipi.it

A. Del Corso, U. Mura
Department of Biology, University of Pisa, Via S. Zeno 51, 56127 Pisa, Italy
e-mail: adelcorso@biologia.unipi.it, umura@biologia.unipi.it

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series
DOI 10.1007/978-88-470-2571-4_6, © Springer-Verlag Italia 2012

have been proposed as biologically inspired computational models and have been later applied to the description of biological systems.

The π -calculus and new calculi based on it [34, 37] have been particularly successful in the description of biological systems, as they allow describing systems in a compositional manner. However, these calculi offer very low-level interaction primitives, and this causes models to become very large and difficult to understand. Calculi such as those proposed in [8, 10, 14] give a more abstract description of systems and offer special biologically motivated operators. However, they are often specialized to the description of some particular kinds of phenomena such as membrane or protein interactions. Finally, P Systems have a simple notation and are not specialized to the description of a particular class of systems, but they are still not completely general. For instance, it is possible to describe biological membranes and the movement of molecules across membranes, and there are some variants able to describe also more complex membrane activities. However, the formalism is not flexible enough to allow describing easily new activities observed on membranes without extending the formalism itself.

Therefore, the need has arisen for a formalism having a simple notation, having the ability of describing biological systems at different levels of abstraction, having some notions of compositionality and being flexible enough to allow describing new kinds of phenomena, without being specialized to the description of a particular class of systems. For this reason in [2, 3, 5, 6] we have introduced the Calculus of Looping Sequences (CLS) with some variants.

CLS is a formalism based on term rewriting with some features, such as a commutative parallel composition operator, and some semantic means, such as bisimulations, which are common in process calculi. All this permits one to combine the simplicity of notation of rewriting systems with the advantage of a form of compositionality.

Given an alphabet of symbols representing basic biological entities, such as genes, proteins and other macro-molecules, CLS terms are constructed by applying to these symbols operators of sequencing, looping, containment and parallel composition. Terms constructed by means of these operators represent biological structures such as DNA sequences and membranes. Rewrite rules can be used to model biological events that permit the system to evolve. In particular, they can be used to model biochemical reactions and structure rearrangements such as membrane fusion and dissolution.

Stochastic CLS [4] allows describing quantitative aspects of the modeled systems such as the frequency of chemical reactions. For Stochastic CLS a simulator has been developed that allows simulating the evolution of biological systems over time.

In this paper, we describe the application of Stochastic CLS to a biochemical pathway, namely EGFR signaling cascade. We show how to model the signaling pathway and we study the behaviour of the modelled system by simulations in which we vary the value of parameters.

2 The Calculus of Looping Sequences (CLS)

In the next sections we introduce the Calculus of Looping Sequences (CLS) and the Stochastic CLS. Stochastic CLS, equipped with a stochastic semantics, permits to consider quantitative aspects of biological systems.

CLS is based on term rewriting, hence a CLS model consists of a term and a set of rewrite rules. The term is intended to represent the structure of the modeled system, and the rewrite rules to represent the events that may cause the system to evolve.

We start with defining the syntax of terms. We assume a possibly infinite alphabet $\mathcal{E}$ of symbols ranged over by $a, b, c, \dots$

Definition 1 (Terms). Let $\mathcal{E}$ be a possibly infinite alphabet of symbols ranged over by $a, b, c, \dots$. *Terms* T and *sequences* S are given by the following grammar:

$$\begin{aligned} T &::= S \mid (T)^L \mid T \mid T \\ S &::= \epsilon \mid a \mid S \cdot S \end{aligned}$$

where a is a generic element of $\mathcal{E}$, and ϵ represents the empty sequence. We denote with $\mathcal{T}$ the infinite set of terms, and with $\mathcal{S}$ the infinite set of sequences.

In CLS we have a sequencing operator $\_ \cdot \_$, a parallel composition operator $\_ \mid \_$, a looping operator $(\_)^L$ and a containment operator $\_ \rfloor \_$. Sequencing can be used to concatenate elements of the alphabet $\mathcal{E}$. The empty sequence ϵ denotes the concatenation of zero symbols. By definition, looping and containment are always applied together, hence we can consider them as a single binary operator $(\_)^L \rfloor \_$. Brackets can be used to indicate the order of application of the operators, and we assume $(\_)^L \rfloor \_$ to have precedence over $\_ \mid \_$.

The biological interpretation of the operators is the following: the main entities which occur in cells are DNA and RNA strands, proteins, membranes, and other macro-molecules. DNA strands (and similarly RNA strands) are sequences of nucleic acids, but they can be seen also at a higher level of abstraction as sequences of genes. Proteins are sequence of aminoacids which usually have a very complex three-dimensional structure. In a protein there are usually (relatively) few subsequences, called domains, which actually are able to interact with other entities by means of chemical reactions. CLS sequences can model DNA/RNA strands and proteins by describing each gene or each domain with a symbol of the alphabet. Membranes are closed surfaces, often interspersed with proteins, which may contain something. Looping and containment allow the representation of membranes with their contents. For example, the term $(a \mid b)^L \rfloor c$ represents a membrane with the elements a and b on its surface and containing the element c . Other macro-molecules can be modeled either as single alphabet symbols, or as short sequences. Finally, juxtaposition of entities can be described by the parallel composition of their representations.

In Fig. 1 we show some examples of CLS terms and their visual representation. Term (i), $(\epsilon)^L \rfloor \text{virus} \mid (a \mid b \cdot c \mid a \cdot b)^L \rfloor (n)^L \rfloor \text{DNA}$, represents an environment

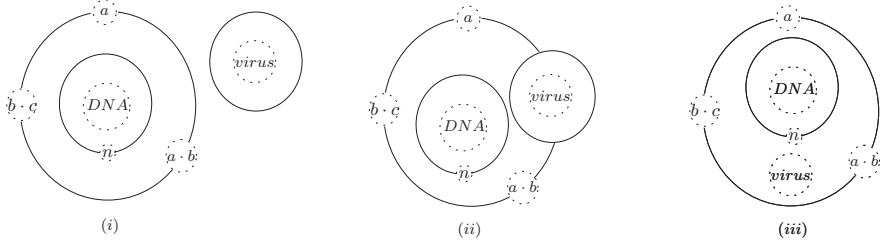


Fig. 1. A visual representation of some examples of CLS terms

where there exist a virus, represented by the membrane $(\epsilon)^L \mid virus$, a membrane with three possible proteins on its surface, represented by the sequences a , $b \cdot c$ and $a \cdot b$, and, as the content of the membrane, another membrane containing a strand of DNA, namely the sequence DNA , surrounded by a nucleus, the membrane with surface n . Term (ii), $(a \mid b \cdot c \mid a \cdot b \mid (\epsilon)^L \mid virus)^L \mid (n)^L \mid DNA$, represents the same membrane of term (i) where the virus, $(\epsilon)^L \mid virus$, is attached to the surface of the external membrane. Finally, term (iii), $(a \mid b \cdot c \mid a \cdot b)^L \mid (virus \mid (n)^L \mid DNA)$, represents the fact that the virus, by dissolving its external membrane, has entered the membrane containing the nucleus and the DNA strand. Notice that this sequence of states represented by terms (i), (ii) and (iii) describe the entrance, by a virus, into a cell containing a strand of DNA. This process is similar to the sporulation process described as example in [6].

In CLS we may have syntactically different terms representing the same structure. We introduce a structural congruence relation to identify such terms.

Definition 2 (Structural Congruence). The *structural congruence relations* $\equiv_S$ and $\equiv$ are the least congruence relations on sequences and on terms, respectively, satisfying the following rules:

$$\begin{aligned}
 S_1 \cdot (S_2 \cdot S_3) &\equiv_S (S_1 \cdot S_2) \cdot S_3 & S \cdot \epsilon &\equiv_S \epsilon \cdot S \equiv_S S \\
 S_1 &\equiv_S S_2 \text{ implies } S_1 &\equiv_S S_2 \\
 T_1 \mid T_2 &\equiv T_2 \mid T_1 & T_1 \mid (T_2 \mid T_3) &\equiv (T_1 \mid T_2) \mid T_3 \\
 T \mid \epsilon &\equiv T & (\epsilon)^L \mid \epsilon &\equiv \epsilon.
 \end{aligned}$$

Rules of the structural congruence state the associativity of $\cdot$ and $\mid$, the commutativity of the latter and the neutral role of ϵ .

Rewrite rules will be defined essentially as pairs of terms, with the first term describing the portion of the system in which the event modeled by the rule may occur, and the second term describing how that portion of the system changes when the event occurs. In the terms of a rewrite rule we allow the use of variables. As a consequence, a rule will be applicable to all terms which can be obtained by properly instantiating its variables. Variables can be of three kinds: two of these are associated

with the two different syntactic categories of terms and sequences, and one is associated with single alphabet elements. We assume a set of term variables TV ranged over by $X, Y, Z, \dots$, a set of sequence variables SV ranged over by $\tilde{x}, \tilde{y}, \tilde{z}, \dots$, and a set of element variables $\mathcal{X}$ ranged over by $x, y, z, \dots$. All these sets are possibly infinite and pairwise disjoint. We denote by $\mathcal{V}$ the set of all variables, $\mathcal{V} = TV \cup SV \cup \mathcal{X}$, and with ρ a generic variable of $\mathcal{V}$. Hence, a pattern is a term that may include variables.

Definition 3 (Patterns). *Patterns* P and *sequence patterns* SP of CLS are given by the following grammar:

$$\begin{aligned} P &::= SP \mid (P)^L \mid P \mid P \mid X \\ SP &::= \epsilon \mid a \mid SP \cdot SP \mid \tilde{x} \mid x \end{aligned}$$

where a is a generic element of $\mathcal{E}$, and $X, \tilde{x}$ and x are generic elements of TV, SV and $\mathcal{X}$, respectively. We denote with $\mathcal{P}$ the infinite set of patterns.

We assume the structural congruence relation to be trivially extended to patterns. An *instantiation* is a partial function $\sigma : \mathcal{V} \rightarrow \mathcal{T}$. An instantiation must preserve the type of variables, thus for $X \in TV, \tilde{x} \in SV$ and $x \in \mathcal{X}$ we have $\sigma(X) \in \mathcal{T}, \sigma(\tilde{x}) \in \mathcal{S}$ and $\sigma(x) \in \mathcal{E}$, respectively. Given $P \in \mathcal{P}$, with $P\sigma$ we denote the term obtained by replacing each occurrence of each variable $\rho \in \mathcal{V}$ appearing in P with the corresponding term $\sigma(\rho)$. With Σ we denote the set of all the possible instantiations and, given $P \in \mathcal{P}$, with $Var(P)$ we denote the set of variables appearing in P . We can now define rewrite rules.

Definition 4 (Rewrite Rules). A *rewrite rule* is a pair of patterns (P_1, P_2) , denoted with $P_1 \mapsto P_2$, where $P_1, P_2 \in \mathcal{P}$, $P_1 \not\equiv \epsilon$ and such that $Var(P_2) \subseteq Var(P_1)$. We denote with $\mathfrak{R}$ the infinite set of all the possible rewrite rules.

A rewrite rule $P_1 \mapsto P_2$ states that a term $P_1\sigma$, obtained by instantiating variables in P_1 by some instantiation function σ , can be transformed into the term $P_2\sigma$. For instance, the rule $a \cdot b \mid b \cdot \tilde{x} \mapsto c \cdot \tilde{x}$ will replace, with respect to an instantiation function σ , one sequence $a \cdot b$ and one sequence $b \cdot \sigma(\tilde{x})$ with one sequence $c \cdot \sigma(\tilde{x})$. If the term to which the rule is applied is $a \cdot b \mid b \cdot a \cdot b \mid c$, given an instantiation function $\sigma = \{\{\tilde{x}, a \cdot b\}\}$, the result of applying the rule to the term is $c \cdot a \cdot b \mid c$. Analogously, the rule $p \mid (m \mid Y)^L \mid X \mapsto (m \mid Y \mid p)^L \mid X$ describes, from a biological point of view, the attachment of sequence p to the surface of a membrane containing on its surface at least a sequence m . As in the previous example, the membrane to which the sequence p gets attached, depends on the instantiation of variables X and Y .

With reference to Fig. 1, the rewrite rule which transforms the state represented by term (i) into the state represented by term (ii) could be $(\epsilon)^L \mid virus \mid (X)^L \mid (Y \mid (n)^L \mid DNA) \mapsto (X \mid (\epsilon)^L \mid virus)^L \mid (Y \mid (n)^L \mid DNA)$. Analogously, the rule which transforms the state represented by term (i) into the state represented by term (ii) could be $(X \mid (\epsilon)^L \mid virus)^L \mid (Y \mid (n)^L \mid DNA) \mapsto (X)^L \mid (Y \mid virus \mid (n)^L \mid DNA)$.

We now define the semantics of CLS as a transition system, in which states correspond to terms, and transitions correspond to rules applications.

Definition 5 (Semantics). Given a set of rewrite rules $\mathcal{R} \subseteq \mathfrak{R}$, the *semantics* of CLS is the least transition relation $\rightarrow$ on terms closed under $\equiv$, and satisfying the following inference rules:

$$\frac{P_1 \mapsto P_2 \in \mathcal{R} \quad P_1 \sigma \not\equiv \epsilon \quad \sigma \in \Sigma}{P_1 \sigma \rightarrow P_2 \sigma} \quad \frac{T_1 \rightarrow T_2}{T \mid T_1 \rightarrow T \mid T_2}$$

$$\frac{T_1 \rightarrow T_2}{(T)^L \mid T_1 \rightarrow (T)^L \mid T_2} \quad \frac{T_1 \rightarrow T_2}{(T_1)^L \mid T \rightarrow (T_2)^L \mid T}$$

where the symmetric rule for the parallel composition is omitted.

Notice that, by the definition of the semantics, a rule cannot be applied, and hence cannot rewrite, a portion of a sequence and, at least, must rewrite a whole sequence because it cannot contain empty reactants.

Finally, a *model* in CLS is given by a term describing the initial state of the system and by a set of rewrite rules describing all the events that may occur.

2.1 A Stochastic Extension

In CLS only qualitative aspects of biological systems are considered, such as their structure and the presence (or the absence) of certain molecules. As a consequence, on CLS models it is only possible to verify properties such as the reachability of particular states or causality relationships between events. It would be interesting to verify also properties such as the time spent to reach a particular state, or the probability of reaching it. To face this problem, in [4] we have developed a stochastic extension of CLS, called *Stochastic CLS*, in which quantitative aspects, such as time and probability are taken into account.

The standard way of extending a formalism to model quantitative aspects of biological systems is by incorporating the stochastic framework developed by Gillespie with its simulation algorithm for chemical reactions [18] in the semantics of the formalism. This has been done, for instance, for the π -calculus [33, 35]. The idea of Gillespie's algorithm is that a rate constant is associated with each chemical reaction that may occur in the system. Such a constant is obtained by multiplying the kinetic constant of the reaction by the number of possible combinations of reactants that may occur in the system. The resulting rate constant is then used as the parameter of an exponential distribution modeling the time spent between two occurrences of the considered chemical reaction.

The use of exponential distributions to represent the (stochastic) time spent between two occurrences of chemical reactions allows the description of the system as a Continuous Time Markov Chain (CTMC), and consequently it allows the verification of properties of the described system by means of analytic tools and by means of stochastic model checkers.

In Stochastic CLS, the incorporation of Gillespie's stochastic framework is not a simple exercise. The main difficulty is counting the number of possible reactant combinations of the chemical reaction described by a rewrite rule. As in CLS, reactants are given by the left pattern in a rewrite rule and, in general, may contain variables. Dependently on the instantiation of the variables and on the number of occurrences of the reactants in the term which describes the state of the system, a rewrite rule can be applied in different positions of the term. Counting the number of different positions where the rewrite rule can be applied, by taking into account that rules may contain variables, is not trivial.

Rewrite rules are enriched, with respect to those of CLS, with a kinetic constant. Intuitively, a rule with reactants $a \mid (c \mid X)^L \mid \epsilon$ modeling an interaction between a sequence a and a sequence c contained on the surface of a membrane, $(c \mid X)^L \mid \epsilon$, should have a rate dependent on the number of sequences c appearing on the surface itself.

The syntax of terms and the structural congruence relations are the same of CLS as defined in Sect. 2.

Definition 6 (Patterns). *Left patterns* P_L and *right patterns* P_R of Stochastic CLS are given by the following grammar:

$$\begin{aligned}
 P_L &::= SP \mid (P_X)^L \mid P_X \mid P_L \mid P_L \\
 P_X &::= P_L \mid P_L \mid X \\
 P_R &::= SP \mid (P_R)^L \mid P_R \mid P_R \mid P_R \mid X \\
 SP &::= \epsilon \mid a \mid SP \cdot SP \mid \tilde{x} \mid x
 \end{aligned}$$

where $X, \tilde{x}$ and x are generic elements of TV, SV and $\mathcal{X}$, respectively. The sets of all left and right patterns are denoted with $\mathcal{P}_L$ and $\mathcal{P}_R$, respectively.

We assume the structural congruence relation to be trivially extended to patterns. Furthermore, note that there exist restrictions on the syntax of patterns. In particular, term variables inside left patterns cannot appear at top-level and they can appear only once inside a looping or a containment operator. Both these restrictions simplify the implementation of the simulator for the formalism. As regards the use of term variables at top-level of a left pattern of a rule, by the definition of the semantics, such a rule could rewrite any portion of the term representing the state of the system. From a biological perspective, this would correspond to modeling a reaction with an uncertain number of reactants and this is not correct with respect to the standard kinetics of chemical reactions. Consequently, as chemical reactions happen usually between a fixed and small number of reactants, the restriction appears to be reasonable. As regards the restriction on the uniqueness of term variables inside the looping and the containment operators we have that, for reasons similar to the previous ones, the restriction appears to be reasonable. In fact, two term variables inside a looping or a containment operator of a rule could be instantiated in many different ways correspondingly to different partitions of the term. From a biological perspec-

tive, this would represent a reaction between, for instance, two arbitrary portions of the content of a membrane.

As an example of well-formed left and right patterns let us consider $(a \cdot b \cdot \widetilde{x})^L \mid (b \mid X)$ and $(a \cdot b \cdot \widetilde{x})^L \mid (X \mid X)$, respectively. The former describes a membrane with a single sequence prefixed by $a \cdot b$ as surface and, as content, at least a sequence b . The latter describes a membrane which contains a single sequence prefixed by $a \cdot b$ on its surface and two copies of the same term as content. Notice that, due to the restrictions on patterns previously outlined, the latter would not be a well-formed left pattern because it contains two term variables X as the content of a membrane.

We now introduce stochastic rewrite rules.

Definition 7 (Stochastic Rewrite Rules). A *stochastic rewrite rule* is a pair of patterns and a kinetic constant (P_1, P_2, k) , denoted with $P_1 \xrightarrow{k} P_2$, where $P_1 \in \mathcal{P}_L$, $P_2 \in \mathcal{P}_R$, $P_1 \not\equiv \epsilon$, $k \in \mathbb{R}$ and such that $\text{Var}(P_2) \subseteq \text{Var}(P_1)$. We denote with $\mathfrak{R}$ the infinite set of all the possible rewrite rules.

For instance, recalling the patterns previously discussed, the rule $(a \cdot b \cdot \widetilde{x})^L \mid (b \mid X) \xrightarrow{0.01} (a \cdot b \cdot \widetilde{x})^L \mid (X \mid X)$ models a replacement of membranes when applied to a term T . Such an event occurs with a frequency given by the kinetic constant 0.01. Along the lines of Gillespie's algorithm, from a state T the rate of application of a rewrite rule is given by its kinetic constant multiplied by the number of positions in the term where the rule can be applied, namely by the number of membranes represented by the term $P\sigma$ inside T , for any instantiation function σ .

Stochastic CLS can be used both for stochastic and deterministic simulations. Deterministic simulations can be performed by the transformation of CLS rewrite rules into differential equations. The coefficients of the differential equations terms can be automatically inferred by the kinetic constants of the Stochastic CLS rules. In the following section we use both kinds of simulations.

3 Modelling the EGFR Signalling Pathway

In order to show how CLS can be used to model and analyze biochemical pathways in this section we propose a CLS model of the signalling pathway originated by the Epidermal Growth Factor Receptor (EGFR). The signalling pathway involving EGFR is a rather complex network able to integrate different signals ending with modulation of cell proliferation through the activation of the MAPKinase (Mitogen Activate Protein Kinase) cascade induced by the activation of Ras (Rat sarcoma proteins) [43]. The CLS model is constructed by following the lines of model based on Ordinary Differential Equations (ODEs) developed by Kholodenko et al in [23].

Evidence suggests that the EGFR is involved in the pathogenesis and progression of different carcinoma types. The EGFR is often over-expressed in human carcinomas, and in vivo and in vitro studies have shown that these proteins are able to induce cell transformation. Amplification of the EGFR gene and mutations of the

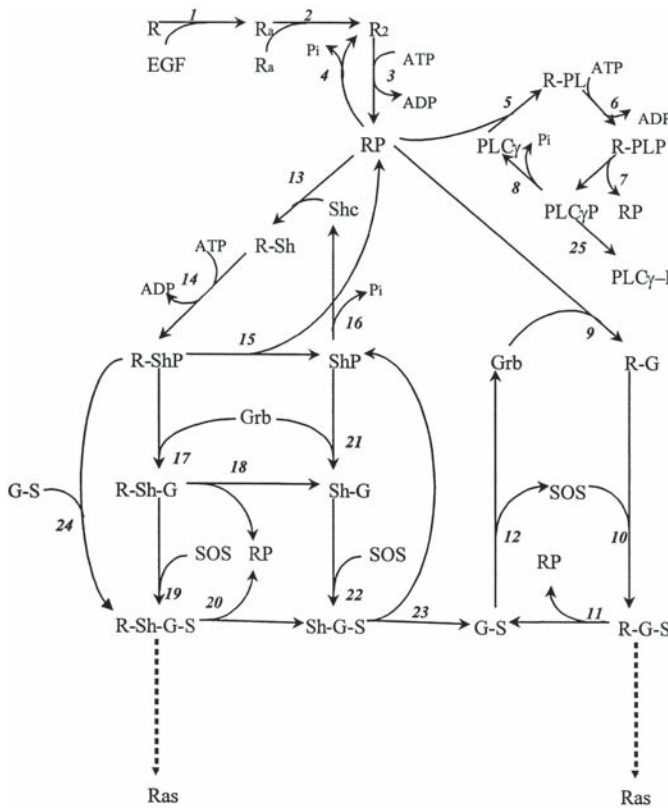


Fig. 2. Kinetic scheme of the EGFR signalling pathway (from [23])

EGFR tyrosine kinase domain have been recently demonstrated to occur in carcinoma patients [11, 25, 42].

The model in [23] aims at giving a quantitative description of the short term EGFR signalling based on a detailed scheme of the interactions of proteins and other signalling molecules involved. The model focuses on the first reactions of the signalling pathway, namely the EGF-induced dimerization of EGFR, its autophosphorylation and the recruitment on the cell membrane, through a sequence of protein-protein interactions, of a number of cytosolic proteins; namely: the growth factor receptor-binding protein 2 (Grb), Src homology and collagen domain protein (Shc), Son of Seveless homolog protein (SOS) and phospholipase C (PLC_{γ}).

The scheme of the modelled reactions is depicted in Fig. 2, where each interaction step is defined by a direct and a reverse reaction. Reactions taking place in the cytosol are organized in three partially independent branches that we call *PLC branch* (reactions 5–7 and 25), *Grb-SOS branch* (reactions 9–12) and *Shc-Grb-SOS branch* (reactions 13–22). The three branches compete for phosphorylated dimers (denoted RP in the scheme). The Grb-SOS and Shc-Grb-SOS branches share part

of the reactants and both produce complexes (denoted R-Sh-G-S and R-G-S) able to activate the Ras protein in the subsequent part of the pathway (not modelled). The PLC branch is completely independent from the other branches, apart from the competition for RP.

The CLS model of the EGFR signalling pathway consists of 53 rewrite rules, each corresponding to a reaction of the scheme in Fig. 2. Rules corresponding to the reactions taking place on the cell membrane are listed in Table 1. The reactions of the three branches taking place in the cytosol faced cytoplasmic membrane are listed in Tables 2, 3 and 4. Rewrite rules include a kinetic constant that allows a quantitative analysis of the pathway dynamics to be performed. Moreover, reactions 4,8 and 16 that are described in [23] by using Michaelis-Menten kinetic parameters are translated each into three rewrite rules describing corresponding reactions through proper kinetic constants.

The initial state of the pathway is described by a CLS term representing the cell membrane with receptors on its surface, extracellular EGF proteins and intracellular proteins. The cell membrane is modelled by a looping sequence in which R represents EGFR. Other elements are named as the proteins they represent.

$$Egf \mid \dots \mid Egf \mid (R \mid \dots \mid R)^L \mid (PLC_\gamma \mid \dots \mid Grb \mid \dots \mid Shc \mid \dots \mid SOS \mid \dots).$$

The number of instances of Egf , R , PLC_γ , Grb , Shc and SOS elements depends on the considered concentrations of the corresponding proteins. We assume to include in the term a number of instances of an element equal to the value of the considered concentration of the corresponding protein expressed in nanomoles/liter. Note that, as in [23], the concentration of EGF is multiplied in the model by a factor ~ 33.3 to formally scale it to the cytoplasmic water volume. The initial concentrations of proteins considered in all simulations (if not differently specified) is the following:

Table 1. Initial reactions for the formation of phosphorylated dimers

<i>id</i>	<i>Rewrite rule</i>	<i>Kinetic constant</i>
r1	$Egf \mid (R \mid X)^L \mid Y \mapsto (Ra \mid X)^L \mid Y$	0.003
rn1	$(Ra \mid X)^L \mid Y \mapsto Egf \mid (R \mid X)^L \mid Y$	0.06
r2	$Ra \mid Ra \mapsto R2$	0.01
rn2	$R2 \mapsto Ra \mid Ra$	0.1
r3	$R2 \mapsto RP$	1.0
rn3	$RP \mapsto R2$	0.01
r4a	$(RP \mid X)^L \mid (E1 \mid Y) \mapsto (RPE1 \mid X)^L \mid Y$	5
r4b	$(RPE1 \mid X)^L \mid Y \mapsto (RP \mid X)^L \mid (E1 \mid Y)$	25
r4c	$(RPE1 \mid X)^L \mid Y \mapsto (R2 \mid X)^L \mid (E1 \mid Y)$	225

Table 2. Reaction in the PLC branch

<i>id</i>	<i>Rewrite rule</i>	<i>Kinetic constant</i>
r5	$(RP X)^L \rfloor (PLCgamma Y) \mapsto (RPL X)^L \rfloor (Y)$	0.06
m5	$(RPL X)^L \rfloor (Y) \mapsto RP (RP X)^L \rfloor (PLCgamma Y)$	0.2
r6	$RPL \mapsto RPLP$	1
m6	$RPLP \mapsto RPL$	0.05
r7	$(RPLP X)^L \rfloor (Y) \mapsto (RP X)^L \rfloor (PLCgammaP Y)$	0.3
m7	$(RP X)^L \rfloor (PLCgammaP Y) \mapsto (RPLP X)^L \rfloor (Y)$	0.006
r8a	$PLCgammaP E2 \mapsto PLCgammaPE2$	0.01
r8b	$PLCgammaPE2 \mapsto PLCgammaP E2$	0.95
r8c	$PLCgammaPE2 \mapsto PLCgamma E2$	0.05
r25	$PLCgammaP \mapsto PLCgammaI$	1
m25	$PLCgammaI \mapsto PLCgammaP$	0.03

Table 3. Reaction in the Grb-SOS branch

<i>id</i>	<i>Rewrite rule</i>	<i>Kinetic constant</i>
r9	$(RP X)^L \rfloor (Grb Y) \mapsto (RG X)^L \rfloor Y$	0.003
m9	$(RG X)^L \rfloor Y \mapsto (RP X)^L \rfloor (Grb Y)$	0.05
r10	$(RG X)^L \rfloor (SOS Y) \mapsto (RGS X)^L \rfloor (Y)$	0.01
m10	$(RGS X)^L \rfloor (Y) \mapsto (RG X)^L \rfloor (SOS Y)$	0.06
r11	$(RGS X)^L \rfloor (Y) \mapsto (RP X)^L \rfloor (GS Y)$	0.03
m11	$(RP X)^L \rfloor (GS Y) \mapsto (RGS X)^L \rfloor (Y)$	0.0045
r12	$GS \mapsto Grb SOS$	0.0015
m12	$Grb SOS \mapsto GS$	0.0001

$[EGFR] = 100nM$, $[EGF] = 680nM$ (corresponding to actual $\sim 20nM$), $[PLC\gamma] = 105nM$, $[Grb] = 85nM$, $[Shc] = 150nM$ and $[SOS] = 34nM$.

We perform both deterministic and stochastic simulation of the CLS model with the aim of better define, in the above concentration network, the relative relevance of the three branches of the signalling pathway. Moreover, the analysis is furthered by varying the concentration of some of the involved proteins to get more insight on their relevance in modulating the pathway.

In Figs. 3 and 4 we give simulation results describing the pathway dynamics in a macroscopic manner. In fact, what is observed in such graphs is the change in the overall concentration with the time of phosphorylated EGFR (including the fraction

Table 4. Reaction in the Shc-Grb-SOS branch

<i>id</i>	<i>Rewrite rule</i>	<i>Kinetic constant</i>
r13	$(RP X)^L \downarrow (Shc Y) \mapsto (RSh X)^L \downarrow Y$	0.09
m13	$(RSh X)^L \downarrow Y \mapsto (RP X)^L \downarrow (Shc Y)$	0.6
r14	$RSh \mapsto RShP$	6
m14	$RShP \mapsto RSh$	0.06
r15	$(RShP X)^L \downarrow Y \mapsto (RP X)^L \downarrow (ShP Y)$	0.3
m15	$(RP X)^L \downarrow (ShP Y) \mapsto (RShP X)^L \downarrow Y$	0.0009
r16a	$ShP E3 \mapsto ShPE3$	0.001
r16b	$ShPE3 \mapsto ShP E3$	0.225
r16c	$ShPE3 \mapsto Shc E3$	0.085
r17	$(RShP X)^L \downarrow (Grb Y) \mapsto (RShG X)^L \downarrow Y$	0.003
m17	$(RShG X)^L \downarrow Y \mapsto (RShP X)^L \downarrow (Grb Y)$	0.1
r18	$(RShG X)^L \downarrow Y \mapsto (RP X)^L \downarrow (ShG Y)$	0.3
m18	$(RP X)^L \downarrow (ShG Y) \mapsto (RShG X)^L \downarrow Y$	0.0009
r19	$(RShG X)^L \downarrow (SOS Y) \mapsto (RShGS X)^L \downarrow Y$	0.01
m19	$(RShGS X)^L \downarrow Y \mapsto (RShG X)^L \downarrow (SOS Y)$	0.00214
r20	$(RShGS X)^L \downarrow Y \mapsto (RP X)^L \downarrow (ShGS Y)$	0.12
m20	$(RP X)^L \downarrow (ShGS Y) \mapsto (RShGS X)^L \downarrow Y$	0.00024
r21	$ShP Grb \mapsto ShG$	0.003
m21	$ShG \mapsto ShP Grb$	0.1
r22	$ShG SOS \mapsto ShGS$	0.03
m22	$ShGS \mapsto ShG SOS$	0.064
r23	$ShGS \mapsto ShP GS$	0.1
m23	$ShP GS \mapsto ShGS$	0.021
r24	$(RShP X)^L \downarrow (GS Y) \mapsto (RShGS X)^L \downarrow Y$	0.009
m24	$(RShGS X)^L \downarrow Y \mapsto (RShP X)^L \downarrow (GS Y)$	0.0429

bound to proteins of the signalling network), by varying the concentration of EGF in the cell environment. Results of both deterministic and stochastic simulations are shown in the figures. The results of deterministic simulations are in line with those in [23], and show that a high concentration of EGF (20 nM) produces a spike in the formation of phosphorylated EGFR. The results of stochastic simulation are actually the average of ten simulations for each value of EGF concentration. In the stochastic simulation the population of proteins is approximated by assuming a concentration of 1 nM as an individual. This approximation is acceptable when concentrations

Fig. 3. Deterministic simulations of the concentration of phosphorylated EGFR (including those bound to proteins in the cytoplasm) by varying the EGF concentration (20 nM, 2 nM and 0.2 nM)

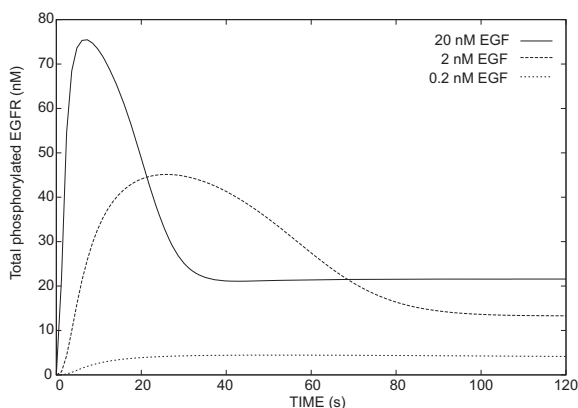
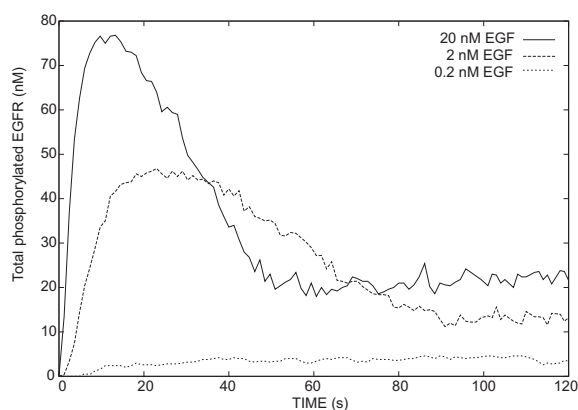


Fig. 4. Stochastic simulations of the concentration of phosphorylated EGFR (including those bound to proteins in the cytoplasm) by varying the EGF concentration (20 nM, 2 nM and 0.2 nM)



are rather high and stochastic simulation can exhibit sensitivity of the pathway to stochastic noise. In this case the macroscopic dynamics obtained by the stochastic approach is essentially the same coming from deterministic simulations. In fact, there is only a slight difference in the shape of the spike caused by the higher concentration of EGF, which in the stochastic simulation has a wider amplitude. This comparison is not repeated for the other simulations where the deterministic approach is adopted.

Fig. 5 shows the distribution of phosphorylated EGFR among the three branches of the pathway. Each panel refers to a different concentration of EGF. In the first two panels, namely for concentrations of EGF which give rise to a spike for the total concentration of phosphorylated EGFR (see Fig. 3), a spike for the curves representing the PLC and Shc-Grb-SOS branches is also observed. On the contrary, the Grb-SOS branch does not contribute to the spike. It slowly leads to an accumulation of phosphorylated EGFR by reaching a steady-state concentration higher than the one of the other two branches. For a more clear comparison the time course of total phosphorylated EGFR already shown in Fig. 3 is reported in the panels of Fig. 5.

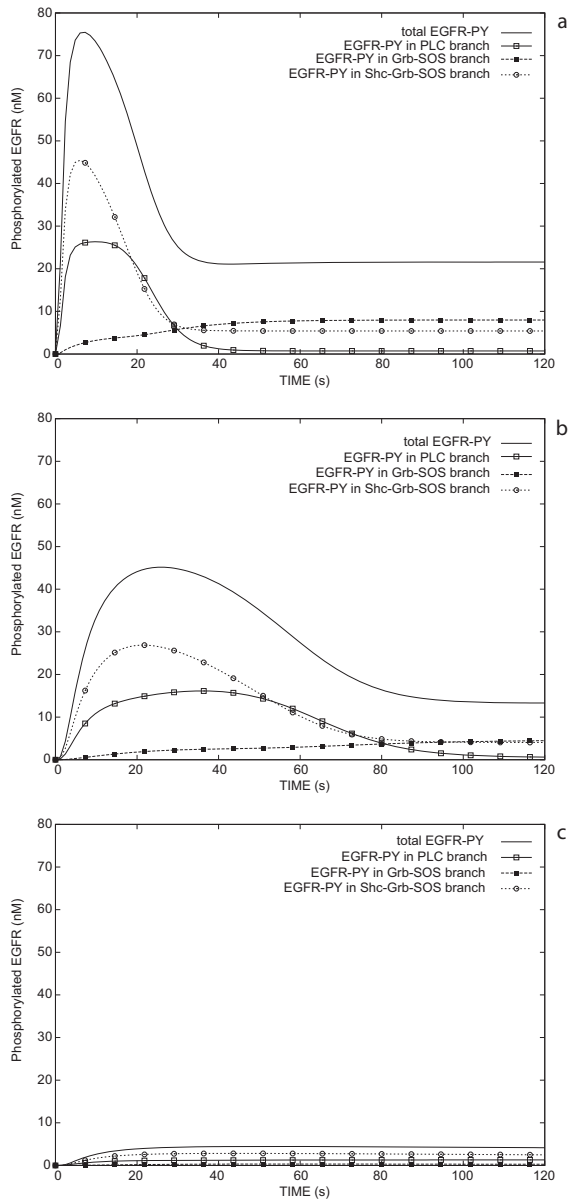


Fig. 5. Phosphorylated EGFR involved in each of the three branches of the pathway by varying the EGF concentration: 20 nM (a), 2 nM (b) and 0.2 nM (c)

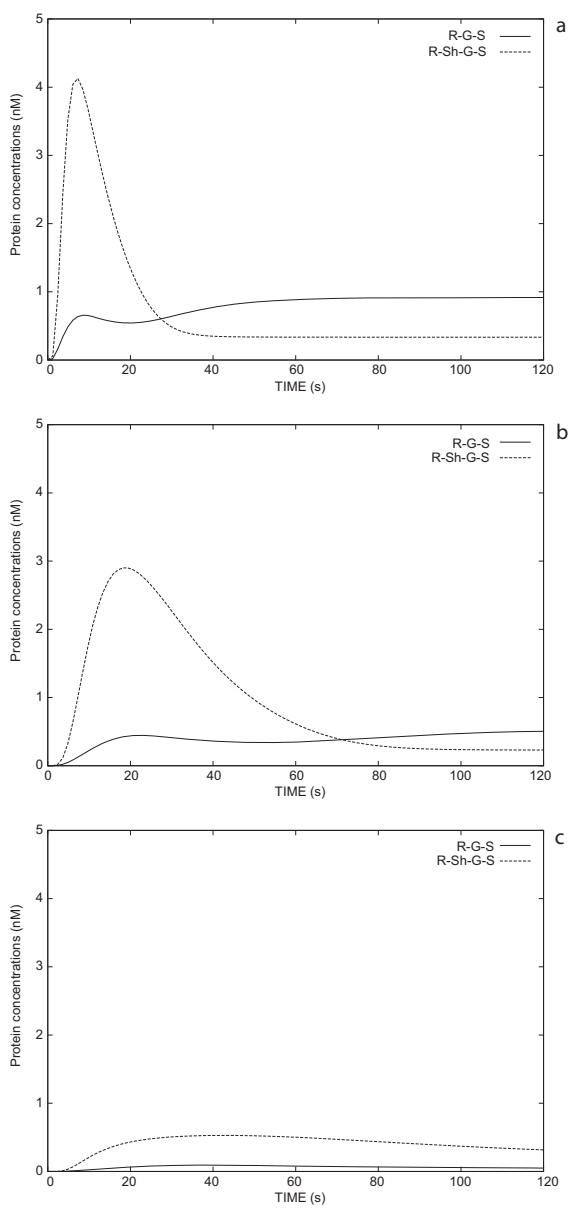


Fig. 6. Production of R-G-S and R-Sh-G-S complexes by varying the EGF concentration: 20 nM (a), 2 nM (b) and 0.2 nM (c)

Fig. 6 shows another important macroscopic result of the pathway, namely the production of R-Sh-G-S and R-G-S complexes. Indeed, such complexes, modulating the activity of the Ras protein, are the connection between the transduction phase of the EGFR signalling network and the true amplification phase of the MAP kinase enzyme cascade. Obviously, the production of the complexes is related with the activity of the Shc-Grb-SOS and of the Grb-SOS branches. In fact, we can observe that higher concentrations of EGF give rise to a spike in the concentration of the R-Sh-G-S complex. On the other hand, R-G-S has a more regular dynamics.

Simulation results show that the Shc-Grb-SOS branch seems to have a more central role in the dynamics of the pathway since in the initial times it attracts a major part of the phosphorylated EGFR and it produces a high quantity of complexes having an effect on the Ras protein in the subsequent steps of the pathway. The PLC branch is also a strong attractor of phosphorylated EGFR, but it does not lead to the production of any substance influencing Ras. Finally, the Grb-SOS branch even though able to intervene on Ras modulation, since its small production of R-G-S appears to have a minor impact on the activation of the cascade.

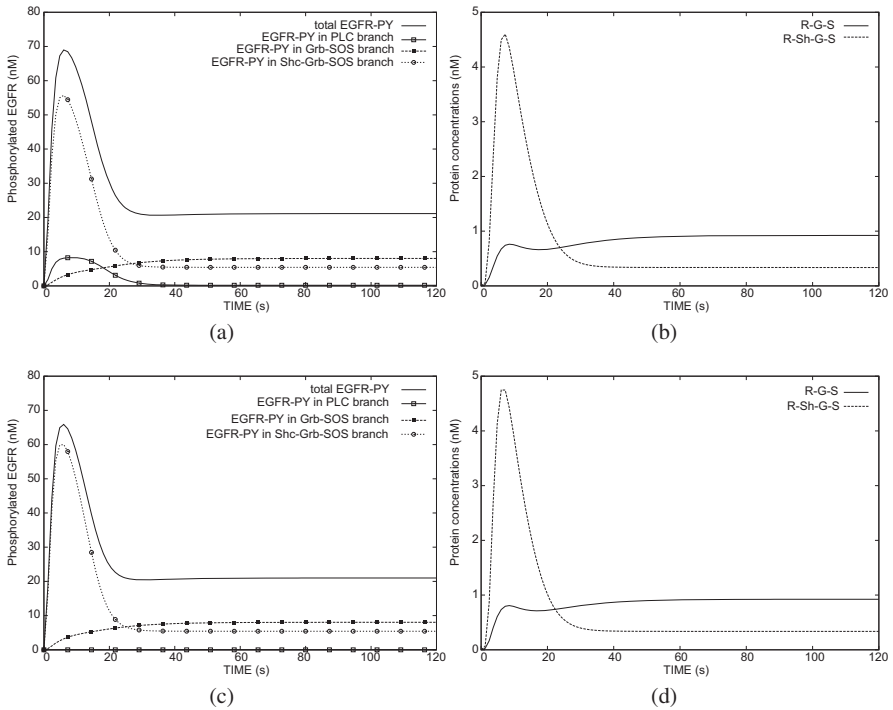


Fig. 7. Comparison by varying PLC γ : reduced to 25% in (a) and (b); absent in (c) and (d)

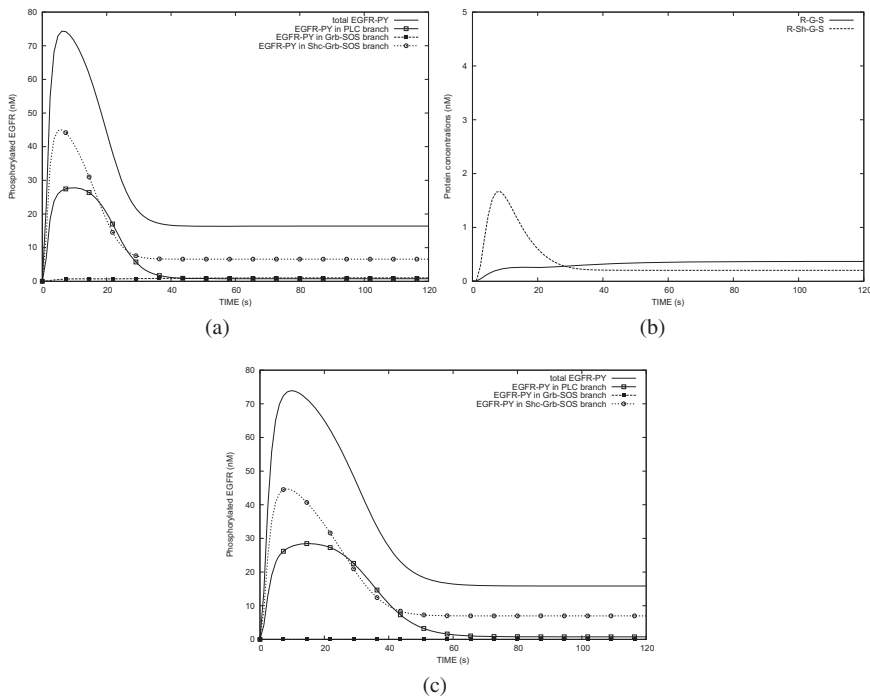


Fig. 8. Comparison by varying Grb: reduced to 25% in (a) and (b); absent in (c)

The modulation of the overall pathway and of single branches by changes in concentration of some of the constitutive proteins was studied. In particular, Fig. 7 shows the simulation results obtained when the concentration of PLC_γ was firstly reduced to 25% of the initial value and then brought to zero. For each condition, the overall concentrations of phosphorylated EGFR, its distribution in the three branches and the concentrations of R-G-S and R-Sh-G-S complexes are reported.

As expected, a reduced competition of the PLC branch for phosphorylated EGFR, following the imposed reduction of PLC_γ level, favours other branches. In particular, the Shc-Grb-SOS branch attracts almost all of the phosphorylated EGFR that is no longer involved in the PLC branch. As a result, the spike of the concentration of the R-Sh-G-S complex reaches higher values, even though slightly. On the other hand, the dynamics of the Grb-SOS branch as well as the level of the R-G-S concentration do not change significantly.

Fig. 8 shows the simulation results obtained when the concentration of Grb was firstly reduced to 25% of the initial value and then brought to zero. For each condition, the overall concentrations of phosphorylated EGFR, its distribution in the three branches and the concentrations of R-G-S and R-Sh-G-S complexes are reported.

As shown in the figure, a reduced concentration of Grb has a significant impact in the generation of the R-G-S and R-Sh-G-S complexes (panel (b) of the figure). Obviously, no complexes are formed when Grb is absent.

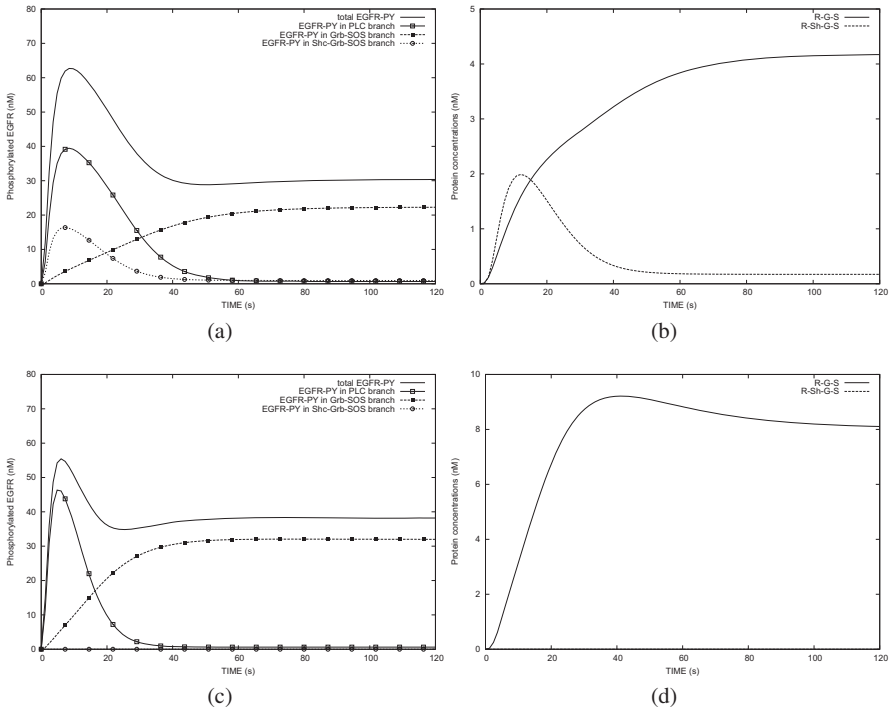


Fig. 9. Comparison by varying Shc: reduced to 25% in (a) and (b); absent in (c) and (d)

Finally, Fig. 9 shows the simulation results obtained when the concentration of Shc was firstly reduced to 25% of the initial value and then brought to zero. For each condition, the overall concentrations of phosphorylated EGFR, its distribution in the three branches and the concentrations of R-G-S and R-Sh-G-S complexes are reported.

A reduced concentration of Shc completely changes the macroscopic dynamics of the pathway. In this case due to the reduced level of Shc, the Shc-Grb-SOS branch loses its central role. In this conditions the Grb-SOS branch gains more importance, leading to a relevant and sustained production of R-G-S complex. Since its direct effect on Ras activation, the maintenance of R-G-S complex at high levels for a long time may induce unpredictable effects. In this case it would be necessary to include in the considered reaction scheme other existing signal ceasing mechanisms.

4 Conclusion

The Calculus of Looping Sequences (CLS) is suitable to describe biological systems and their evolution. Terms of the calculus are constructed by basic constituent elements and operators of sequencing, looping, containment and parallel composition. The looping operator, together with the containment one, permits the description of

arbitrarily nested membranes and elements outside, inside or on the surface of the membranes themselves. The evolution of a term is modeled by a set of rewrite rules and the semantics is a transition system, in which states correspond to terms, and transitions correspond to rules applications.

Stochastic CLS allows the description of quantitative aspects of the modeled systems, such as the frequency of chemical reactions. As examples of application to real biological systems, we showed the simulation of the activity of the EGFR signalling cascade.

Stochastic CLS fulfills the wanted request of having a simple notation with the ability of describing biological systems at different abstraction levels, and being flexible enough to allow the description of new kinds of phenomena. In fact rewrite rules as those of CLS are more similar to the reaction notation employed by biologists and a model may be easy understood.

References

1. Alur, R., Belta, C., Ivancic, F., Kumar, V., Mintz, M., Pappas, G.J., Rubin, H., Schug, J.: Hybrid Modeling and Simulation of Biomolecular Networks. In: Hybrid Systems: Computation and Control, LNCS 2034, pp. 19–32, Springer, Heidelberg (2001)
2. Barbuti, R., Cataudella, S., Maggiolo-Schettini, A., Milazzo, P., Troina, A.: A Probabilistic Model for Molecular Systems. *Fund. Inform.* **67**, 13–27 (2005)
3. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P.: Extending the Calculus of Looping Sequences to Model Protein Interaction at the Domain Level. In: Int. Symposium on Bioinformatics Research and Applications (ISBRA'07), LNBI 4463, pp. 638–649, Springer, Heidelberg (2006)
4. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Tiberi, P., Troina, A.: Stochastic Calculus of Looping Sequences for the Modeling and Simulation of Biological Systems. *T. Comp. Sys Biology* **9**, 86–113 (2008)
5. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Troina, A.: Bisimulation Congruences in the Calculus of Looping Sequences. In: Int. Colloquium on Theoretical Aspects of Computing (ICTAC'06), LNCS 4281, pp. 93–107, Springer, Heidelberg (2006)
6. Barbuti, R., Maggiolo-Schettini, A., Milazzo, P., Troina, A.: A Calculus of Looping Sequences for Modelling Microbiological Systems. *Fund. Inform.* **72**, 21–35 (2006)
7. Brodo, L., Degano, P., Priami, C.: A Stochastic Semantics for BioAmbients. In: Parallel Computing Technologies (PaCT'07), LNCS 4671, pp. 22–34, Springer, Heidelberg (2007)
8. Cardelli, L.: Brane Calculi. Interactions of Biological Membranes. In: Computational Methods in Systems Biology (CMSB'04), LNCS 3082, pp. 257–280, Springer, Heidelberg (2005)
9. Cardelli, L., Gordon, A.D.: Mobile Ambients. *Theor. Comput. Sci.* **240**, 177–213 (2000)
10. Chabrier-Rivier, N., Chiaverini, M., Danos, V., Fages, F., Schachter, V.: Modeling and Querying Biomolecular Interaction Networks. *Theor. Comput. Sci.* **325**, 25–44 (2004)
11. Ciardiello, F., Tortora, G.: EGFR Antagonist in Cancer Treatment. *N. Engl. J. Med.* **358**, 1160–1174 (2008)
12. Curti, M., Degano, P., Priami, C., Baldari, A.D.: Modelling Biochemical Pathways through Enhanced pi-calculus. *Theor. Comput. Sci.* **325**, 111–140 (2004)
13. Damm, W., Harel, D.: LSCs: Breathing Life into Message Sequence Charts. *Form. Methods Syst. Des.* **19** (2001)
14. Danos, V., Laneve, C.: Formal Molecular Biology. *Theor. Comput. Sci.* **325**, 69–110 (2004)

15. Danos, V., Pradalier, S.: Projective Brane Calculus. In: Computational Methods in Systems Biology (CMSB'04), LNCS 3082, pp. 134–148, Springer, Heidelberg (2005)
16. Degano, P., Prandi, D., Priami, C., Quaglia, P.: Beta-binders for Biological Quantitative Experiments. In: Quantitative Aspects of Programming Languages (QAPL 2006), ENTCS 164(3), pp. 101–117 (2006)
17. Efroni, S., Choen, I.R., Harel, D.: Toward Rigorous Comprehension of Biological Complexity: Modeling, Execution and Visualization of Thymic t-cell Maturation. *Genome Res.* **13** 2485–2497 (2003)
18. Gillespie, D.: Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **81**, 2340–2361 (1977)
19. Harel, D.: Statecharts: A Visual Formalism for Complex Systems. *Sci. Comput. Program.* **8**, 231–274 (1987)
20. Harel, D.: A Grand Challenge: Full Reactive Modeling of a Multi-cellular Animal. *Bulletin of the EATCS* **81**, 226–235 (2003)
21. Kam, N., Cohen, I.R., Harel, D.: The Immune System as a Reactive System: Modeling t-cell Activation with Statecharts. In: Symposia on Human Centric Computing Languages and Environments (HCC'01), pp. 15, IEEE Computer Society (2001)
22. Kam, N., Harel, D., Kugler, H., Marelly, R., Pnueli, A., Hubbard, E.J.A., Stern, M.J.: Formal Modeling of *C. elegans* Development: A Scenario-Based Approach. In: Computational Methods in Systems Biology (CMSB'03), LNCS 2602, pp. 4–20, Springer, Heidelberg (2003)
23. Kholodenko, B.N., Demin, O.V., Moehren, G., Hoek, J.B.: Quantification of Short Term Signaling by the Epidermal Growth Factor Receptor. *J. Biol. Chem.* **274**, 30169–30181 (1999)
24. Kwiatkowska, M., Norman, G., Parker, D.: Probabilistic Symbolic Model Checking with PRISM: a Hybrid Approach. *Int. Journal on Software Tools for Technology Transfer* **6**, 128–142 (2004)
25. Han, W., Lo, H.W.: Landscape of EGFR signaling network in human cancers: biology and therapeutic response in relation to receptor subcellular locations. *Cancer Lett.* **28**, 124–134 (2012)
26. Laneve, C., Tarissan, F.: A Simple Calculus for Proteins and Cells. In: Workshop on Membrane Computing and Biological Inspired Process Calculi (McBIC'06), ENTCS 171, pp. 139–154 (2007)
27. Matsuno, H., Doi, A., Nagasaki, M., Miyano, S.: Hybrid Petri Net Representation of Gene Regulatory Network. In: Pacific Symposium on Biocomputing, World Scientific Press, pp. 341–352 (2000)
28. Milazzo, P.: Qualitative and Quantitative Formal Modeling of Biological Systems. PhD Thesis, Università di Pisa (2007)
29. Păun, G.: Computing with Membranes. *J. Comput. Syst. Sci.* **61**, 108–143 (2000)
30. Păun, G.: Membrane Computing. An Introduction. Springer, Heidelberg (2002)
31. Pérez-Jiménez, M.J., Romero-Campero, F.J.: A Study of the Robustness of the EGFR Signalling Cascade Using Continuous Membrane Systems. In: IWINAC'05, LNCS 3561, pp. 268–278, Springer, Heidelberg (2005)
32. Pérez-Jiménez, M.J., Romero-Campero, F.J.: P Systems, a New Computational Modelling Tool for Systems Biology. *Trans. Comput. Systems Biology VI*, LNBI 4220, 176–197, Springer, Heidelberg (2006)
33. Priami, C.: Stochastic π -Calculus. *Comput. J.* **38**, 578–589 (1995)
34. Priami, C., Quaglia, P.: Beta Binders for Biological Interactions. In: Computational Methods in Systems Biology (CMSB'04), LNCS 3082, pp. 20–33, Springer, Heidelberg (2005)

35. Priami, C., Regev, A., Silvermann, W., Shapiro, E.: Application of a Stochastic Name-Passing Calculus to Representation and Simulation of Molecular Processes. *Inform. Process. Lett.* **80**, 25–31 (2001)
36. Prusinkiewicz, P., Lindenmayer, A.: *The Algorithmic Beauty of Plants*. Springer, Heidelberg (1990)
37. Regev, A., Panina, E.M., Silverman, W., Cardelli, L., Shapiro, E.: BioAmbients: An Abstraction for Biological Compartments. *Theor. Comput. Sci.* **325**, 141–167 (2004)
38. Regev, A., Shapiro, E.: Cells as Computation. *Nature* **419**, 343 (2002)
39. Regev, A., Silverman, W., Shapiro, E.Y.: Representation and Simulation of Biochemical Processes Using the pi-calculus Process Algebra. In: *Pacific Symposium on Biocomputing*, World Scientific Press, pp. 459–470 (2001)
40. The P Systems web page: <http://psystems.disco.unimib.it/>
41. Wiley, H.S., Shvartsman, S.Y., Lauffenburger, D.A.: Computational Modeling of the EGF-Receptor System: a Paradigm for Systems Biology. *Trends in Cell Biology* **13**, 43–50 (2003)
42. Yasuda, H., Kobayashi, S., Costa, D.B.: EGFR exon 20 insertion mutations in non-small-cell lung cancer: preclinical data and clinical implications. *Lancet Oncol.* **13**, e23–31 (2012)
43. Zwick, E., Hackel, P.O., Prenzel, N., Ullrich, A.: The EGF receptor as central transducer of heterologous signalling systems. *Trends Pharmacol. Sci.* **20**, 408–412 (1999)

Dynamic Simulations of Pathways Downstream of TGF β , Wnt and EGF-Family Growth Factors, in Colorectal Cancer, including Mutations and Treatments with Onco-Protein Inhibitors

Lorenzo Tortolina, Nicoletta Castagnino, Cristina De Ambrosi, Annalisa Barla, Alessandro Verri, Gabriele Zoppoli, Luca Bagnasco, Daniela Piras, Franco Patrone, Alberto Ballestrero, and Silvio Parodi

Abstract With reference to colorectal cancer, we have reconstructed a Molecular Interaction Map downstream of TGF β , Wnt and EGF-family. Based on an extensive and systematic direct/indirect data extrapolation from several dozens of published experimental papers, and some data interpolation that could fit with the general behavior of this signaling-network region, we were able to obtain an operative

L. Tortolina · N. Castagnino · S. Parodi (✉)

Department of Internal Medicine (Di.M.I.), Research Center for Computational Learning (CRAC), IRCCS Azienda Ospedaliera Universitaria San Martino-IST, Istituto Superiore di Oncologia (ISO), Largo Benzi 10, 16132 Genoa, Italy
e-mail: lorenzo.tortolina@unige.it, nicoletta.castagnino@unige.it, silvio.parodi@unige.it

C. De Ambrosi

Department of Internal Medicine (Di.M.I.), Department of Computer and Information Science (DISI), Research Center for Computational Learning (CRAC), IRCCS Azienda Ospedaliera Universitaria San Martino-IST, Largo Benzi 10, 16132 Genoa, Italy
e-mail: cristina.deambrosi@unige.it

A. Barla · A. Verri

Department of Computer and Information Science (DISI), Research Center for Computational Learning (CRAC), University of Genoa, via Dodecaneso 35, 16146 Genoa, Italy
e-mail: annalisa.barla@unige.it, alessandro.verri@unige.it

G. Zoppoli

Department of Internal Medicine (Di.M.I.), Research Center for Computational Learning (CRAC), IRCCS Azienda Ospedaliera Universitaria San Martino-IST, Largo Benzi 10, 16132 Genoa, Italy
e-mail: gabriele.zoppoli@unige.it

L. Bagnasco · D. Piras

Department of Internal Medicine (Di.M.I.), IRCCS Azienda Ospedaliera Universitaria San Martino-IST, Istituto Superiore di Oncologia (ISO), Largo Benzi 10, 16132 Genoa, Italy
e-mail: luca.bagnasco@unige.it, piras-daniela@tiscali.it

F. Patrone · A. Ballestrero

Department of Internal Medicine (Di.M.I.), IRCCS Azienda Ospedaliera Universitaria San Martino-IST, Largo Benzi 10, 16132 Genoa, Italy
e-mail: fpatrone@unige.it, aballestrero@unige.it

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_7, © Springer-Verlag Italia 2012

mathematical simulation model. We could simulate normal conditions of the network, behavior in the presence of important colorectal cancer mutations, behavior in the presence of virtual drug inhibitors of different specifically altered onco-proteins affected by excess of function. The dynamic behavior of the simulation seems quite reasonable, in terms of what is known about the physiology and the pathology of this signaling-network region. Preliminary experimental verification experiments look encouraging.

1 Introduction

Modern medicine and biomedical research is based on a detailed understanding of the development, function, maintenance, and disease of our organs. While (in some instances and in some respects) we have accumulated a vast amount of knowledge, we are still insufficiently capable of reconstructing, modeling and simulating how cells, tissues and organs, become diseased or how they would respond to potential therapies. A predictive simulation could perhaps be the ultimate test of our basic understanding of a very complex set of phenomena. We cannot work at the same level of detail in front of increasing complexities. Moving from a network fragment of biochemical interactions involving a limited aspect of the cell cycle, to the multicellular level of a basic tissue unit like a colon crypt (proliferation, differentiation, tissue architecture involved), to even higher levels of integration, we are paradoxically forced to increase the shallowness of our approach, and the type of mathematical modeling. At the level of networks of biochemical interactions we can simulate molecular changes relevant in controlling different phases of the cell cycle (for instance for stem cells at the bottom of a colon crypt). A typical example is the G0 – G1 transition, where cells decide to stay quiescent or to start the replication process. At the level of biochemical networks concerning a reference individual cell, we can try to test our real comprehension of what is going on, introducing virtual mutations, alterations, aberrant behaviors, that progressively occur during the process of malignant transformation, with the goal of a better understanding of the biological rules governing health and cancer disease. The activity of target-selective drugs and / or siRNAs capable of inhibition of *mutated genes/proteins affected by excess of function*, can also be simulated in our mathematical modeling. To predict the final fate of a cell (for instance proliferation, differentiation or death) could be too risky, because it includes many more pathways of what have been considered here. Prediction of effects immediately adjacent to or inside our network fragment (a phosphorylated protein or the activation of a specific transcription factor) is probably a more conservative but safer approach. New altered functions introduced via gene transfection in experimental models of reference can be simulated well. The same is true for the activity of onco-protein inhibitors. An important point is the distinction between *pre-clinical/(sometimes clinical) information used during the training phase* of a given model (introduction of multiple constraints), usually a rather extensive and multifaceted (both direct and indirect) set of information obtained from literature data, and *subsequent independent wet lab verification experiments*. It seems appropriate that

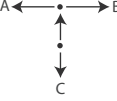
during the training phase we tend to consider satisfactory only behaviors in agreement with whatever known and sufficiently sound preclinical and clinical datum, utilized directly or via reasonable extrapolations/deductions. The molecular oncologist specific expertise concerning a given cancer becomes important at this stage. An extensive and successful training phase, followed by dynamic simulations, will generate predictions concerning the subsequent verification experiments. It is however of note that these expectations become possible only a posteriori of the mathematical modeling: *the size of our signaling-network is largely beyond the breaking point of a direct a priori intuitive capability of an expert human observer (shared opinion [2, 28])*. We adopted the approach of reconstructing the molecular anatomy of a network relevant in colorectal cancer through a Molecular Interaction Map (MIM), drawn according to the syntactic rules devised by Kurt W. Kohn [1, 19–21] (Fig. 1). A MIM is also accompanied by a glossary offering a synthetic information about the molecular species represented in the MIM (not reported here). A MIM is also accompanied by an *Annotation List*, in which all the references that have allowed the MIM reconstruction and direct or indirect assignment of many parameters are listed (see Supplemental Material of [39]. For us, a MIM is a descriptive heuristic guide for subsequent writing of all the reported reactions (in terms of ODEs).

For a mathematical modeling of signaling networks, quantitative data are required (to be obtained directly or indirectly from wet-lab works, within the limits of acceptable approximations and extrapolations). Despite a continuous increase in the amount and quality of experimental data, direct quantitative measurements of concentrations and interaction rates, concerning many molecular components, in time and space, are still relatively scarce, *at least as a direct punctual information*.

Reaction symbols

	Non-covalent binding
	Covalent bond
	Transcription
	Stoichiometric conversion
	Cleavage of a covalent bond
	Degradation
	Conformational modification

Molecular interaction symbols

	We can add a node between two interacting A and B molecules, to indicate the dimer (A-B). A node between a molecule C and a dimer-node indicates a trimer (A-B-C), and so on. Structural details about the interactions are omitted.
	Symbol for an A:A homo dimer

Contingency symbols

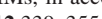
	Stimulation
	Requirement
	Requirement of one A event for one B event
	Inhibition
	Enzymatic catalysis

Fig. 1. The syntactic rules for drawing MIMs, in according to K.W. Kohn. Modified from [39]: Tortolina et al. *Curr Cancer Drug Targets* **12** 339–355 (2012) © Bentham Science Publishers

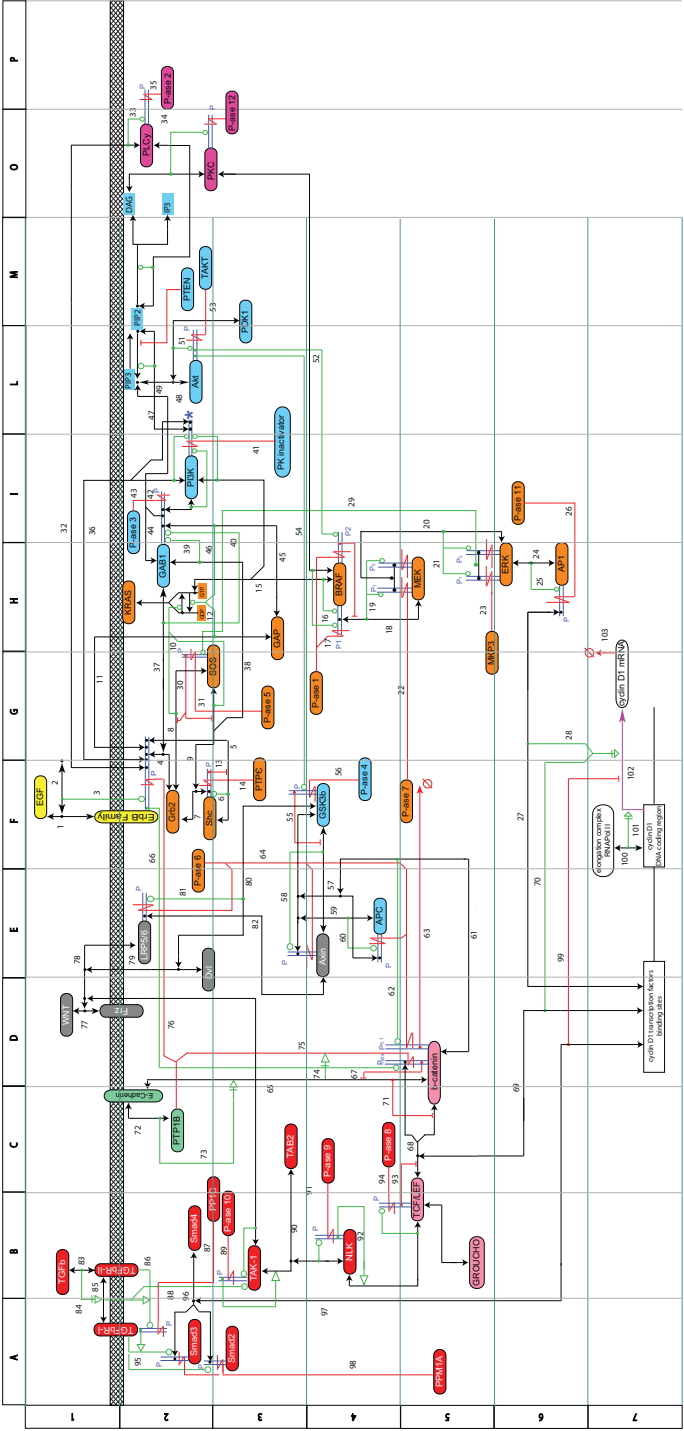


Fig. 2. A heuristic MIM downstream of TGFβ, WNT and EGF-family growth factors which are relevant in colorectal cancer. Relatively distinct pathways are represented in different colors. Modified from [39]: Tortolina et al. Curr Cancer Drug Targets 12 339–355 (2012) © Bentham Science Publishers

Model output is in principle dependent on parameter values, typically total protein concentrations and rate constants (forward and reverse rate constants for complexes formation, and catalytic rate constants for enzymatic activities). Approximate values and estimations could still be useful, suggesting *semi-quantitative relative response ranges*, capable of offering for instance *rank correlations with the results of wet lab experiments*. Correlations of this type were obtained by us in a previous work using a similar approach, centered on breast cancer [6]. Concentrations and rates can be derived from available literature (but in the majority of cases only through indirect extrapolations). A positive aspect is that the trained software must satisfy a multiplicity of direct and indirect constraints coming from a large set of published experimental results. It is much easier to find in literature (directly or through an elaboration of reported results) an equilibrium dissociation constant (K_d), rather than k_{diss} or k_{ass} rates (notice that $K_d = k_{diss}/k_{ass}$). Often k_{ass} is much less variable than k_{diss} . k_{diss} reflects the strength of the binding (which can vary dramatically), k_{ass} reflects the time required to the molecules to find each other in a given cellular compartment. We could try to give by default, for instance always a constant $k_{ass} = 10^7 \cdot M^{-1} \cdot sec^{-1}$ or a $k_{ass} = 10^8 \cdot M^{-1} \cdot sec^{-1}$ [38].

Training is a long patient process (a vast multiplicity of reported behaviors has to be satisfied). Perhaps a positive property of the behavior of this type of biochemical-interactions networks (graphs) could be associated with the fact that we repeatedly observed that 10 times x or / random parameter changes have negligible effects at more than three edges random distances (an edge separates two adjacent nodes), in directions not involving or opposite to signal flow. This observation coexists with the presence of very few non-random preferential walks in the propagation of the information. These seem mostly the pathways in which important mutations are involved (biologic hubs at the level of malignant transformation).

Following the empirical (multiple trials and errors) patient patchwork of adjustments required during the training phase, there could still be some *risk of retrotting the data*, a risk that we have difficulty to assess precisely, but that probably decreases with the incremental consideration of more and more reported literature results, which precisely decrease the freedom of a retrofitting of the data. In addition, it remains for the moment unanswered the question of the possible co-existence of more than one constellation of interpolated (experimentally unknown) parameters that still could satisfy the multiplicity of literature and our own experimental data.

For the moment, we have reached the provisional convincement that a subsequent robust (involving multiple pathways) experimental verification phase is essential to distinguish retrofitting from predictive capabilities.

Preferably these experimental verifications should concern outcomes not quasi-directly reported in previous experimental studies. We applied apparently successfully this strategy in a previous similar study centered on breast cancer [6].

2 Our Mathematical Model

Even if this is an oversimplification, in our simulations we were careful to move close to a short term stationary state of our biochemical network. It should be underlined that our network (see our MIM) simulates a biological cell time of probably less than 30 minutes. We stop at the beginning of c-myc and cyclin-D1 transcription, which take a longer time (roughly 1–4 hours), perhaps longer for cyclin-D1 than for c-myc. Slower processes of molecular neo-synthesis / degradation were not considered in our model, except in an oversimplified form: for instance a convenient constant rate of formation and degradation for PIP2 or β -catenin.

The dynamics of the signaling-network were numerically simulated by solving a system of ordinary differential equations (ODEs), with the help of dedicated software, such as the SimBiology toolbox of Matlab. The numerical solver adopted was *ode23tb*, a solver for stiff differential equations that is an implementation of the TR-BDF2 algorithm, an implicit Runge-Kutta formula.

Similar numerical solution approaches have been pursued by many authors [5, 7, 17, 27]. The pathways reconstructed in our MIM are downstream of TGF β , Wnt and EGF-family growth factors. They include 62 basic molecular species (proteins and other small molecules). In our signaling-network we considered 397 modified species and complexes (complexes with inhibitors included). Our mathematical modeling implies 811 reactions (reversible and catalytic reactions). To our knowledge, this is probably one of the largest (perhaps the largest) signaling-network ever simulated at the biochemical-interaction scale level. The pathways reconstructed in our MIM are not just parallel walks but, to some extent, they also intersect each other forming a veritable signaling-network structure. We have summarized the eight involved pathways (in some cases converging pathways) as follows:

1. Pathway [ErbB-family receptors – PI3K – PTEN – Akt – GSK3 γ – APC – β -catenin – TCF/LEF – DNA binding site 2, transcription agonist];
2. Pathway [ErbB-family receptors – Grb2 – Shc – SOS – GAP – KRAS – BRAF – MEK – ERK – AP1 – DNA binding site 3, transcription agonist];
3. Pathway [ErbB-family receptors-E-Cadherin (Cadherin/Catenin adhesive complex)];
4. Pathway [ErbB-family receptors – PLC γ – PIP2 – PKC – BRAF – MEK – ERK – AP1 – DNA binding site 3, transcription agonist, converging with pathway 2];
5. Pathway [WNT – Frz/LRP5/6 – Dvl – AXIN – APC – GSK3 γ – β -catenin];
6. Pathway [TGF β -receptors – SMAD2/3 – SMAD4 – DNA binding site 1, transcription antagonist];
7. Pathway [TGF β -receptors – TAK-1 – TAB2 – NLK – TCF/LEF, converging with 8];
8. Pathway [WNT – Frz/LRP5/6 – TAK-1 – TAB2 – NLK – TCF/LEF, converging with 7].

3 What Happens in the Presence of a Work in Progress?

As an example, we make reference to the pathway 5. Pathway [WNT – Frz/LRP5/6 – Dvl – AXIN – APC – GSK3 β – β -catenin]. Some reports suggest a more complicated picture, where an additional protein family is involved: CK1 (Casein Kinase 1). CK1 is an additional modulator of the phosphorylation (S45) of β -catenin. This β -catenin phosphorylation is the first step for β -catenin ubiquitination and degradation [31].

It is not unusual that MIMs can be further enriched of additional modulator proteins.

In our opinion, an actual MIM (at a given stage of reconstruction) has been implicitly parametrically adjusted by the constructor of the MIM, to be in a permissive condition toward additional finer regulations. In other words, the training of a less expanded MIM with the available experimental data, gives origin to a parameter fitting still allowing an overall functioning of the network, but expectedly at a coarser degree of refinement (an evolutionary process in MIM reconstruction?).

This empirical behavior of our MIMs, that we have repeatedly observed, is probably compatible with subsequent stages of MIM enrichment and refinement (*compatible with a work in progress*). Without this property of the MIMs, we could start to build mathematical predictive models only when in possession of a complete knowledge of any detail. We would never start.

We reconstructed a simplified, first approximation, formula, for cyclin D1 (or c-myc) mRNA levels, through the following algorithm [notice that, in the algorithm, TFBS is an abbreviation for **T**ranscription **F**actor **B**inding **S**ite. We considered *three adjacent, but distinct, DNA binding sites*. K1-5 are constants that avoid our formula becoming 0 or ∞ .

mRNA levels =

$$\frac{(\text{TCF/LEF}:\beta\text{-catenin:TFBS2} + \text{TCF/LEF}:\beta\text{-catenin\_PY654:TFBS2} + \text{K1})}{(\text{SMAD2-P:SMAD4:TFBS1} + \text{SMAD3-P:SMAD4:TFBS1} + \text{K4})} \cdot \frac{(\text{AP1:TFBS3}) + (\text{AP1-P:TFBS3} + \text{K2}) \cdot (\text{TCF/LEF:SMAD4:TFBS2} + \text{K3})}{(\text{GROUCHO:TCF/LEF:TFBS2} + \text{K5})}$$

The different DNA-binding complexes have been deduced from several experimental published papers, and utilized for a first approximation formula [8, 12, 18, 23, 26, 30, 33, 34, 36].

We have preliminary evidence that TCF/LEF related regulations of transcription could start to be significant already at two hours, and precede AP1 regulations of transcription via an intermediate role of Elk-1 [11]. The AP1 transcriptional role could start to be significant only at four hours. For c-myc could be more important the first group of transcription regulations and a direct dependence from the transcription factor Elk-1, for cyclin D1 could be more relevant the subsequent role of AP1 (David Duffy, Dublin, personal communication).

4 Results from our Mathematical Model

In the following figures we report some results generated by our trained / constrained mathematical model. We ‘learned’, from the reading of a large set of experimental reports, from the reconstruction of our MIM, and from multiple reiterated trainings of our mathematical modeling software, an ‘a posteriori feeling’ (reasonable expectations) of results that could make sense. In general, we could find (a posteriori) explanations of what is going on.

With reference to column 9 of Fig. 6, an additional comment is perhaps of interest. We start from the prediction (Fig. 6) that, in the presence of a double alteration (excess of function at the level of the EGFR-family + PI3K excess of function), a Pan-ERB inhibitor (a dual inhibitor that blocks both EGFR and ErbB2 simultaneously) alone is poorly effective (column 7). A PI3K inhibitor alone could even display a paradoxical stimulatory effect (column 8). The association of both inhibitors is predicted to be strongly effective (column 9)!

The medicinal chemistry (academy + pharmaceutical industry) is steadily generating, year after year, new inhibitors of new onco-proteins affected by excess of function.

At the same time, new-generation instruments for medium throughput DNA re-sequencing (50–100 pre-selected potential oncogenes), have started to reach the market.

Their operative cost could soon become compatible with routine clinical diagnostic applications.

Presently, the most frequent solid tumors are (at best) sequenced for 3–5 potential oncogenes (probability of discovering one mutated oncogene roughly about 50%).

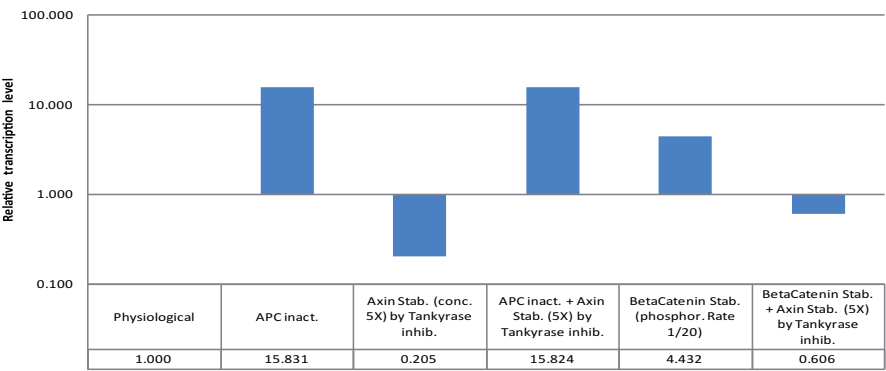


Fig. 3. Axin concentration versus β -catenin degradation. Tankyrase is a poly-ADP-ribose enzyme favoring axin degradation (ubiquitin-proteasome pathway). XAV939 (Novartis) is a tankyrase inhibitor [13]. Axin concentration can be increased by a tankyrase inhibitor. An increased axin concentration stimulates β -catenin degradation (reduced β -catenin:TCF/LEF transcription). Within an inactivated APC complex, an increased axin remains however unable to restore β -catenin phosphorylation and degradation (a simplified model)

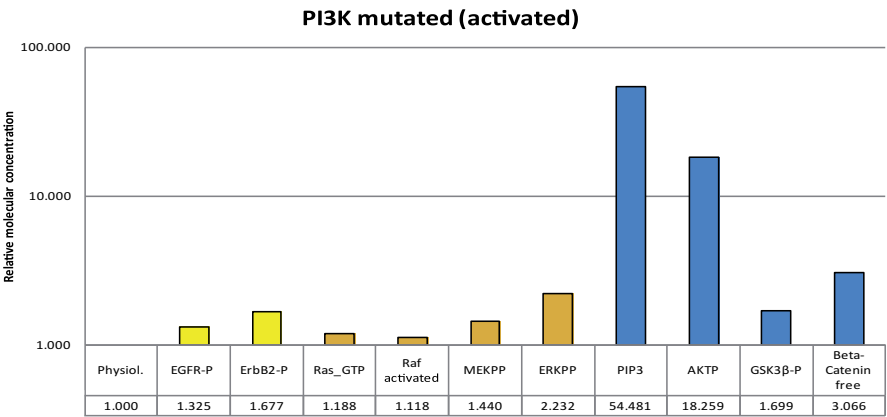


Fig. 4. A mutated PI3K-P (PI3K activated, strongly phosphorylated phosphatase susceptibility lost) has complex effects: the PIP2 → PIP3 transition is strongly increased. A nonlinear stimulation of the downstream pathway AKTP → GSK3γ-P → *free* β-catenin (essential for TCF/LEF: β-catenin formation + downstream transcription) can be observed. In addition (much less intuitive), PI3K-P is not binding anymore EGFR-family activated (-P) receptors. They are more available for other pathways, for instance the KRAS → BRAF → MEKPP → ERKPP pathway

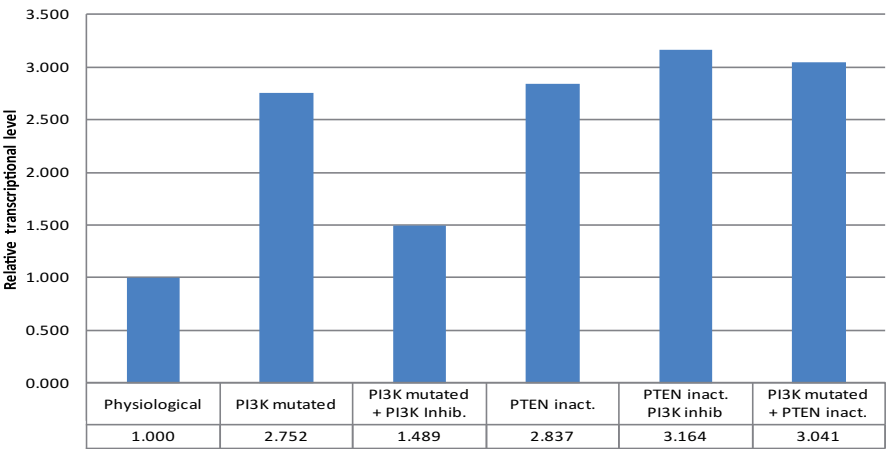


Fig. 5. Role of a PI3K-inhibitor along the PI3K → PTEN pathway. Notice that PTEN is downstream of PI3K. Notice that PI3K mutation and PTEN inactivation (along the same pathway) tend to be mutually exclusive (no additional gain, 6th column to the right)

Re-sequencing 50–100 potential oncogenes (the most frequently mutated oncogenes, for instance in a colorectal cancer), we could expect to detect in an individual tumor an average of 5–10 mutated oncogenes [35, 40].

At this point the characterization of the molecular pathology of an individual tumor will be much better (more thoroughly) defined. Therapeutic indications for associations of 2–3 selective drug inhibitors of 2–3 altered oncogenes affected by

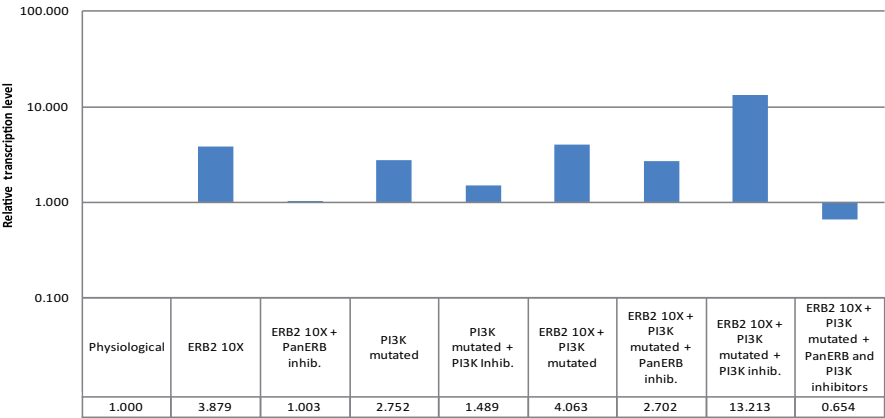


Fig. 6. Interplay of ErbB2 and PI3K alterations, with their respective inhibitors. Notice that ErbB-family is upstream of four different pathways: 1. Pathway [ErbB-family receptors – PI3K – PTEN – Akt – GSK3 γ – APC – β -catenin – TCF/LEF – DNA binding site 2, tr. agonist]; 2. Pathway [ErbB-family receptors – Grb2 – Shc – SOS- GAP- KRAS – BRAF – MEK – ERK – AP1 – DNA binding site 3, tr. agonist]; 3. Pathway [ErbB-family receptors-E-Cadherin (Cadherin/Catenin adhesive complex)]; 4. Pathway [ErbB-family receptors – PLC γ – PIP2 – PKC – BRAF – MEK – ERK – AP1 – DNA binding site 3, tr. agonist, converging with pathway 2]; this could explain the paradoxical activity observed in column eight (for instance stimulation of pathway 4 via an increased PIP2 and decreased PIP3, in the presence of PI3K inhibitor). Notice that the alteration of two distinct pathways requires the inhibition of both (column 9)

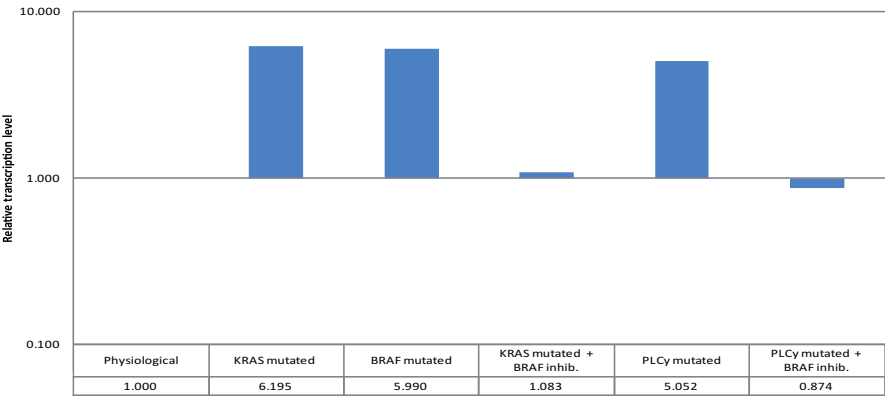


Fig. 7. Convergence on BRAF protein of two distinct pathways. Notice the convergence on BRAF of pathways two and four: Pathway 2. [ErbB-family receptors – Grb2 – Shc – SOS- GAP- KRAS – BRAF – MEK – ERK – AP1 – DNA binding site 3, tr. agonist]; pathway 4. [ErbB-family receptors – PLC γ – PIP2 – PKC – BRAF – MEK – ERK – AP1 – DNA binding site 3, tr. agonist, converging with pathway 2]

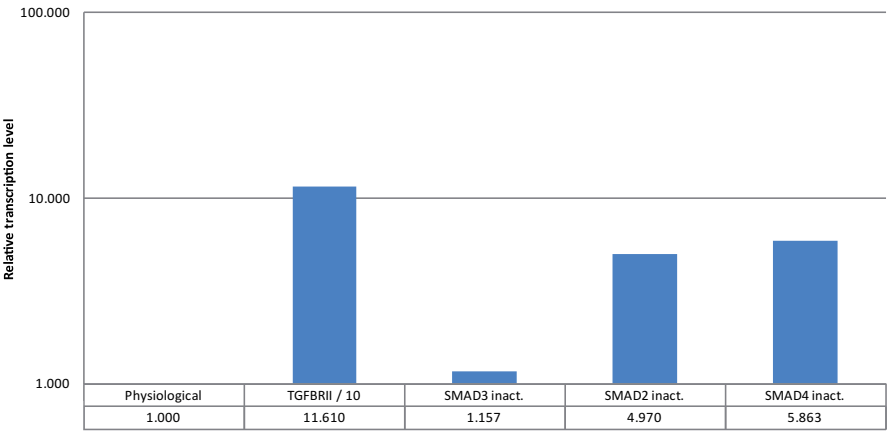


Fig. 8. Malfunctioning of $TGF\beta$ inhibitory pathway. Notice that SMAD2 has been reported to be 4–5 times more abundant than SMAD3. Also remember that the formed heterodimers are SMAD2-P:SMAD4 and SMAD3-P:SMAD4. Notice however that, in the transcription formula, the component (*TCF/LEF:SMAD4 :TFBS2+ K3*) stimulates transcription; SMAD4 inactivation could have complex different effects, in different cancer lines

excess of function (belonging to distinct pathways), could become a realistic feasible possibility (*see previous and subsequent gures*).

As also shown in Fig. 8, all the pathways described previously, converge on three distinct DNA-binding sites, exerting cooperative and / or antagonistic transcriptional controls.

We had in mind the final transcription of two important mRNAs (among others): c-myc and cyclin D1. These two molecules play key roles in the regulation of cell-cycle G0–G1 transition, and their expression is carefully regulated. Preliminary results in cancer lines carrying a β -catenin mutation or an APC inactivation, seem to suggest that a negative modulation of c-myc transcription with MEK inhibitors is already well detectable at two hours, while a negative modulation of cyclin D1 becomes evident only at four hours (David Duffy, Dublin, personal communication).

In an individual tumor, mutations belonging to a single pathway tended to be mutually exclusive [24, 37, 41], especially if they were relatively close within the pathway. In principle, there is no significant evolutionary Darwinian advantage in having deregulated twice in a very similar way the same pathway. In our model, this is also true in the oncologic reality [9, 10, 14, 16, 25], only inhibitors acting at, or downstream (not upstream), of a given mutation were active. This phenomenon was faithfully reproduced in our virtual simulations.

A double alteration along two distinct pathways, for a maximum effect, required the inhibition of both pathways, in order to restore an efficient control of a multi-altered network. We had previously observed a similar trend in a breast-cancer focused dynamic-simulation paper [15].

5 Robustness/Sensitivity of our Network to Perturbations

We started an analysis of sensitivity/robustness of our network, and we systematically introduced several individual fluctuations of total concentrations (or reaction rates), concerning distinct independent molecular species. Perhaps not unexpectedly, a perturbed species is affected by perturbing variations in protein concentration (or reaction rate), usually with a rapidly decreasing intensity according to distance among perturbing and perturbed nodes (distance measured in terms of intervening edges) [3]. The effects are extinguished especially rapidly along non biologically significant paths along the graph (Fig. 9). Rare exceptions could reflect very specific biologic pathway functions, in a biochemical interaction graph displaying different kinds of asymmetry and un-homogeneity, and essentially an uni-directionality of signal transmission.

Practically no perturbations effects of a ten times change could be observed when a perturbing node was separated more than three edges from a perturbed node. This also confirms previous observations by our group [6].

The arbitrary choice of a range of perturbations of ten times ($\times$ or $/$) was made with reference to the range of uncertainties that can be present in extrapolations obtained

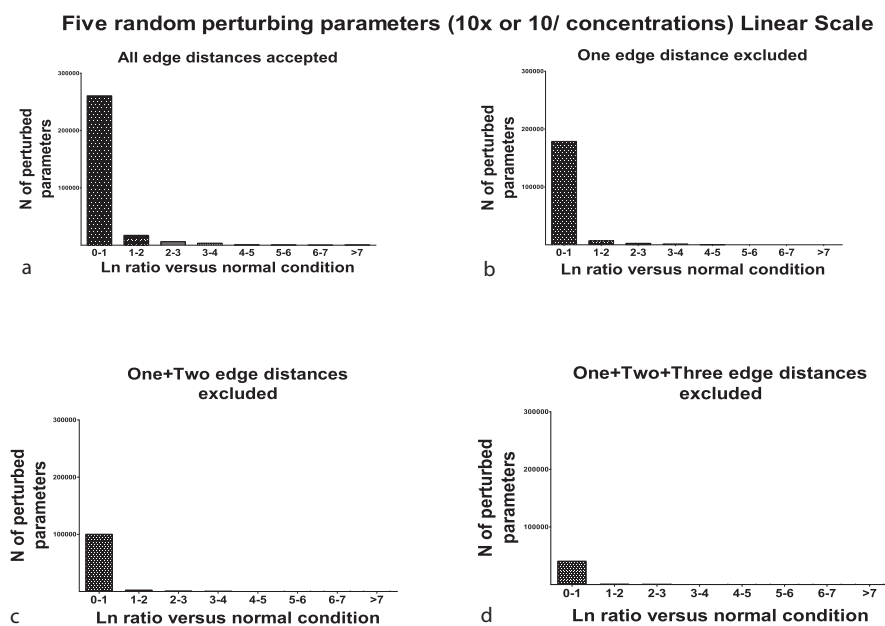


Fig. 9. Five random perturbing parameters (10x or 10/ concentrations). Linear Scale in a breast cancer network (data are from [8]). The total N of randomly perturbed parameters was $= 10,000 \times (34 - 5) = 290,000$. Notice that most of the perturbations are located in a range $e^0 - e^1$. For an increasing distance in terms of edges significantly perturbed parameters become progressively more infrequent (rare). Very rare, pathway and direction compatible, cases of significant perturbations at > 3 edges distance, don't show up in our linear scale

from experimental papers. A range of uncertainty of 100 times seems in general larger of what can be extrapolated from reported experimental results and connected methods. A 100 times range of variation is indeed closer to a mutational change (a drastic loss or gain of function).

We had incomplete information about parameters (concentrations and rates).

We obtained this information more or less indirectly from published experimental papers.

We patiently filled up the full set of parameters (concentrations and rates) with additional extrapolations compatible with the general picture and behavior of the network, to the best of our competence and knowledge in the field (from experimental literature data).

However, for a fraction of parameters we had no direct or indirect information. We decided a given interpolation because empirically they would seem to fit best with what was known about the general behavior of our signaling network region, especially in the neighborhood of the missing datum.

An intriguing question (Dr Alberto d'Onofrio personal communication) is the following: The subset of missing parameters could have admitted more than one significantly different solution, still fitting our imposed constraints? In the future we have planned to explore this important open question taking advantage of his collaboration (European Inst. of Oncology, Milan, Italy).

6 Final Comments and Conclusion

At a biochemical-interaction level, a dozen dynamic-simulation studies of limited signaling-network regions, have been published during the last ten years (six of them just in 2009–2010).

These works tend to confirm the feasibility, the interest and the innovative potential of this type of studies [4–7, 22, 27, 29, 32, 42, 43]. As we have also partially mentioned before, new technologies are entering the market at the clinical molecular-diagnostic level (*DNA re-sequencing and other approaches*).

The world of Medicinal Chemistry is steadily generating new selective inhibitors of onco-proteins affected by excess of function.

During the next 5–10 years, in the perspective of clinical applications, thanks to mathematical modeling of signaling networks, we will become more and more capable of entering the intricacies of a given signaling network region, and attributing a correct positioning (in progressively larger and more comprehensive networks) to both altered gene functions and the activity of selective drug inhibitors.

Mathematical modeling of the behavior of these networks *in physiological conditions, in the presence of altered / mutated onco-proteins, in the presence of one or two or three different selective inhibitors (a not too distant future)*, could be the most reasonable approach for trying to understand the non-intuitive intricacies of a not too limited signaling-network region, and taking the best possible advantage of new selective drugs.

Acknowledgements This paper was partially supported by: Liguria Region Project (N. 280 – 2010–2011): “Study and simulation of molecular interaction networks relevant in malignant transformation; search and study of inhibitors of the onco-proteins c-Myc and Bcl-XL”; CARIGE Foundation Project (N. 2010/228–16): “Analysis of molecular alterations in signal transduction networks downstream of EGFR-family receptors, in HER2 positive breast cancers and triple negative cancers. Rationalization at the clinical level of personalized antineoplastic therapies with onco-protein inhibitors”; Compagnia di San Paolo Project (1471 SD/CC N. 2009.1822): “Models and computational methods in the study of physiology and pathology of signaling networks of oncologic interest”; Istituto Superiore di Oncologia (ISO): Grant 2006 from Istituto Superiore di Sanità: “Development of new drugs capable of modifying the cancer micro-environment”. Grant RF-CAM-2006-353005 Regione Campania, from Italian Ministry of Health: “Molecular Diagnostic and Prognostic Markers of Thyroid Neoplasias”. Finally, we are deeply grateful to Kurt W. Kohn for the continuous exchange of ideas in the field of Molecular Interaction Maps, that he pioneered already a dozen years ago.

References

1. Aladjem, M.I., Pasa, S., Parodi, S., Weinstein, J.N., Pommier, Y., Kohn, K.W.: Molecular interaction maps a diagrammatic graphical language for bioregulatory networks. *Sci. STKE*. **222**, 1–5 (2004)
2. Aldridge, B.B., Burke, J.M., Lauffenburger, D.A., Sorger, P.K.: Physicochemical modelling of cell signalling pathways. *Nat. Cell Biol.* **8**, 1195–1203 (2006)
3. Barabasi, A.L., Oltvai, Z.N.: Network biology: understanding the cell’s functional organization *Nat Rev Genet.* **5**, 101–113 (2004)
4. Birtwistle, M.R., Hatakeyama, M., Yumoto, N., Ogunnaike, B.A., Hoek, J.B., Kholodenko, B.N.: Ligand-dependent responses of the ErbB signaling network: experimental and modeling analyses. *Mol. Syst. Biol.* **3**, 144 (2007)
5. Borisov, N., Aksamitiene, E., Kiyatkin, A., Legewie, S., Berkhout, J., Maiwald, T., Kaimachnikov, N.P., Timmer, J., Hoek, J.B., Kholodenko, B.N.: Systems-level interactions between insulin-EGF networks amplify mitogenic signaling. *Mol. Syst. Biol.* **5** e256 (2009)
6. Castagnino, N., Tortolina, L., Balbi, A., Pesenti, R., Montagna, R., Ballestrero, A., Soncini, D., Moran, E., Nencioni, A., Parodi, S.: Dynamic Simulations of Pathways Downstream of ERBB-Family, Including Mutations and Treatments. Concordance with Experimental Results. *Curr. Cancer Drug Targets* **10**, 737–757 (2010)
7. Chen, W.W., Schoeberl, B., Jasper, P.J., Niepel, M., Nielsen, U.B., Lauffenburger, D.A., Sorger, P.K.: Input-output behavior of ErbB signaling pathways as revealed by a mass action model trained against dynamic data. *Mol. Syst. Biol.* **5**, e239 (2009)
8. Daniels, D.L., Weis, W.I.: β -catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nature Structural & Molecular Biology* **12**, 364–371 (2005)
9. Eichhorn, P.J., Gili, M., Scaltriti, M., Serra, V., Guzman, M., Nijkamp, W., Beijersbergen, R.L., Valero, V., Seoane, J., Bernards, R., Baselga, J.: Phosphatidylinositol 3-kinase hyperactivation results in lapatinib resistance that is reversed by the mTOR/phosphatidylinositol 3-kinase inhibitor NVP-BE235. *Cancer Research* **68**, 9221–30 (2008)
10. Engelman, J.A.: Targeting PI3K signalling in cancer: opportunities, challenges and limitations. *Nat. Rev. Cancer* **9**, 550–562 (2009)
11. Hazzalin, C.A., Mahadevan, L.C.: MAPK-regulated transcription: a continuously variable gene switch? *Nat Rev Mol Cell Biol.* **3**, 30–40 (2002)
12. He, T.C., Sparks, A.B., Rago, C., Hermeking, H., Zawel, L., da Costa, L.T., Morin, P.J., Vogelstein, B., Kinzler, K.W.: Identification of c-MYC as a Target of the APC Pathway. *Science* **281**, 1509–1512 (1998)

13. Huang, S.M., Mishina, Y.M., Liu, S., Cheung, A., Stegmeier, F., Michaud, G.A., Charlat, O., Wielle, E., Zhang, Y., Wiessner, S., Hild, M., Shi, X., Wilson, C.J., Mickanin, C., Myer, V., Fazal, A., Tomlinson, R., Serluca, F., Shao, W., Cheng, H., Shultz, M., Rau, C., Schirle, M., Schlegl, J., Ghidelli, S., Fawell, S., Lu, C., Curtis, D., Kirschner, M.W., Lengauer, C., Finan, P.M., Tallarico, J.A., Bouwmeester, T., Porter, J.A., Bauer, A., Cong, F.: Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling. *Nature* **461**, 614–20 (2009)
14. Hynes, N.E., Dey, J.H.: PI3K inhibition overcomes trastuzumab resistance: blockade of ErbB2/ErbB3 is not always enough. *Cancer Cell* **15**, 353–5 (2009)
15. Iadevaia, S., Lu, Y., Morales, F.C., Mills, G.B., Ram, P.T.: Identification of optimal drug combinations targeting cellular networks: integrating phospho-proteomics and computational network analysis. *Cancer Research* **70**, 6704–14 (2010)
16. Junttila, T.T., Akita, R.W., Parsons, K., Fields, C., Lewis Phillips, G.D., Friedman, L.S., Sam-path, D., Sliwkowski, M.X.: Ligand-independent HER2/HER3/PI3K complex is disrupted by trastuzumab and is effectively inhibited by the PI3K inhibitor GDC-0941. *Cancer Cell* **15**, 429–440 (2009)
17. Kiyatkin, A., Aksamitiene, E., Markevich, N.I., Borisov, N.M., Hoek, J.B., Kholodenko, B.N.: Scaffolding protein GAB1 sustains Epidermal Growth Factor-induced mitogenic and survival signaling by multiple positive feedback loops. *J Biol Chem.* **281**, 19925–19938 (2006)
18. Klein, E.A., Assoian, R.K.: Transcriptional regulation of the cyclin D1 gene at a glance. *J. Cell. Sci.* **121**, 3853–3857 (2008)
19. Kohn, K.W.: Molecular interaction map of the mammalian cell cycle control and DNA repair systems. *Mol. Biol. Cell* **10**, 2703–2734 (1999)
20. Kohn, K.W., Aladjem, M.I., Kim, S., Weinstein, J.N., Pommier, Y.: Depicting combinatorial complexity with the molecular interaction map notation. *Mol Syst Biol.* **2**, 51 (2006)
21. Kohn, K.W., Aladjem, M.I., Weinstein, J.N., Pommier, Y.: Molecular Interaction Maps of Bioregulatory Networks: A General Rubric for Systems Biology. *Mol. Biol. Cell* **17**, 1–13 (2006)
22. Lee, E., Salic, A., Krueger, R., Heinrich, R., Kirschner, M.W.: The roles of APC and Axin derived from experimental and theoretical analysis of the Wnt pathway. *PLoS Biol.* **1**, E10 (2003)
23. Lim, S.K., Hoffmann, F.M.: Smad4 cooperates with lymphoid enhancer-binding factor 1/T cell-specific factor to increase c-myc expression in the absence of TGF-beta signaling. *Proc. Natl. Acad. Sci. U.S.A* **103**, 18580–18585 (2006)
24. Lin, J., Gan, C.M., Zhang, X., Jones, S., Sjoblom, T., Wood, L.D., Parsons, D.W., Papadopoulos, N., Kinzler, K.W., Vogelstein, B., Parmigiani, G., Velculescu, V.E.: A multidimensional analysis of genes mutated in breast and colorectal cancers. *Genome Res.* **17**, 1304–1318 (2007)
25. Liu, P., Cheng, H., Roberts, T.M., Zhao, J.J.: Targeting the phosphoinositide 3-kinase pathway in cancer. *Nature reviews Drug discovery* **8**, 627–644 (2009)
26. Massagué, J., Seoane, J., Wotton, D.: Smad transcription factors Smad transcription factors. *Genes&Development* **19**, 2783–2810 (2005)
27. Nakakuki, T., Birtwistle, M.R., Saeki, Y., Yumoto, N., Ide, K., Nagashima, T., Brusch, L., Ogunnaike, B., Okada-Hatakeyama, M., Kholodenko, B.N.: Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. *Cell* **141**, 884–896 (2010)
28. Nurse, P.: A long twentieth century of the cell cycle and beyond. *Cell* **100**, 71–78 (2000)
29. Orton, R.J., Adriaens, M.E., Gormand, A., Sturm, O.E., Kolch, W., Gilbert, D.R.: Computational modelling of cancerous mutations in the EGFR/ERK signalling pathway. *BMC Syst Biol.* **3** art. n. 100 (2009)
30. Piedra, J., Martinez, D., Castano, J., Miravet, S., Dunach, M., de Herreros, A.G.: Regulation of beta-catenin structure and activity by tyrosine phosphorylation. *The Journal of biological chemistry* **276**, 20436–50443 (2001)
31. Price, M.A.: CKI, there's more than one: casein kinase I family members in Wnt and Hedgehog signaling. *Genes&Development* **20**, 399–410 (2006)

32. Schmierer, B., Tournier, A.L., Bates, P., Hill, C.S.: Mathematical modeling identifies Smad nucleocytoplasmic shuttling as a dynamic signal-interpreting system. *Proc. Natl. Acad. Sci. USA* **105**, 6608–6613 (2008)
33. Shaulian, E., Karin, M.: AP-1 in cell proliferation and survival. *Oncogene* **20**, 2390–2400 (2001)
34. Shutman, M., Zhurinsky, J., Simcha, I., Albanese, C., D’Amico, M., Pestell, R., Ben-Ze’ev, A.: The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. *Proc. Natl. Acad. Sci. U.S.A* **96**, 5522–5527 (1999)
35. Sjoblom, T., Jones, S., Wood, L.D., Parsons, D.W., Lin, J., Barber, T.D., Mandelker, D., Leary, R.J., Ptak, J., Silliman, N., Szabo, S., Buckhaults, P., Farrell, C., Meeh, P., Markowitz, S.D., Willis, J., Dawson, D., Willson, J.K.V., Gazdar, A.F., Hartigan, J., Wu, L., Liu, C., Parmigiani, G., Park, B.H., Bachman, K.E., Papadopoulos, N., Vogelstein, B., Kinzler, K.W., Velculescu, V.E.: The consensus coding sequences of human breast and colorectal cancers. *Science* **314**, 268–274 (2006)
36. Tetsu, O., McCormick, F.: Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* **398**, 422–426 (1999)
37. Thomas, R.K., Baker, A.C., Debiase, R.M., Winckler, W., Laframboise, T., Lin, W.M., Wang, M., Feng, W., Zander, T., MacConaill, L., Macconnaill, L.E., Lee, J.C., Nicoletti, R., Hatton, C., Goyette, M., Girard, L., Majumdar, K., Ziaugra, L., Wong, K.K., Gabriel, S., Beroukhim, R., Peyton, M., Barretina, J., Dutt, A., Emery, C., Greulich, H., Shah, K., Sasaki, H., Gazdar, A., Minna, J., Armstrong, S.A., Mellinghoff, I.K., Hodi, F.S., Dranoff, G., Mischel, P.S., Cloughesy, T.F., Nelson, S.F., Liao, L.M., Mertz, K., Rubin, M.A., Moch, H., Loda, M., Catalona, W., Fletcher, J., Signoretti, S., Kaye, F., Anderson, K.C., Demetri, G.D., Dummer, R., Wagner, S., Herlyn, M., Sellers, W.R., Meyerson, M., Garraway, L.A.: High-throughput oncogene mutation profiling in human cancer. *Nat Genet.* **39**, 347–351 (2007)
38. Tortolina, L., Castagnino, N., Montagna, R., Pesenti, R., Balbi, A., Parodi, S.: Simulations of biochemical interactions. Applications to molecular oncology in Functional Proteomics and Nanotechnology-based Microarrays. Eds. Nicolini C., LaBaer J. *Pan Stanford Series on Nanobiotechnology*, London – New York – Singapore, **II** chapter 11 (2010)
39. Tortolina L., Castagnino N., De Ambrosi C., Moran E., Patrone F., Ballestrero A. and Parodi S.: A multi-scale approach to colorectal cancer: from a biochemical-interaction signaling network level, to multi-cellular dynamics of malignant transformation. Interplay with mutations and onco-protein inhibitor drugs. *Curr Cancer Drug Targets* **12**, 339–355 (2012)
40. Wood, L.D., Parsons, D.W., Jones, S., Lin, J., Sjoblom, T., Leary, R.J., Shen, D., Boca, S.M., Barber, T., Ptak, J., Silliman, N., Szabo, S., Dezso, Z., Ustyanksky, V., Nikolskaya, T., Nikolsky, Y., Karchin, R., Wilson, P., Kaminker, J.S., Zhang, Z., Croshaw, R., Willis, J., Dawson, D., Shipitsin, M., Willson, J.K.V., Sukumar, S., Polyak, K., Park, B.H., Pethiyagoda, C.L., Pant, P.V.K., Ballinger, D.G., Sparks, A.B., Hartigan, J., Smith, D.R., Suh, E., Papadopoulos, N., Buckhaults, P., Markowitz, S.D., Parmigiani, G., Kinzler, K.W., Velculescu, V.E., Vogelstein, B.: The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108–1113 (2007)
41. Yeang, C.-H., McCormick, F., Levine, A.: Combinatorial patterns of somatic gene mutations in cancer. *FASEB J.* **22**, 2605–2622 (2008)
42. Zi, Z., Klipp, E.: Constraint-based modeling and kinetic analysis of the Smad dependent TGF-beta signaling pathway. *PLoS ONE* **2**, e936 (2007)
43. Zielinski, R., Przytycki, P.F., Zheng, J., Zhang, D., Przytycka, T.M., Capala, J.: The crosstalk between EGF, IGF, and Insulin cell signaling pathways—computational and experimental analysis. *BMC Syst Biol.* **4**, 88 (2009)

Some Results on the Population Behavior of Cancer Stem Cells

Edoardo Beretta, Vincenzo Capasso, Annick Harel-Bellan, and Nadya Morozova

Abstract Recent discovery of cancer stem cells in tumorigenic tissues has raised many questions about their nature, origin, function and their behavior in cell culture. Most of current experiments reporting a dynamics of cancer stem cell populations in culture show the eventual stability of the percentages of these cell populations in the whole population of cancer cells, independently of the starting conditions. In this report we present a review about a class of mathematical models of cancer stem cell population behavior, based on specific features of cancer stem cell divisions and including, as a mathematical formalization of cell-cell communications, an underlying field concept. We compare the qualitative behavior of mathematical models of stem cells evolution, without and with an underlying signal. In presence of an underlying field the model is described by a system of delay differential equations, by admitting a possibly delayed signal originated by existing cells. Under a variety of assumptions on the parameters, in all cases we show that the stability of percentages can be recovered, provided that the delay is sufficiently small. Further, for the DDE case we show the possible existence of, either damped or standing, oscillations in the cell populations. The outcomes of the analysis may offer to experimentalists a tool for addressing the issue regarding the possible non-stem to stem cells transition. Further, they may stimulate further experiments for elucidating the nature of “instructive” signals for cell divisions.

E. Beretta · V. Capasso (✉)

CIMAB, c/o Department of Mathematics, University of Milan, 20133 Milan, Italy

e-mail: edoardo.beretta@uniurb.it, vincenzo.capasso@unimi.it

A. Harel-Bellan · N. Morozova

CNRS FRE 3377, Laboratoire Epigénétique et cancer, CEA Saclay, 91191 Gif-sur-Yvette, France

e-mail: ahbellan@vjf.cnrs.fr, morozova@vjf.cnrs.fr

A. d’Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_8, © Springer-Verlag Italia 2012

1 Introduction

The current concept about stem cells of adult organisms is that, unlike differentiated cells, they may undergo cell division producing stem cells (for renewing) and differentiated cells with the features of the corresponding tissue. Generally speaking, the differentiated cell from first division is named progenitor cell, and also has some limited self-renewal potential; progenitor cells may go through several rounds of cell division before being terminally differentiated. *Cancer stem cells* (CSCs) constitute a subpopulation of cells within tumors that could actively drive tumor growth and recurrence.

Recent discovery of *cancer stem cells* in tumorigenic tissues has raised plenty of debates about their nature, origin, possible connection with “normal stem cells” and even about the possibility to use this term (cancer stem cells), if its biological nature is not completely elucidated yet, and there is a possibility that they may have nothing in common with “normal stem cells”. At the moment, cancer stem cells can be determined only operationally, by their ability to seed new tumors, and for this reason, they have also been termed “tumor-initiating cells” [1]. However, for some cancer cell lines promising results were obtained, for elucidating the proper cancer stem cell markers, which can allow biochemical determination of this population and its further analysis. This allowed recent identification of CSCs in hematopoietic, breast, colon, ovary, brain, pancreas, and prostate cancers [3, 7, 8, 11, 13, 15, 18, 24].

A very important observation made in works on several cancer cell lines is that *the percentage of cancer stem cell population is maintained at the same level during several years of passages*. Another important observation is that when a cancer stem cell population is isolated, purified and maintained in culture, the percentage of stem cells rapidly decreased and finally stabilized at the initial level, characteristic for this given cell culture (personal communications with C. Ginestier, A. Harel-Bellan, V. Maguer-Satta, and M.S. Wicha).

Fig. 1. Schematic representation of the cancer stem cell subpopulation (M+) dynamics in the whole cancer stem cell population (black), and in purified M+ subpopulation (gray squared line), according to numerous experimental data in cell culture (personal communications with C. Ginestier, A. Harel-Bellan, V. Maguer-Satta). Cancer stem cells were marked using specific cancer stem cell markers (M)

These experimental data raise two questions:

- Do there exist “instructive signals” for maintaining this population behavior?
- Which molecular mechanisms are involved in the realization of this process?

Various mathematical models have been recently proposed for describing the dynamics of stem cell populations (see e.g. [5, 19], and references therein), and some of them (see [4, 16, 17, 25, 26]), also consider cell-cell communication. However, these works do not address the question about cancer stem cell population stability, and suggest different approaches for the description of cell population dynam-

ics. On the other hand a recent paper by [10] proposes a simple stochastic Markov chain model to address the existence of a stochastic equilibrium, based on a similar assumption as ours, that cancer stem-like cells can arise from non-stem-like cells.

In our work [2] we have additionally addressed the first question, and tried to find some solutions using mathematical modelling of cancer stem cells evolution, based on the main biological features of cancer stem cells population behavior. In this paper we have included an additional naive mathematical model, which may explain the existence of a stable equilibrium just with a little number of technical ingredients. Hopefully the results of mathematical modelling can enable experimentalists to design the proper search of the molecular mechanisms adequately solving the second question.

1.1 The main assumptions

Our main goal is to create a mathematical model describing the phenomena of cancer stem cell population stabilization. Our first statement is that the evolution of the structure of a whole cell population (in our case, the population structure means just a percentage of stem cells in it) depends on the pattern of cell divisions, on the speed of cell divisions in both stem and non-stem cell populations, and on the rates of cell death. Our main assumption is the existence of an underlying field, carrying the information about the population structure and influencing the pattern and speed of cell divisions in the given conditions. Thus, there are two biological points which we considered as the main ones for building the mathematical model of cells population behavior:

- the type of cell divisions;
- the speed of cell divisions.

It is very important to mention that, with all differences in normal (non cancer) and cancer stem cells nature, and independently from the experimental methods of their investigation (mostly in vitro in cell culture for cancer stem cell, and mostly in vivo for normal stem cells), the two main statements, which we consider as the basic ones for our modelling, are the same in both cases:

- as a rule, stem cells undergo asymmetric divisions with relatively slow rates, while non-stem cells undergo symmetric divisions with relatively quick rates;
- the main feature of stem cells population behavior is the tendency of maintaining the proper (very small) percentage of stem cells in the whole tissue (culture).

Thus it is clear that, though our main objective is the modeling of cancer cell population behavior, the main general idea of the model – the mathematical description of the main features of the “instructive signals” from the whole system (tissue or culture) to its cells for maintaining the proper population pattern – can be applied also to normal (non cancer) stem cells population behavior. In this cases “non-stem cells” in our model should be understood as differentiated cells of normal tissues. It is important to note, that in cancer tissues/cultures non-stem cells can not be named “differentiated”, as carcinogenesis strongly influences the proper cell differentiation.

Considering the types of cell divisions, we can point out that, among several possible scenarios, the asymmetric cell divisions, providing self-renewing, is considered to be the main one. However, other possible scenarios of cancer stem cells division are not completely excluded (see Fig. 2) – the symmetric division giving rise to two identical daughter cells with stem cell properties; the symmetric division with two identical daughter cells with non-stem cell properties, and also the direct development of initial stem cell into non-stem one. All these scenarios are currently under hot discussion, as for normal, so for cancer stem cells. Non-stem (daughter) cells, as a rule undergo the symmetric division producing two identical daughter cells with non-stem cell properties. The question about other possible types of non-stem cell division, namely, the possibility for non-stem cancer cells to undergo under some specific conditions the process of non-stem to stem cell transition (analog of dedifferentiation for normal cell case), is one of the main open problems in current molecular biology (see Fig. 3). Considering the speed of cell divisions, it has been well documented that the speed of stem cells division are much slower than the speed of cell divisions of non-stem cells (see Fig. 4). The current theory for normal stem cells evolution is that stem cells remain undifferentiated due to environmental cues in their particular niche, and differentiate when they leave that niche or no longer receive corresponding signals [22]. However, this theory still leaves open the main question: what is the nature of the signals for maintaining the proper population behavior of stem cells within the given niche? Thus, our mathematical model describing the main features of possible “instructive signals” maintaining the proper pattern of a system, can be valuable for addressing this important question not only for cancer, but also for normal stem cells population behavior.

2 The Concept Model

We may then consider a mathematical model for the evolution of stem cells (S), and non-stem cells (D) populations. Based on the discussion above, we will consider that stem cells may divide either into two stem cells or into one stem cell and one non-stem cell, or into two non-stem cells, or into one non-stem cell (with a total rate $\lambda_1 > 0$), according to the following scheme (see Fig. 2)

$$S \Longrightarrow S + D \quad \text{with probability } p_1, \quad (1)$$

$$S \Longrightarrow D + D \quad \text{with probability } p_2, \quad (2)$$

Fig. 2. Possible modes of stem cell division

Fig. 3. Possible modes of non-stem cell division

Fig. 4. The speed of cell divisions



such that

$$p_i \geq 0, \quad i = 1, \dots, 4; \quad p_1 + p_2 + p_3 + p_4 = 1.$$

We assume here that non-stem cells may divide either into two non-stem cells or into one stem cell and one non-stem cell (with a rate λ_2)



such that

$$q_i \geq 0, \quad i = 1, 2; \quad q_1 + q_2 = 1.$$

We have ignored the possible case 3 for non-stem cells as from Fig. 3. Its inclusion in the model would only lead to additional mathematical technicalities.

Stem cells may die at a rate γ_1 , while non-stem cells may die at a rate γ_2 .

We assume a time scale much larger than the typical cell cycle time and that cell populations are large enough to ignore randomness [12]. We may then assume a one point (in space) model with a continuous time evolution, just to specify the main ingredients of the model of evolution of stem cells (S) and non-stem cells (D).

As a consequence, to start with, the mathematical model that we propose is given by the following system of ordinary differential equations:

$$\frac{d}{dt}S(t) = (-\gamma_1 + (-1 + 2p_3 + p_1)\lambda_1)S(t) + q_2\lambda_2D(t), \quad (7)$$

$$\frac{d}{dt}D(t) = (p_1 + 2p_2 + p_4)\lambda_1S(t) + (-\gamma_2 + q_1\lambda_2)D(t), \quad (8)$$

subject to initial conditions $S(0) = S_0$, and $D(0) = D_0$, such that $(S_0, D_0) \in \mathbf{R}_+^2$, $(S_0, D_0) \neq (0, 0)$.

System (7), (8) is an homogeneous system of two linear ODE equations whose matrix of (constant) coefficients is

$$\mathbf{A} = \begin{pmatrix} \alpha_S & \beta_D \\ \beta_S & \alpha_D \end{pmatrix}, \quad (9)$$

where

$$\alpha_S = -\gamma_1 + (-1 + 2p_3 + p_1)\lambda_1, \quad \beta_D = q_2\lambda_2$$

$$\beta_S = (p_1 + 2p_2 + p_4)\lambda_1, \quad \alpha_D = -\gamma_2 + q_1\lambda_2,$$

with $\beta_S \geq 0$, and $\beta_D \geq 0$.

For the time being we assume that $\beta_D > 0$, and $\beta_S > 0$; under this circumstance, matrix A has positive off-diagonal elements. Later the case $q_2 = 0$ will be analyzed in more detail.

Then, all the solutions of (7), (8), subject to the above initial conditions, will satisfy positivity i.e. $(S(t), D(t)) \in \mathbf{R}_+^2$, with $S(t) > 0$, and $D(t) > 0$, for $t > 0$.

Let $r_i, i = 1, 2$ denote the eigenvalues of matrix A ; they are the roots of the characteristic equation

$$r^2 - (Tr\mathbf{A})r + Det\mathbf{A} = 0,$$

where $Tr\mathbf{A} = \alpha_S + \alpha_D$, $Det\mathbf{A} = \alpha_S \alpha_D - \beta_S \beta_D \neq 0$, and the discriminant is $\Delta = (Tr\mathbf{A})^2 - 4Det\mathbf{A} > 0$. Therefore the two eigenvalues of A are real; they are given by

$$r_1 = \frac{1}{2} \left\{ Tr\mathbf{A} - \sqrt{\Delta} \right\}, \quad r_2 = r_1 + \sqrt{\Delta}. \quad (10)$$

Since experimental data are expressed in terms of percentages of stem cells, as opposed to percentages of non-stem cells, with respect to the total number of cells

$$N(t) = S(t) + D(t), \quad (11)$$

we are interested in the evolution of the fractions

$$s(t) = \frac{S(t)}{N(t)}, \quad d(t) = \frac{D(t)}{N(t)} \quad (12)$$

such that

$$s(t) + d(t) = 1,$$

at any time $t \geq 0$.

2.1 Time Evolution of $s(t) = \frac{S(t)}{N(t)}$, and Its Asymptotic Behavior

In terms of fractions $s(t)$ and $d(t)$, system (7), (8) can be rewritten in the following form

$$\frac{d}{dt}s(t) = P + Qs(t) + R s^2(t); \quad (13)$$

$$d(t) = 1 - s(t), \quad t \geq 0; \quad (14)$$

subject to an initial condition $s(0) \in [0, 1]$; the above system is complemented by

$$\frac{d}{dt}N(t) = [-R s(t) + (\lambda_2 - \gamma_2)]N(t), \quad (15)$$

subject to an initial condition

$$N(0) = N_0 > 0, \quad t \geq 0.$$

Here

$$P = q_2 \lambda_2, \quad (16)$$

$$Q = (p_3 - (p_2 + p_4))\lambda_1 - (1 + q_2)\lambda_2 + \gamma_2 - \gamma_1, \quad (17)$$

$$R = \lambda_2 - (1 - p_4)\lambda_1 - (\gamma_2 - \gamma_1). \quad (18)$$

The differential equation (13) is a Riccati equation of the first type with constant coefficients, which can be explicitly solved (but we omit here its explicit solution). The main properties of system (13), (14) depend upon the function $f : [0, 1] \rightarrow \mathbb{R}$, given by

$$f(s) := P + Qs + Rs^2, \quad s \in [0, 1]. \quad (19)$$

We may notice that, if we assume $q_2 > 0$, then

$$\dot{s}(t)|_{s=0} = f(0) = q_2 \lambda_2 > 0, \quad (20)$$

$$\dot{s}(t)|_{s=1} = f(1) = (-1 - p_2 + p_3)\lambda_1 < 0, \quad (21)$$

which imply that for any initial condition $s(0) \in [0, 1]$, the corresponding solution of (13) is such that $s(t) \in (0, 1)$ for all $t > 0$.

The equilibria of Eq. (13) are the roots of the algebraic equation

$$f(s) = 0, \quad s \in [0, 1]. \quad (22)$$

Because of (20), and (21), f must admit at least one zero within $(0, 1)$, and because of the structure of the function f , there exists exactly a unique root $s^* \in (0, 1)$ of (22) such that

$$\dot{s}(t) = f(s) > 0, \quad \text{for } s \in [0, s^*), \quad \text{and} \quad \dot{s}(t) = f(s) < 0, \quad \text{for } s \in (s^*, 1].$$

Of course no solution may cross the equilibrium s^* . Finally we may state the following result.

Theorem 1. *A unique $s^* \in (0, 1)$ exists such that $(s^*, d^* = 1 - s^*)$ is a globally asymptotically stable equilibrium for system (13), (14); specifically*

$$\left(s^* = \frac{\beta_D}{r_2 + \beta_D - \alpha_S}, d^* = \frac{r_2 - \alpha_S}{r_2 + \beta_D - \alpha_S} \right). \quad (23)$$

Here r_2 is one of the eigenvalues of the matrix A as defined in (10). The global stability of the nontrivial equilibrium (s^*, d^*) can be derived via a simple Lyapunov function.

2.2 A Discussion about the Parameters

From all previous discussion it follows that the positivity of $s^* \in (0, 1)$ is related to the positivity of $P = q_2 \lambda_2 = \beta_D$, i.e. to the positivity of the probability q_2 that normal cells D may divide into a stem cell S and a non-stem cell D. If the probability of this elementary process is equal to zero, we see that also s^* in Theorem 1 goes to zero. Therefore it becomes biologically interesting to see what happens when $q_2 = 0$, i.e. when it is not allowed non-stem cells to produce stem cells. However, before analyzing the case $q_2 = 0$ it can be useful to discuss more about the parameters.

Since the branching processes of non-stem cells (D) occur at a global rate faster than the ones of stem cells (S), it is reasonable to assume that $\lambda_2 > \lambda_1$. Moreover, it seems reasonable to assume that the death rates γ_1 , and γ_2 are about the same for both stem and non-stem cells, i.e. $\gamma_1 \simeq \gamma_2$. Finally we additionally assume that the total branching rate λ_2 of non-stem cells exceeds the related death rate γ_2 , i.e. $\lambda_2 > \gamma_2 \simeq \gamma_1$. By summarizing we assume

$$\lambda_2 > \lambda_1; \quad \gamma_1 \simeq \gamma_2; \quad \gamma_i < \lambda_2, \quad i = 1, 2. \quad (24)$$

The above assumptions (24) enable to define the sign of the quantities P, Q, R defined by (16)–(18). Namely we have

$$P = q_2 \lambda_2 > 0, \quad (25)$$

$$Q \simeq (p_3 - (p_2 + p_4))\lambda_1 - (1 + q_2)\lambda_2 < 0, \quad (26)$$

$$R \simeq \lambda_2 - (1 - p_4)\lambda_1 > 0; \quad (27)$$

2.3 About the Total Population

From (15) we know that

$$N(t) = N_0 \exp \left\{ \int_0^t (-R s(\tau) + \lambda_2 - \gamma_2) d\tau \right\}. \quad (28)$$

In the case $\lambda_2 \gg \lambda_1$, and $\gamma_2 \ll \lambda_2$, we may use the approximation

$$N(t) \gtrsim N_0 \exp \left\{ \int_0^t (-\lambda_2 s(\tau) + \lambda_2) d\tau \right\} = N_0 \exp \left\{ \int_0^t \lambda_2 d(\tau) d\tau \right\}. \quad (29)$$

Because of the monotone convergence of the solutions to the equilibrium, it is not difficult to check that, in all cases,

$$N(t) \uparrow +\infty, \quad \text{as } t \rightarrow +\infty. \quad (30)$$

The eventual explosion of the total population is not a surprise since we have not imposed any saturation (such as a logistic growth), since we may assume that the culture is continuously fed by nutrients.

3 Underlying Field

Again based on the introductory remarks, one may conjecture that, given constant proliferation rates (λ_1 and λ_2), all proportions p_i and q_j of the elementary processes depend upon some underlying field u , (a biochemical signal) which is produced by the same population of cells

$$p_i = p_i(u), \quad i = 1, 2, 3, 4,$$

such that

$$\sum_{i=1}^4 p_i(u) = 1, \quad \text{for any } u \geq 0;$$

$$q_j = q_j(u), \quad j = 1, 2,$$

$$\sum_{j=1}^2 q_j(u) = 1, \quad \text{for any } u \geq 0.$$

For an updated mathematical model, we introduce the functional response $g(S, D)$ for the production of the underlying field from existing stem cells and non-stem cells, and a decay rate ε of the underlying field itself.

As far as the underlying field u is concerned, we additionally consider the possibility that the response to existing cells may be delayed by a constant delay $r \geq 0$, so that the mathematical model finally becomes

$$\frac{d}{dt}u(t) = -\varepsilon u(t) + g(S(t-r), D(t-r)), \quad (31)$$

$$\begin{aligned} \frac{d}{dt}S(t) = & -\gamma_1 S(t) + (-1 + 2p_3(u(t)) + p_1(u(t)))\lambda_1 S(t) \\ & + q_2(u(t))\lambda_2 D(t), \end{aligned} \quad (32)$$

$$\begin{aligned} \frac{d}{dt}D(t) = & -\gamma_2 D(t) + (-1 + 2q_1(u(t)) + q_2(u(t)))\lambda_2 D(t) \\ & + (p_1(u(t)) + 2p_2(u(t)) + p_4(u(t)))\lambda_1 S(t). \end{aligned} \quad (33)$$

We may anticipate here that this extended model may now lead to a much larger variety of dynamical behaviors (see later in the section on simulations), so to stimulate further experiments under various scenarios.

For a simpler handling of the above system, we consider relative amounts, as for the case independent of u . Define

$$s(t) := \frac{S(t)}{S(t) + D(t)},$$

so that s satisfies the equation

$$s'(t) = P(u(t)) + Q(u(t))s(t) + R(u(t))s(t)^2, \quad (34)$$

with

$$P(u) = q_2(u)\lambda_2, \quad (35)$$

$$Q(u) = (p_3(u) - (p_2(u) + p_4(u)))\lambda_1 - \lambda_2 - q_2(u)\lambda_2 + \gamma_2 - \gamma_1, \quad (36)$$

$$R(u) = \lambda_2 - (1 - p_4(u))\lambda_1 - (\gamma_2 - \gamma_1). \quad (37)$$

Note that Eq. (34) depends on p_4, q_2 and only on the combination $(p_3 - (p_2 + p_4))$. For a technical simplification let us assume that

$$g(S(t), D(t)) = g\left(\frac{S(t)}{S(t) + D(t)}\right),$$

so that we may take

$$g(S(t), D(t)) = g(s(t)) \quad (38)$$

with some function $g(s)$.

Based on previous assumptions, if we let $r \in \mathbb{R}_+$ be the constant delay, we take the mathematical model as follows.

$$\frac{d}{dt}u(t) = -\varepsilon u(t) + g(s(t-r)), \quad (39)$$

$$\frac{d}{dt}s(t) = f(u(t), s(t)), \quad (40)$$

$$d(t) = 1 - s(t), \quad (41)$$

for $t \geq 0$. complemented by the equation for the total population

$$N'(t) = [-R(u(t))s(t) + \lambda_2 - \gamma_2]N(t), \quad (42)$$

and subject to suitable initial conditions

$$(u(\cdot), s(\cdot)) \in C_b([-r, 0], \mathbb{R}_+ \times [0, 1]).$$

Here

$$f(u, s) := P(u) + Q(u)s + R(u)s^2. \quad (43)$$

About the structure of the system we assume that

- (i) $\varepsilon > 0$, a constant;
- (ii) $\lambda_i, \gamma_i \in \mathbb{R}_+, i = 1, 2$ satisfy the assumptions in (24);
- (iii) $g : [0, 1] \rightarrow \mathbb{R}_+, g \in C^1((0, 1));$ with $g'(s) > 0$ for any $s \in (0, 1)$;
- (iv) $p_i : \mathbb{R}_+ \rightarrow (0, 1), p_i \in C^1(\mathbb{R}_+), i = 1, 2, 3, 4$, such that $\sum_{i=1}^4 p_i(u) = 1$, for any $u \in \mathbb{R}_+$;

- (v) $q_j : \mathbb{R}_+ \rightarrow (0, 1)$, $q_j \in C^1(\mathbb{R}_+)$, $j = 1, 2$, such that $\sum_{j=1}^2 q_j(u) = 1$, for any $u \in \mathbb{R}_+$;

Under the assumptions (ii), (iv), (v) we have that

$$f(u, 0) = P(u) = q_2(u)\lambda_2 > 0, \quad (44)$$

$$f(u, 1) = P(u) + Q(u) + R(u) = (-1 - p_2(u) + p_3(u))\lambda_1 < 0, \quad (45)$$

for all $u > 0$.

Furthermore, the assumptions (ii), (iv), (v) also imply that

$$Q(u) = (p_3(u) - (p_2(u) + p_4(u)))\lambda_1 - \lambda_2 - q_2(u)\lambda_2 + \gamma_2 - \gamma_1 \leq \quad (46)$$

$$\leq p_3(u)\lambda_1 - \lambda_2 < 0; \quad (47)$$

$$R(u) = \lambda_2 - (1 - p_4(u))\lambda_1 - (\gamma_2 - \gamma_1) \geq \lambda_2 - \lambda_1 > 0, \quad (48)$$

for all $u > 0$, that is the signs of $P(u)$, $Q(u)$, and $R(u)$ remain constant for all $u > 0$.

On the other hand, we wish that the underlying field u has to provide a delayed feedback on the observed percentage level of stem cells; i.e. an increase of $s(t - r)$ has to imply an increase of $u(t)$ thanks to Eq. (39). Such an increase of $u(t)$ must imply a decrease of $s(t)$ via Eq. (40). This can be obtained by imposing the following additional assumption.

- (vi) $f \in C^{1,1}(\mathbb{R}_+ - \{0\} \times (0, 1))$, must satisfy

$$\frac{\partial f(u, s)}{\partial u} < 0 \quad \text{for all } (u, s) \in \mathbb{R}_+ \times (0, 1). \quad (49)$$

3.1 Remarks

- (P.1) Properties (44) and (45) ensure that $s(t) \in (0, 1)$, for all $t > 0$.
(P.2) The positivity of $g(s)$, for all $s \in (0, 1)$, implies that $u(t) > 0$, for all initial conditions $u(0) \geq 0$.
(P.3) Since $s(t) \in (0, 1)$, for all $t > 0$, and $g(s)$ is strictly increasing with s , we know that

$$g(s(t)) < g(1),$$

and therefore, from (39) we get

$$\frac{d}{dt}u(t) < -\varepsilon u(t) + g(1), \quad (50)$$

$$\limsup_{t \rightarrow +\infty} u(t) \leq \frac{g(1)}{\varepsilon}. \quad (51)$$

- (P.4) Thanks to (P.1)–(P.3) we may claim that the compact set

$$\Omega := [0, \frac{g(1)}{\varepsilon}] \times [0, 1]$$

is a global attractor and a positively invariant set for the solution of system (39)–(40); moreover any solution staying on the boundary of Ω will eventually enter into the interior of Ω .

The equilibria of system (39)–(40) are independent of the delay value r , so that we may study the equilibria of the ODE system associated to (39)–(40) when $r = 0$, for which the compact set Ω is still an invariant set. By the Poincaré-Bendixon Theorem and the Bendixon Criterion (see e.g. [9, page 44]) applied to the planar flow of the ODE system, we may claim the following.

Proposition 1. *If in system (39)–(40) we choose*

$$\varepsilon > \sup_{(u,s) \in \Omega} (2R(u)s + Q(u)), \quad (52)$$

then, for any choice of $r \geq 0$ there exists at least one (positive) equilibrium (u^*, s^*) of system (39)–(40), in $\overset{\circ}{\Omega}$.

As far as the stability of such an equilibrium is concerned, for the case $r = 0$ we proceed as follows.

By defining

$$A(u^*, s^*) = \frac{\partial f(u, s)}{\partial u} \Big|_{(u^*, s^*)}; \quad B(u^*, s^*) = -\frac{\partial f(u, s)}{\partial s} \Big|_{(u^*, s^*)}, \quad (53)$$

the Jacobi matrix of system (39)–(40) for $r = 0$ (no delay) at (u^*, s^*) is given by

$$J = \begin{bmatrix} -\varepsilon & g'(s^*) \\ A(u^*, s^*) & -B(u^*, s^*) \end{bmatrix}$$

As a consequence of known results, we may then state that necessary and sufficient conditions for the (local) asymptotic stability of the equilibrium (u^*, s^*) are

$$\varepsilon + B(u^*, s^*) > 0 \quad \text{and} \quad \varepsilon B(u^*, s^*) - g'(s^*)A(u^*, s^*) > 0. \quad (54)$$

Since the equilibrium s^* is the smallest root of $P(u^*) + Q(u^*)s^* + R(u^*)s^{*2} = 0$, because of the negativity of $Q(u)$ we obtain

$$s^* < \frac{1}{2} \frac{|Q(u^*)|}{R(u^*)}.$$

Accordingly $-B(u^*, s^*) = Q(u^*) + 2R(u^*)s^* < 0$, implying $B(u^*, s^*) > 0$.

Since by (iii) and (iv) we have $g'(s^*)A(u^*, s^*) < 0$, it is clear that conditions (54) hold true whenever a positive equilibrium (u^*, s^*) exists. Hence the following holds.

Proposition 2. *If in system (39)–(40) we choose ε satisfying (52) then, for $r = 0$ there exists one (positive) equilibrium (u^*, s^*) of system (39)–(40), in $\overset{\circ}{\Omega}$, which is asymptotically stable.*

3.2 Delay Induced Hopf Bifurcation

The problem we are considering in this section is the following one; in the delay system (39)–(40) let us choose the delay r as bifurcation parameter.

According to Proposition 2, when $r = 0$, the equilibrium $X^* = (u^*, s^*)^T$ is locally asymptotically stable, that is $\operatorname{Re} \lambda < 0$, for all the characteristic roots λ of the Jacobi matrix J .

We now investigate the possible occurrence of delay induced stability switches at the equilibrium X^* by increasing the delay from the value $r = 0$.

Simple computations show that the characteristic equation associated with system (39)–(40), for $r \geq 0$ at (u^*, s^*) , is given by

$$\lambda^2 + a\lambda + c + de^{-\lambda r} = 0 \quad (55)$$

with coefficients

$$a = \varepsilon + B(u^*, s^*), \quad (56)$$

$$c = \varepsilon B(u^*, s^*), \quad (57)$$

$$d = -g'(s^*)A(u^*, s^*). \quad (58)$$

In the following we apply the Hopf bifurcation Theorem as from [20, see Theorem 4.8. p. 55, and references therein] (the solution $\lambda(r)$ of (55) is parameterized in terms of $r \geq 0$).

Here we omit the question about the stability of the bifurcating periodic solutions. Searching for the value of the delay at which the equilibrium X^* switches from asymptotic stability to instability, by an increasing delay, we follow [6].

The inequalities (54), which hold true for all delays $r \geq 0$, imply that in the characteristic equation the coefficients must satisfy

$$a > 0, \quad c + d > 0, \quad \text{for all } r \geq 0. \quad (59)$$

Inequalities (59) imply that for any delay $r \geq 0$, $\lambda = 0$ cannot be a root of (55); hence, a stability switch from asymptotic stability to instability can only occur if at some delay value r_0 a pair of simple imaginary roots $\pm i\omega$, $\omega \in \mathbf{R}_+$, cross the imaginary axis towards $\mathbf{C}^+$ (i.e. the right hand side of the complex plane $\mathbf{C}$).

According to the Hopf bifurcation Theorem, the stability switch value r_0 too is a bifurcation value of the delay.

Easy computations show that

- (i) if $c^2 - d^2 \geq 0$, then there are no positive real solutions for ω , i.e. the equilibrium (u^*, s^*) remains asymptotically stable for all delays $r > 0$;
- (ii) if $c^2 - d^2 < 0$, then there is one real positive root for ω , say ω_+ , given by

$$\omega_+ = \left\{ \frac{1}{2} \left[(2c - a^2) + \sqrt{\Delta} \right] \right\}^{\frac{1}{2}}, \quad (60)$$

where

$$2c - a^2 = -(\varepsilon^2 + B(u^*, s^*)^2) < 0,$$

and

$$\Delta := (2c - a^2)^2 - 4(c^2 - d^2) > 0.$$

The pair of characteristic roots $\lambda = \pm i\omega_+$, $\omega_+ \in \mathbb{R}_+$ cross the imaginary axis according to

$$\text{sign} \left\{ \left(\frac{d(\text{Re}(\lambda))}{dr} \right)_{\lambda=\pm i\omega_+} \right\} = \text{sign}(\sqrt{\Delta}) > 0, \quad (61)$$

i.e. towards $\mathbb{C}^+$. The values of the delay at which such a crossing of the imaginary axis occurs are given by

$$r_n = \frac{\theta + 2n\pi}{\omega_+}, \quad n \in \mathbb{N}, \quad (62)$$

where $\theta \in [0, 2\pi)$ is the solution of the following system

$$\sin \theta = \frac{a \omega_+}{d}, \quad \cos \theta = \frac{\omega_+^2 - c}{d}. \quad (63)$$

Since $c + d > 0$, the occurrence (or not) of stability switches depends on the sign of

$$c - d = \varepsilon B(u^*, s^*) + g'(s^*)A(u^*, s^*).$$

By increasing the delay r , the “stability switch” may only occur at the smallest of the values r_n for which the characteristic Eq. (55) admits solutions. Further, according to the Hopf bifurcation Theorem, we have the following (see [6]),

Theorem 2. (a) If $c - d \geq 0$, the nontrivial equilibrium $X^* = (u^*, s^*)$ is (locally) asymptotically stable for any delay $r \geq 0$.
 (b) If $c - d < 0$, there exists a delay $r_0 = \frac{\theta}{\omega_+} > 0$ such that the nontrivial equilibrium $X^* = (u^*, s^*)$ is (locally) asymptotically stable for any delay $0 \leq r < r_0$, unstable for $r > r_0$, and the delay system (39) (40) admits periodic solutions, the period of which is $T(r) \simeq \frac{2\pi}{\omega_+}$, for r sufficiently close to r_0 , and an Hopf bifurcation towards instability takes place at $r = r_0$.

Note that a necessary condition for $c - d < 0$ to hold is that $g'(s^*)A(u^*, s^*) < 0$, as already assumed above.

Remark 1. In case (b), due to continuous dependence of the solution upon the delay parameter r , we may expect that for $r < r_0$, but sufficiently close to r_0 , stability of the nontrivial equilibrium may occur via damped oscillations.

It may be of interest to notice that a similar behaviour has been obtained for the dynamical behaviour of a population of stem-nonstem cells in [25], by admitting a different kind of model for cell-cell communication.

4 Simulations

In the simulations, as far as the relevant parameters are concerned, we have chosen values in accordance with the usual biological assumptions (24), i.e.

$$\lambda_1 = 1, \quad \lambda_2 = 3; \quad \gamma_1 = \gamma_2 = 0.1.$$

As a possible dependence of the parameters p_i and q_j upon the underlying field u we take the logistic one; we assume

$$\begin{aligned} p_k(u) &= p_{kk}h_1(u), \quad k = 1, 2, 3, \\ p_4(u) &= 1 - (p_3(u) + p_1(u) + p_3(u)); \\ p_{11} &= 0.75, \quad p_{22} = 0.5, \quad p_{33} = 1. \end{aligned}$$

Moreover,

$$\begin{aligned} q_1(u) &= q_{11}h_2(u), \\ q_2(u) &= 1 - q_1(u), \end{aligned}$$

with

$$q_{11} = 1.$$

where

$$h_1(u) = \frac{p p_0}{(p - p_0)e^{-\alpha_p u} + p_0}, \quad p = 0.4, \quad p_0 = 0.1, \quad \alpha_p = 1. \quad (64)$$

and

$$h_2(u) = \frac{q q_0}{(q - q_0)e^{-\alpha_q u} + q_0}, \quad q = 0.7, \quad q_0 = 0.1, \quad \alpha_q = 2. \quad (65)$$

respectively, for a given set of parameters.

In this way the parameter $p_k(u)$ may range from $p_{kk}p_0$ to $p_{kk}p$, $k = 1, 2, 3$; similarly the parameter $q_1(u)$ may range from $q_{11}q_0$ to $q_{11}q$. The specific values of p , p_0 , α_p , and q , q_0 , α_q , have been arbitrarily chosen for the time being; as for the many free other parameters, they will be the subject of an inverse problem, once reliable data will be available.

The other relevant parameter ε has been chosen so to guarantee the existence of a nontrivial equilibrium (u^*, s^*) of the delay differential system (39)–(40). If we take, as an example, $u^* = 0.5$, from the second equation we get $s^* = 0.38056$; hence

$$\varepsilon = \frac{g(s^*)}{u^*}.$$

The last element to be assigned is the function g in (39); as a very arbitrary choice we have taken $g(s) = \exp(ws)$, with $w \in \mathbb{R}_+$, to be chosen later in order to guarantee a possible Hopf bifurcation.

As a consequence of the above choices,

$$A(u^*, s^*) = -0.50835; \quad B(u^*, s^*) = 3.16015,$$

so that the equilibrium (u^*, s^*) is asymptotically stable when $r = 0$.

The condition of stability switch for an increasing delay r becomes

$$c - d = \exp(ws^*) \left\{ \frac{B(u^*, s^*)}{u^*} - w|A(u^*, s^*)| \right\} < 0,$$

that is

$$w > w_c := \frac{B(u^*, s^*)}{u^* |A(u^*, s^*)|} = 12.4329.$$

It is therefore sufficient to choose, for example $w = 13$ to imply $c - d < 0$.

With this choice we get

$$\varepsilon = \frac{\exp(13 * 0.38056)}{0.5} = 281.5829,$$

and with the parameters chosen as above, we get that, for increasing r , the first stability switch occurs at r_0 given by

$$r_0 = \frac{\theta}{\omega_*} = 3.093.$$

In accordance with Theorem 2 (ii), we may expect that, for any delay $r < r_0$, we observe damped oscillations, while for $r > r_0$ sustained oscillations occur.

In Fig. 5 we have taken $r = 1 < 3.093$, while in Fig. 6 we have taken $r = 4 > 3.093$. Both numerical simulations confirm our theoretical results.

The delay differential system (39)–(40) has been initialized by the history during the time interval $[-r, 0]$, obtained by imposing $u(\tau) = 0$, for any $\tau \in [-r, 0]$, and solving equation (40) with $s(-r) = 0.60$.

5 The case $q_2 = 0$

We are now ready to discuss the case $q_2 = 0$; in this case $P = \beta_D = \lambda_2 q_2 = 0$, which implies that the function f defined in (19) now becomes

$$f(s) := Qs + Rs^2, \quad s \in [0, 1]; \quad (66)$$

as a consequence

$$\dot{s}(t)|_{s=0} = f(0) = 0; \quad (67)$$

$$\dot{s}(t)|_{s=1} = f(1) = (-1 - p_2 + p_3)\lambda_1 < 0; \quad (68)$$

accordingly, one equilibrium is $(s^* = 0, d^* = 1)$.

Since, by (18), $R > 0$, and, by (68), $f(1) < 0$, then

$$f(s) = Qs + Rs^2 < 0, \quad s \in (0, 1], \quad (69)$$

thus implying that no other equilibrium may exist in $(0, 1]$. Further, since

$$\frac{d}{dt}s(t) = f(s(t)) < 0 \quad \text{for all } t > 0 \quad \text{such that } s(t) \in (0, 1], \quad (70)$$

we may claim the following.

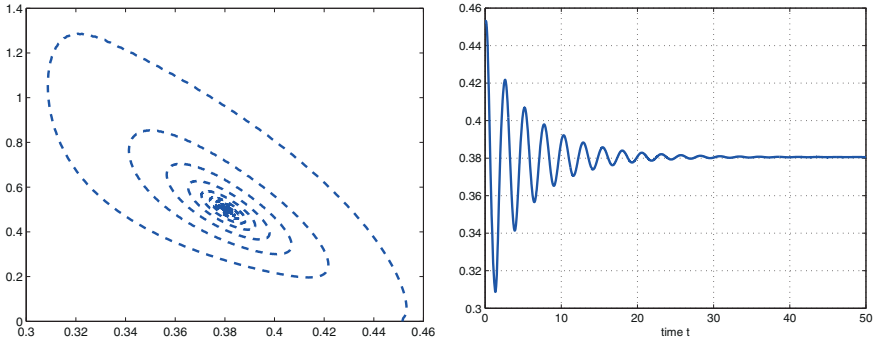


Fig. 5. Delay system (39)–(40): $r = 1$ – damped oscillations can be observed as from the theory

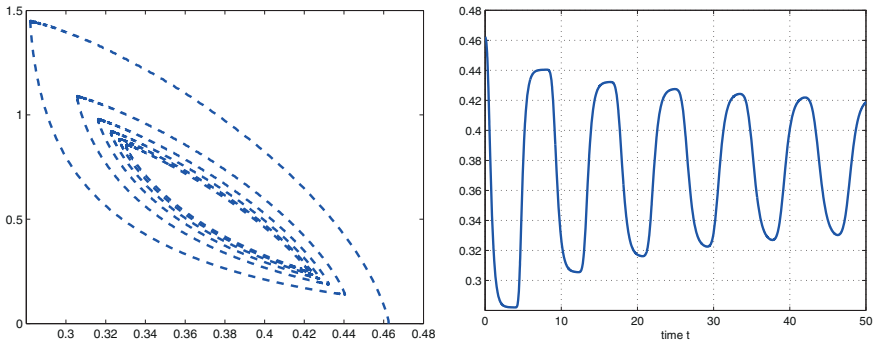


Fig. 6. Delay system (39)–(40): $r = 4$ – convergence to sustained oscillations can be observed as from the theory

Theorem 3. *If $q_2 = 0$, the unique equilibrium $(s^* = 0, d^* = 1)$ is globally asymptotically stable in $[0, 1] \times [0, 1]$, and the convergence to the equilibrium is monotone.*

Remark 2. We may conclude by observing that, if in the parameter values we drop the second and the third assumptions in (24), then it becomes possible, for the case $q_2 = 0$, to have $Q > 0$, and $R < 0$, maintaining true both (67), and (68). In this case, besides $(s^* = 0, d^* = 1)$, it must exist a nontrivial equilibrium too.

For a sake of technical simplifications, we will analyze this case by assuming that stem cells may branch either into two stem cells or into one stem cell and one non-stem cell, at a rate λ_1 ,

$$S \implies S + S, \quad \text{with probability } p_3, \quad (71)$$

$$S \implies S + D, \quad \text{with probability } p_1, \quad (72)$$

such that

$$p_1 + p_3 = 1.$$

As far as non-stem cells we assume here that they may branch only into two non-stem cells (at a rate λ_2)

$$D \implies D + D, \quad \text{with probability } q_1 = 1, \quad (73)$$

so that

$$q_1 = 1, q_2 = 0.$$

As a consequence in this case the evolution equation for the fraction of stem cells $s(t)$ is given by

$$\frac{d}{dt}s(t) = Q(p_3)s(t) + R s^2(t); \quad (74)$$

subject to an initial condition $s(0) \in [0, 1]$, with

$$Q(p_3) = (p_3 \lambda_1 - \gamma_1) - (\lambda_2 - \gamma_2), \quad (75)$$

$$R = -[(\lambda_1 - \gamma_1) - (\lambda_2 - \gamma_2)]. \quad (76)$$

From now on we assume that

$$\lambda_i > \gamma_i, \quad i = 1, 2, \text{ and } \lambda_1 - \gamma_1 > \lambda_2 - \gamma_2. \quad (A1)$$

Since $Q(p_3)$ is a continuous strictly increasing function of p_3 ,

$$\exists! p_3^* \in (0, 1) : Q(p_3^*) = 0; \quad p_3^* = \frac{(\lambda_2 - \gamma_2) + \gamma_1}{\lambda_1}. \quad (77)$$

Accordingly we have that

$$R < 0, \quad \text{and} \quad Q(p_3) \begin{cases} > 0, & \text{if } p_3 > p_3^*; \\ \leq 0, & \text{if } p_3 \leq p_3^*. \end{cases} \quad (78)$$

Eq. (74) can be written as

$$\frac{d}{dt}s(t) = f(s(t)); \quad (79)$$

with

$$f(s) = s[Q(p_3) - |R| s], \quad s \in (0, 1]. \quad (80)$$

Therefore, we can see that, whenever $p_3 \leq p_3^*$, $Q(p_3) \leq 0$, so that (79) admits only the trivial equilibrium solution $s_0^* = 0$, which is globally asymptotically stable, since in this case $f(s) < 0$ for all $s \in (0, 1]$.

On the other hand, suppose that $p_3^* < p_3 \leq 1$; in this case, besides $s_0^* = 0$, there exists an additional nontrivial feasible equilibrium $s_+^* \in (0, 1]$, given by

$$s_+^* = \frac{Q(p_3)}{|R|} = \frac{(p_3 \lambda_1 - \gamma_1) - (\lambda_2 - \gamma_2)}{(\lambda_1 - \gamma_1) - (\lambda_2 - \gamma_2)}.$$

We may easily notice that in this case we have

$$\frac{d}{dt}s(t) = -|R|(s - s_+^*)s \begin{cases} > 0, & \text{if } s \in (0, s_+^*); \\ < 0, & \text{if } s \in (s_+^*, 1], \end{cases} \quad (81)$$

and vanishes if and only if $s = s_+^*$. The property in (81) implies the global asymptotic stability of s_+^* in $(0, 1]$, and the instability of s_0^* .

By summarizing we may state the following theorem.

Theorem 4. (a) *A necessary and sufficient condition in order that our system admits a nontrivial equilibrium $s_+^* \in (0, 1)$, is that*

$$p_3 > p_3^*;$$

This equilibrium is globally asymptotically stable in $(0, 1]$, whereas the trivial equilibrium $s_0^* = 0$ is unstable.

(b) *If $0 \leq p_3 \leq p_3^*$, then s_+^* is no more feasible ($s_+^* \leq 0$), and the trivial equilibrium $s_0^* = 0$ is globally asymptotically stable in $[0, 1]$.*

We may like to notice the following.

Corollary 1. *If $R \geq 0$, then $Q(p_3) < 0$, and no additional nontrivial feasible equilibrium exists; the trivial equilibrium $s_0^* = 0$ is globally asymptotically stable in $[0, 1]$,*

Let us go back to conditions (A1) i.e. $R < 0$, together with the possibility that the parameter p_3 may depend upon an underlying signal u as in previous sections.

Such a dependence $p_3(u)$ may act on the parameter p_3 by allowing it to shift from a value below the threshold value p_3^* , to a value above it, so to shift the behavior of the system from the case of eventual extinction of stem cells to the case of eventual tendency to a nontrivial steady state, as discussed in the introduction. All other parameters are taken as independent of u .

We will then consider the following model including the coupling of the evolution of the cell population s with the evolution of an underlying field u , but in a first instance ignore possible delays,

$$\frac{d}{dt}u(t) = -\mu u(t) + g(s(t)), \quad (82)$$

$$\frac{d}{dt}s(t) = f(u(t), s(t)). \quad (83)$$

$$f(u, s) = Q(u)s + Rs^2, u \in \mathbb{R}_+, s \in [0, 1],$$

where

$$Q(u) = p_3(u)\lambda_1 - \gamma_1 - (\lambda_2 - \gamma_2);$$

$$R = -[(\lambda_1 - \gamma_1) - (\lambda_2 - \gamma_2)].$$

We assume that $\mu > 0$, and

$$g : [0, 1] \rightarrow \mathbb{R}_+; \quad g(0) = 0; \\ g \in C^1((0, 1)), \quad \text{with} \quad g'(s) > 0, \quad \text{for} \quad s \in (0, 1).$$

Under the above assumptions

$$f(u, 0) = P(u) = 0, \quad (84)$$

$$f(u, 1) = Q(u) + R = \lambda_1(p_3(u) - 1), \quad (85)$$

for all $u \geq 0$; note that $f(u, 1) < 0$ whenever $p_3(u) < 1$.

It can be easily seen that

$$\limsup_{t \rightarrow +\infty} u(t) \leq \bar{u} := \frac{g(1)}{\mu}. \quad (86)$$

Since $E_0 = (u^* = 0, s^* = 0)$ is a trivial equilibrium for system (82)–(83), we may claim that the compact set

$$\Omega := [0, \bar{u}] \times [0, 1] \subset \mathbb{R}_+^2$$

is positively invariant for the solutions of our system.

As far as the possible functional dependence of p_3 upon the underlying field u is concerned, we assume that

$$p_3 : [0, \bar{u}] \mapsto [0, 1]; \quad p_3 \in C^1((0, \bar{u})); \quad p'_3(u) < 0, \quad \text{for } u \in (0, \bar{u}).$$

Assume further that

$$\begin{aligned} p_3(\bar{u}) = 0 &\Leftrightarrow S \nrightarrow S + S \text{ is switched off} \\ p_3(0) = 1 &\Leftrightarrow S \nrightarrow S + D \text{ is switched off.} \end{aligned}$$

Since $p_3(u)$ is a continuous function, and $p'_3(u) < 0$ for $u \in (0, \bar{u})$, then

$$\exists! \hat{u} \in (0, \bar{u}) \quad \text{such that} \quad p_3(\hat{u}) = p_3^*, \quad \text{and} \quad Q(\hat{u}) = 0;$$

therefore

$$p_3(u) \begin{cases} \leq p_3^*, \text{ and } Q(u) \leq 0, & \text{if } u \in [\hat{u}, \bar{u}] \\ > p_3^*, \text{ and } Q(u) > 0, & \text{if } u \in [0, \hat{u}). \end{cases}$$

Under the above assumptions, the following theorem holds

Theorem 5. In $\overset{\circ}{\Omega}_B := (0, \hat{u}) \times (0, 1)$ there exists a unique nontrivial positive equilibrium point $E_+ = (u_+^*, s_+^*)$ of our system, which is locally asymptotically stable.

Proof. We may rewrite system (82)–(83) as follows

$$\frac{d}{dt}u(t) = \mu(\widetilde{u}(s(t)) - u(t)) \quad (87)$$

$$\frac{d}{dt}s(t) = -|R|(s(t) - \widetilde{s}(u(t))s(t)), \quad (88)$$

where

$$\widetilde{u}(s) = \frac{1}{\mu}g(s), \quad s \in [0, 1] \quad (89)$$

is the continuous isocline curve on which $\frac{d}{dt}u(t) = 0$, and

$$\widetilde{s}(u) = \frac{1}{|R|}Q(u), \quad u \in [0, \widehat{u}] \quad (90)$$

is the continuous isocline curve on which $\frac{d}{dt}s(t) = 0$.

We may notice that $\widetilde{u}(s)$ is strictly increasing from $\widetilde{u}(0) = 0$ to $\widetilde{u}(1) = \bar{u}$, as s increases from 0 to 1, whereas $\widetilde{s}(u)$ is strictly decreasing from $\widetilde{s}(0) = 1$ to $\widetilde{s}(\widehat{u}) = 0$, as u increases from 0 to $\widehat{u}$.

Therefore the two isoclines must intersect exactly in one point, say $E_+ = (u_+^*, s_+^*)$ in $\overset{\circ}{\Omega}_B$, at which is the unique positive equilibrium.

Due to the fact that $p'_3(u_+^*)$ and $g'(s_+^*)$ have opposite signs, the local stability of E_+ easily follows from the structure of the Jacobian

$$J(u_+^*, s_+^*) = \begin{pmatrix} -\mu & g'(s_+^*) \\ \lambda_1 p'_3(u_+^*) s_+^* & -Q(u_+^*) \end{pmatrix}, \quad (91)$$

since

$$\text{Tr}J(u_+^*, s_+^*) = -(\mu + Q(u_+^*)) < 0,$$

and

$$\text{Det}J(u_+^*, s_+^*) = \mu Q(u_+^*) - \lambda_1 p'_3(u_+^*) s_+^* g'(s_+^*) > 0,$$

do have opposite signs. $\square$

Global stability requires additional analysis.

Theorem 6. *Assume that the derivatives $g'(s)$ and $p'_3(u)$ are upper and lower bounded according to the following:*

$$0 < g' \leq g'(s) \leq G', \quad s \in (0, 1), \quad (92)$$

where

$$g' := \inf_{s \in (0,1)} g'(s); \quad G' := \sup_{s \in (0,1)} g'(s);$$

and

$$0 < p' \leq |p'_3(u)| \leq P', \quad u \in (0, \bar{u}), \quad (93)$$

where

$$p' := \inf_{u \in (0, \bar{u})} |p'_3(u)|; \quad P' := \sup_{u \in (0, \bar{u})} |p'_3(u)|.$$

If these bounds satisfy the following condition

$$\lambda_1 (G' P' - g' p') < 2\mu |R|, \quad (94)$$

then the unique nontrivial positive equilibrium point $E_+ = (u_+^*, s_+^*)$ of our system, is globally asymptotically stable in the whole region $\Omega \setminus \{E_0\}$.

Proof. We may center system (82)–(83) around the nontrivial equilibrium as follows

$$\frac{d}{dt}u(t) = -\mu(u(t) - u_+^*) + g(s(t)) - g(s_+^*) \quad (95)$$

$$\frac{d}{dt}s(t) = s(t)[Q(u(t)) - Q(u_+^*) - |R|(s(t) - s_+^*)]. \quad (96)$$

By the Lagrange Theorem, System (95)–(96) can be rewritten as follows

$$\frac{d}{dt}u(t) = -\mu(u(t) - u_+^*) + g'(v(t))(s(t) - s_+^*), \quad (97)$$

$$\frac{d}{dt}s(t) = s(t)[- \lambda_1 |p'_3(\eta(t))| (u(t) - u_+^*) - |R|(s(t) - s_+^*)], \quad (98)$$

for suitable functions $v(t) \in (s(t), s_+^*)$, $\eta(t) \in (u(t), u_+^*)$.

System (97)–(98) is taken subject to an initial condition $(u(0), s(0)) \in \Omega - \{E_0\}$.

Consider the following Lyapunov function,

$$V(u(t), s(t)) = \frac{1}{2}w_1(u(t) - u_+^*)^2 + w_2s_+^* \left[\frac{s(t)}{s_+^*} - 1 - \ln \left(\frac{s(t)}{s_+^*} \right) \right], \quad (99)$$

where $w_i > 0, i = 1, 2$ are constants, to be suitably chosen later.

The Lyapunov function (99) is positive definite with respect to E_+ , and its time derivative along the solutions of System (97)–(98) is

$$\frac{d}{dt}V(u(t), s(t)) = w_1(u(t) - u_+^*)[-\mu(u(t) - u_+^*) + g'(v(t))(s(t) - s_+^*)] \quad (100)$$

$$+ w_2 \frac{(s(t) - s_+^*)}{s(t)} s(t) \quad (101)$$

$$\times [-\lambda_1 |p'_3(\eta(t))| (u(t) - u_+^*) - |R|(s(t) - s_+^*)] \quad (102)$$

$$= -\mu w_1(u(t) - u_+^*)^2 - w_2 |R|(s(t) - s_+^*)^2 \quad (103)$$

$$+ (w_1 g'(v(t)) - w_2 \lambda_1 |p'_3(\eta(t))|)(u(t) - u_+^*)(s(t) - s_+^*) \quad (104)$$

$$\leq -\mu w_1(u(t) - u_+^*)^2 - w_2 |R|(s(t) - s_+^*)^2 \quad (105)$$

$$+ |w_1 g'(v(t)) - w_2 \lambda_1 |p'_3(\eta(t))|| |u(t) - u_+^*| |s(t) - s_+^*|. \quad (106)$$

It can be rewritten as the quadratic form of the real symmetric matrix

$$A(t) = \begin{pmatrix} -\mu w_1 & \frac{1}{2} |w_1 g'(v(t)) - w_2 \lambda_1 |p'_3(\eta(t))|| \\ \frac{1}{2} |w_1 g'(v(t)) - w_2 \lambda_1 |p'_3(\eta(t))|| & -w_2 |R| \end{pmatrix}.$$

as follows

$$\frac{d}{dt}V(u(t), s(t)) \leq (|u(t) - u_+^*|, |s(t) - s_+^*|)A(t) \begin{pmatrix} |u(t) - u_+^*| \\ |s(t) - s_+^*| \end{pmatrix}.$$

Hence, if we show that, under the assumptions of the theorem, the matrix $A(t)$ is negative definite, uniformly in time, the time derivative of the Lyapunov function will be itself negative definite, uniformly in time, thus implying the required global asymptotic stability of E_+ .

To this aim a necessary and sufficient condition is that the principal minors of the matrix $A(t)$, say $\Delta^{(i)}$, $i = 1, 2$, satisfy

$$\Delta^{(1)} \leq -c_1 < 0, \quad \Delta^{(2)} = \text{Det}A(t) \geq c_2 > 0,$$

for some positive constants c_i , $i = 1, 2$.

Of course the condition $\Delta^{(1)} \leq -c_1 < 0$ is trivially satisfied, whereas

$$\Delta^{(2)} = w_1 w_2 \mu |R| - \frac{1}{4} (w_1 g'(\nu(t)) - w_2 \lambda_1 |p'_3(\eta(t))|)^2,$$

requires the choice

$$w_1 = \frac{\lambda_1 P'}{G'} > 0,$$

so that it can be written as:

$$\begin{aligned} \Delta^{(2)} &= w_2 \mu |R| \frac{\lambda_1 P'}{G'} - \frac{1}{4} (\lambda_1 P')^2 \left(\left(\frac{g'(\nu)}{G'} \right)^2 + w_2^2 \left(\frac{|p'_3(\eta)|}{P'} \right)^2 \right) \\ &\quad + \frac{1}{2} w_2 (\lambda_1 P')^2 \frac{g'(\nu)}{G'} \frac{|p'_3(\eta)|}{P'} \\ &\geq (\lambda_1 P')^2 \left\{ -\frac{1}{4} w_2^2 + \frac{1}{2} \left[2 \frac{\mu |R|}{\lambda_1 G' P'} + \frac{g' p'}{G' P'} \right] w_2 - \frac{1}{4} \right\} > 0. \end{aligned} \quad (107)$$

The positivity of $\Delta^{(2)}$ is equivalent to say that there exists a choice for $w_2 > 0$ for which the inequality:

$$w_2^2 - 2 \left[2 \frac{\mu |R|}{\lambda_1 G' P'} + \frac{g' p'}{G' P'} \right] w_2 + 1 < 0$$

holds true. It is trivial to see that this choice is admissible provided that

$$2 \frac{\mu |R|}{\lambda_1 G' P'} + \frac{g' p'}{G' P'} > 1$$

or, in other words, if the derivatives $g'(s)$ and $p'_3(u)$ satisfy the assumption of the theorem

$$\lambda_1 (G' P' - g' p') < 2 \mu |R|. \quad \square$$

Remark 3. By making the simplest choice for both $g(s)$, and $p_3(u)$, that is

$$g(s) = \beta s \quad s \in [0, 1]; \quad p_3(u) = 1 - \frac{u}{\bar{u}}, \quad u \in [0, \bar{u}],$$

with $\beta > 0$, we find that the coordinates of the positive equilibrium E_+ are

$$u_+^* = \bar{u} \frac{|R|}{|R| + \lambda_1}, \quad s_+^* = \frac{|R|}{|R| + \lambda_1}.$$

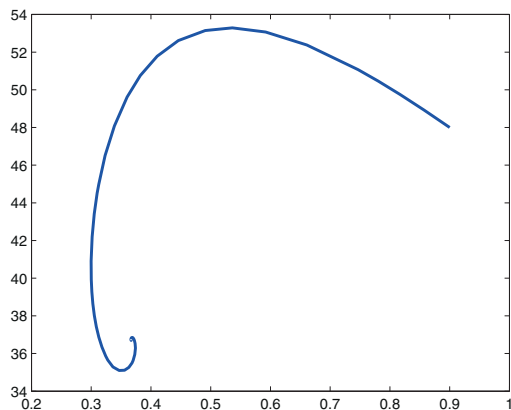


Fig. 7. Existence of a nontrivial attractor for model (82)–(83), with choices as in Remark 3

In this case we have that $G' = g'$ and $P' = p'$, and therefore the sufficient condition for the global stability trivially holds true.

The above theorem provides a sufficient condition for the global asymptotic stability of E_+ (and of the instability of E_0). However, numerical simulations (see Fig. 7) suggest the conjecture that in our model the equilibrium E_+ is always globally asymptotically stable in $\Omega - \{E_0\}$.

Remark 4. We wish to mention that an extension of model (82)–(83) including a constant time delay r in the functional response to the underlying signal u would be appropriate; i.e.

$$\frac{d}{dt}u(t) = -\mu u(t) + g(s(t-r)), \quad (108)$$

$$\frac{d}{dt}s(t) = Q(u(t))s(t) + R s^2(t), \quad (109)$$

and then proceed as in Sect. 3, but we omit further analysis of this case.

6 Conclusion

The main motivation of this research was to establish a mathematical model which might explain an intriguing experimental fact, i.e. that the percentage of cancer stem cell population is maintained at the same level during many years, and also, that this percentage starting from around 100 percent in isolated population, rapidly decreases and finally stabilizes at the same characteristic level for this given cell line. To address the question about cancer stem cell population stability, we have proposed a frame ODE model based on a catalogue of possible cell divisions of

stem and non-stem cells, and have extended our model so to include an additional variable, as a mathematical formalization of cell-cell communications in terms of a possible physical entity that we call underlying field. Next, as far as the underlying field was concerned, we additionally considered the possibility that a realistic response of this field to existing cells may be delayed by a constant time delay.

An assumption of regulation of tissue structure by an underlying (morphogenetic) field is a crucial point in biology since the concept of “positional information” was suggested by [23], and the mathematical evidence of this hypothesis is extremely important, especially for problems of pattern formation; indeed in a spatially structured model the proposed underlying signal may carry additional information regarding specific locations in space, thus driving specific behaviors accordingly. The underlying field concept, provided here, allows to give a new insight into the problem of cell-cell communication issue influencing the dynamics of population behavior of stem and non-stem cells. We may note that the description of the behavior of this variable (u) corresponding to the underlying field, which we provide here, may help to elucidate a specific biochemical nature of a substance(s) responsible for such a field formation.

From a mathematical point of view the inclusion of the third variable (u) in the equations describing the dynamics of cell populations, allows a larger variety of possible dynamical behaviors, including the possibility of an oscillatory behavior in the time evolution of stem cells, either with damped or with standing oscillations, as predicted in [25] too. We wish to notice that the stability of a characteristic percentage level of stem cells in this case can still be proven, but only for sufficiently small delays in the response of the underlying field to modifications in the cells concentrations.

Also, our modelling helps to address a very important question in current molecular biology, which is under a hot discussion and is still open: is it possible for cancer non-stem cells under some specific conditions to undergo cell division producing cancer stem cells? All experiments reporting this possibility, were considered to be proved not well enough, or to be doubted due to the specific details of stem cells markers used for these investigations.

Here we have proposed a couple of cases under which the stability of cancer stem cells population level is possible.

In the first case we have considered the possibility of dedifferentiation of non-stem cells, i.e. $q_2 > 0$, with a set of realistic assumptions on the parameters.

In the second case we have ignored dedifferentiation of non-stem cells, i.e. $q_2 = 0$, and via an oversimplified model, together with different assumptions on the parameters, we have recovered the possibility that the cancer stem cells population level eventually stabilizes, in this case too.

In conclusion, we may say that available experimental data at the moment are not decisive to validate this or that model, but seem to be in contrast with simple models which do not include communication among cells. On the other hand, existing experimental knowledge has suggested the model that we present here; our hope is that in turn our model may suggest new experiments to validate the assumptions on which we have based the model itself. As a matter of example, a crucial experiment

would be to validate the existence of an oscillatory behavior in the time evolution of stem cells, either with damped or with standing oscillations.

Acknowledgements The research contribution by Beretta and Capasso has been performed within the Italian PRIN project “Mathematical Theory of Populations: Methods, Models, Comparison with Experimental Data” (grant 2007.77BWEP-003). The research contribution by Harel-Bellan and Morozova was performed within the project “Cancérôpole” – Île-de-France, n. 2007-1-ACI-CNRS EST-1.

References

1. Bao, S., Wu, Q., McLendon, R.E., Hao, Y., Shi, Q., Hjelmeland, A.B., Dewhirst, M.W., Bigner, D.D., Rich, J.N.: Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756–760 (2006)
2. Beretta, E., Capasso, V., Morozova, N.: Mathematical modelling of cancer stem cells. Population behavior. *Math. Model. Nat. Phenom.* **7**, 306–336 (2012)
3. Bonnet, D., Dick, J.E.: Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell. *Nature Med.* **3**, 730–737 (1997)
4. Doumic, M., Marciniak-Czochra, A., Perthame, B., Zubelli, J.F.: A structured population model of cell differentiation. *SIAM J. Appl. Math.* **71**, 1918–1940 (2011)
5. d’Onofrio, A., Tomlison, I.P.M.: A nonlinear mathematical model of cell renewal, turnover and tumorigenesis in colon crypts. *J. Theor. Biol.* **244**, 367–374 (2007)
6. Freedman, H.I., Kuang, Y.: Stability switches in linear scalar neutral delay equations. *Funkcial. Ekvac.* **34**, 187–209 (1991)
7. Gimble, J.M., Katz, A.J., Bunnell, B.A.: Adipose-derived stem cells for regenerative medicine. *Circ. Res.* **100**, 1249–1260 (2007)
8. Ginestier, C., Wicha, M.S.: Mammary stem cell number as a determinate of breast cancer risk. *Breast Cancer Res.* **9**, 109 (2007)
9. Guckenheimer, J., Holmes, Ph.: Nonlinear oscillations, dynamical systems, and bifurcation of vector fields. Springer, New York (1983)
10. Gupta, P.B., Fillmore, C.M., Jiang, G., Shapira, S.D., Tao, K., Kuperwasser, C., Lander, E.S.: Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146** 633–644 (2011)
11. Gupta, P.B., Onder, T.T., Jiang, G., Tao, K., Kuperwasser, C., Weinberg, R.A., Lander, E.S.: Identification of selective inhibitors of cancer stem cells by high-throughput screening. *Cell* **138**, 645–659 (2009)
12. Johnston, M.D., Edwards, C.M., Bodmer, W.F., Maini, P.K., Chapman, S.J.: Mathematical modelling of cell population dynamics in the colonic crypt and in colorectal cancer. *PNAS* **104**, 4008–4013 (2007)
13. Li, C., Heidt, D.G., Dalerba, P., Burant, C.F., Zhang, L., Adsay, V., Wicha, M., Clarke, M.F., Simeone, D.M.: Identification of pancreatic cancer stem cells. *Cancer Res.* **67**, 1030–1037 (2007)
14. Li, X., Lewis, M.T., Huang, J., Gutierrez, C., Osborne, C.K., Wu, M.F., Hilsenbeck, S.G., Pavlick, A., Zhang, X., Chamness, G.C., Wong, H., Rosen, J., Chang, J.C.: Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J. Natl. Cancer Inst.* **100**, 672–679 (2008)
15. Maitland, N.J., Collins, A.T.: Prostate cancer stem cells: a new target for therapy. *J. Clin. Oncol.* **26**, 2862–2870 (2008)

16. Marciniak-Czochra, A., Stiehl, T., Ho, A.D., Jaeger, W., Wagnar, W.: Modeling of asymmetric cell division in hemapoietic stem cells: Regulation of self-renewal is essential for efficient repopulation. *Stem Cells Dev.* **18**, 377–386 (2009)
17. Michor, F.: Mathematical models of cancer stem cells. *J. Clin. Oncol.* **26**, 2854–2861 (2008)
18. O'Brien, C.A., Pollett, A., Gallinger, S., Dick, J.E.: A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* **445**, 106–110 (2007)
19. Roeder, I., Herberg, M., Horn, M.: An “age” structured model of hemapoietic stem cell organization with application to chronic myeloid leukemia. *Bull. Math. Biol.* **71**, 602–626 (2009)
20. Smith, H.: An introduction to delay differential equations with applications to the life sciences. Springer, New York (2010)
21. Weinberg, R.A.: The biology of cancer. Garland Science, New York (2007)
22. Whetton, A.D., Graham, G.J.: Homing and mobilization in the stem cell niche. *Trends Cell Biol.* **9**, 233–238 (1999)
23. Wolpert, L.: Positional information and the spatial pattern of cellular differentiation. *J. Theor. Biol.* **25**, 1–47 (1969)
24. Zhang, S., Balch, C., Chan, M.W., Lai, H.C., Matei, D., Schilder, J.M., Yan, P.S., Huang, T.H., Nephew, K.P.: Identification and characterization of ovarian cancer-initiating cells from primary human tumors. *Cancer Res.* **68**, 4311–4320 (2008)
25. Zhdanov, V.P.: Effect of cell-cell communication on the kinetics of proliferation and differentiation of stem cells. *Chem. Phys. Letters* **437**, 253–256 (2007)
26. Zhdanov, V.P.: Signal propagation in stem cell niches. *Physica A* **388**, 3791–3797 (2009)

Glucose Metabolism in Multicellular Spheroids, ATP Production and Effects of Acidity

Antonio Fasano

Abstract In hypoxic conditions cells switch to the anaerobic pathway of glucose metabolism. As a byproduct lactate is produced together with H^+ , lowering pH. Acidity beyond some threshold may cause cell death. Moreover, anaerobic metabolism is far less efficient in terms of ATP production, and a scarce availability of glucose may result in insufficient ATP production rate to sustain the cell life. In this paper first of all we review some results concerning the structure of multicellular tumor spheroids in which the onset of a necrotic core is caused by the drop of ATP production rate below the viability threshold. Then we shortly discuss the effects of combined insufficient ATP production rate and excess of acidity on both vascular and avascular tumors. The last part of the paper is devoted to the phenomenon of acid mediated tumor invasion, which exploits the different resistance to acidity of tumor and of normal cells. In particular, a one dimensional model is discussed in which invasion takes the form of a travelling wave front.

1 Introduction

The importance of glucose metabolism in connection with tumor development and invasion is well known, to the point that O.H. Warburg as early as 1924 postulated the famous hypothesis bearing his name [30], according to which a cell becomes malignant by switching its glucose metabolism from the normal oxidative pathway to the anaerobic pathway. The anaerobic pathway, far less efficient from the energy point of view, may keep the cell alive at a quiescent level in the absence of oxygen. For long time Warburg's hypothesis has been criticized, since it looked obvious that

A. Fasano (✉)

Department of Mathematics U. Dini, University of Florence, Viale Morgagni 67/A, 50134 Florence, Italy

CNR Institute for Systems Analysis and Computer Science A. Ruberti, Viale Manzoni 30, 00185 Rome, Italy

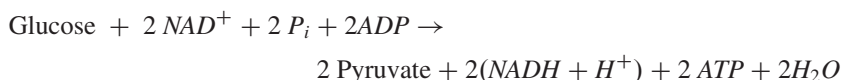
e-mail: fasano@math.unifi.it

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

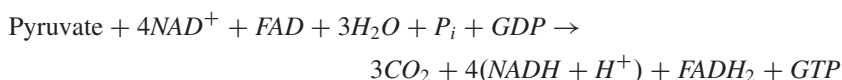
DOI 10.1007/978-88-470-2571-4_9, © Springer-Verlag Italia 2012

such a switch is rather the consequence than the origin of the tumor growth, as a natural reaction to growth related hypoxia. However, in some cases tumor cells differentiate to the point that they possess less mitochondria, actually reducing their respiratory ability. Glucose metabolism has the goal of producing the nucleotides ATP (adenosine-triphosphate) and GTP (guanosine- triphosphate), providing the energy for proteins synthesis. The whole process is sketched in Fig. 1, where the following two stages are emphasized¹.

1. Glycolysis:



2. Krebs cycle:



Krebs cycle takes place in cells mitochondria. The transformation $\text{NADH} \rightarrow \text{NAD}^+$ provides energy for the formation of an average (efficiency is variable, depending on

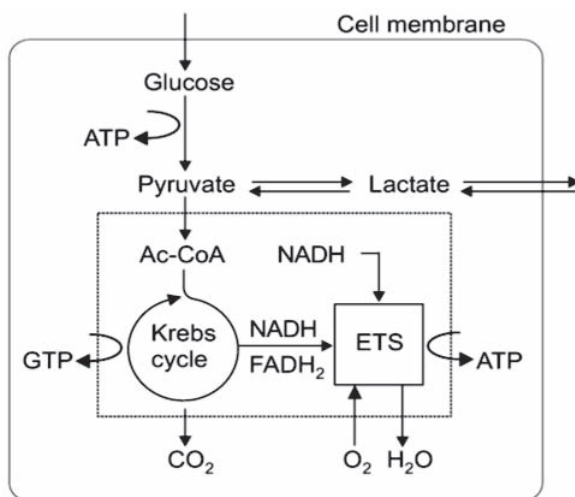


Fig. 1. Sketch of glucose metabolism

¹ Chemical notations are as follows: ADP, adenosine diphosphate; P_i , inorganic orthophosphate; GDP, guanosine diphosphate; NAD^+ , nicotinamide adenine dinucleotide (oxidized form); NADH, nicotinamide adenine dinucleotide (reduced form); FAD, flavin adenine dinucleotide (oxidized form); FADH_2 , flavin adenine dinucleotide (reduced form).

cell species and status) of 2.5 ATP molecules (by phosphorylation of ADP), while an average of 1.5 ATP molecule is generated by each $FADH_2$ molecule [19], such reactions being carried out in the so-called Electron Transport System (ETS).

After glycolysis, Pyruvate has some probability to enter the full oxidation cycle (depending on the availability of oxygen), or to reversibly decay into Lactate, which may eventually leave the cell. In the anaerobic pathway Krebs cycle is ignored with two consequences: a) ATP production is kept at a minimal level; b) since all pyruvate is engaged in the reversible conversion to lactate, the concentration of the latter increases.

The pyruvate-lactate transition has the following scheme



The level of lactate is responsible for the degree of local acidity.

We may conclude that hypoxia raises the concentration of H^+ ions. The corresponding pH lowering turns out to be more harmful for surrounding healthy cells than for the tumor cells themselves, creating a situation which can be advantageous for tumor invasion. For an extended illustration of glucose metabolism see [27].

In this paper we want to illustrate two interesting trends in the literature of cancer modeling, related to the ATP production rate as an index of cell viability, and to the effects of hypoxia-induced acidity. More precisely, we will deal with two main subjects: 1) The onset of a necrotic core in multicellular spheroids due to insufficient ATP production rate; 2) Tumor acidity as a tool for invasion.

Models including pH evolution in connection with lactate production have been formulated in [7, 8]. We are going to use data from those papers, and we will make reference to the classical paper [15], whose collection of experimental data is still of great value, and to [23].

2 ATP-Regulated Cell Apoptosis

Most of mathematical models for cancer growth including spatial structure (hence based on partial differential equations) deal with just one limiting nutrient (typically oxygen). Only recently intracellular ATP production has been incorporated in continuous models for the growth of spheroids [29] and of tumor cords [1]. Also [18, 24] have included it in their models utilizing cellular automata. The model proposed in [3, 4], also encompassing the dynamics of glucose, oxygen and pyruvate, was based on the assumption that apoptosis occurs not in correspondence of some threshold concentration, but when the ATP production rate falls below some minimal value. This is the approach we want to summarize here. The viewpoint of [3, 4] is rather partial, since in those papers the authors only considered a multicellular spheroid, *at the steady state, supposing that its radius R is known*, with the aim of showing that, in the selected framework for necrosis, sufficiently large spheroids have to contain a necrotic core, whose radius is uniquely determined in terms of R . This setting of the problem bypasses the difficulty of finding R in correspondence

of the boundary data for the nutrients (see the review paper [11] for the relevant literature). In that way we only have to consider the reaction-diffusion processes for all the involved nutrients and metabolites in the presence of a sharp interface separating the necrotic core from the external viable shell. Looking for R is out of our objectives and it requires the full analysis of the motion of the tumor components (cells, extracellular liquid, etc.) and of their mechanical behavior. Despite such a substantial simplification, the problem remains highly nontrivial.

In the following the consumption rates of oxygen, glucose, and lactate will be denoted by f_i , $i = O, G, L$ ($\text{mol}^{-1} \text{ cell}^{-1} \text{ sec}^{-1}$). Independently of how they are related with intra- and extracellular concentrations, the stoichiometric balance in the chemical reactions illustrated in the previous section implies

$$f_O = 6f_G + 3f_L. \quad (1)$$

The proof of (1) is lengthy (see [3, 4]). It must be pointed out that, while f_O, f_G are always positive, f_L can be negative, meaning that lactate can be produced. Similarly, denoting now by f_{ATP} the ATP production rate, we have

$$f_{ATP} = 2f_G + \frac{1}{3}(5\eta_1 + \eta_2 + 1)f_O, \quad (2)$$

where η_1, η_2 are the efficiency conversion coefficients defined above, namely: $\eta_1 = 2.5$, $\eta_2 = 1.5$, so that the coefficient of f_O in (2) is 5 (it is anyway an approximate value).

On the basis of the motivations provided in [3, 4], in turn substantiated by the already quoted papers [7, 8, 29] and by [2, 6, 10, 22, 26], the following scheme has been selected for the evolution of the concentrations $\bar{\sigma}_i$ at the cell interior, together with appropriate expressions for the f_i 's. The symbol σ_i denote concentrations at the cell exterior. We set $\bar{\sigma}_O = \sigma_O$, since the cell membrane is practically transparent to oxygen.

$$v \frac{d\bar{\sigma}_P}{dt} = 2f_G(\sigma_G) - \psi(\bar{\sigma}_P, \bar{\sigma}_L) - \phi_P(\bar{\sigma}_P, \sigma_O), \quad (3)$$

$$v \frac{d\bar{\sigma}_L}{dt} = \psi(\bar{\sigma}_P, \bar{\sigma}_L) + v h(\sigma_L - \bar{\sigma}_L), \quad (4)$$

$$f_G = \frac{\sigma_G}{K_G + \sigma_G}, \quad (5)$$

$$\psi(\bar{\sigma}_P, \bar{\sigma}_L) = nu(k_+ \bar{\sigma}_P - k_- \bar{\sigma}_L), \quad (6)$$

$$\phi_P(\bar{\sigma}_P, \sigma_O) = F_P \frac{\bar{\sigma}_P}{K_P + \bar{\sigma}_P} \frac{\sigma_O}{K_O + \sigma_O}, \quad (7)$$

$$f_O = 3\phi_P \quad (8)$$

$$f_O(\sigma_G, \sigma_L, \sigma_O) \approx \gamma(2f_G + p_1 \sigma_L) \frac{\sigma_O}{K_O + \sigma_O}, \quad (9)$$

where $\gamma = 3F_P/p_2$, with $p_1 = k_- h v / (k_- + h)$, $p_2 = k_+ h v K_P / (k_+ + h)$. In the system above ψ denotes the pyruvate-lactate flow rate, while ϕ_P is the pyruvate flow rate

towards the Krebs cycle. $F_P, F_G, K_P, K_G, k_-, k_+, h$ are the constants, v denotes the cell volume. The numerical values of the parameters are listed in [4]. Eq. (9) is written in an approximate simplified form: for its limit of validity see [4].

Remark. In the scheme above we have neglected both the Crabtree and the Pasteur effect. The Crabtree effect consists in the increase of f_O when σ_G decreases. The Pasteur effect is instead the increase of f_G when σ_O decreases. Attempts to introduce Pasteur effect were unsuccessful, since the fit with experimental data was not improved.

In writing (3) we have supposed that pyruvate exchange with the cell exterior is negligible. At the cell steady state, which we are interested in, the intracellular concentrations can be expressed in terms of the extracellular concentrations as follows:

$$\bar{\sigma}_L = \frac{k_+ \bar{\sigma}_P + h \sigma_L}{k_- + h} \quad (10)$$

$$\bar{\sigma}_P = \frac{K_P}{2p_2} \left((w^2 + 4p_2 (2f_G(\sigma_G) + p_1 \sigma_L))^{1/2} - w \right), \quad (11)$$

where:

$$w = p_2 + F_P \frac{\sigma_O}{K_O + \sigma_O} - 2f_G(\sigma_G) - p_1 \sigma_L.$$

In the viable region the spatial profile of the extracellular concentrations σ_i is governed by the equations

$$D_O \Delta \sigma_O = \frac{v^*}{v} f_O(\sigma_G, \sigma_L, \sigma_O), \quad (12)$$

$$D_L \Delta \sigma_L = \frac{v^*}{v} f_L(\sigma_G, \sigma_L, \sigma_O), \quad (13)$$

$$D_G \Delta \sigma_G = \frac{v^*}{v} f_G(\sigma_G), \quad (14)$$

where Δ is the Laplace operator in radial symmetry, $\Delta = \partial_r^2 + (2/r)\partial_r$, and v^* is the fraction of volume occupied by cells. As we said, we define the viable region according to the inequality $f_{ATP} > f_N$, i.e. as the region in which the ATP production rate exceeds the given threshold f_N holds true. The scalar fields $\sigma_G, \sigma_L, \sigma_O$ are flat in the necrotic core and no mass exchange takes place at the boundary, meaning that

$$\sigma'_i(\rho_N) = 0, \quad i = G, L, O. \quad (15)$$

At the external boundary of the spheroid we assume that $\sigma_G, \sigma_L, \sigma_O$ are known:

$$\sigma_i(R) = \sigma_i^*, \quad i = G, L, O \quad (16)$$

where σ_G^* and σ_O^* are strictly positive, and σ_L^* is nonnegative.

The analysis of the differential system (12), (13), (14) with the boundary conditions (16), has been performed in [3], showing in particular that if (15) is substituted by $\sigma'(0) = 0$, the condition $f_{ATP} > f_N$ cannot be maintained throughout the spheroid,

if R is sufficiently large. Thus spheroids at the steady state exceeding some critical size (depending on the boundary data) must have a necrotic core. A further step in the study of the glucose profile within the spheroid stems from the remark that, for fixed R , the boundary value problem for (14) with data $\sigma_G(R) = \sigma_G^*$, $\sigma'_G(0) = 0$ provides a solution with a limiting slope $\sigma'_G(R) = \Sigma_G^* > 0$. If we solve the Cauchy problem for (14) with data $\sigma_G(R) = \sigma_G^*$, $\sigma'_G(R) = \Sigma_G < \Sigma_G^*$, the corresponding profile exhibits a minimum at some positive radius ρ . In the limit $\Sigma_G \rightarrow 0^+$ the radius ρ approaches R and obviously it goes to zero when $\Sigma_G \rightarrow (\Sigma_G^*)^-$. The existence of ρ is proved by showing that both σ_G and σ'_G depend monotonically on $\Sigma_G \in (0, \Sigma_G^*)$. Moreover the value $\sigma_G(\rho)$ is a decreasing function of Σ_G . The function $\rho(\Sigma_G)$ is defined implicitly via the equality

$$\Sigma_G = \frac{\Gamma_G}{R^2} \int_{\rho(\Sigma_G)}^R r^2 f_G(\sigma_G) dr, \quad (17)$$

where $\Gamma_G = D_G^{-1} v^* / (1 - v^*)$. Though the solutions of this family have generally no physical meaning, ρ is candidate to be identified with the interface ρ_N when additional conditions are satisfied. For this to occur it is necessary that at that same point both the derivatives σ'_L , σ'_O vanish, and that f_{ATP} takes precisely the value f_N . Thus we adopt the following strategy. We introduce the parameters Σ_L , Σ_O as the slopes σ'_L , σ'_O at $r = R$, respectively. Then we consider the Cauchy problem for the system (12), (13) with the natural conditions (16) and the just mentioned slopes. To the solutions we impose the conditions $\sigma'_L(\rho) = \sigma'_O(\rho) = 0$, exactly at the point $\rho = \rho(\Sigma_G)$ defined before. The procedure just sketched leads to the pair of equations

$$\Sigma_L = \frac{\Gamma_G}{R^2} \int_{\rho(\Sigma_L)}^R r^2 f_L(\sigma_L) dr, \quad \Sigma_O = \frac{\Gamma_O}{R^2} \int_{\rho(\Sigma_O)}^R r^2 f_O(\sigma_O) dr, \quad (18)$$

where $\Gamma_O = v^* / D_O$, $\Gamma_L = v^* / (D_O(1 - v^*))$. Such a system can be interpreted as defining a curve in the three-dimensional space of the parameters $(\Sigma_G, \Sigma_L, \Sigma_O)$, parametrized by the variable Σ_G :

$$\Sigma_L - \Omega_L(\Sigma_G, \Sigma_L, \Sigma_O) = 0, \quad (19)$$

$$\Sigma_O - \Omega_O(\Sigma_G, \Sigma_L, \Sigma_O) = 0. \quad (20)$$

A key point in the existence proof is to show that the Jacobian of the system (19)–(20) is bounded and bounded away from zero. Such a result has been demonstrated in [3] and the proof is rather long. Once the above curve has become available, we can evaluate f_{ATP} along it with Σ_G varying over the whole interval $(0, \Sigma_G^*)$ and check that it takes the threshold value f_N exactly once, in correspondence to some $\hat{\Sigma}_G$, provided the spheroid is sufficiently large. The knowledge of $\hat{\Sigma}_G$ allows to construct the solution of the original free boundary problem. We refer to [4] for the comparison of the numerically simulated necrotic radius with experiments. Experimental data were taken from [16]. They refer to two sets of four experiments each. In the first set $\sigma_O^* = 0.07 \text{ mM}$ (low oxygen concentration), while $\sigma_O^* = 0.28 \text{ mM}$ in the second set

(non-hypoxic state). In the four experiments σ_G^* takes the respective values $0.8mM$, $1.7mM$, $5.5mM$, $16.5mM$. All the parameters in the simulations were fixed once for all. The selection of p_1 and of f_N was made on the basis of a best fitting. The parameter γ was chosen with the criterion of matching the measured oxygen consumption rate for the case $\sigma_O^* = 0.28 mM$ and $\sigma_G^* = 5.5 mM$. The agreement was certainly very good, with the only exception of the extreme case of hypoxia in the presence of the largest glucose concentration, in which ρ_N slightly deviates from experiments for small and for large spheroids, in opposite directions. This is the case in which lactate concentration takes large values, inducing a high acidity level, suggesting that a possible improvement of the model can be the inclusion of acidity effects on necrosis.

3 Necrosis by Acidity and Tumor Invasiveness

As we said, in the previous analysis we neglected the fact that pH can drop below the threshold tolerated by cells. Acidity increase in tumors was pointed out already in the classical paper [14]. A very interesting feature concerning resistance to acidity is that the above said threshold is lower for tumor cells than for normal cells. Quantitatively, a threshold $pH = 6.8$ has been reported for tumor cells [7], where the complex mechanism of H^+ ions buffers in cells is described, while the corresponding value for normal cells is $pH = 7.1$ [9]. It has been postulated [25] that the existence of this gap gives the tumor an advantage to spread at the expenses of the host tissue. According to this theory necrotic layer may develop at the tumor boundary due to the death of the host tissue expose to acid environment. Experimental evidence of such a layer has been reported in the paper [24]. In the paper [25] a highly simplified model is presented in which a multicellular spheroid grows within a host tissue, which is well perfused. Owing to the two different acidity threshold a gap between the two kinds of tissue is present. Two cases are considered: the tumor can be vascularized or not. In both cases the gap is vascularized. In [25] the H^+ production is prescribed and not correlated with glucose metabolism. A scheme based on the double acidity threshold was studied in [5], which included many more aspects. First of all the H^+ ions production rate was fully related with glucose consumption, thus incorporating the glucose dynamics. Then both in the avascular and in the vascular case, the vasculature in the gap or in the tumor and the gap was affected by acidity. Moreover, the death of tumor cells could be due to excessive acidity, or to defective glucose. Finally, living cells could be proliferating or quiescent according to the local value of nutrient concentration. In correspondence of that complex framework, a large number of different cases could occur and all the possible steady states were analyzed, thus providing a rather precise picture of the influence of acidity on the structure of multicellular spheroids. We are not going to report the long list of different cases and we just refer to the paper [5] for the detailed analysis of all possible configurations.

4 A Model of Acid-Mediated Tumor Invasion Based on Travelling Waves

An interesting approach to acid-mediated tumor invasion in the spirit of the previous section was proposed in [17]. The authors considered a one dimensional setting in which a population of tumor cells invades a healthy tissue by exploiting the different tolerance to acidity. In synthesis, the model proposed in [17] has the following structure:

$$\frac{\partial u}{\partial t} = u(1 - u) - auw, \quad (21)$$

$$\frac{\partial v}{\partial t} = d \frac{\partial}{\partial x} \left[(1 - u) \frac{\partial v}{\partial x} \right] + bv(1 - v), \quad (22)$$

$$\frac{\partial v}{\partial t} = \frac{\partial^2 w}{\partial x^2} + c(v - w). \quad (23)$$

In the scheme above all variables are dimensionless² and the concentrations u , v , w are normalized to vary between 0 and 1. They refer to host tissue cells, tumor cells, and to H^+ ions, respectively. The space coordinate x varies from $-\infty$ to $+\infty$. The terms $u(1 - u)$ and $bv(1 - v)$ express logistic growth for the respective population. Of course the coefficient b is greater than 1, since tumor cells have a larger replications rate. The negative term in (21) models the aggression of H^+ ions on healthy cells, weighted by the constant a which plays a very important role. In the third equation diffusivity of H^+ ions is normalized to 1, c is a positive constant, and the positive contribution is due to production by tumor cells, while the negative term expresses the ions decay. In the second equation the dimensionless diffusivity coefficient d is the ratio between cells diffusivity and ions diffusivity, so it is expected to be very small (though the ions diffusivity is not as large as one can imagine [28]), more precisely of the order 10^{-5} . The factor $(1 - u)$ in brackets says that tumor cells diffusion is hindered by the healthy cells, so that no penetration can occur without the intervention of acidity. Healthy cells are not supposed to move, thus forming a coherent structure. Thus the mechanism for tumor invasion depicted in (21)–(22)–(23) consists in the killing of healthy cells by the ions, which provides space for the tumor cell to advance by slowly diffusing forward. The absence of diffusivity of healthy cells can be justified by the fact that they strongly adhere to each other, while tumor cells have a higher mobility. It is rather obvious that the model (21)–(22)–(23) is an oversimplified description of the invasion phenomenon, ignoring for instance the role of extracellular fluid and, above all, the glucose metabolism leading to the increase of acidity. The material used to build the new cells can be supposed to be taken e.g. from the necrotic material in the gap, but such details are not specified and the model keeps just the main processes. In [17] travelling wave solutions have been found for (21)–(22)–(23), setting

$$u(x, t) = U(x - \theta t), v(x, t) = V(x - \theta t), w(x, t) = W(x - \theta t), \quad (24)$$

² For the dimensional formulation and the values of the physical parameters see [12].

thus transforming the p.d.e.'s system to an o.d.e's system

$$0 = \theta U' + U(1 - U) - aVW, \quad (25)$$

$$0 = d(1 - U)V'' - U'V' + \theta V' + bV(1 - V), \quad (26)$$

$$0 = W'' + \theta W' + c(V - W). \quad (27)$$

For an invasion problem ($\theta > 0$) the natural boundary conditions are

$$U(-\infty) = \max(0, 1 - a), V(-\infty) = 1, W(-\infty) = 1, \quad (28)$$

$$U(+\infty) = 1, V(+\infty) = 0, W(+\infty) = 0. \quad (29)$$

Thus (28) expresses total elimination of the host tissue if $a \geq 1$, or a partial survival if $a < 1$. The system above has been studied numerically in [17] choosing a speed θ which the authors supposed to be the minimal possible wave speed, in the spirit of the celebrated theories by Fisher [13] and by Kolmogorov, Petrovsky, and Piskunov [20]. We will return very soon to this basic aspect. As a matter of fact, for particular values of the parameters they found travelling waves characterized by a steep advancing front (where V jumps from a value close to 1 to a value close to 0), a steep regressing front (with a symmetric behavior of U) with a gap between the two. Actually, the parameters were chosen so to obtain a gap width much larger than the experimental one, to make it visible on the graphs. The great merit of [17] was to show that a very simple scheme of acid-mediated invasion could generate a travelling wave. In the paper [12] all possible travelling waves associated to (25)–(26)–(27), (28), (29) have been analyzed. Let us summarize the main results. First of all there are two classes of waves, related to the behavior of the wave speed as d goes to zero:

(I) *Fast Waves*: $\theta = O(1)$

(II) *Slow Waves*: $\theta = \theta_0 d^\alpha$

the coefficient θ_0 remaining finite as $d \rightarrow 0$. As we shall see, the precise definition of 'slow' is that

$$\theta_0 d^\alpha \ll \sqrt{c}. \quad (30)$$

The slow waves are the ones of physical interest, because they are sensitive to diffusivity, which is the advancement mechanism in the model. Therefore we confine our attention to them, referring the reader to [12] for the properties of the fast waves. The first result concerning slow waves is that the exponent α can vary in the interval $(0, 1/2]$, as we shall see very soon. Thus, considering the fact that $d < 1$, $\alpha = 1/2$ corresponds to the slowest possible waves class, as in the already recalled travelling wave solutions of the classical problems in [13] and [20]. Moreover for $\alpha = 1/2$ the coefficient θ_0 has a minimum acceptable value:

$$\theta_0 \geq \sqrt{b \min\left(\frac{a}{2}, 1\right)} \quad (31)$$

which increases with the proliferation rate, i.e. b , and also with the sensitivity to acid aggression, when $a < 2$. It is interesting to remark that results of this kind (the existence of a minimal wave speed) are recurrent in the literature about invasion processes by cell populations. Besides the classical papers quoted above, we recall also [21] as an example of invasion process related with wound healing. The technique for studying the slow waves consists in distinguishing the outer solution (i.e. far from the wave front) and the inner solution, close to the front. This is suggested by the rapid change that V undergoes across the front. In the following we will use the notations $U(z, d)$, $V(z, d)$, $W(z, d)$, emphasizing the dependence on the relative diffusivity d , where $z = x - \theta t$ is the coordinate identifying the front position as $z = 0$. A first remarkable aspects of (25) is that it is a Bernoulli type equation, so that $U(z, d)$ can be found in terms of $W(z, d)$

$$U(z, d) = \frac{\theta \Phi(z, d)}{\int_z^\infty \Phi(s, d) ds}, \quad (32)$$

where

$$\Phi(z, d) = \exp \left(-\frac{1}{\theta} \int_0^z (1 - W(s, d)) ds \right). \quad (33)$$

The properties of the function Φ have been investigated in [12], checking that U satisfies the correct boundary conditions. At the same time (27) can be integrated, expressing W as a functional of V

$$W(z, d) = \frac{c}{r_1 - r_2} \left(e^{r_1 z} \int_z^\infty e^{-r_1 s} V(s, d) ds + e^{r_2 z} \int_z^\infty e^{-r_2 s} V(s, d) ds \right), \quad (34)$$

where

$$r_1 = \frac{-\theta + \sqrt{\theta^2 + 4c}}{2} > 0, \quad \frac{-\theta - \sqrt{\theta^2 + 4c}}{2} < 0.$$

4.1 Outer Solution

Since far from the front the tumor cells diffusivity has no influence, the outer solution can be obtained by simply setting $d = 0$:

$$U_{out} = U(z, 0), V_{out} = V(z, 0), W_{out} = W(z, 0). \quad (35)$$

Taking into account (28), (29) the simplified equation (26) provides the simple step-wise solution

$$V_{out}(z) = \begin{cases} 1 & \text{if } z < 0 \\ 0 & \text{if } z > 0. \end{cases} \quad (36)$$

Integrating (27) leads now to two exponential branches for W_{out} :

$$W_{out}(z) = \begin{cases} 1 - \frac{1}{2} e^{\sqrt{c}z} & \text{if } z < 0 \\ \frac{1}{2} e^{-\sqrt{c}z} & \text{if } z > 0. \end{cases} \quad (37)$$

The analysis for U takes different ways, depending on the value of a and is much more complicated. We just summarize the conclusions, still dealing with the outer solutions.

1) Case $0 < a < 1$:

$$U_{out}(z) = \begin{cases} 1 - a + \frac{1}{2}e^{\sqrt{c}z} & \text{if } z < 0 \\ 1 - \frac{a}{2}e^{-\sqrt{c}z} & \text{if } z > 0 \end{cases} \quad (38)$$

showing the partial survival of healthy cells in the acidic region.

2) Case $1 < a < 2$:

is the same as in the previous case for $z > 0$. A substantial difference arises for $z < 0$: we have to analyze $U(z, t)$ separately in the intervals $(z_-, 0)$ and $(-\infty, z_-)$, where

$$z_- = \frac{1}{\sqrt{c}} \log \frac{2(1-a)}{a} < 0. \quad (39)$$

Indeed in the first interval $U_{out}(z)$ drops to zero:

$$U_{out}(z) = 1 - a + \frac{a}{2}e^{+\sqrt{c}z} \quad (40)$$

i.e. it has the same structure as the previous case, and it vanishes at $z = z_-$. At the left of z_- the outer solution is basically zero. However the leading term in d can be found:

$$U(z, d) = \sqrt{\frac{\sqrt{c}(a-1)\theta_0}{2\pi}} d^{\alpha/2} \exp\left(\frac{\phi_-(z) - \phi_-(z_-)}{d^\alpha}\right), \quad (41)$$

where the function

$$\phi_-(z) = \frac{1}{\theta_0} \left(\frac{a}{2\sqrt{c}}(1 - e^{\sqrt{c}z}) + (a-1)z \right) \quad (42)$$

has a positive relative maximum in z_- . Formula (41) emphasizes how the wave speed determines the rate at which u goes to zero at $-\infty$ through θ_0, d , and α .

3) Case $a > 2$:

Here we need to define the point z_+ having a role symmetric to z_-

$$z_+ = \frac{1}{\sqrt{c}} \log \frac{a}{2}. \quad (43)$$

For $z > z_+$ the profile of the healthy cells concentration is the following

$$U_{out}(z) = 1 - \frac{a}{2}e^{-\sqrt{c}z} \quad (44)$$

vanishing at $z = z_+$. To the left of that point, at the zero order in d the solution is zero, and to the leading order in d we have

$$U(z, d) = \sqrt{\frac{\theta_0 \sqrt{c}}{2\pi}} d^{\alpha/2} \exp\left(\frac{\phi_+(z) - \phi_+(z_+)}{d^\alpha}\right) \quad (45)$$

which parallels the expression (41) of the previous case, where now

$$\phi_+(z) = \frac{1}{\theta_0} \left(\frac{a}{2\sqrt{c}} (1 - e^{-\sqrt{c}z} - z) \right). \quad (46)$$

4) Limit cases $a = 1, a = 2$:

For $a = 1$ we can take the limit $a \rightarrow 1$ in (38) and the limit $a \rightarrow 2$ in (40) (note that $z_- \rightarrow -\infty$), obtaining in both cases

$$U_{out}(z) = \begin{cases} \frac{1}{2} e^{\sqrt{c}z} & \text{if } z < 0 \\ 1 - \frac{1}{2} e^{-\sqrt{c}z} & \text{if } z > 0. \end{cases} \quad (47)$$

For $a \rightarrow 2^\pm$ the limit of z_+ is zero. Hence $U_{out}(z)$ does not vanish only for $z > 0$. The leading order term in d on the negative axis is

$$U_{out}(z) = \sqrt{\frac{\theta_0 \sqrt{c}}{2\pi}} d^{\alpha/2} \exp\left(\frac{\phi_-(z)}{d^\alpha}\right). \quad (48)$$

4.2 Inner Solutions

In order to find out the structure of the travelling wave in the neighborhood of the front we have to put a magnifying lens in the vicinity of $z = 0$, constructing the so called inner solution. Since w , and consequently u , vary slowly across the front, we expect the corresponding inner solutions to be constant, assuming the value of the outer solution as z tends to zero. Investigating the inner profile of v is instead a more difficult and interesting task. Once more we distinguish the three intervals $(0, 1)$, $(1, 2)$, $(2, +\infty)$ for the sensitivity parameter a . Let us define the variable $\xi = z/d^\alpha$, stretching the wave profile near the front, so that ξ is $O(1)$. Setting

$$U(z, d) = X(\xi, d), V(z, d) = Y(\xi, d), W(z, d) = Z(\xi, d),$$

Eq. (26) transforms to

$$d^{1-2\alpha} ((1-X)Y'' - X'Y') + \theta_0 Y' + bY(1-Y) = 0 \quad (49)$$

and from (32), (34) we get

$$X(\xi, d) = \theta_0 \frac{d^\alpha \Phi(d^\alpha \xi, d)}{\int_{d^\alpha \xi}^{\infty} \Phi(s, d) ds} \quad (50)$$

$$Z(\xi, d) = \frac{c}{r_1 - r_2} \left(e^{r_1 d^\alpha \xi} \int_{d^\alpha \xi}^{\infty} e^{-r_1 s} Y\left(\frac{s}{d^\alpha}, d\right) ds + e^{r_2 d^\alpha \xi} \int_{-\infty}^{d^\alpha \xi} e^{-r_2 s} Y\left(\frac{s}{d^\alpha}, d\right) ds \right). \quad (51)$$

The function Φ is defined in (33). The appropriate inner solution is obtained by matching X, Y, Z with the outer solution on both sides of $z = 0$. That means

$$X(\pm\infty, d) = U_{out}(0^\pm, d), Y(\pm\infty, d) = V_{out}(0^\pm, d), Z(\pm\infty, d) = W_{out}(0^\pm, d) \quad (52)$$

Remark. Eq. (49) shows very clearly that $\alpha = 1/2$ is the minimum value allowed for α , since otherwise the influence of d would be amplified when d decreases. It also emphasized that $\alpha = 1/2$ is a very special case, owing to the fact that, for d very small and $\alpha < 1/2$, (49) is basically a first order differential equation, while for $\alpha = 1/2$ it becomes similar to Fisher's equation (and identical to it if X is constant).

It is immediately realized that when $d \ll 1$ it has very little influence in (4.2). Indeed, according to (52) we can put $Y \approx 0$ in the first integral, and $Y \approx 1$ in the second, thus obtaining

$$Z(\xi, d) = \frac{1}{r_2} \frac{c}{\sqrt{\theta_0^2 d^{2\alpha} + 4c^2}}$$

so that we can neglect the term containing d to eventually find

$$Z_{in}(\xi) = \frac{1}{2}. \quad (53)$$

Note that in the argument above the coefficient θ_0 has been assumed to be not too large, in the spirit of (30). Similarly, using the same approximation in (49), we can get

$$Y_{in}(\xi) = \frac{1}{1 + e^{b\xi/\theta_0}}, \quad (54)$$

which describes (with the data for the specific problem we are dealing with) the rapid transition of tumor cells across the front. In order to find X_{in} , we go back to the distinction of the three intervals for a .

Case 1) $0 < a < 1$

From the approximation

$$U(d^\alpha \xi, d) = \begin{cases} 1 - a + \frac{a}{2} e^{\sqrt{c} d^\alpha \xi} & \text{if } \xi < 0 \\ 1 - \frac{a}{2} e^{-\sqrt{c} d^\alpha \xi} & \text{if } \xi > 0 \end{cases} \quad (55)$$

taking the limit $d \rightarrow 0$ we deduce

$$X_{in}(\xi) = 1 - \frac{a}{2}. \quad (56)$$

Case 2) $1 < a < 2$

In this case (56) still holds true, but the meaning is different. Indeed, while in the previous case the population of healthy cells never becomes extinct, the fact that X_{in} does not vanish indicates that there is an interval in which the two populations u, v overlap while they are both finite. The interval of overlapping turns out to be $(z_-, 0)$.

Case 3) $1 < a < 2$

Using the approximation

$$U(d^\alpha \xi, d) = \sqrt{\frac{\theta_0 \sqrt{c}}{2\pi}} d^{\alpha/2} \cdot \begin{cases} \exp\left(\frac{\phi_-(d^\alpha \xi) - \phi_-(z_-)}{d^\alpha}\right) & \text{if } \xi < 0 \\ \exp\left(\frac{\phi_+(d^\alpha \xi) - \phi_+(z_+)}{d^\alpha}\right) & \text{if } \xi > 0 \end{cases} \quad (57)$$

we note that for $d \rightarrow 0$ we have the limits

$$\lim_{d \rightarrow 0} \frac{\phi_\pm(d^\alpha \xi)}{d^\alpha} = \frac{(a-2)\xi}{2\theta_0}$$

and we conclude that we have simply

$$X_{in}(\xi) = 0, \quad (58)$$

indicating the existence of the gap ahead the front, already pointed out in a special case in [17]. Thus, according to our theory, the gap width is nothing but z_+ , which is not the same as the one proposed in [17], but it is very close to it in the cases considered there. It is remarkable that (49) reduces to Fisher-KPP equation when $\alpha = 1/2$. The consequent implications have been discussed in [12]. Both the estimates of the finite overlapping region ($1 < a < 2$) and of the gap agree very well with the numerical simulations illustrated in [12]. Figs. 2 and 3 show numerical simulations, taken from [12], with an overlapping or of a gap, respectively.

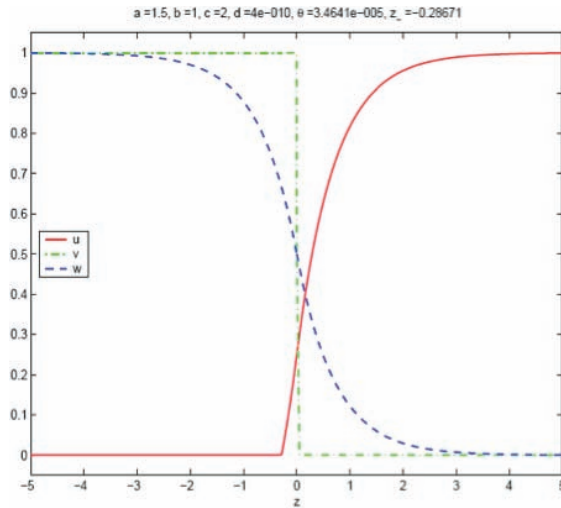


Fig. 2. Front propagation with overlapping of the two cell species ($0 < a < 2$). From [12]: A. Fasano et al. *Math. Biosci.* **220**, 45–56 (2009) © Elsevier Science

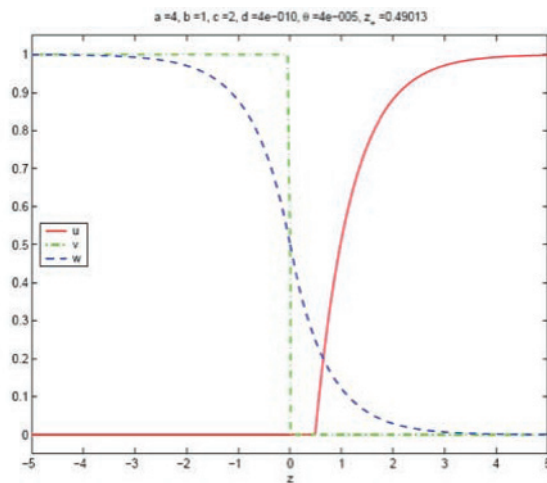


Fig. 3. Front propagation with a gap separating the two cell species ($a > 2$). From [12]: A. Fasano et al. *Math. Biosci.* **220**, 45–56 (2009) © Elsevier Science

5 Conclusion

We have reviewed very shortly some topics related with the glucose metabolism in tumors, with particular emphasis on hypoxic conditions and the consequent switch from respiratory to anaerobic pathway. The latter is accompanied by a sharp decrease of ATP production rate and an increase of acidity. The former effect may become so

severe to lead to cell apoptosis, while the latter can give an advantage for invasion to tumor cell, which are generally more tolerant to acidity than the host tissue. We have summarized the theory of [3, 4] concerning the formation of a necrotic core in a multicellular tumor spheroid, including the metabolism of glucose, oxygen consumption and the production rate of lactate (to which the pH lowering is related). The combined action of acidity and of low ATP production rate on shaping the steady state of a vascularized or an avascular tumor spheroid surrounded by a healthy host tissue has been just mentioned in our exposition, since the number of different situations that can arise is very large (see [5]). The last problem we have examined was the acidic aggression of a tumor on the surrounding host tissue, which is gradually eroded so to create a necrotic gap between the two regions, through which the tumor cell can diffuse. This advancement mechanism has been described in [17] by means of a model in which the concentrations of three components (tumor cells, healthy cells, H^+ ions) take a rigid profile travelling at a constant speed. In this paper we have briefly discussed the analysis performed in [12] about the structure of a class of travelling waves associated to the model of [17], namely the slow waves, whose speed tends to zero as some power of the cell diffusivity (a class of waves preserving a finite speed in the same limit has also been found, but we have ignored it here because of its lesser biological interest). The main feature of such slow waves is that they do exhibit the gap between the two kinds of populations, provided the effect of acidity on healthy cells is strong enough. In the opposite case an interval in which the two populations overlap is present. The width of the two intervals can be calculated. The structure of the asymptotic tails of the two populations depend on the wave speed. The wave of main physical interest is the slowest possible one, in the same spirit as the classical theory of [13] and [20].

Acknowledgements The author is grateful to the editors of this volume for their kind invitation to write the paper.

References

1. Astanin, S., Tosin, A.: Mathematical model of tumor cord growth along the source of nutrient. *Math. Model. Nat. Phenom.* **2**, 153–177 (2007)
2. Aubert, A., Costalat, R.: Interaction between astrocytes and neurons studied using a mathematical model of compartmentalized energy metabolism. *J. Cereb. Blood Flow Metabol.* **25**, 1476–1490 (2005)
3. Bertuzzi, A., Fasano, A., Gandolfi, A., Sinisgalli, C.: ATP production and necrosis formation in a tumor spheroid model. *Math. Model. Nat. Phenom.* **2**, 30–46 (2007)
4. Bertuzzi, A., Fasano, A., Gandolfi, A., Sinisgalli, C.: Necrotic core in EMT6/Ro tumor spheroids: is it caused by an ATP deficit? *J. Theor. Biol.* **262**, 142–150 (2010)
5. Bianchini, L., Fasano, A.: A model combining acid-mediated tumor invasion and nutrient dynamics. *Nonlinear Anal. Real World Appl.* **10**, 1955–1975 (2009)
6. Carpenter, L., Halestrap, A.P.: The kinetics, substrate and inhibitor specificity of the lactate transporter of Ehrlich-Lette tumor cells studied with the intracellular pH indicator BCECF. *Biochem. J.* **304**, 751–760 (1994)

7. Casciari, J.J., Sotirchos, S.V., Sutherland, R.M.: Mathematical modelling of microenvironment and growth in EMT6/Ro multicellular tumor spheroids. *Cell Prolif.* **25**, 1–22 (1992)
8. Casciari, J.J., Sotirchos, S.V., Sutherland, R.M.: Variations in tumor cell growth rates and metabolism with oxygen concentration, glucose concentration, and extracellular pH. *J. Cell. Physiol.* **151**, 386–394 (1992)
9. Dairkee, S.H., Deng, G., Stampfer, M.R., Waldman, F.M., Smith, H.S.: Selective cell culture of primary breast carcinoma. *Cancer Res.* **55**(12), 2516–2519 (1995)
10. Dimmer, K.S., Friedrich, B., Lang, F., Deitmer, W., Braer, S.: The low-affinity monocarboxylate transporter MCT4 is adapted to the export of lactate in highly glycolytic cells. *Biochem. J.* **350**, 219–227 (2000)
11. Fasano, A., Gandolfi, G.: The steady state of multicellular tumor spheroids: a modelling challenge. In Friedman A., Kashdan E., Ledzewicz U., Schaettler H. (eds.) *Mathematical Methods and Models in Biomedicine*, Springer, New York (in press)
12. Fasano, A., Herrero, M.A., Rocha Rodrigo, M.: Slow and fast invasion waves in a model of acid-mediated tumor growth. *Math. Biosci.* **220**, 45–56 (2009)
13. Fisher, R.A.: The wave of advance of advantageous genes, *Ann. Eugenics* **7**, 353–369 (1937)
14. Folkman, J., Hochberg, M.: Self-regulation of growth in three dimensions. *J. Exp. Med.* **138**, 745–753 (1973)
15. Freyer, J.P., Sutherland, R.M.: A reduction in the in situ rates of oxygen and glucose consumption of cells in EMT6/Ro spheroids during growth. *J. Cell. Physiol.* **124**, 516–524 (1985)
16. Freyer, J.P., Sutherland, R.M.: Regulation of growth saturation and development of necrosis in EMT6/Ro multicellular spheroids by the glucose and oxygen supply. *Cancer Res.* **46**, 3504–3512 (1986)
17. Gatenby, R.A., Gawlinski E.T.: A reaction-diffusion model of cancer invasion, *Cancer Res.*, **56**, 5745–5753 (1996)
18. Gerlee, P., Anderson, A.R.: A hybrid cellular automaton model of clonal evolution in cancer: the emergence of the glycolytic phenotype. *J. Theor. Biol.* **250**, 705–722 (2008)
19. Hinkle, P.C., Kumar, M.A., Resetar, A., Harris, D.L.: Mechanistic stoichiometry of mitochondrial oxidative phosphorylation. *Biochemistry* **30**, 3576–3582 (1991)
20. Kolmogorov, A.A., Petrovsky, I.G., Piskunov, N.S.: Study of the diffusion equation with growth of the quantity of matter and its application to a biology problem, *Acta Physiokimica URSS* **9**, 105–130 (1938)
21. Olsen, L., Sherrat, J.A., Maini, Ph.K., Arnold, F.: A mathematical model for the capillary endothelial cell-extracellular matrix interaction in wound-healing angiogenesis. *IMA J. Math. Appl. Med. Biol.* **14**, 261–281 (1997)
22. Poole, R.C., Halestrap, A.P.: Transport of lactate and other monocarboxylates across mammalian plasma membranes. *Am. J. Physiol.* **264**, C761–C782 (1993)
23. Schroeder, T., Yuan, H., Viglianti, B.L., Peltz, C., Asopa, S., Vujaskovic, Z., Dewhirst, M.W.: Spatial heterogeneity and oxygen dependence of glucose consumption in R3230Ac and fibrosarcomas of the Fischer 344 rat. *Cancer Res.* **65**, 5163–5171 (2005)
24. Smallbone, K., Gatenby, R.A., Gillies, R.J., Maini, P.K., Gavaghan, D.J.: Metabolic changes during carcinogenesis: potential impact on invasiveness. *J. Theor. Biol.* **244**, 703–713 (2007)
25. Smallbone, K., Gavaghan D.J., Gatenby R.A., Maini P.K.: The role of acidity in solid tumor growth and invasion, *Br. J. Radiol.* **76**, S11 (2005)
26. Spencer, T.L., Lehninger, A.L.: L-lactate transport in Ehrlich ascites-tumor cells. *Biochem. J.* **154**, 405–414 (1976)
27. Stryer, L.: *Biochemistry*. W.H. Freeman, New York, Chaps. 15–17, (1988)
28. Vaughan-Jones, R.D., Peercy, B.E., Keener, J.P., Spitzer, K.W.: Intrinsic H⁺ ion mobility in the rabbit ventricular myocyte. *J. Physiol.* **541**, 139–158 (2002)

29. Venkatasubramanian, R., Henson, M.A., Forbes, N.S.: Incorporating energy metabolism into a growth model of multicellular tumor spheroids. *J. Theor. Biol.* **242**, 440–453 (2006)
30. Warburg O.H, Posener K., Negelein, E.: Über den Stoffwechsel der Tumoren. *Biochemische Zeitschrift*, **152**, 319–344 (1924). Reprinted in English in: Warburg O.: On metabolism of tumors. Constable, London, (1930)

Cell-Cell Interactions in Solid Tumors – the Role of Cancer Stem Cells

Xuefeng Gao, J. Tyson McDonald, Lynn Hlatky, and Heiko Enderling

Abstract It is increasingly argued that solid tumors follow a cellular hierarchy comparable to normal tissues, with so-called cancer stem cells on top of the hierarchy. In this model, cancer stem cells have the unique capability to initiate and propagate solid tumors. Non-stem cancer cells will form the bulk of the tumor population, but are by themselves incapable of giving rise to a continuously growing tumor. The two distinct phenotypes interact with one another and compete for common resources such as oxygen, nutrients, or available space. Single cell kinetics are parameterized with *in vitro* data and the interplay between cancer stem cells and their non-stem cancer cell counterpart is studied using two different modeling approaches: a cellular automaton model and a cellular Potts model. Simulations of tumor growth with both techniques reveal that cancer stem cell-driven solid tumors grow as conglomerates of self-metastases, suggesting a robust biological phenomenon. The growth rate of the tumor is dependent on the complex interplay of the underlying model parameters.

1 Introduction

Tumors are populations of individual cells that escaped orchestrated tissue-level organization and execute basic cellular programs like migration and proliferation. These cells, although evading growth suppressors [25, 26], are still subject to external constraints as they compete with one another, and with host cells, for essential resources. When the local carrying capacity is exhausted, continued tumor growth may be forestalled by lack of support for expansion, resulting in a dormant state [24].

X. Gao · J.T. McDonald · L. Hlatky · H. Enderling (✉)

Center of Cancer Systems Biology, Tufts University School of Medicine, St. Elizabeth's Medical Center, 736 Cambridge Street, Boston, MA 02135, USA

e-mail: Xue-feng.Gao@tufts.edu, Tyson.McDonald@tufts.edu, Lynn.Hlatky@tufts.edu, Heiko.Enderling@tufts.edu

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_10, © Springer-Verlag Italia 2012

Two prominent examples of such microenvironment-induced tumor dormancy are vascular dependence [3, 19, 24] and immune system surveillance [10, 11, 40]. Both these means of regulating tumor dormancy were recently addressed from a systems perspective, integrating experiments with their theoretical underpinnings [15]. But beyond these vascular and immune regulations which can induce full dormancy, we find that even if growth factor supply is plentiful and host inhibition negligible, the tumor growth rate may still decelerate with tumor mass increase. In the interior of the tumor, cells experience physical stress and overwhelming inbound forces from their neighboring cells that disallow increase in cell volume, which is required for cell division. Although viable and metabolically active, a large fraction of cells in the interior of many solid tumors are forced into a quiescent state. This can be approximated by assuming a homogeneous cancer cell population, where the tumor grows radially symmetric with proliferating cells on the outer rim and quiescent cells in the interior.

But it has long been apparent that tumors are not homogeneous but rather contain a vast array of cells of different fates and kinetics. More recently, the cancer stem cell hypothesis proposed that there is a subpopulation of cells in a tumor – so-called cancer stem or tumor initiating cells – that is uniquely able to initiate and re-initiate heterogeneous tumors at a secondary site or after cytotoxic treatment [29, 43, 44, 52]. Non-stem cancer cells, although transformed, lack replicative and physiologic immortality. Therefore, non-stem cancer cells have only a finite lifespan and proliferative capacity. Pedantically, the cancer stem cell hypothesis is in some sense also a “cancer non-stem cell hypothesis”. If only a minority of cells in a tumor are putative cancer stem cells, as it is often reported [50], and the bulk of tumor cells lack tumor initiation potential and longevity, what are the implications for tumor growth? Although the phenotypic differences of cancer stem cells and non-stem cancer cells are appreciated conceptually, experimental and clinical validations are still in their infancy as reliable makers and experimental protocols are yet to be established. The problems of experimental validation pave the way for alternative approaches, and the interplay of heterogeneous cell types lends itself to theoretical modeling. Agent-based modeling is a computational methodology to simulate emerging properties of a dynamic system in time and space that has been widely used in a variety of fields including cancer and tumor biology [2, 4, 9, 12, 21, 27, 35, 39, 42, 48, 51]. In such models the system is represented as a collection of autonomous, decision-making cells that have a set of intrinsic state variables and predefined instructions that determine how they behave and interact with each other and their local environment. Cell properties and rules of interaction are based on empirical observations from wet-lab experiments or hypothesized biological features. The full details of cells and their interactions as the simulation progresses are visualizable, measurable, and accessible to intervention [48].

2 A Cellular Automaton Model of Cancer Stem Cell-driven Tumor Growth

Based on experimental findings [1] we assume cancer stem cells (CSC) to be immortal and have unlimited replicative potential attributable, in part, to up-regulation of telomerase that prevent telomeric DNA erosion during mitosis [32, 46]. Non-stem cancer cells (CC) have a finite replicative potential that decrements with each division and is inherited by the daughter cells. A CC dies when its replicative potential is exhausted. Different models have been proposed for cellular division fates that account for tumor heterogeneity [16, 20, 30, 31]. In the model herein we ignore fate plasticity and assume a constant cellular phenotype. CSCs divide either symmetrically which yields two CSCs, or asymmetrically which yields one CSC and one CC. Let p_s denote the frequency of symmetric division and $1 - p_s$ the frequency of asymmetric division. A CC that is derived from a CSC inherits a proliferation potential ρ_{max} . A dividing CC with proliferation potential ρ gives rise to two CC with decremented proliferation potential $\rho - 1$. Both cell types, CSC and CC, are assumed to divide with rate k . These kinetics are summarized in Fig. 1. The cell division rate is cell type dependent and must be estimated for the tumor type of interest. Glioblastoma multiforme is the most common and most aggressive primary human brain tumor, and model parameters are estimated using the U87-MG cell line derived from a human primary glioblastoma. Using live microscopy, single cells are tracked and the time between proliferation events is recorded. The average proliferation time of U87-MG cells is estimated to be 25 ± 5 hours, which is in good agreement with the literature [28]. It has been argued previously that cell migration is a pivotal factor in tumor growth [5, 16, 37]. Using single cell tracking in live microscopy the migration speed of U87-MG glioma cells is estimated to be $23 \pm 10 \mu\text{m hr}^{-1}$ (Fig. 2) again in good agreement with the literature [36, 45]. This data can be used to calibrate the model. As our *in silico* cells live on a computational square lattice with $(10 \mu\text{m})^2$ grid points, a single migration step to a neighboring location is $10 \mu\text{m}$ in a 4-cell von Neumann neighborhood, and 10 or $14 \mu\text{m}$ in an 8-cell Moore neighborhood

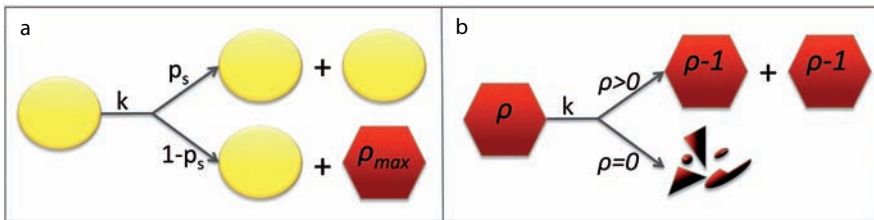


Fig. 1. Schematic of cell division fates in the model. (a) Division fate of a cancer stem cell (CSC, yellow circle). The cell divides with rate k , either symmetrically with probability p_s or asymmetrically with probability $1 - p_s$, where one daughter cell is a CC with proliferation potential ρ_{max} . (b) Division fate of a CC (red hexagon) with proliferation potential ρ . The cell divides with rate k , and produces two CC with decremented proliferation potential $\rho - 1$ if $\rho > 0$, or dies if the proliferation potential is exhausted ($\rho = 0$)

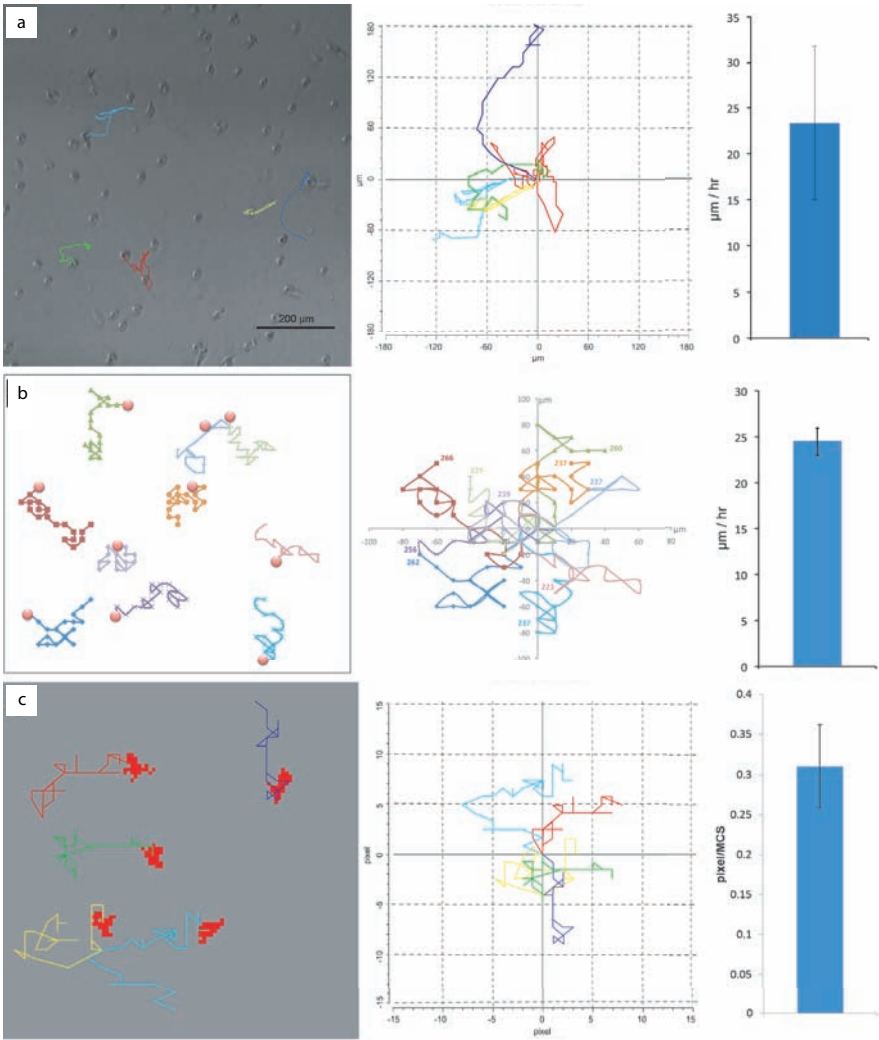


Fig. 2. (a) *In vitro* tracking and quantification of U87-MG glioma cell migration. (b) Cellular automaton model realization and quantification of cell migration for 10 cells over 10 hours. 1 *pixel* $\approx 10 \mu\text{m}$. (c) Cellular Potts model realization and quantification of cell migration for five cells for 30 Monte Carlo Steps (MCS) equivalent of 1 hour. 1 *pixel* $\approx 2.5 \mu\text{m}$

(Fig. 3) [7]. Using the 8-cell Moore neighborhood, the estimated cell displacement in a random walk at each time step is $10.7 \mu\text{m}$. With cells moving $23 \mu\text{m hr}^{-1}$, 2.15 migration events in the simulation equate one hour. Therefore, the simulation step is set to $60/2.15 = 28$ minutes. The random migration tracks of 10 cells in the cellular automaton model simulated for 10 hours demonstrates the applicability of such approach to mimic cellular behavior (Fig. 2). As alluded to above, although acting

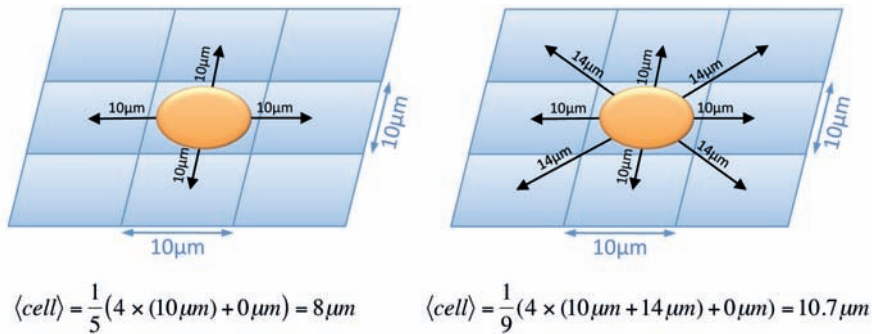


Fig. 3. Schematic of cell migration on a two-dimensional lattice in the cellular automaton model for a 4-cell (von Neumann) and 8-cell (Moore) neighborhood. The expected cell displacement including the first trivial non-movement per time step is calculated for both neighborhoods

individually, cancer cells influence one another by competing for the same environment. In early tumor growth, when oxygen and nutrients are plentiful, the competition is for space. In the cellular automaton model such competition is described by vacancy or occupation of neighboring lattice points. If a cell is completely surrounded by other cells, that is eight cells in the 8-cell Moore neighborhood, a cell is assumed to be unable to migrate or proliferate and hence forced into a quiescent state. The cell enters proliferation and migration again when any of the neighboring lattice points becomes vacant after cells migrated away or died. Tumor growth is initialized with a single cell in the center of the computational domain and monitored over time. The initial cell performs a random walk and divides. The two daughter cells also perform random walks and divide. These dynamics repeat and a large population of migrating and proliferating cells forms. A dense area in the interior of the radially symmetric population emerges where cells become quiescent, and proliferating and migrating cells are exclusively found at the population periphery where space is available. The morphology of the growing population *in silico* mimics *in vitro* tumor cluster formation (Fig. 4). The cellular automaton model can be used to compare tumor growth of homogenous populations of either CSCs or CCs (Fig. 5). If all cells are CSCs, tumor growth initiated from a single CSC is, as one would expect, radially symmetric. A dense population arises where cells in the interior become quiescent and only cells on the outer rim can proliferate and migrate. As all cells on the outer rim (as well as anywhere else in the tumor) are CSCs, the outermost cells continue to divide and expand the population. Due to the increasing quiescent region, however, tumor growth is inhibited by tumor mass [41] and progression follows decremented exponential growth [38]. On the other hand, if tumor growth is initiated from a CC with limited proliferation potential, for illustration purposes set to $\rho_{max} = 15$, the tumor cannot grow indefinitely. Initial exponential growth is followed by a short plateau phase and inevitable population shrinkage and extinction (Fig. 5). When competition for space and quiescence arises in the population, proliferation is limited to cells on the outer rim, which eventually exhaust their

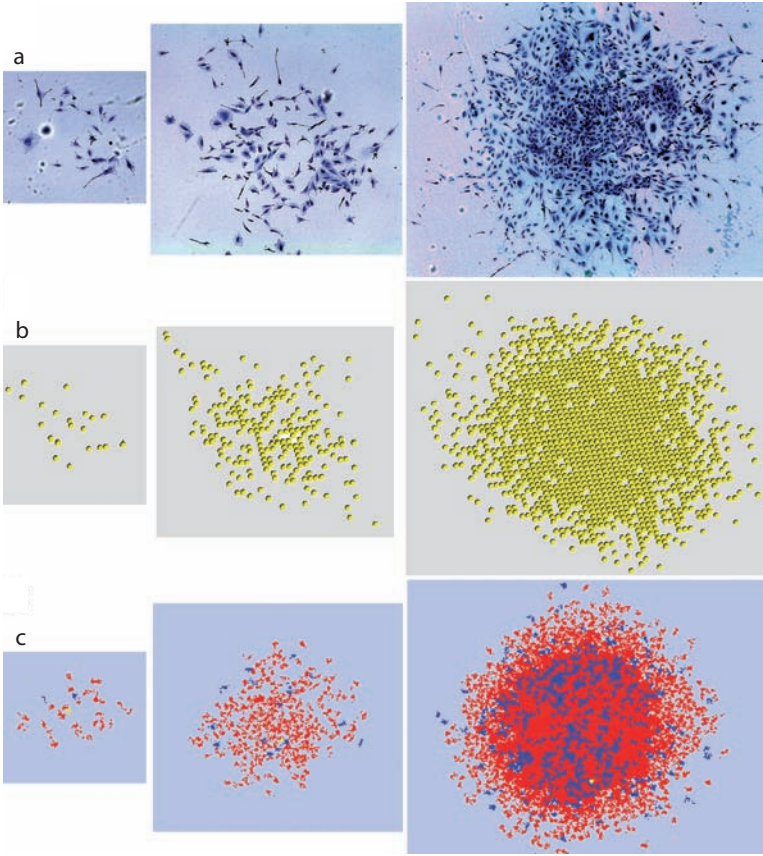


Fig. 4. (a) *In vitro* growth of a U87-MG cell population. (b) *In silico* growth of a cell population using the cellular automaton model parameterized with U87-MG data. (c) *In silico* growth of a cell population using the cellular Potts model parameterized with U87-MG data. Shown are proliferating cells in red and emergent quiescent cells in blue

proliferation capacity and die. Previously quiescent cells behind the proliferating rim then become active again until their proliferation potential is exhausted as well. The tumor cannot grow beyond a limited size defined by the proliferation potential of its initiating cell, with larger initial ρ_{max} values yielding larger population sizes. Regardless of the proliferation capacity, however, all tumors initiated by a CC will eventually shrink and die out.

The purpose of the model is to simulate tumor growth of a heterogeneous population of CSC and CC. Simulations are initiated with a single CSC in the center of the computational domain. Initiation with a CC will lead to tumor population extinction after initial growth as outlined above. To introduce heterogeneity, the symmetric division probability of CSC is set to $p_s = 0.01$ (i.e., 1%), which reflects the low frequency of CSC reported in the literature [50]. It follows that with probability

$1 - p_s = 0.99$ CSC division is asymmetric that maintains the CSC and produces a CC. Let the maximum proliferation potential of CC be $\rho_{max} = 10$. With the low p_s value the initial CSC produces predominantly CC offspring that themselves generate more CC. The CSC is quickly outnumbered by their non-stem counterparts and trapped in the population core where it is forced into quiescence. Population expansion depends on the CC located at the outer rim and, as described above, these cells will cease to divide when their proliferation potential is exhausted (i.e., $\rho = 0$). The wave of cell death on the boundary then allows for previously quiescent cells to re-enter the division cycle and divide until their proliferation potential is exhausted as well. These dynamics yield oscillations in total number until the CSC can become active again. If the following CSC divisions are asymmetric the above-described dynamics repeat until the CSC stochastically produces a new CSC via symmetric division. Then both CSC can opportunistically spatially separate and form independent tumor populations in the vicinity of one another. Both populations are then subject to the same dynamics until the next CSC arises and a third cluster is formed. The waiting time for the formation of a new CSC decreases with increasing number of CSCs and the overall tumor population grow rate accelerates (Fig. 6). The distinct morphological expansion of cancer stem cell-driven solid tumors has been termed “self-metastatic tumor progression” [16, 17, 37]. The primary tumor population is self-limited in size, and through shedding of CSC to initiate formation of new tumor cluster nearby it forms metastases to itself. The primary tumor is not metastatic because it is large; rather it is large because it is self-metastatic [37].

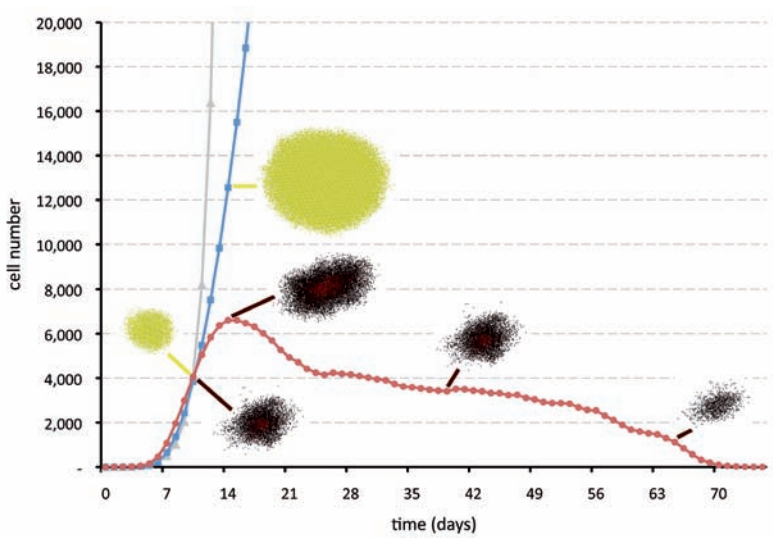


Fig. 5. Tumor growth initiated from a cancer stem cell (blue squares, yellow tumor mass) or a non-stem cancer cell with limited proliferation potential $\rho_{max} = 15$ (red circles, red to black tumor mass color-coded for cells’ remaining proliferation potential; red: large ρ value and black: low ρ value). Exponential growth curve shown for comparison (grey triangles)

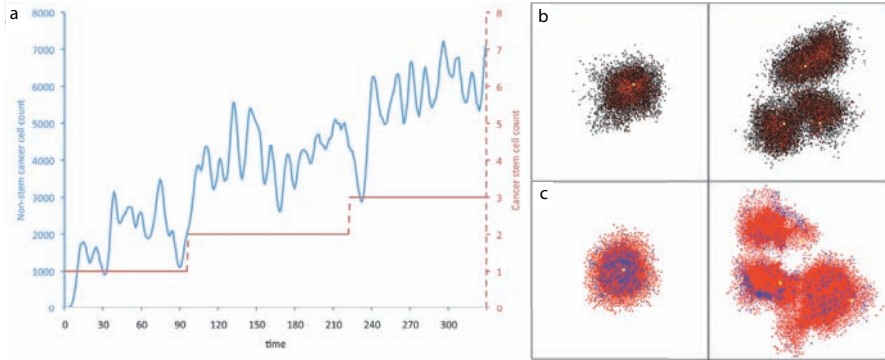


Fig. 6. Self-metastatic tumor progression for $\rho_{max} = 10$ and $p_s = 0.01$. An initial cluster of a single cancer stem cell surrounded by non-stem cancer cells evolves into multiple clusters seeded nearby the primary with similar morphologies. (a) Temporal evolution of cancer stem cell (red) and non-stem cancer cell (blue) populations. (b) Self-metastatic tumor progression in the cellular automaton model. Color coded are cancer stem cells (yellow) and non-stem cancer cells (red-to-black gradient representing large to low remaining proliferation potential, ρ). (c) Self-metastatic tumor progression in the cellular Potts model. Color-coded are cancer stem cells (yellow) and non-stem cancer cells (red). Emergent quiescent cells are shown in blue

The cellular automaton model can simulate emerging biological phenomena in early tumor growth. The applicability, however, is limited by constraints of the modeling technique. Cells live on a square lattice and quiescence is enforced on surrounded cells, without considering cells' ability to push and exert forces onto one another. The applicability of the simulation results can be increased if similar conclusions can be drawn from results of different modeling techniques that address the shortcomings of the cellular automaton approach.

3 A Cellular Potts Model of Cancer Stem Cell-Driven Tumor Growth

The cellular Potts model (CPM) (also known as Glazier-Graner-Hogeweg model) [22, 23] is a lattice-based computational modeling method to simulate multicellular systems through capturing key cellular behaviors like cell growth, division and death, intracellular adhesion, as well as volume and surface area constraints. The CPM approach facilitates simulations by defining spatially generalized cells, which are able to represent cell clusters, single cells, sub-compartments of single cells or small subdomains of non-cellular materials. Unlike the cellular automaton approach, generalized cells in CPM are spatially-extended domains residing on a collection of connected lattice sites at locations i , and sharing the same cell index σ . Cell growth, deformation, taxis, as well as cell-cell interactions including adhesion energies, volume and surface area constraints are embodied primarily in the *effective energy* E

(Hamiltonian function):

$$\begin{aligned}
 E = & \overbrace{\sum_{i,j \text{ neighbors}} J \left(\tau(\sigma(\vec{i})), \tau(\sigma(\vec{j})) \right) \left(1 - \delta(\sigma(\vec{i}), \sigma(\vec{j})) \right)}^{\text{Boundary energy}} \\
 & + \overbrace{\sum_{\sigma} \lambda_{\text{volume}}(\sigma) \left(v(\sigma) - V_{\text{target}}(\sigma) \right)^2}^{\text{Volume constraint energy}} + \overbrace{\sum_{\sigma} \lambda_{\text{surface}}(\sigma) \left(s(\sigma) - S_{\text{target}}(\sigma) \right)^2}^{\text{Surface-area constraint energy}},
 \end{aligned} \tag{1}$$

where for each cell σ , $J \left(\tau(\sigma(\vec{i})), \tau(\sigma(\vec{j})) \right)$ denotes the contact energy per unit area between neighboring cells $\vec{i}$ and $\vec{j}$; $v(\sigma)$ and $s(\sigma)$ are the volume (number of pixels in the cell) and surface area, which are constrained to be close to the target volume V_{target} and target surface area S_{target} respectively; λ_{volume} and λ_{surface} denote the inverse volume compressibility and inverse membrane compressibility of the cell; and $-2\lambda_{\text{volume}}(\sigma) \left(v(\sigma) - V_{\text{target}}(\sigma) \right)$ is the internal pressure in the cell. To model cytoskeleton-driven cell motility, the cell lattice evolves through attempts by cells to extend their boundaries into neighboring cells' lattice sites, either replacing the extra-cellular environment or displacing the neighboring cells which currently occupy those sites. The cell rearrangements evolve through relaxational Monte-Carlo-Boltzmann-Metropolis dynamics [6, 33]. For each index-copy attempt, a cell boundary pixel (source) is randomly selected and attempted to displace a randomly chosen neighboring pixel (target). The *effective energy* change ΔE of the source cell displacing the target cell at that pixel is calculated. If ΔE is negative, i.e., the change is energetically favorable, the replacement will occur. If ΔE is positive the replacement can happen with probability

$$P = e^{-\Delta E/T_m}, \tag{2}$$

where T_m denotes intrinsic cytoskeleton-driven motility. For any N sites lattice, N displacement attempts are made in each Monte Carlo Step (MCS).

We use the data derived from the U87-MG cell line to calibrate the CSC/CC kinetics model. The average size of U87-MG glioma cell is approximately 10 μm in diameter, and the initial size of a cancer cell in our CPM is 4×4 pixels. Therefore, the length scale of the simulation lattice is set to 1 *pixel* = 2.5 μm . As cell motility cannot be explicitly defined in the CPM construct, we obtained an average random displacement of 30 individual cancer cells of 0.31 ± 0.05 *pixel MCS*⁻¹ for $T_m = 50$. Equating the experimentally derived U87-MG glioma cell migration and the simulated random cell displacement yields that one *MCS* represents 2 minutes. Therefore, one hour can be simulated in 30 *MCS* (Fig. 2), and one day (the time frame of cell proliferation) is simulated in 720 *MCS*.

In CPM, cell growth can become energetically unfavorable and thus stalled for congested cells. Instead of setting a cell to be quiescent dependent on the number of cell neighbors as in the cellular automaton model, growth arrest emerges in the CPM approach with a growing population. We label growth-arrested cells as quiescent

when they failed to grow (i.e., increase the number of pixels) for a continuous period of $t = 6$ hours. The simulation of tumor growth is initialized with a single CSC in the center of the lattice domain and monitored through $t = 15$ days ($p_s = 0.01$, $\rho_{\max} = 10$). Simulating tumor growth from a single cell mimics population kinetics and morphologies observed *in vitro* and in the cellular automaton model (Fig. 3). Cells are migrating and dividing until a large population of cells emerges. Frequent transition between proliferating and quiescent states can already be observed early during tumor formation. After $t = 11$ days a population of quiescent cells emerges in the interior of the tumor mass, and proliferating and migrating cells are primarily found at the population periphery.

We introduce cellular heterogeneity into the simulation with CSC and CC. With a low symmetric division probability of $p_s = 0.01$, initial tumor expansion from a single CSC is dominated by the produced non-stem cancer cells that produce more CC offspring themselves. The CSC is quickly inhibited by the CC population and becomes growth arrested and thus quiescent. Proliferating CC are mainly located at the tumor periphery, and when their proliferation potential is exhausted (i.e., $\rho < 0$) they will die. Cell death is simulated by setting the target volume of the cell to zero (i.e., $V_{\text{target}}(\sigma) = 0$). The death of proliferating cells loosens energy pressure on previously quiescent cells that now continue growing and dividing. In accordance with the cellular automation model simulation results, the initial growth dynamics result in pseudo-stable cluster size due to CC dying and previously quiescent CC becoming proliferative again. The expansion of CSC pool depends on symmetrical division events. Once a new CSC is produced it may spatially separate from the parent CSC and seed an independent colony, which yields overall tumor population growth. Fig. 6 shows the emergence of self-metastases in the CPM model, which despite the difference in modeling technique resemble the self-metastatic progression observed in the cellular automaton approach.

4 Self-Metastatic Tumor Progression as a Function of Model Parameters

Using conceptually different modeling techniques based on identical biological assumptions indicates that self-metastatic tumor progression is a robust biological phenomenon of cancer stem cell-driven solid tumors. The speed of tumor growth and self-metastatic progression, however, is dependent on the model parameters that must be estimated depending on the tumor type of interest. Crucial model parameters like p_s and $\rho_{\max}$, however, are yet to be reliably identified by biological experiments. Exploration of the parameter space in the model allows for the following conclusions. A larger probability of symmetric CSC division yields a larger fraction of cancer stem cells and thus accelerates tumor growth [18]. Despite cancer stem cells being the engine of tumor progression [47], their activity is surprisingly modulated by the non-stem cancer cells [34]. Any population-level kinetics that spatially impede CSC slow down tumor growth and might even yield prolonged phases of

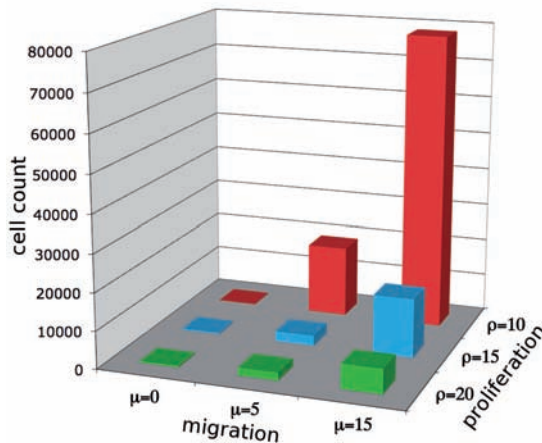


Fig. 7. Evaluation of tumor growth dependent on model parameters in the cellular automaton model ($p_s = 0.01$, time $t = 38$ months). Reproduced from [16]: Enderling et al. Br. J. Cancer **100**, 1917–1925 (2009), DOI 10.138/sj.bjc.6605071, © Nature Publishing Group

tumor dormancy [13]. Slow cell migration and, counter intuitively a large proliferation potential of non-stem cancer cells impair tumor growth [16, 34]. Similarly, kinetics that liberate CSC, such as short non-stem cancer cell lifespan or introduction of cell death in the non-stem subpopulation, promotes tumor growth [14] (Fig. 7).

5 Discussion

The cancer stem cell hypothesis for solid tumors is an attractive concept that has received great attention over the past decade. If tumors follow a cellular hierarchy with cancer stem cells (CSC) on the top, than eradication of the CSC compartment alone would be sufficient to treat the tumor [8, 18, 43]. The silver bullet against CSC, however, is yet to be identified. Whilst the properties of CSC are well articulated, their experimental validation proves to be difficult. After the era of “how to identify CSC”, the next question to address will be “how do tumors grow if they are driven by CSC?” Mathematical and theoretical models can be utilized to map this uncharted territory. By defining the kinetics of CSC and, equally important, non-stem cancer cells, their interaction can be simulated and population-level dynamics can be derived from a large number of independent stochastic simulations. The outcome of such simulations, however, is heavily dependent on the model parameters. In close dialog with experimental biologists, clinicians, and quantitatively trained researchers, e.g. computer scientists and mathematicians, we are able to derive biologically educated rules for single-cell-behavior and obtain basic cell kinetics parameters to reduce the number of uncertainties in such a model of tumor growth. We have previously reported that cell migration is the pivotal mechanism in

CSC-driven tumor progression and assumed migration rates from the literature [16]. When focusing on a specific tumor, like in our model of glioblastoma multiforme, we experimentally measured cell migration to be four to ten-fold larger than we previously assumed [16]. Despite the larger migration rate, the predicted self-metastatic growth morphology is preserved and assumed to be a robust feature of tumors initiated and driven by CSC. To strengthen this observation we utilized with cellular Potts modeling a different simulation technique [23]. Unlike the cellular automaton approach, cellular Potts model simulates cellular growth and cell-cell interaction via energy optimization functions. Cells can actively exert forces on the local environment as well as on one another, which enables, amongst other features, emergence of cellular growth arrest and quiescence. With two conceptually and mathematically distinct modeling approaches we are able to describe CSC-driven tumor growth as conglomerates of self-metastases [16, 17]. This characteristic morphology is yet to be validated experimentally, but clinical observations in inflammatory breast carcinoma, for example, are encouraging [49].

Acknowledgements The authors wish to thank Alberto d’Onofrio, Paola Cerrai, and Alberto Gandolfi for inviting us to contribute to this book, and Philip Hahnfeldt and James Glazier for their contribution to the original development of the different models that are summarized in this chapter. This project was supported by the National Cancer Institute under Award Number U54CA149233 to L.H. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

References

1. Al-Hajj, M., Wicha, M.S., Benito-Hernandez, A., Morrison, S.J., Clarke, M.F.: Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. USA* **100**, 3983–3988 (2003)
2. Alarcón, T., Byrne, H.M., Maini, P.K.: A cellular automaton model for tumour growth in inhomogeneous environment. *J. Theor. Biol.* **225**, 257–274 (2003)
3. Almog, N., Ma, L., Raychowdhury, R., Schwager, C., Erber, R., Short, S., Hlatky, L., Vajkoczy, P., Huber, P.E., Folkman, J., Abdollahi, A.: Transcriptional switch of dormant tumors to fast-growing angiogenic phenotype. *Cancer Res.* **69**, 836–844 (2009)
4. Anderson, A.R.A.: A hybrid mathematical model of solid tumour invasion: the importance of cell adhesion. *Math. Med. Biol.* **22**, 163–186 (2005)
5. Basanta, D., Hatzikirou, H., Deutsch, A.: Studying the emergence of invasiveness in tumours using game theory. *Eur. Phys. J. B* **63**, 393–397 (2008)
6. Cipra, B.: An introduction to the Ising model. *Am. Math. Monthly* **94**, 937–959 (1987)
7. Dewri, R., Chakraborti, N.: (2005) Simulating recrystallization through cellular automata and genetic algorithms. *Modelling Simul. Mater. Sci. Eng.* **13**, 173–183
8. Dingli, D., Michor, F.: Successful therapy must eradicate cancer stem cells. *Stem Cells* **24**, 2603–2610 (2006)
9. Dionysiou, D.D., Stamatakis, G.S., Uzunoglu, N.K., Nikita, K.S., Marioli, A.: A four-dimensional simulation model of tumour response to radiotherapy in vivo: parametric validation considering radiosensitivity, genetic profile and fractionation. *J. Theor. Biol.* **230**, 1–20 (2004)
10. Dunn, G.P., Bruce, A.T., Ikeda, H., Old, L.J., Schreiber, R.D.: Cancer immunoediting: from immunosurveillance to tumor escape. *Nat. Immunol.* **3**, 991–998 (2002)

11. Dunn, G.P., Old, L.J., Schreiber, R.D.: The three Es of cancer immunoediting. *Annu. Rev. Immunol.* **22**, 329–360 (2004)
12. Enderling, H., Alexander, N.R., Clark, E.S., Branch, K.M., Estrada, L., Croke, C., Jourquin, J., Lobdell, N., Zaman, M.H., Guelcher, S.A., Anderson, A.R., Weaver, A.M.: Dependence of invadopodia function on collagen fiber spacing and cross-linking: computational modeling and experimental evidence. *Biophys. J.* **95**, 2203–2218 (2008)
13. Enderling, H., Anderson, A.R., Chaplain, M.A., Beheshti, A., Hlatky, L., Hahnfeldt, P.: Paradoxical dependencies of tumor dormancy and progression on basic cell kinetics. *Cancer Res.* **69**, 8814–8821 (2009)
14. Enderling, H., Hahnfeldt, P.: Cancer stem cells in solid tumors: Is evading apoptosis a hallmark of cancer? *Prog. Biophys. Mol. Biol.* **106**, 391–399 (2011)
15. Enderling, H., Hahnfeldt, P., Hlatky, L., Almog, N.: Systems Biology of tumor dormancy: linking biology and mathematics on multiple scales to improve cancer therapy. *Cancer Res.* **71**, 2172–2175 (2012)
16. Enderling, H., Hlatky, L., Hahnfeldt, P.: Migration rules: tumours are conglomerates of self-metastases. *Br. J. Cancer* **100**, 1917–1925 (2009)
17. Enderling, H., Hlatky, L., Hahnfeldt, P.: Tumor morphological evolution: directed migration and gain and loss of the self-metastatic phenotype. *Biol. Direct* **5**, 23 (2010)
18. Enderling, H., Park, D., Hlatky, L., Hahnfeldt, P.: The importance of spatial distribution of stemness and proliferation state in determining tumor radioresponse. *Math. Model. Nat. Phenom.* **4**, 117–133 (2009)
19. Folkman, J.: Tumor angiogenesis: therapeutic implications. *New Engl. J. Med.* **285**, 1182–1186 (1971)
20. Ganguly, R., Puri, I.K.: Mathematical model for the cancer stem cell hypothesis. *Cell Prolif.* **39**, 3–14 (2006)
21. Gerlee, P., Anderson, A.R.A.: An evolutionary hybrid cellular automaton model of solid tumour growth. *J. Theor. Biol.* **246**, 583–603 (2007)
22. Glazier, J., Balter, A.: Magnetization to morphogenesis: A brief history of the Glazier-Graner-Hogeweg model. In: Anderson, A.R.A., Chaplain, M.A.J., Rejniak, K.A. (eds) *Single-Cell-Based Models in Biology and Medicine*. Birkhauser, Basel (2007)
23. Graner, F., Glazier, J.: Simulation of biological cell sorting using a two-dimensional extended Potts model. *Phys. Rev. Lett.* **69**, 2013–2016 (1992)
24. Hahnfeldt, P., Panigrahy D., Folkman J, Hlatky L.: Tumor development under angiogenic signaling: a dynamical theory of tumor growth, treatment response, and postvascular dormancy. *Cancer Res.* **59**, 4770–4775 (1999)
25. Hanahan, D., Weinberg, R.A.: The hallmarks of cancer. *Cell* **100**, 57–70 (2000)
26. Hanahan, D., Weinberg, R.A.: Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011)
27. Hatzikirou, H., Basanta, D., Simon, M., Schaller, K., Deutsch, A.: “Go or Grow”: the key to the emergence of invasion in tumour progression? *Math. Med. Biol.* **29**, 49–65 (2010)
28. Hegedüs, B., Czirók, A., Fazekas, I., B’abel, T., Madar’asz, E., Vicsek, T.: Locomotion and proliferation of glioblastoma cells in vitro: statistical evaluation of videomicroscopic observations. *J. Neurosurg.* **92**, 428–434 (2000)
29. Hermann, P.C., Bhaskar, S., Cioffi, M., Heeschen, C.: Cancer stem cells in solid tumors. *Semin. Cancer Biol.* **20**, 77–84 (2010)
30. Leder, K., Holland, E.C., Michor, F.: The therapeutic implications of plasticity of the cancer stem cell phenotype. *PLoS ONE* **5**, e14366 (2010)
31. Marciniak-Czochra, A., Stiehl, T., Ho, A.D., Jaeger, W., Wagner, W.: Modeling of asymmetric cell division in hematopoietic stem cells—regulation of self-renewal is essential for efficient repopulation. *Stem Cells Dev.* **18** 377–385 (2009)
32. Marian, C.O., Wright, W.E., Shay, J.W.: The effects of telomerase inhibition on prostate tumor-initiating cells. *Int. J. Cancer* **127**, 321–331 (2010)
33. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953)

34. Morton, C.I., Hlatky, L., Hahnfeldt, P., Enderling, H.: Non-stem cancer cell kinetics modulate solid tumor progression. *Theor. Biol. Med. Model.* **8**, 48 (2011)
35. Mukhopadhyay, R., Costes, S.V., Bazarov, A.V., Hines, W.C., Barcellos-Hoff, M.H., Yaswen, P.: Promotion of variant human mammary epithelial cell outgrowth by ionizing radiation: an agent-based model supported by in vitro studies. *Breast Cancer Res.* **12**, R11 (2010)
36. Nakada, M., Anderson, E.M., Demuth, T., Nakada, S., Reavie, L.B., Drake, K.L., Hoelzinger, D.B., Berens, M.E.: The phosphorylation of ephrin-B2 ligand promotes glioma cell migration and invasion. *Int. J. Cancer* **126** 1155–1165 (2010)
37. Norton, L.: Conceptual and practical implications of breast tissue geometry: toward a more effective, less toxic therapy. *The Oncologist* **10**, 370–381 (2005)
38. Norton, L.: Cancer stem cells, self-seeding, and decremented exponential growth: theoretical and clinical implications. *Breast Dis.* **29**, 27–36 (2008)
39. Piotrowska, M.J., Angus, S.D.: A quantitative cellular automaton model of in vitro multicellular spheroid tumour growth. *J. Theor. Biol.* **258**, 165–178 (2009)
40. Prehn, R.T.: Immunomodulation of tumor growth. *Am. J. Pathol.* **77**, 119–122 (1974)
41. Prehn, R.T.: The inhibition of tumor growth by tumor mass. *Cancer Res.* **51**, 2–4 (1991)
42. Rejniak, K.A., Anderson, A.R.A.: A computational study of the development of epithelial acini: I. Sufficient conditions for the formation of a hollow structure. *Bull. Math. Biol.* **70**, 677–712 (2008)
43. Reya, T., Morrison, S.J., Clarke, M.F., Weissman, I.L.: Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001)
44. Rich, J.N.: Cancer stem cells in radiation resistance. *Cancer Res.* **67**, 8980–8984 (2007)
45. Sabari, J., Lax, D., Connors, D., Brotman, I., Mindrebo, E., Butler, C., Entersz, I., Jia, D., Foty, R.A.: Fibronectin matrix assembly suppresses dispersal of glioblastoma cells. *PLoS ONE* **6**, e24810 (2011)
46. Shay, J.W., Wright, W.E.: Telomeres and telomerase in normal and cancer stem cells. *FEBS Lett.* **584**, 3819–3825 (2010)
47. Solé, R.V., Rodríguez-Caso, C., Deisboeck, T.S., Saldaña, J.: Cancer stem cells as the engine of unstable tumor progression. *J. Theor. Biol.* **253**, 629–637 (2008)
48. Tang, J., Enderling, H., Becker-Weimann, S., Pham, C., Polyzos, A., Chen, C.Y., Costes, S.V.: Phenotypic transition maps of 3D breast acini obtained by imaging-guided agent-based modeling. *Integr. Biol. (Camb)* **3**, 408–421 (2011)
49. Vermeulen, P.B., van Laere, S.J., Dirix, L.Y.: Inflammatory breast carcinoma as a model of accelerated self-metastatic expansion by intravascular growth. *Br. J. Cancer* **101**, 1028–1029, author reply 1030 (2009)
50. Visvader, J.E., Lindeman, G.J.: Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat. Rev. Cancer* **8**, 755–768 (2008)
51. Wang, Z., Zhang, L., Sagotsky, J., Deisboeck, T.S.: Simulating non-small cell lung cancer with a multiscale agent-based model. *Theor. Biol. Med. Model.* **4**, 50 (2007)
52. Wicha, M.S., Liu, S., Dontu, G.: Cancer stem cells: an old idea—a paradigm shift. *Cancer Res.* **66**, 1883–1890, discussion 1895–1896 (2006)

Hybrid Cellular Potts Model for Solid Tumor Growth

Marco Scianna and Luigi Preziosi

Abstract We present a hybrid computational framework, the aim of which is to reproduce and analyze the early growth of a solid tumor. The model couples an extended version of the discrete Cellular Potts Model, used to represent the phenomenological behavior of malignant cells, with a continuous approach of reaction-diffusion equations, employed to describe the evolution of microscopic variables, as the growth factors and the matrix proteins present in the host tissue and the proteolytic enzymes secreted by the tumor. The behavior of each cancer cell is determined by a balance of interaction forces, such as homotypic (cell-cell) and heterotypic (cell-matrix) adhesions and haptotaxis, and is mediated by its molecular state, which regulates the motility and proliferation rate. The resulting model captures the different phases of the development of the tumor mass, i.e. its exponential growth and the subsequent stabilization in a steady-state due to limitations in vital molecules. The proposed approach also predicts the influence on the cancer morphology of changes in specific intercellular adhesive mechanisms.

1 Introduction

Carcinogenesis is a multistage process thought to arise from a genetic mutation or a epigenetic alteration, localized in one or in small number of cells (which are typically somatic stem cells, as recently suggest in [36]) and able to escape from DNA repair mechanisms [43]. The damaged individuals acquire over time further malfunctions, that allow them to ignore growth-inhibiting signals from the neighbors or from the environment, and therefore to form a sphere-like hyper-proliferative colony [2].

M. Scianna · L. Preziosi (✉)

Department of Mathematics, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Turin, Italy

e-mail: marcosci1@alice.it, luigi.preziosi@polito.it

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_11, © Springer-Verlag Italia 2012

Once this primary tumor has established a foothold in the host tissue, it undergoes a first period of growth *in situ*, during which it has neither a vascular system of its own, nor a direct access to the existing vasculature. Indeed, the cancer mass has to rely upon the host for crucial substances, such as oxygen, nutrients, and growth factors, which are supplied by diffusion from the local environment [2, 41, 64]. Substrate gradients therefore form within the growing tumor, causing a well-localized phenotypic specification of malignant individuals, which typically differentiate in an outer viable rim of highly metabolic and proliferative cells, an interior band of quiescent individuals and a central necrotic core, formed by cells dead for deprivation of vital chemicals [9]. In particular, in this stage, the increment in cell mass due to the aberrant mitosis of external cells is only able to balance the loss of volume of the interior necrotic region. This limits the size of the tumor spheroid, which reaches a well-characterized steady state [19, 45]. Indeed, in order to grow further, the cancer mass will have to coopt the existing vasculature, or acquire a new one. During these early phases of cancer development, the local microenvironment is completely modified, as the activity of tumor proteolytic enzymes (such as Matrix Metalloproteinases) degrades matrix components, both providing a space in which malignant individuals can move and a gradient which can direct their motion via haptotactic mechanisms [42, 59].

The above-described diffusional limitations of growth can be overcome by the carcinoma with morphological instabilities which result from many factors, including the downregulation of intercellular adhesion, specific mechanical stresses and/or the enhancement in cell motility, and increment the surface area of its interface with the host, thereby allowing internal individuals to have greater access to vital molecules. Understanding the mechanisms able to control the morphological evolution of the malignant mass is therefore important for controlling also its invasive ability. This topic has been indeed addressed in the last decades by a variety of theoretical approaches, which include cellular automata and agent-based models, and continuous single or multiphase mixture models, as pointed out by the books [36, 45] and by the excellent reviews [3, 7, 12, 33, 46, 62].

We here present a hybrid computational framework, which aims to reproduce the early development of a multicellular tumor spheroid. In particular, our modeling environment is based on the extensions of the Cellular Potts Model, a lattice-based Monte Carlo technique employing a stochastic energy minimization [4, 28–30, 37], recently described in [52]. Indeed, the proposed model couples the standard discrete CPM for the phenomenological behavior of the malignant mass with reaction-diffusion (RD) systems for the evolution of molecular microenvironmental variables. In particular, we integrate and interface the two model components by linking, with specific constitutive laws, variations in the cell phenotypic behavior (i.e., motility and proliferation rate) with changes in their microscopic chemical conditions (i.e., intracellular level of growth factors). The resulting model, accounting for both cell-based mesoscopic processes and for subcellular dynamics, is able to reproduce the dynamical evolution of the malignant spheroid, characterizing its morphology and quantifying its malignancy in terms of invasive radius. Moreover, our approach describes the dependence of tumor invasiveness on intercellular contact interactions,

predicting critical adhesive conditions separating a compact, non-invasive growth from an unstable infiltrative progression.

2 Mathematical Model

We elaborate a hybrid, multilevel modeling framework for solid tumor invasive growth, which is based on the extensions of the Cellular Potts Model reviewed in [52]. The aim is a simulation environment in which the cellular and multicellular behavior autonomously emerges, rather than being imposed *a priori* by a set of phenomenological rules. In particular, the malignant cells are represented by physical discrete objects, that locally interact with each other and with the microenvironment through their membranes. The molecular biology is instead incorporated with a macroscopic description of the evolution of nutrients, ECM proteins and tumor matrix metalloproteinases, see Fig. 1 for a schematic representation. As a key feature of our discrete-continuum composite approach, the different scales affect each other, as the distribution of nutrients and ECM proteins in the extracellular environment influence cells' properties and phenomenology (with carefully-calibrated constitutive relations).

2.1 Extended Cellular Potts Model

The invasion of the tumor mass is modeled at the mesoscopic level using an extended Cellular Potts Model, a grid-based stochastic approach, which realistically preserves the identity of single individuals and describes their behavior and mutual interac-

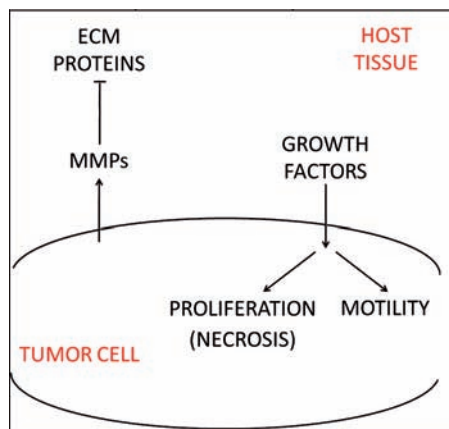


Fig. 1. Schematic diagram of the key processes included in the mathematical model. The extracellular environment provides to tumor cells soluble growth factors, whose uptake enhances cell intrinsic motility and proliferation. If the intracellular level of growth factor drops below a certain threshold, malignant individuals enter in the necrotic state. Tumor cells, in turn, secrete matrix metalloproteinases (MMPs), which degrade ECM proteins

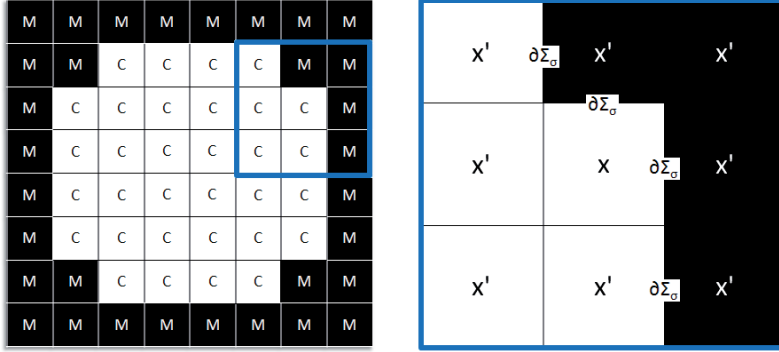


Fig. 2. For representative purposes, section of a simulated tumor cell Σ_σ of type C . The extracellular matrix is labeled with $\tau = M$, and is colored in black. In the zoom view it is possible to see a sub-plasmamembrane lattice site, $\mathbf{x}$, and its first-nearest neighbors, $\mathbf{x}' \in \Omega'_x$

tions in energetic terms and constraints. The simulation domain is a lattice $\Omega \subset \mathbb{R}^3$, formed by identical closed grid sites that, with an abuse of notation, will be identified by their center $\mathbf{x} \in \mathbb{R}^3$. Each grid site is labeled by an integer number, $\sigma(\mathbf{x}) \in \mathbb{N}$, that can be interpreted as a degenerate *spin* originally coming from statistical physics [35, 44]. The border of $\mathbf{x}$ is identified as $\partial\mathbf{x}$, one of its neighbors by $\mathbf{x}'$, while its overall neighborhood by Ω'_x , i.e. $\Omega'_x = \{\mathbf{x}' \in \Omega : \mathbf{x}' \text{ is a neighbor of } \mathbf{x}\}$. A cancer cell, denoted by Σ_σ , consists of a subdomain of contiguous sites with identical spin σ (i.e., $\Sigma_\sigma = \{\mathbf{x} \in \Omega : \sigma(\mathbf{x}) = \sigma\}$), and has an associated type $\tau(\Sigma_\sigma) = C$. Its plasma-membrane (PM) is defined as $\partial\Sigma_\sigma = \bigcup_{\mathbf{x} \in \Sigma_\sigma, \mathbf{x}' \in \Omega'_x : \mathbf{x}' \notin \Sigma_\sigma} (\partial\mathbf{x} \cap \partial\mathbf{x}')$ (see

Fig. 2). The tumor mass resides in an extracellular matrix, a special generalized cell of type $\tau = M$, which is assumed to be isotropically distributed throughout the simulation domain, forming no large-scale structures. It reproduces in fact the mixture of soluble components (among others, long carbohydrate polymers, and non-proteoglycan polysaccharides), which, together with the water solvent, compose the so-called interstitial medium.

Changes in cell configuration are governed by an iterative and stochastic reduction of the free energy of the overall system, given by the *Hamiltonian* H , whose expression will be clarified below. The core algorithm is a modified Metropolis method for Monte Carlo dynamics [30, 39], which is able to implement the natural exploratory behavior of cells, reproducing their cytoskeletally-driven membrane fluctuations and extensions of pseudopods. Procedurally, at each time step t , called Monte Carlo Step (MCS, the basic unit of time of the discrete model), a lattice site, $\mathbf{x}_{source}$ belonging to cell Σ_σ , is selected at random and assigns its spin, $\sigma(\mathbf{x}_{source})$, to one of its unlike neighbors, $\mathbf{x}_{target} \in \Omega'_{\mathbf{x}_{source}} : \mathbf{x}_{target} \notin \Sigma_\sigma$, also randomly selected. The net energy difference due to the proposed change of domain configuration, $\Delta H|_{\sigma(\mathbf{x}_{source}) \rightarrow \sigma(\mathbf{x}_{target})} = H_{(after \text{ spin copy})} - H_{(before \text{ spin copy})}$, is then evaluated.

The trial spin update is finally accepted with a Boltzmann-like probability function:

$$P(\sigma(\mathbf{x}_{source}) \rightarrow \sigma(\mathbf{x}'_{target}))(t) = \tanh(\varepsilon T_{\Sigma_{\sigma(\mathbf{x}_{source})}}(t)) \min\{1, e^{-\Delta H/T_{\Sigma_{\sigma(\mathbf{x}_{source})}}(t)}\}, \quad (1)$$

where $T_{\Sigma_{\sigma(\mathbf{x}_{source})}}(t) \in \mathbb{R}_+$ is a Boltzmann temperature and measures the agitation rate of the moving individual $\Sigma_{\sigma(\mathbf{x}_{source})}$. The specific form of (1) is a modification of the Boltzmann probability function, used in published CPM applications and recovered in the limit $\varepsilon \rightarrow \infty$. As commented in [52], the classical law has a significant weakness in the fact that, for non positive energy gradients, each cell Σ_{σ} is going to move, regardless of its intrinsic motility, defined by $T_{\Sigma_{\sigma}}$. This lacks of biological realism: for example a frozen individual, pretreated by cytochalasin B or held at 4 °C, has a null probability to move in all situations. In order to address this issue (i.e., to take into account the cell agitation rate also in the case of energetically unfavorable displacement attempts), we propose to multiply the standard Boltzmann law by a sort of *maximal transition probability* function, which depends on $T_{\Sigma_{\sigma}}$ and, in particular, is zero for $T_{\Sigma_{\sigma}}=0$ and goes to one in the case of high values of $T_{\Sigma_{\sigma}}$. Among the wide range of laws that satisfy these characteristics, and that therefore can be chosen, we opt to use $\tanh$. Specifically, the intrinsic motility of the malignant cancer individuals, a measure of the agitation of their cytoskeleton (i.e., of the random actin polymerization/depolarization dynamics), is typically stimulated by the intracellular level of growth factors, as provided by experimental observations through classical wound healing assays with different tumor cell lines (see for example [20, 21] for the Hepatocyte Growth Factor, [8, 60] for the Vascular Endothelial Growth Factor families, and [55] for the Fibroblast Growth Factor). Therefore, we set:

$$T_{\Sigma_{\sigma}}(t) = T_0 \left[\frac{n(\Sigma_{\sigma}, t)}{n_0(\Sigma_{\sigma}) + h(n(\Sigma_{\sigma}, t) - n_0(\Sigma_{\sigma}))} \right], \quad (2)$$

where the quantity $n(\Sigma_{\sigma}, t) = \sum_{\mathbf{x} \in \Sigma_{\sigma}} n(\mathbf{x}, t)$ is the total amount of growth factors inside cell Σ_{σ} , as $n(\mathbf{x}, t)$ is their local concentration, defined in Eq. (9), while $n_0(\Sigma_{\sigma}) = \sum_{\mathbf{x} \in \Sigma_{\sigma}} n_0$ is instead their overall basal level. T_0 therefore corresponds to the basal motility of tumor cells, while T_0/h is their asymptotic motility for saturating concentrations of chemicals. For any given time t the free energy, whose minimization, as seen, drives the system evolution, is composed of three contributions:

$$H(t) = H_{shape}(t) + H_{adhesion}(t) + H_{haptotaxis}(t). \quad (3)$$

H_{shape} models the geometrical attributes of tumor cells. They are written as relative deformations in the following form:

$$\begin{aligned} H_{shape}(t) &= H_{volume}(t) + H_{surface}(t) \\ &= \sum_{\Sigma_{\sigma}} \left[\kappa_{\Sigma_{\sigma}} \left(\frac{v_{\Sigma_{\sigma}}(t) - V_{\tau(\Sigma_{\sigma})}(t)}{v_{\Sigma_{\sigma}}(t)} \right)^2 + v_{\Sigma_{\sigma}} \left(\frac{s_{\Sigma_{\sigma}}(t) - S_{\tau(\Sigma_{\sigma})}(t)}{s_{\Sigma_{\sigma}}(t)} \right)^2 \right], \end{aligned} \quad (4)$$

where v_{Σ_σ} and s_{Σ_σ} represent actual volume and surface of cell Σ_σ , and $V_{\tau(\Sigma_\sigma)}$ and $S_{\tau(\Sigma_\sigma)}$ the relative target measures. The formulation of (4) allows to have finite energetic contributions, as well as a blow up in the case of $v_{\Sigma_\sigma}(t), s_{\Sigma_\sigma}(t) \rightarrow 0$. This means that for instance an infinite energy is needed to shrink a cell to a point. $\kappa_{\Sigma_\sigma}(t)$ and $\nu_{\Sigma_\sigma}(t)$ are mechanical moduli in units of energy. In particular, $\kappa_{\Sigma_\sigma}(t) \in \mathbb{R}_+$ refer to cell volume changes, while $\nu_{\Sigma_\sigma}(t) \in \mathbb{R}_+$ relate to their resistance to compression (i.e., to their elasticity). In the presence of an adequate intracellular level of growth factors, malignant cells keep almost the same volume during tumor growth: therefore, for any Σ_σ , $V_{\tau(\Sigma_\sigma)}(t) = V_C$ is their initial volume and $\kappa_{\Sigma_\sigma} = \kappa$ is a high constant value. On the contrary, when the internal level of growth factors drops below a certain threshold, here defined as n_l , a cell enters in an irreversible necrotic state, and starts to lose volume at a constant rate. This process, the model counterpart of the biological lysing of death cells, is modelled by setting:

$$\frac{dV_{\tau(\Sigma_\sigma)}(t)}{dt} = -k \quad (5)$$

if $n(\Sigma_\sigma, t) < n_l$. Zero volume cells are finally removed.

Moreover, given that $S_{\tau(\Sigma_\sigma)}(t) = S_C$ is the initial cell surface, because tumor cells are typically deformable, as they are able to significantly remodel to invade their surroundings more efficiently, for any individual Σ_σ , we use a sufficiently low $\nu_{\Sigma_\sigma} = \nu$.

$H_{adhesion}$, deriving from the Steinberg's Differential Adhesion Hypothesis (DAH) [30, 57, 58], measures the adhesive interactions between different cells or between a cell and the external environment:

$$H_{adhesion}(t) = \sum_{(\partial x \in \partial \Sigma_\sigma) \cap (\partial x' \in \partial \Sigma_{\sigma'})} J_{\tau(\Sigma_\sigma(x)), \tau(\Sigma_{\sigma'}(x'))}(t), \quad (6)$$

where x and x' are two neighboring sites and Σ_σ and $\Sigma_{\sigma'}$ two neighboring individuals. The coefficients $J_{\tau(\Sigma_\sigma(x)), \tau(\Sigma_{\sigma'}(x'))} \in \mathbb{R}_+$ are binding forces per unit area, describe the relative preference of tumor cells to be in contact with other cells or with matrix components, and are obviously symmetric w.r.t. the indices. In particular, $J_{C,C}$ represents the adhesive strength between the membranes of two nearby cells, a measure of the quantity of active and exposed cadherins, while $J_{C,M}$ evaluates the heterophilic adhesive bonds between the integrins on the cell surface and suitable ligands (such as laminin and fibronectin) in the extracellular matrix. By setting constant and homogeneous values for both $J_{C,C}$ and $J_{C,M}$, we assume a uniform distribution of adhesion molecules on the cell surface and a homogenous density of ligands in the external microenvironment. Finally, supposing that tumor cells prefer to adhere to one another rather than to the host (the strong homotypic interactions within solid tumor masses are typically regulated by the over-expression of E-cadherins and by other intercellular mechanisms), $J_{C,C}$ has to be lower than $J_{C,M}$. In particular, as given in Table 1, we set $J_{C,C} = 1/2 J_{C,M}$: this is an arbitrary choice, however a complete screening of the role played by the cell-cell adhesive strength in determining the tumor phenotype will be performed in the Results section (cf. Figs. 7 and 8).

$H_{haptotaxis}$ reproduces the effect of cell preferential movement in the direction of zones with higher concentrations of extracellular matrix proteins (such as fibrin, vitronectin, and some components of the collagen family) and is implemented with a local linear-type relation:

$$\Delta H_{haptotaxis} = \mu [p_t(\mathbf{x}_{target}, t) - p_t(\mathbf{x}_{source}, t)], \quad (7)$$

where $\mathbf{x}_{source}$ and $\mathbf{x}_{target}$ are, respectively, the source and the final lattice site randomly selected during a trial update in a MCS, see (1), and $p_t(\mathbf{x}, t) = p(\mathbf{x}, t) + \sum_{\mathbf{x}' \in \Omega'_t} p(\mathbf{x}', t)$, where $\mathbf{x} \in \{\mathbf{x}_{source}, \mathbf{x}_{target}\}$, evaluates the local level of ECM proteins sensed by the moving cell membrane site, as $p(\mathbf{x}, t)$ is their amount at site $\mathbf{x}$ (evolving according to Eq. (10) below). Finally, $\mu \in \mathbb{R}_+$ represents the local strength of haptotaxis. The term (7) is similar to that used in [51] for the chemotactic processes in Dictyostelium Discoideum aggregation.

In addition to the rules enforced by the *Hamiltonian*, we also take into account of the cell proliferation. In particular, as suggested by a number of experimental works as early as [54], the time between cell divisions has a stochastic distribution, which depends on both the internal state of each cell and on the time from its last mitotic process (i.e., except for extremely rapidly dividing cells, those who recently divided will be still growing and the likelihood of their reentering the S phase is extremely small, see [2]). For each cell Σ_σ , we therefore define its present probability P^{Σ_σ} to undergo mitosis with the following functional form, which resembles that used in a similar approach [63]:

$$P^{\Sigma_\sigma}(t) = \begin{cases} 0 & \text{if } (t - t^{\Sigma_\sigma}) \leq t_0; \\ \left[\frac{n(\Sigma_\sigma, t) - n_0(\Sigma_\sigma)}{n(\Sigma_\sigma, t)} \right]^+ \frac{(t - t^{\Sigma_\sigma})^2}{(t - t^{\Sigma_\sigma})^2 + 1} & \text{if } (t - t^{\Sigma_\sigma}) > t_0, \end{cases} \quad (8)$$

where, as already seen, $n(\Sigma_\sigma, t)$ and $n_0(\Sigma_\sigma)$ are, respectively, the actual and initial intracellular amount of growth factors, and $[\cdot]^+$ is the positive part of the function. Indeed, t^{Σ_σ} is the last time that cell Σ_σ underwent mitosis, while t_0 is a dormant period during which the cells are prohibited to proliferate (i.e., it correspond to the G1 phase, during which the cells are metabolically active and grow). In particular, relation (8) takes into account that, once the time since last division exceeds t_0 , the probability of another mitosis slowly increases, approaching 1. The duration of the G1 phase is a characteristic of each cell, that depends on its type and on intra- and extracellular factors (for example, the conditions in which it grows or the signals it receives from the external environment [2]). However, the duration of such a dormant period is usually assumed constant for all individuals of the same mass, even though some variations can occur. In particular, we set t_0 equal to 7 hours: this choice is a compromise between very slow growing and very aggressive tumor lines, and is consistent with the mean duration of the G1 phase measured in different types of gliomas, see [25, 38] and references therein. Indeed, the specific mitotic rate of an individual is mostly due to the duration of the other phases of its cell cycle, which are

more closely dependent on the intracellular quantity of growth factors, as taken into account by the first factor in (8) and widely demonstrated in literature with mitogenic assays (refer for example to [40, 55, 66]). Procedurally, the mechanism of cell division is implemented by dividing the proliferating individual into two identical daughter cells, with a volume halved with respect to their parent: however, due to the shape constraints in Eq. (4), they will gradually "mature" into full-size cells. We further assume that both daughter cells evenly inherit all the parent's biophysical properties (i.e., its motility and adhesive properties). Finally, the newly formed individuals are placed symmetrically about the parent cell center of mass with a random orientation.

2.2 Continuous Model for Microenvironmental Variables

As a matter of fact, the growth of an avascular solid tumor is strongly dictated by the availability of growth factors (and nutrients) diffusing in the extracellular environment. In particular, they are supplied to the host tissue, diffuse and decay at a constant rate, and are consumed by tumor cells. The spatial profile of environmental growth factors, $n(\mathbf{x}, t)$, therefore satisfies the following equation:

$$\begin{aligned} \frac{\partial n}{\partial t}(\mathbf{x}, t) = & \underbrace{D_n \nabla^2 n(\mathbf{x}, t)}_{\text{diffusion}} - \underbrace{\lambda_n n(\mathbf{x}, t) \delta(\tau(\Sigma_{\sigma(\mathbf{x})}), M)}_{\text{decay}} + \\ & - \underbrace{\min\{n_{\max}, \chi_n n(\mathbf{x}, t)\} \delta(\tau(\Sigma_{\sigma(\mathbf{x})}), C)}_{\text{uptake}} + \underbrace{S}_{\text{production}}, \end{aligned} \quad (9)$$

where for instance $\delta(x, y) = \{1 \text{ if } x = y; 0 \text{ if } x \neq y\}$. D_n is the characteristic diffusion coefficient, homogeneous throughout domain Ω and λ_n is the decay rate in the ECM. The third term in (9) models the local uptake by tumor cells, which follows a piecewise-linear approximation of a Michaelis-Menten law. In particular, $\chi_n \gg \lambda_n$, as we assume that the nutrient natural decay is negligible compared to the uptake by tumor cells [48]. S describes the production of chemical factors at a constant rate per unit of time at the border of the domain, as we will described in the next section.

The substrate contains matrix soluble proteins (i.e., we neglect their production, assuming a uniform distribution at the beginning of each simulation, see next section for more details), that are degraded upon contact by the metalloproteinases (MMPs) secreted by malignant cells. The change in the local amount of ECM proteins, $p(\mathbf{x}, t)$, is therefore described by:

$$\frac{\partial p}{\partial t}(\mathbf{x}, t) = - \underbrace{\chi_p p(\mathbf{x}, t) m(\mathbf{x}, t)}_{\text{degradation}}, \quad (10)$$

where χ_p is the constant degradation rate.

Tumor proteolytic enzymes are produced by malignant cells, diffuse throughout the tissue and undergo some form of decay (passive, active, or due to the neutraliza-

tion by endogenous matrix inhibitors). Their local concentration, defined as $m(\mathbf{x}, t)$, hence is governed by:

$$\frac{\partial m}{\partial t}(\mathbf{x}, t) = \underbrace{D_m \nabla^2 m(\mathbf{x}, t)}_{\text{diffusion}} - \underbrace{\lambda_m m(\mathbf{x}, t) \delta(\tau(\Sigma_{\sigma(\mathbf{x})}), M)}_{\text{decay}} + \underbrace{\pi_m \delta(\tau(\Sigma_{\sigma(\mathbf{x})}), C)}_{\text{production}}, \quad (11)$$

where λ_m is the decay rate in the ECM and D_m the diffusion coefficient, whose low value models the fact that proteolysis is strongly localized in the regions close to the cell membrane, in agreement with experimental evidence in [42]. π_m is the constant production rate of degrading enzymes, active at each cell site.

3 Simulation Details and Results

The simulation domain Ω is a square lattice of $350 \times 350 \times 350$ sites. The characteristic length of each site is $1 \mu\text{m}$, and therefore Ω represents a tissue with a volume of 0.04 mm^3 . One MCS is set to correspond to 20 sec: the overall simulations stop after 56250 MCS, so that they reproduce a time-lapse of nearly 15 days. The PDE for the evolution of nutrients is numerically solved with a finite difference scheme on a grid with the same spatial resolution as Ω , characterized by 30 diffusion steps per MCS. This temporal scale is sufficiently small to guarantee the stability of the numerical method.

As represented in Fig. 3, we start all simulations with a cluster of 24 cells in the center of the lattice. Each cell is initially a sphere with a volume of $3380 \mu\text{m}^3$, which is consistent with the average dimensions of human glioma cells [10]. The specific initial configuration reproduces an avascular solid tumor spheroid, which is invading the surrounding host tissue. The parameters describing the biophysical properties of the cells, such as their basal motility, elasticity and adhesive strengths, have been evaluated, consistently with biological considerations clarified along the text, through preliminary simulations, that showed the model consistency in a wide range of values.

The boundary conditions for the proteolytic enzymes, m , are zero flux at the border of the entire domain (i.e., $\mathbf{n}_{\partial\mathbf{x}} \cdot \nabla m(\partial\mathbf{x} \in \partial\Omega, t) = 0$, where $\mathbf{n}_{\partial\mathbf{x}}$ is the outward normal to $\partial\mathbf{x}$), assuming that these parts of the tissue are far enough from the cancer mass. The borders of the domain represent instead an extended source for the environmental growth factors, which are supposed to be supplied by the host: consequently, we set $n(\partial\mathbf{x} \in \partial\Omega, t) = \phi_n$. The intracellular basal level of chemical factors, n_0 , is set equal to $0.5 \mu\text{M}$, while its threshold for the necrotic transition to $0.2 \mu\text{M}$. These values have been experimentally measured for glioma spheroids grown from glioblastoma multiforme (GBM) cells [6, 13, 23]. The extracellular environment is initially seeded with a low level of growth factors, $n_{\text{ext},0} = 0.4 \mu\text{M}$, and with a saturating level of ECM, $p_{\text{ext},0} = 4 \mu\text{M}$. Finally, there are no secreted MMPs in the extracellular medium at the beginning of the simulations.

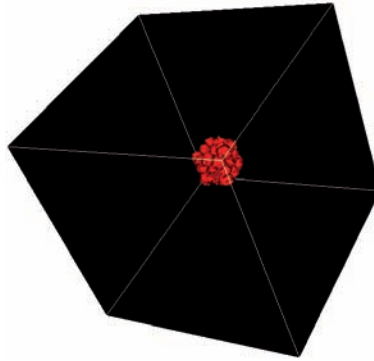


Fig. 3. Initial condition (i.e., at $t = 0$) of the tumor mass for all sets of simulations. We start with a cluster of 24 cells placed in the center of the lattice. The dimensions of cells, as well as all the other model parameters, are given in Table 1

Table 1. Model Parameters

<i>Parameter</i>	<i>Description</i>	<i>Value</i>	<i>Reference</i>
V_C	cell volume	$3380 [\mu\text{m}^3]$	[10]
S_C	cell surface	$1130 [\mu\text{m}^2]$	[10]
T_0	basal motility of cells	4.5	
ε	coefficient of Boltzmann-like probability	1	
h	Michaelis-Menten coefficient for T	1/3	
κ	cell volume compressibility	10	
k	loss rate of cell volume	1/5	
ν	cell elasticity	5	
$J_{C,C}$	cell-cell adhesive strength	2.5	
$J_{C,M}$	cell-matrix adhesive strength	5	
μ	haptotaxis strength	5	
D_n	diffusion constant of nutrients	$10^3 [\mu\text{m}^2\text{s}^{-1}]$	[22]
λ_n	on-rate constant of nutrient decay	$2 \cdot 10^{-4} [\text{s}^{-1}]$	fit to exp. [22, 23]
χ_n	on-rate constant of nutrient uptake	$1 [\text{h}^{-1}]$	fit to exp. [22, 23]
ϕ_n	on-rate constant of growth factor production	$0.78 [\text{h}^{-1}]$	fit to exp. [22, 23]
$n_{\max}$	maximal growth factor consumption	$0.08 [\text{pg/cell/h}]$	fit to exp. [24]
$n_{\text{ext},0}$	initial extracellular level of growth factors	$0.4 [\mu\text{M}]$	
n_0	basal intracellular level of growth factors	$0.5 [\mu\text{M}]$	[6, 13, 23]
n_l	threshold level for necrotic transition	$0.2 [\mu\text{M}]$	[6, 13, 23]
χ_p	on-rate constant of ECM protein degradation	$3 [\text{h}^{-1}]$	fit to exp. [42]
$p_{\text{ext},0}$	initial extracellular level of ECM protein	$4 [\mu\text{M}]$	
D_m	diffusion constant of MMPs	$5^{-4} [\mu\text{m}^2\text{s}^{-1}]$	[42]
λ_m	on-rate constant of MMP decay	$2 \cdot 10^{-3} [\text{s}^{-1}]$	fit to exp. [42]
π_m	on-rate constant of MMP production	$5 \cdot 10^{-3} [\text{s}^{-1}]$	fit to exp. [42]

To quantify the effects of the different model parameters on the tumor phenotype, we concentrate on the evolution on the invasive depth of the mass, defined as $d(t)$ and measured as the radius of the sphere circumscribing the 95 per cent of malignant cells. This choice is done to avoid biases toward accounting for few shed individuals. d has a clear clinical relevance, since it quantitatively estimates the severity of the disease and characterizes the presence of a significant number of metastasis delocalized with respect to the main body of the tumor.

We first simulate the growth of the tumor in the standard parameter setting. As reproduced in the time series of Fig. 4, the malignant mass grows exponentially for the first 9-10 days. Due to the abundance of chemical supply, external cells (i.e., those located near the tumor/host interface) can in fact freely proliferate, while, in this stage of development, the basal intracellular level of growth factors is sufficiently high for the survival of more internal individuals. As the evolution proceeds, the cancer cells near the center of the tumor begin the irreversible necrotic transition. Without exposition to the external environment, their internal amount of nutrients in fact drops below the quantity needed for viability. Indeed, a core of dead individuals, which loose volume, forms at the center of the mass. In this phase, the proliferation of external cells is only sufficient to balance the loss of mass in the central region of the tumor, whose growth therefore slows down until it reaches a diffusion-limited steady state, characterized by $d(t = 15 \text{ days}) \approx 90 \mu\text{m}$.

In the steady configuration, a small fraction of cells is escaped from the bulk of the spheroid, see again Fig. 4. The increased metabolism of these individuals, due to the availability of environmental growth factors, has in fact resulted in a dramatic increment in their motility, which has overcome the intercellular adhesive forces and the physiological regulation of the contact-inhibition of movement, inducing a sort of mesenchimal transition. These shed individuals, which are difficult to be clinically detected, have the greatest potential to invade the host and to further metastasize. They in fact display an evident ability to wonder in the close proximity and to spread in the surrounding tissue, mainly driven by haptotactic mechanisms via the local ECM gradients created by the activity of their proteolytic enzymes. The scattered individuals, evading destruction by the immune system, may subsequently enter the host bloodstream or lymphatics, extravasate at a distant site, and establish secondary colonies with devastating consequences for the wellbeing of the patient, as the likelihood of success of therapeutic interventions strongly decreases, as analyzed in [27, 49].

As reproduced in Fig. 5, the overall growth of the tumor, quantified by the evolution of both the invasive radius d and of the total number of cells, observes a Gompertzian law, which typically provides an accurate description of the early stages of avascular cancer development, where the observed saturation is due to the competition of single individuals for vital nutrients.

A specific description of the cell phenotypic differences within the spheroid is obtained by considering the distribution of the intracellular levels of growth factors, given in Fig. 6 (left panel). Cells in the more external ring of the tumor have an abundant supply of chemicals and, as seen, can freely proliferate and eventually shed in the extracellular environment, as a consequence of their increased metabolism and

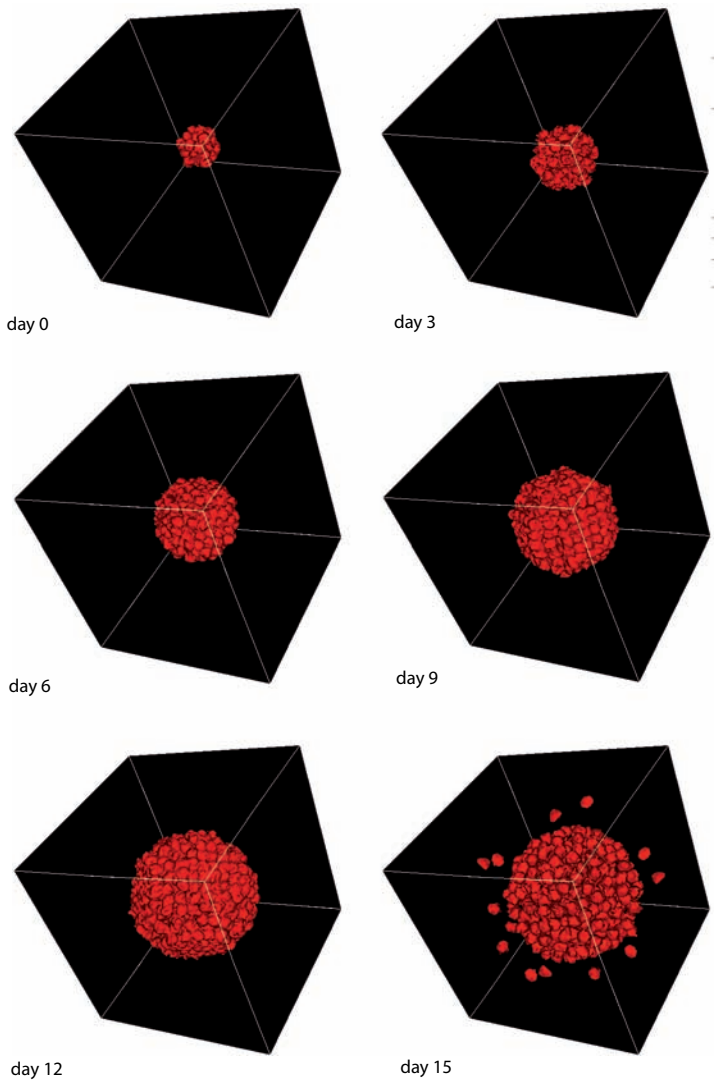


Fig. 4. Time-series of the simulated tumor evolution with the standard parameter setting given in Table 1. For the first 9-10 days, the malignant spheroid exponentially grows. It then stabilizes in a diffusion limited steady-state, where the proliferation of external viable cells is only enough to exactly balance the lost of mass due to the necrosis of the inner region, consequence of nutrient deprivation. In particular, in the final configuration, few single individuals are separated from the bulk of the tumor, as their motility has overcome intercellular adhesive interactions. These shed cells are the most aggressive, being able to further invade the host tissue

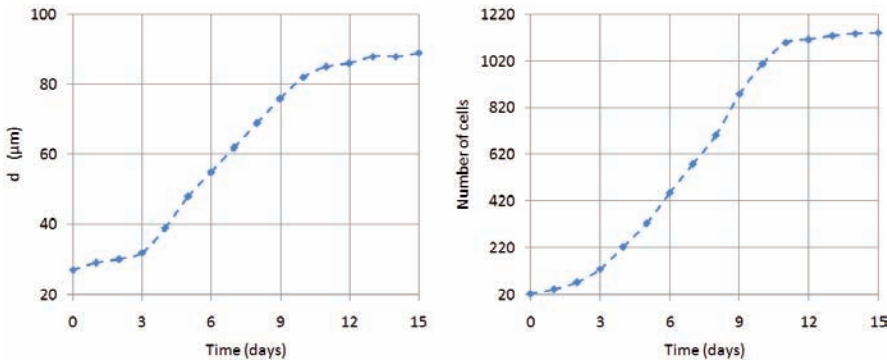


Fig. 5. Regression of the Time evolution of both the invasive radius d (left panel) and the number of cells (right panel) for the tumor spheroid, developing under the standard parameter settings, as depicted in Fig. 4

motility. As the distance between the spheroid border increases, the local availability of growth factors progressively falls. Indeed, a sort of equilibrium point is reached at which cells have an internal amount of vital factors too low to undergo mitosis, but sufficiently high to survive. These individuals are therefore in a so-called quiescent state. Finally, in the core of the mass, the supply of chemicals is negligible, so that inner cells, once consuming their basal level of growth factors, are unable to stay alive. They therefore die due to nutrient deprivation forming, as seen, a region of necrotic debris, which loose volume and are eventually removed. Summing up, malignant individuals within a solid, avascular mass display three different phenotypic state: *proliferating*, *quiescent*, and *necrotic*.

Fig. 6 represents also the final distribution of tumor MMPs and of matrix proteins. In particular, it can be clearly observed that the host tissue is significantly modified, as the matrix components have been dramatically degraded by the activity of tumor

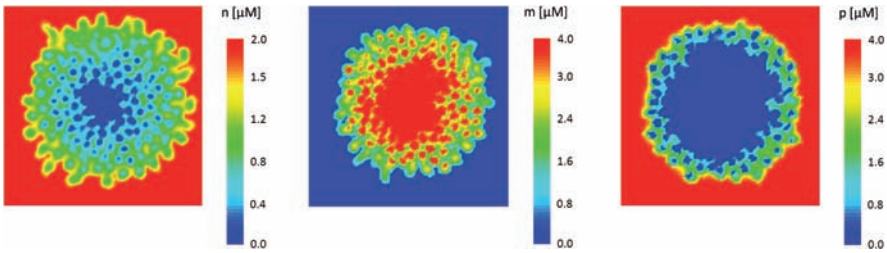


Fig. 6. Final patterns (i.e., at $t=15$ days) of the microscopic variables included in the model, taken in a representative section of the domain. From left to right panel: concentration of growth factors (n), tumor proteolytic enzymes (m), and matrix proteins (p). It is straightforward to notice that cells in the center of the spheroid have a negligible amount of chemicals and therefore undergo the necrotic transition. The more external individuals have instead an abundant supply of vital nutrients, which increase their metabolism. Indeed, the extracellular environment is completely modified by the tumor, as the matrix components are degraded by the localized activity of malignant MMPs

proteolytic enzymes. Indeed, steep protein gradients have formed, which eventually drive, via haptotactic mechanisms, a further invasion of the single, aggressive individuals.

The evolution of the malignant mass captured in the model reproduces the first stages of the growth of several tumors grown as spheroids in spinner cultures, such as ovarian [11, 53] or breast [26] carcinomas. Moreover, the model results are consistent with the development of avascular gliomas, both embedded *in vitro* in collagenous gels [32, 56] and implanted *in vivo* in mice [1, 14]. It is useful to underline that such a qualitative agreement occurs even without an exact bookkeeping of diffusing growth factors and a detailed inclusion of the cell cycle. Indeed, a similar growth of primary solid tumors have been also predicted by different types of theoretical models, see again the reviews [3, 12, 17, 47] and the comprehensive books [36, 45].

As briefly sketched in the previous section, the values of the parameters J s have a clear biological relevance, as they describe the relative preference of tumor cells to be in contact with other cells or with matrix components. At a molecular level,

Fig. 7. Invasiveness of the tumor mass in response to alterations of cell-cell adhesive strength, $J_{C,C}$ (i.e., all the other model parameters are the same as in the simulation in Fig. 4). Regression of the final depth of invasion $d(t = 15 \text{ days})$ vs $J_{C,C}^{ext}$. Downregulations of intercellular contact interactions enhance the invasive potential of the tumor

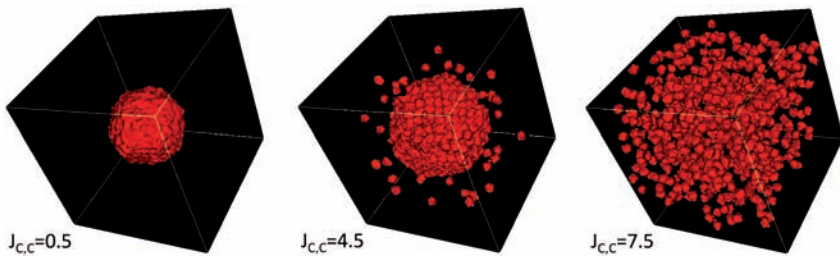
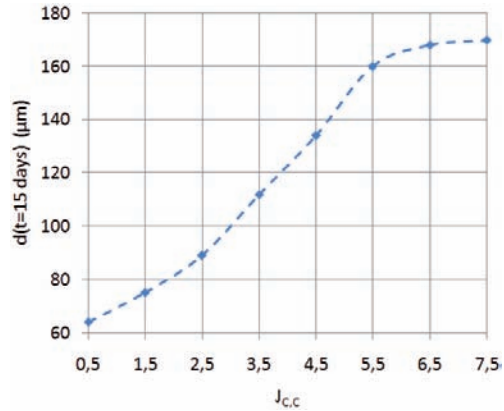


Fig. 8. Invasiveness of the tumor mass in response to alterations of cell-cell adhesive strength, $J_{C,C}^{ext}$ (i.e., all the other model parameters are the same as in the simulation in Fig. 4). By enforcing cell-cell adhesion, the spheroid remains densely compact and smooth with a low invasive potential. On the contrary, decrements in cell-cell adhesion lead to increments in the malignancy of the lesion, as external cells quickly shed in the host tissue and, consequently, also internal individuals have access to vital growth factors, surviving and proliferating. Examples of final configurations for $J_{C,C}^{ext} = 0.5$ (left panel), $J_{C,C}^{ext} = 4.5$ (central panel), and $J_{C,C}^{ext} = 7.5$ (right panel)

they are in fact a measure of the expression and the engagement of the different types of cell adhesion molecules, cadherins and integrins, respectively. As the variation of the J s may be expected to have a substantial impact on the overall development of the tumor, we analyze the model outcomes obtained by varying the cell-cell adhesive strength, defined by $J_{C,C}$. As reproduced in Fig. 7, at low values of $J_{C,C}$, which mean overexpressions of intercellular adhesive molecules, the final invasive radius of the spheroid slightly decreases. In particular, the tumor remains compact and strictly packed, without single individuals shed in the surrounding tissue, see left panel of Fig. 8. An enhanced activity of cadherins in fact is able to stabilize the morphology of the disease, reducing its aggressiveness: a smoother carcinoma mass can be in fact more efficiently treated surgically. As $J_{C,C}$ starts to increase, the tumor expands and invades deeper. The progressive reduction of cell-cell adhesive interactions results in fact in an increasing number of single individuals able to escape from the main mass into the external environment (see also central panel of Fig. 8). Finally, at even large values of $J_{C,C}$ (i.e., $> J_{C,M}$), not only the external cells quickly loose contact, dissociate and spread away from the rest of the mass creating a dispersed ring (fewer of them adhere with each other within 2 days, not shown), but a repulsion occurs also among individuals within the core of the mass. The surface area of the tumor/host interface therefore increases, thereby allowing internal cells greater access to vital growth factors needed to survive and eventually to proliferate. The overall tumor is therefore able to overcome the diffusional limitations and continues to expand both in the number of cells and in their dispersion, which further increases the overall bias toward invasion, as reproduced in the right panel of Fig. 8. Interestingly, there is not a critical value of the cell homotypic adhesion that clearly separates a compact phenotype of the tumor from a scattered configuration: rather the metastatic potential of the disease increases progressively with the downregulation of cell-cell contact strength. These results represent a definitive confirmation that an alteration in cell-cell adhesive properties is an important consideration in regulating the morphology of the disease and, eventually, in determining its severity. As a support of the model outcomes, a number of experimental studies have recently demonstrated that the down-regulation of cadherin molecules is implicated in a variety of metastatic cancers [15, 16, 18, 61, 65]. In particular, glioma cell lines with low N-cadherin expressions have been observed to aggressively invade matrix gels, whereas the same populations with a high N-cadherin activity have been instead shown to grow slower and to expand less significantly in the host tissues [31, 34].

4 Model Developments

The proposed model has been able to analyze the early stages of avascular tumor growth, reproducing its morphological evolution and quantifying its malignancy. In its present form, it can be applied to predict biologically relevant mechanisms able to influence the development of the cancerous mass, in particular its stability, spatial configuration and overall invasiveness.

However, there are many ways to extend our approach. For example, it would be possible to use more detailed and experimentally-determined laws for the rate of

cell proliferation and for the transition between cell phenotypic states (i.e., *mitotic-quiescent-necrotic*). Equally, additional diffusive fields describing the action of other chemical substances, present in the tumor environment, may be included. In particular, these molecular variables may be externally supplied by the host (such as oxygen or other nutrients) or expressed/secreted by the malignant mass itself (i.e., waste, products of cell degradation, or angiogenic factors such as hypoxia-inducible factors HIFs).

It would be also relevant to analyze the influence on the cancer evolution of specific interactions with components of the extratumoral environment not explicitly reproduced in the present version of the model. First, the malignant mass not only chemically degrades the soluble part of the ECM (via the secretion of MMPs), but also remodels its fibrous component, formed by networks of collagenous threads. At this regard, it would be interesting to introduce in the simulation domain specific meshes of matrix bundles, which surround the tumor mass, and to analyze how their structural and topological characteristics affect tumor growth, as done by other CPM authors [5, 50]. Moreover, malignant spheroids have also the potential to displace and compress the surrounding tissue, including the basement membrane and, in the case of carcinomas, the normal epithelium (from which they originate). Indeed, they create extraspaces for further growth and invade the tissue with fingers of cells along paths of least mechanical resistance. Finally, recent evidences provide the fact that malignant cells from epithelial origin induce apoptosis in the neighboring normal individuals (via acidosis) and alter gene expression in the surrounding stroma. All these tumor-environment interactions could be taken into account by defining in the model framework other types of cell populations, with specific rules for their influence on the evolution of the malignant colony, as done in [27] in the case of the ovarian cancer transmesothelial migration.

Distinct types of cells regularly characterize also the tumor mass itself. In fact, if the original node of malignant individuals stems from a common bulk of cells of the concerned organ or tissue, it is well-known that further mutations, loss of control of division and many other genetic defects subsequently bring to the formation of different tumor subpopulations. Indeed, the same multicellular spheroid displays the presence of differentiated individuals, with specific features and properties, for example different responses to analogous microenvironment signals or different migratory, proliferative and adhesive capabilities. The behavior and the interactions of these different subpopulations may affect the development of the overall tumor: for instance, a hyper-proliferation of more motile individuals may increase the metastatic potential of the disease, whereas a majority of highly metabolic cells may result in an early consumption of environmental substrates, which may lead to a quicker stabilization of the mass, as investigated both experimentally and theoretically in [36]. The presented approach could be easily improved in this respect by defining selected distinct types of tumor cells (i.e., labeled by different τ) and giving for each subgroup specific biophysical properties (i.e., such basal motility or adhesive strengths J_s) and constitutional laws (i.e., duplication rates) regulating the phenotypic behavior and evolution.

References

1. Abramovitch, R., Meir, G., Neeman, M.: Neovascularization induced growth of implanted C6 glioma multicellular spheroids: magnetic-resonance microimaging. *Cancer Res.* **55**, 1956–1962 (1995)
2. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., Walter, P.: *Molecular Biology of the Cell*, 4th edn. Garland Science, New York, (2002)
3. Araujo, R., McElwain, D.: A history of the study of solid tumour growth: the contribution of mathematical modelling. *Bull. Math. Biol.* **66**, 1039–1091 (2004)
4. Balter, A., Merks, R.M.H., Poplawski, N.J., Swat, M., Glazier, J.A.: The Glazier-Graner-Hogeweg model: extensions, future directions, and opportunities for further study. In: Anderson, A.R.A., Chaplain, M.A.J., Rejniak, K.A. (eds.) *Single-Cell-Based Models in Biology and Medicine*, pp. 151–167. *Mathematics and Biosciences in Interactions*, Birkhäuser, Basel, (2007)
5. Bauer, A.L., Jackson, T.L., Jiang, Y.: A cell-based model exhibiting branching and anastomosis during tumor-induced angiogenesis. *Biophys. J.* **92**, 3105–3121 (2007)
6. Bell, H.S., Whittle, I.R., Walker, M., Leaver, H.A., Wharton, S.B.: The development of necrosis and apoptosis in glioma: experimental findings using spheroid culture systems. *Neuropathol. Appl. Neurobiol.* **27**, 291–304 (2001)
7. Bellomo, N., Li, N.K., Maini, P.K.: On the foundations of cancer modelling: selected topics, speculations, and perspectives. *Math. Models Methods Appl. Sci.* **18**, 593–646 (2008)
8. Bock, J.M., Sinclair, L.L., Bedford, N.S., Jackson, R.E., Lee, J.H., Trask, D.K.: Modulation of cellular invasion by VEGF-C expression in squamous cell carcinoma of the head and neck. *Arch. Otolaryngol. Head. Neck. Surg.* **134**, 355–362 (2008)
9. Brown, J.M.: Tumor microenvironment and the response to anticancer therapy. *Cancer Biol. Ther.* **1**, 453–458 (2002)
10. Burgess, P.K., Kulesa, P.M., Murray, J.D., Alvord, E.C.: The interaction of growth rates and 544 diffusion coefficients in a three-dimensional mathematical model of gliomas. *J. Neuropathol. Exp. Neurol.* **56**, 704–713 (1997)
11. Burleson, K.M., Boente, M.P., Parmabuccian, S.E., Skubitz, A.P.: Disaggregation and invasion of ovarian carcinoma ascites spheroids. *J. Transl. Med.* **4**, 1–16 (2006)
12. Byrne, H., Alarcón, T., Owen, M., Webb, S., Maini, P.: Modeling aspects of cancer dynamics: a review. *Philos. Trans. R. Soc. A* **364**, 1563–1578 (2006)
13. Casciari, J.J., Sotirchos, S.V., Sutherland, R.M.: Variations in tumor-cell growth-rates and metabolism with oxygen concentration, glucose-concentration, and extracellular pH. *J. Cell. Physiol.* **151**, 386–394 (1992)
14. Castro, M.A.A., Klamt, F., Grieneisen, V.A., Grivicich, I., Moreira, J.C.F.: Gompertzian growth pattern correlated with phenotypic organization of colon carcinoma, malignant glioma and non-small cell lung carcinoma cell lines. *Cell Prolif.* **36**, 65–73 (2003)
15. Cavallaro, U., Christofori, G.: Cell adhesion in tumor invasion and metastasis: loss of the glue is not enough. *Biochim. Biophys. Acta.* **1552**, 39–45 (2001)
16. Cavallaro, U., Christofori, G.: Cell adhesion and signalling by cadherins and Ig-CAMs in cancer. *Nat. Rev. Cancer.* **4**, 118–132 (2004)
17. Chaplain, M.: Avascular growth, angiogenesis and vascular growth in solid tumours: the mathematical modelling of the stages of tumour development. *Math. Comput. Modelling* **23**, 47–87 (1996)
18. Christofori, G.: Changing neighbours, changing behaviour: cell adhesion molecule-mediated signalling during tumour progression. *EMBO J.* **22**, 2318–2323 (2003)
19. Clauss, M., Breier, G. (eds.): *Mechanisms of angiogenesis*, pp. 65–76. Birkhäuser, Basel (2005)

20. De Luca, A., Arena, N., Sena, L.M., Medico, E.: Met overexpression confers HGF-dependent invasive phenotype to human thyroid carcinoma cells in vitro. *J. Cell Physiol.* **180**, 365–371 (1999)
21. Di Renzo, M.F., Oliviero, M., Narsimhan, R.P., Bretti, S., Giordano, S., Medico, E., Gaglia, P., Zara, P., Comoglio, P.M.: Expression of the Met/HGF receptor in normal and neoplastic human tissues. *Oncogene* **6**, 1997–2003 (1991)
22. Freyer, J.P., Sutherland, R.M.: Determination of diffusion constants for metabolites in multi-cell tumor spheroids. *Adv. Exp. Med. Biol.* **159**, 463–475 (1983)
23. Freyer, J.P., Sutherland, R.M.: Regulation of growth saturation and development of necrosis in EMT6-Ro multicellular spheroids by the glucose and oxygen-supply. *Cancer Res.* **46**, 3504–3512 (1986)
24. Freyer, J.P.: Rates of oxygen consumption for proliferating and quiescent cells isolated from multicellular tumor spheroids. *Adv. Exp. Med. Biol.* **345**, 335–342 (1994)
25. Gallagher, H.C., Bacon, C.L., Odumeru, O.A., Gallagher, K.F., Fitzpatrick, T., Regan, C.M.: Valproate activates phosphodiesterase-mediated cAMP degradation: relevance to C6 glioma G1 phase progression. *Neurotoxicol Teratol.* **26**, 73–81 (2004)
26. Gatenby, R., Smallbone, K., Maini, P., Rose, F., Averill, J., Nagle, R., Worrall, L., Gillies, R.: Cellular adaptations to hypoxia and acidosis during somatic evolution of breast cancer. *Br. J. Cancer* **97**, 646–653 (2007)
27. Giverso, C., Scianna, M., Preziosi, L., Lo Buono, N., Funaro, A.: Individual cell-based model for in-vitro mesothelial invasion of ovarian cancer. *Math. Model. Nat. Phenom.* **5**, 203–223 (2010)
28. Glazier, J.A., Balter, A., Poplawski, N.J.: Magnetization to morphogenesis: a brief history of the Glazier-Graner-Hogeweg model. In: Anderson, A.R.A., Chaplain, M.A.J., Rejniak, K.A. (eds.) *Single-Cell-Based Models in Biology and Medicine, Mathematics and Biosciences in Interactions*, pp. 79–106. Birkhäuser (2007)
29. Glazier, J.A., Graner, F.: Simulation of the differential adhesion driven rearrangement of biological cells. *Physical. Rev. E* **47**, 2128–2154 (1993)
30. Graner, F., Glazier, J.A.: Simulation of biological cell sorting using a two dimensional extended Potts model. *Phys. Rev. Lett.* **69**, 2013–2017 (1992)
31. Kaufman, L.J., Brangwynne, C.P., Kasza, K.E., Filippidi, E., Gordon, V.D., Deisboeck, T.S., Weitz, D.A.: Glioma expansion in collagen I matrices: analyzing collagen concentration-dependent growth and motility patterns. *Biophys. J.* **89**, 635–650 (2005)
32. Khaitan, D., Chandna, S., Arya, M.B., Dwarakanath, B.S.: Establishment and characterization of multicellular spheroids from a human glioma cell line: implications for tumor therapy. *J. Transl. Med.* **4**, 12–25 (2006)
33. Hatzikirou, H., Deutsch, A., Schaller, C., Simon, M., Swanson, K.: Mathematical modeling of glioblastoma tumour development: a review. *Math. Models Methods Appl. Sci.* **15**, 1779–1794 (2005)
34. Hegedus, B., Marga, F., Jakab, K., Sharpe-Timms, K.L., Forgacs, G.: The interplay of cell-cell and cell-matrix interactions in the invasive properties of brain tumors. *Biophys. J.* **91**, 2708–2716 (2006)
35. Ising, E.: Beitrag zur theorie des ferromagnetismus. *Z. Physik.* **31**, 253 (1925)
36. Lowengrub, J., Cristini, V.: *Multiscale modeling of cancer: an integrated experimental and mathematical modeling approach*. Cambridge University Press (2010)
37. Marée, A.F.M., Grieneisen, V.A., Hogeweg, P.: The Cellular Potts Model and biophysical properties of cells, tissues and morphogenesis. In: Anderson, A.R.A., Chaplain, M.A.J., Rejniak K.A. (eds.) *Single-Cell-Based Models in Biology and Medicine*, pp. 107–136. *Mathematics and Biosciences in Interactions*. Birkhäuser (2007)

38. Martin, M.L., Regan, C.M.: The anticonvulsant valproate teratogen restricts the glial cell cycle at a defined point in the mid-G1 phase. *Brain Res.* **554**, 223–228 (1991)
39. Metropolis, N., Rosenbluth, A.E., Rosenbluth, M.N., Teller, A.H., Teller, E.: Equation of state calculations by fast computing machines. *J. Chem. Phys.* **21**, 1087–1092 (1953)
40. Montero, E., Abreu, C., Tonino, P.: Relationship between VEGF and p53 expression and tumor cell proliferation in human gastrointestinal carcinomas. *Journal of Cancer Research and Clinical Oncology* **134**, 193–201 (2007)
41. Mueller-Klieser, W.: Tumor biology and experimental therapeutics. *Crit. Rev. Oncol. Hematol.* **36**, 123–139 (2002)
42. Murphy, G., Gavrilovic, J.: Proteolysis and cell migration: creating a path? *Curr. Opin. Cell Biol.* **11**, 614–621 (1999)
43. Osada, H., Takahashi, T.: Genetic alterations of multiple tumor suppressors and oncogenes in the carcinogenesis and progression of lung cancer. *Oncogene* **21**, 7421–7434 (2002)
44. Potts, R.B.: Some generalized order-disorder transformations. *Proc. Camb. Phil. Soc.* **48**, 106–109 (1952)
45. Preziosi, L. (ed.): *Cancer modeling and simulation. Mathematical Biology and Medicine Sciences.* Chapman and Hall/CRC, London (2003)
46. Preziosi, L., Tosin, A.: Multiphase and multiscale trends in cancer modelling. *Math. Model. Nat. Phenom.* **4**, 1–11 (2009)
47. Quaranta, V., Weaver, A., Cummings, P., Anderson, A.: Mathematical modeling of cancer: the future of prognosis and treatment. *Clin. Chim. Acta* **357**, 173–179 (2005)
48. Ramanathan, A., Wang, C., Schreiber, S.: Perturbational profiling of a cell-line model of tumorigenesis by using metabolic measurements. *Proc. Natl. Acad. Sci. USA* **102**, 5992–5997 (2005)
49. Ramis-Conde, I., Drasdo, D., Anderson, A.R.A., Chaplain, M.A.J.: Modeling the influence of E-cadherin-beta-catenin pathway in cancer cell invasion: a multiscale approach. *Biophys. J.* **95**, 155–165 (2008)
50. Rubenstein, B.M., Kaufman, L.J.: The role of extracellular matrix in glioma invasion: a Cellular Potts Model approach. *Biophys. J.* **95**, 5661–5680 (2008)
51. Savill, N.J., Hogeweg, P.: Modelling morphogenesis: from single cells to crawling slugs. *J. Theor. Biol.* **184**, 118–124 (1997)
52. Scianna, M., Preziosi, L.: Multiscale Developments of the Cellular Potts Model. *Multiscale Model Simul.* **10**, 342–382 (2012)
53. Shield, K., Ackland, M.L., Ahmed, N., Rice, G.E.: Multicellular spheroids in ovarian cancer metastases: biology and pathology. *Gynec. Oncol.* **113**, 143–148 (2008)
54. Smith, J.A., Martin, L.: Do cells cycle? *Proc. Natl. Acad. Sci. U.S.A.* **70**, 1263–1267 (1973)
55. Steele, I.A., Edmondson, R.J., Leung, H.Y., Davies, B.R.: Ligands to FGF receptor 2-IIIb induce proliferation, motility, protection from cell death and cytoskeletal rearrangements in epithelial ovarian cancer cell lines. *Growth Factors* **24**, 45–53 (2006)
56. Stein, A.M., Demuth, T., Mobley, D., Berens, M., Sander, L.M.: A mathematical model of glioblastoma tumor spheroid invasion in a three-dimensional in vitro experiment. *Biophys. J.* **92**, 356–365 (2007)
57. Steinberg, M.S.: Reconstruction of tissues by dissociated cells. Some morphogenetic tissue movements and the sorting out of embryonic cells may have a common explanation. *Science* **141**, 401–408 (1963)
58. Steinberg, M.S.: Does differential adhesion govern self-assembly processes in histogenesis? Equilibrium configurations and the emergence of a hierarchy among populations of embryonic cells. *J. Exp. Zool.* **173**, 395–433 (1970)

59. Stetler-Stevenson, W.G., Aznavoorian, S., Liotta, L.A.: Tumor cell interactions with the extra-cellular matrix during invasion and metastasis. *Ann. Rev. Cell Biol.* **9**, 541–573 (1993)
60. Su, J.L., Yang, P.C., Shih, J.Y., Yang, C.Y., Wei, L.H., Hsiah, C.Y., Chou, C.M., Jeng, Y.M., Wang, M.Y., Kuo, M.L., Chang, K.J., Hung, M.C.: The VEGF-C/Flt-4 axis promotes invasion and metastasis of cancer cells. *Cancer Cell* **9**, 209–223 (2006)
61. Sundfeldt, K.: Cell-cell adhesion in the normal ovary and ovarian tumors of epithelial origin; an exception to the rule. *Molecular and Cellular Endocrinology* **202**, 89–96 (2003)
62. Tracqui, P.: Biophysical models of tumour growth. *Rep. Prog. Phys.* **72**, 056701 (2009)
63. Turner, S., Sherratt, J.A.: Intercellular adhesion and cancer invasion: A discrete simulation using the extended Potts model. *J. Theor. Biol.* **216**, 85–100 (2002)
64. Vaupel, P., Hockel, M.: Blood supply, oxygenation status and metabolic micromilieu of breast cancers: characterization and therapeutic relevance. *Int. J. Oncol.* **17**, 869–879 (2000)
65. Vleminckx, K., Vakaet, L., Mareel, M., Fiers, W., Vanroy, F.: Genetic manipulation of E-cadherin expression by epithelial tumor cells reveals an invasion suppressor role. *Cell* **66**, 107–119 (1991)
66. Zhang, Y.W., Vande Woude, G.F.: HGF/SF-Met signaling in the control of branching morphogenesis and invasion. *J. Cell Biochem.* **88**, 408–417 (2003)

Computational Models as Novel Tools for Cancer Vaccines

Filippo Castiglione, Pier Luigi Lollini, Santo Motta, Arianna Palladini,
Francesco Pappalardo, and Marzio Pennisi

Nessuna umana investigazione si può dimandare vera scienza,
se essa non passa per le matematiche dimostrazioni.

Leonardo da Vinci,
Trattato della pittura, I,1

Abstract Prevention of tumor growth by immunological approaches is based on the assumption that the immune system, if adequately stimulated before tumor onset, could be able to protect from specific cancers. In the last decade active immunization strategies effectively prevented some virus-related cancers in humans.

An immunopreventive cell vaccine for the non-virus-related human breast cancer has been recently developed. This vaccine, called Triplex, targets the HER-2-neu oncogene in HER-2/neu transgenic mice and has shown to almost completely prevent HER-2/neu-driven mammary carcinogenesis when administered with an intensive and life-long schedule. To better understand the preventive efficacy of the Triplex vaccine in reduced schedules we employed a computational approach. The computer model developed allowed us to test specific vaccination schedules in the quest for optimality. Furthermore, another computational model was developed to simulate

F. Castiglione (✉)

Institute for Computing Application “M. Picone”, National Research Council of Italy, Viale Manzoni 30, 00185 Rome, Italy

e-mail: f.castiglione@iac.cnr.it

P.L. Lollini · A. Palladini

Department of Hematology and Oncologic Sciences “L. e A. Seragnoli”, University of Bologna, Via Massarenti 9, 40138 Bologna, Italy

e-mail: pierluigi.lollini@unibo.it, arianna.palladini@unibo.it

S. Motta · M. Pennisi

Department of Mathematics and Computer Science, University of Catania, V.le A. Doria 6, 95125 Catania, Italy

e-mail: motta@dm.unict.it, mpennisi@dm.unict.it

F. Pappalardo

Department of Pharmaceutical Sciences, University of Catania, V.le A. Doria 6, 95125 Catania, Italy

e-mail: francesco.pappalardo@unict.it

A. d’Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_12, © Springer-Verlag Italia 2012

the scenario arising from the immunotherapy experiments with the Triplex vaccine as a therapeutic approach against lung metastases derived by mammary carcinoma. This chapter describes the trail we followed starting from the problem of evaluating immunopreventive schedules with a generic computer model for the immune system response to a model of metastasis passing through an in-silico detailed model of the cancer-immune system interaction in HER-2/neu transgenic mice. Altogether it provides an example of the successful use of a combination of animal and computational modeling to speed up the way from lab to the bedside and even the patient.

1 Introduction

Cancer vaccines are a real application of the extensive knowledge of immunology to the field of oncology. Tumors are dynamic complex systems in which several entities, events and conditions interact among them resulting in growth, invasion and metastases. The immune system includes many cells and molecules that cooperatively act to protect the host organism from foreign agents. Interactions between the immune system and the tumor mass include a huge number of biological factors.

The evaluation of the antitumor efficacy of cancer vaccines in mouse models (also here referred to as *biological models*) is a required prelude to the clinical use of these treatments. Testing of some cancer vaccine features, such as the best conditions for vaccine administration, can be very difficult or even impossible only through experiments with biological models simply because a high number of variables need to be considered at the same time. This is where computational models can prove handy as they have shown to be able to reproduce enough biological complexity to be of use in suggesting new experiments. This characteristic makes computer models suited to perform “what-if” analyses to elucidate relationships between different phenomena and to aid in the validation or rejection of working hypotheses. Indeed, computational models can be used in *addition* to biological models.

This chapter describes the trail we followed starting from the problem of evaluating immunopreventive schedules with a generic computer model for the immune system response to a model of metastasis passing through an in-silico detailed model of the cancer-immune system interaction in HER-2/neu transgenic mice. It provides altogether an example of the successful use of a combination of animal and computational modeling to speed up the way from lab to the bedside and even the patient.

2 Immune System, Tumor Immunology and Cancer Vaccines: A Brief Priming

For the sake of consistency, we briefly sketch in this section the main features of the human immune system and the basic concepts of tumor immunology and cancer vaccines. Furthermore we focus on the cancer immunoprevention potential of a special vaccine tested on HER-2/neu transgenic mice.

2.1 The Immune System

The immune system reacts to molecules recognized as foreign (mainly components of microbes) to prevent infectious diseases, by various mechanisms altogether named immune response [1]. A first line of defense of the immune system is supported by the innate immunity that includes physical barriers, soluble mediators and specialized killer cells. The innate immune response remains essentially unaltered by repeated infections.

The adaptive immune system provides a second line of defense against infections as it recognizes in a specific way distinct components called antigens. Lymphocytes are the cellular players of this elaborated response and are able to store information on the acquired antigen recognition, to improve the immune response to repeated exposures. Finally, the adaptive immunity is specific for foreign antigens and tolerant to autologous (self) components.

The lymphocytes population includes millions of clones, each one with a different specific antigen receptor. This variability among lymphocytes receptors is the reason for lymphocytes ability to recognize a high number of different antigens.

Lymphocytes are mainly divided in T and B cells, and bear antigen receptor molecules on their cell surface. B cells present immunoglobulins (Ig) on cellular membrane; antibodies, secreted by plasma cells (antigen-stimulated B cells), are a soluble form of the B cell receptor. Antibodies make insoluble antigen-antibody complex binding to their cognate antigens; complexes are rapidly removed from the blood. Antibodies also activate *complement* components inducing lysis of the bound cell and promote phagocytosis and cell lysis by leukocytes.

The surface antigen receptor of T cells (TCR), recognizes small peptides bound to a cellular membrane protein called the major histocompatibility complex (MHC). T cells include three main subpopulations with different functions: cytotoxic T cells (named Tc, CTL or CD8+) directly kill cells expressing the antigen; regulatory T cells (Treg) inhibit immune responses; helper T cells (named Th or CD4+) positively regulate the activity of B and Tc cells. Th cells are further divided into type 1 and type 2 cells (Th1 and Th2). Th1 cells release interferon gamma (IFN- γ) and other cytokines to stimulate immune responses directed against viruses and intracellular bacteria, whereas Th2 cells release interleukin 4 (IL-4) to stimulate the immune response against parasites.

2.2 Tumor Immunology

Clinical evidences in immunodeficient individuals and preclinical studies in immunodepressed mice highlighted a strong correlation between immune system weakness and disorderly cell growth. The immune system physiologically prevents tumor onset, but the incidence of neoplastic diseases proves that cancer immune surveillance is not completely effective. Reasons for tumor progression could be related to transient immunodepression, reduced efficacy of the immune system response with aging and tumor cell acquisition of the capability to exploit immunological mechanisms and evade immune surveillance [12, 15, 16].

Immune counterattack to tumors is moved by both innate and adaptive immunity, including many molecules and cellular entities that cooperatively act to limit neoplastic growth. Briefly, phagocytes (granulocytes and macrophages), actors of the innate immunity, directly destroy tumor cells and produce cell fragments. Antigen presenting cells (APC) pick up and process these fragments ultimately presenting tumor antigens for lymphocyte recognition. Dendritic cells, which are professional APCs, uptake tumor antigens in the periphery then migrate to lymph nodes. Moreover, natural killer (NK) cells kill tumor cells with a low MHC expression and play a key role in the defense against circulating metastatic cells.

The T helper cell population is the play-maker of the adaptive immunity team against tumors. Th cells, activated by antigen recognition on APCs, proliferate and activate, by cytokine secretion, Tc, phagocytes, NK cells and B cells. Most solid tumors are protected from antibody or complement dependent lysis, consequently in the antitumor immune response the role of B cells is (mistakenly) considered marginal. Moreover B cells can even downregulate T cell responses promoting tumor growth. Finally also Treg cells can inhibit antitumor responses [21].

Hundreds of different approaches clinically tested over the years have disappointed the expectations of effective therapies for oncologic patients based on Tc cells induction [42]. On the other hand, the development of monoclonal antibodies against cell surface molecules led to clinical success in patients with lymphomas and also with breast cancer and other solid tumors.

2.3 Cancer Vaccines

The idea of developing strategies to support the immune system against tumors has been producing several immunological approaches effectively able to limit tumor growth. Antigen-nonspecific treatments, such as alpha-galactosyl ceramide, Treg depletion, cytokines and others, have induced a broad immune reactivity including antitumor response. Antigen-specific treatments are addressed to specific tumor associated molecules. These strategies can be passive as monoclonal antibodies administration, or active as vaccines [28].

The cure of established tumor masses by immunological strategies (immunotherapy) has produced poor results suggesting to address efforts to adequately stimulate immune system before tumor onset (immunoprevention), to protect the organism from specific cancers. Preclinical studies have shown that prevention is more effective than cure in the tumor immunology field [21].

Vaccines against human hepatitis B and C viruses and human papilloma virus, closely related to liver and cervical carcinoma respectively, are the first cancer immunopreventive approaches used in clinic. The idea of antigen-specific immunological active strategies against infectious tumors has been upgraded to immunopreventive approaches to non-infectious tumors.

Cancer vaccines actively enhance a specific immune response against target tumor antigens. Tumor antigens include a huge number of tumor-associated molecules mostly recognized by the immune system of the host as self, as they are also expressed by normal cells [30]. Consequently, a successful antitumor immune

response against such self antigens requires to break the immune tolerance. Among many described tumor antigens, only a few molecules proved to be good target antigens. Tumor associated molecules that are essential for tumor growth and progression could be suitable cancer vaccine targets, since they cannot be easily downmodulated or negatively selected in precancerous lesions under the pressure of a specific immune attack. Lollini and colleagues have defined these molecules as *oncoantigens* [21, 24]. Oncoantigens should be poorly expressed by normal adult tissues, but overexpressed by tumor [11]. Moreover the presence of target antigens on the cell membrane is an optimal condition to obtain both a cell-mediated and an antibody-mediated immune response.

Vaccines can be based on live, irradiated or genetically modified tumor cells, dendritic cells, proteins, peptides or naked DNA. The most successful vaccines in preclinical studies consisted of combinations of the target tumor antigen with other immunological signals acting as adjuvants. Adjuvant molecules, inducing an aspecific immune response, contribute to break the immune tolerance to tumor antigens. Adjuvants can be microbial sequences, cytokines, allogeneic glycoproteins of the major histocompatibility complex and/or co-stimulatory molecules. Overall the efficacy of a tumor vaccine depends both to a large extent on high concentrations of antigen and on the use of potent adjuvants. Therefore the efficacy of a vaccine is based on its ability to reverse a tolerant state and activate low-avidity immune reaction mechanisms that escape the central tolerance [21].

The vaccine protective ability in preclinical models includes induction of Th cells, Tc lymphocytes and, above all, antibodies that seem to cause vaccine protective ability in preclinical models. In order to ensure a long-term immune response, vaccines should be able to elicit immune memory mechanisms by repeated boost vaccinations.

Although cancer immunopreventive vaccines effectively reduced or delayed carcinogenesis in a variety of different model systems, translation into clinical trials is still far and several questions need to be answered [24].

2.4 Cancer Immunoprevention in HER-2/Neu Transgenic Mice

The human epidermal growth factor receptor 2 (referred to as HER-2 or ErbB2) is a membrane tyrosine kinase overexpressed in 25-30% of human breast cancers [46]. HER-2 has been widely used as target for immunopreventive strategies often evaluated against mammary carcinogenesis in HER-2/neu transgenic mice. A large number of studies have found treatments able to delay and/or reduce tumor onset up to a complete protection [24, 47].

Among HER-2/neu transgenic mice, BALBneuT have been used in our cancer immunopreventive studies. Female mice develop multifocal mammary carcinomas, resembling the main aspects of human counterpart, with a short latency, about 20 weeks of age. Within 33 weeks of age lobular carcinomas are palpable in all 10 mammary glands [3].

The Triplex cellular vaccine that we have developed, is one of the most effective preclinical preventive vaccine [24]. The vaccine is called Triplex because it has three

main components: the target antigen, HER-2/neu, and two adjuvant stimuli, IL-12 and allogeneic MHC molecules. IL-12 is needed to improve antigen presentation and consequently increase Th cell activation. Allogeneic MHC molecules are relevant to break the tolerance to HER-2/neu self antigen by stimulating multiple T cell clones and causing a broad production of immunostimulatory cytokines [13, 28].

BALBneuT mice were completely protected from mammary tumor onset by repeated administrations of the Triplex vaccine, starting at an early age (6 weeks of age). Untreated mice had multiple mammary carcinomas at six months of age while almost all vaccinated mice were tumor-free at one year of age doubling the life expectancy of these mice. IFN- γ and anti-HER-2/neu antibodies were the main actors of the immune response elicited by the Triplex vaccine. IFN- γ production, by Th1 and Tc lymphocytes, was induced through IL-12 administration. IFN- γ promotes isotypic switch to Th1-type immunoglobulins G (IgG) and causes cell proliferation block and apoptotic cell death. Tumor cells were killed by classical immune mechanisms as well as by downmodulating the HER-2/neu receptor. Anti-HER-2/neu antibodies, specifically IgG2a and IgG2b murine subclasses, showed antitumor activities mediated both by immune mechanisms, such as antibody-dependent cell-mediated cytotoxicity and complement-mediated cytotoxicity, and by induction of HER-2/neu recycling and inhibition of receptor dimerization. Finally, IFN- γ induced production of antiangiogenic factors [13, 28]. BALBneuT mice lacking antibodies or IFN- γ production confirmed the key role of IFN- γ and anti-HER-2/neu antibodies in tumor prevention. For these mice the vaccine failed to prevent mammary carcinoma onset [27].

The efficacy of the Triplex vaccine was related to the number and distribution of administrations along the mouse life. The Triplex vaccine was administered in BALBneuT mice according to *chronic* protocol based on 4-week vaccination cycles, starting from 6 week of age for the entire lifetime of mouse or until one year of age, at least. Mice received four vaccine administrations over the first 2 weeks of each 4-week cycle [13]. Such a high number of vaccinations actually limits the clinical use of the vaccine because it reduces patients compliance and increases the risk of side effects. Only three vaccination cycles were insufficient [29].

Note that the Triplex vaccine was not effective in the cure of tumor growth as therapeutic treatment: studies of the efficacy of Triplex vaccine against different phases of tumor progression (from hyperplasia to carcinoma in situ to invasive carcinoma) in BALBneuT mice showed that the chronic vaccination lost progressively its protective action in relation to the delay in the treatment start [30].

Surprisingly, the vaccine showed to be effective against lung metastases induced in BALBneuT mice by intravenous injection of syngeneic mammary carcinoma cells. Mice treated by Triplex vaccine for about a month had a 87 or 99% reduction in the number of lung metastases starting vaccination on day 7 or on day 1 after metastases induction respectively. The vaccine elicited the activation of lung granulocytes and macrophages and of Th cells and their cytokines (IFN- γ and IL-4); moreover the depletion of Treg cells increased the vaccine efficacy against metastases. Finally cancer preventive vaccines that fail to curative purposes can be effective in the therapy of metastases and so potentially effective in adjuvant therapy in clinic [29].

3 The Computational Method

We now move from the real world to the simulated one, describing the computational model of the immune system that is at the core of the one described in the following sections and that have been used to test new vaccine administration schedules of the Triplex vaccine *in silico*.

3.1 The Original Automaton

The original computational model has been around for a while now (see [2] and references therein). It makes use of bit strings to represent the *binding site* of cells and molecules (i.e., lymphocytes receptors, BCRs, TCRs, Major Histocompatibility Complexes MHC, antigen peptides and epitopes, immunocomplexes, etc.). A recent enhanced version of the simulation tool employing amino acid strings (sequences) for T cell peptides and lymphocyte receptors or antibody binding site has developed [41]. However, this new feature was not used in the immunoprevention study described in this article.

The underlying modeling paradigm is of the agent-based modeling (ABM) type, in that all entities are individually represented [6, 7, 48]. It includes the major classes of cells of the lymphoid lineage, that is, T helper lymphocytes, cytotoxic T lymphocytes, B lymphocytes, antibody-producer plasma cells, and natural killer cells (NK) and some of the myeloid lineage i.e., macrophages ($M\phi$), and dendritic cells (DC). All these entities interact each other following a set of rules describing the different phases of the recognition and response of the immune system against a pathogen. In particular, the model takes into account phagocytosis, antigen presentation, cytokine release, cell activation from inactive or anergic states to active states, cytotoxicity, and antibody secretion.

The model simulates a simplified form of innate immunity and a more elaborate form of adaptive immunity (including both humoral and cytotoxic immune responses). For example in the case of endogenous *danger signal* (e.g., coming from apoptotic cells) the innate immune response is modeled using the following rules acting consequentially:

- NK cells release $IFN-\gamma$ in presence of danger signals released by apoptotic cells [49]. The $IFN-\gamma$ is released when a cytotoxic response is required, that is, when the danger signal is released by dying cells (by lysis or by direct killing by CTLs); In the case of innate immune response by *exogenous signals* the activation sequence will begin with antigen presenting cells (DC or $M\phi$) stimulation;
- macrophage activation is mediated by $IFN-\gamma$;
- dendritic cells activation is mediated by $TNF-\alpha$ release by macrophages.

The model of adaptive immunity follows the widely accepted *Clonal Selection Theory* of Burnet that states that the immune response is based on specific clones of B and T lymphocytes that are selected for destruction of the antigens invading the body [5]. The humoral response consists in B cell maturation into plasma cells secreting anti-

bodies. Cell-mediated immunity in the form of cytotoxic T-cell activity is modeled through the cytotoxicity of CD8+ T cells on malignant or infected cells.

In the model, a single human lymph node (or a portion of it) is mapped onto a three-dimensional ellipsoid Cartesian lattice. The primary lymphoid organs thymus and bone marrow are modeled apart: the thymus [10] is implicitly represented by the positive and negative selection of immature thymocytes before they enter into the lymphatic system, while the bone marrow generates already mature B lymphocytes. Hence, only immunocompetent lymphocytes are represented on the primary lymphoid organ modeled. At variance with classical cellular automata models, there is no direct correlation among entities residing on different sites at a fixed time step, and the deterministic character of automata dynamics is replaced by a stochastic behavior. However, at the end of each time step entities diffuse from site to site introducing spatial correlations.

This simulation tool can be considered as a collection of immunological models since it incorporates a few working assumptions or theories, most of which are regarded as established immunological mechanisms.

In details, the model includes:

- the clonal selection theory of Burnet [5];
- the idiotypic network theory of Jerne, [19];
- the Clonal deletion theory (i.e., thymus education of T lymphocytes, [20]);
- the hypermutation of antibodies [4];
- the danger theory of Matzinger [25];
- the replicative senescence of T cells, or the Hayflick limit (i.e., a limit in the number of cell divisions, [18]);
- T cell anergy [44];
- ag-dose induced tolerance in B cells [31].

As in most computer models, these can be toggled on or off. Moreover, other biological processes can be added with relatively few code changes. For example, customizations of the basic model have been used to simulate different phenomena ranging from viral infection (e.g., HIV, EBV [8, 9]) to type I hypersensitivity [43].

Each time step of the simulation corresponds to eight hours. The interactions among the cells determine their functional behavior (Table 1). Interactions are coded as probabilistic rules defining the transition of each cell entity from one state to another. Each interaction requires cell entities to be in a specific state choosing in a set of possible states (e.g., naïve, active, resting, duplicating) that is dependent on the cell type. Once this condition is fulfilled, the interaction probability is directly related to the effective level of binding between ligands and receptors (see below).

Unlike the many immunological models, the present one not only simulates the cellular level of the inter-cellular interactions but also the intra-cellular processes of antigen uptake and presentation. Both the cytosolic and endocytic pathways are modeled. In the model, endogenous antigen is fragmented and combined with MHC class I molecules for presentation on the cell surface to CTLs receptors, whereas the exogenous antigen is degraded into smaller parts (i.e., peptides), which are then bound to MHC class II molecules for presentation to the Ths receptors (Table 1).

Table 1. Biological rules coding for interactions between cells or among cells and molecules and other specific mechanisms of the immune system. Each of the entries of this list corresponds to an algorithm implementing a specific activity of the immune cells. Legend: B=B cell, M ϕ = Macrophage, DC = Dendritic cell, TC = cytotoxic CD8+ T cell, Th = CD4+ T cell

<i>Interactions</i>	<i>Activations</i>
B phagocytosis of antigen	Activation of M ϕ
M ϕ phagocytosis of antigen	B cells anergy
DC phagocytosis of antigen	Th cells anergy
B presentation to Th	Priming of Th cells
M ϕ presentation to	TC cells anergy
DC presentation to Th	Activation of TC cells
Formation of immuno-complexes (IC)	
M ϕ phagocytosis	
Infection of EP cells	
Cytotoxicity of infected cells by TC	
<i>Antigen ingestion and presentation</i>	<i>Other procedures</i>
B exogenous pathway	Clone divisions
M ϕ exogenous pathway	Hematopoiesis
DC exogenous pathway	Plasma secretion of immunoglobulins
EP endogenous pathway	Entity movement
	Hypermutation of antibody

While macroscopic entities like cells are individually represented (i.e., they are considered as agents), low molecular weight molecules, such as interleukins or chemokines, are represented in terms of only their spatial concentration. The corresponding dynamics is modeled by the following parabolic partial differential equation that describes a uniform diffusion process with the addition of a degradation term that takes into account the finite half-life of molecules:

$$\frac{\partial c}{\partial t} = D \nabla^2 c - \lambda c + s(x, t),$$

(1)

where $c = c(x, t)$ is the concentration of chemokines, $s(x, t)$ is the source term (e.g., macrophages), D is the diffusion coefficient and $\lambda = \ln 2/\tau$ where τ is the half-life. We assume $D = 3000 \mu\text{m}^2/\text{min}$ and $\tau = 3 \text{ hrs}$ [17, 45]. Differences in cell mobility also are taken into account. Th cells are the fastest with an average velocity of $11 \mu\text{m}/\text{min}$, followed by B cells with $6 \mu\text{m}/\text{min}$ and DC with a velocity of $3 \mu\text{m}/\text{min}$ [45].

The rules listed in Table 1 are executed for each time step. The stochastic execution of these rules, as in a Monte Carlo method, produces a logical causal/effect sequence of events culminating in the immune response and development of immunological memory. The starting point of this series of events is the injection of

antigen (the priming) at time step t_0 . This may take place anytime after the simulation starts.

The system is designed to maintain a steady state of the global population of cells (*homeostasis*) if no infection is applied. This is achieved by modeling the birth/death process as a mean reverting process of the type

$$\frac{dx_i(t)}{dt} = \frac{\log_2}{\tau_i} (x_i(0) - x_i(t)) + \sigma, \quad (2)$$

where $x_i(t)$ is the population i at time t , τ_i is the specific half-life parameter and σ is a Gaussian random noise.

Initially the system is naïve in the sense that there are neither T and B memory cells nor plasma cells and antibodies. The various steps of the simulated immune response depends on what is actually injected, i.e., recombinant virus or bacteria. For example, in the specific case of protein vaccine, the various steps of the dynamic evolution are:

- t_0 : injection of antigens (the host has been infected);
- $t_1 = t_0 + \delta_1$: if the antigenic molecule contains B epitopes, then B cells bind the antigen; also antigen presenting cells (APC) i.e., macrophages and dendritic cells, non-specifically bind the antigen molecules;
- $t_2 = t_1 + \delta_2$: B cells and APC process the antigen (MHCs bind antigen-peptides if any); they expose the MHC/peptide groove on the surface;
- $t_3 = t_2 + \delta_3$: T cells bind APC and/or B cells which expose the MHC-complex; both B and T cells get stimulated;
- $t_4 = t_3 + \delta_4$: stimulated T cells start to clone; part of the daughter cells become memory cells; stimulated B cells divide into B memory and plasma cells;
- $t_5 = t_4 + \delta_5$: plasma cells secrete antibodies;
- $t_6 = t_5 + \delta_6$: antibodies bind the antigen to create immune complexes.

This sequence of events models the humoral response with production of specific antibodies. A similar sequence of events driven by infected cells presenting viral peptides together with class I HLAs on their surface triggers the cytotoxic activation of CD8+ T cells, the components of the cellular arm of the immune system.

In the description above, the role of signaling cytokines such as IL-2, IL-12, IFN- γ , IL-4 and TNF- α were not included. These cytokines have an impact on macrophage activation, lymphocyte division, [26].

Fig. 2 shows a typical immunization and cross reactivity experiment performed injecting two different antigen at different time steps (see figure caption for details).

The computation of molecular affinity. Strings of 0s and 1s are used to represent specificity elements. Each different bit-string defines a receptor (or molecular) binding specificity. The length of this string is specified as a parameter ℓ . Two bit-strings complement each other (or are a perfect match) if every 0 in one corresponds to a 1 in the other and conversely (see Fig. 1). More generally, a m -bit match is defined as a pair where exactly m bits complement each other. Therefore, we first define the

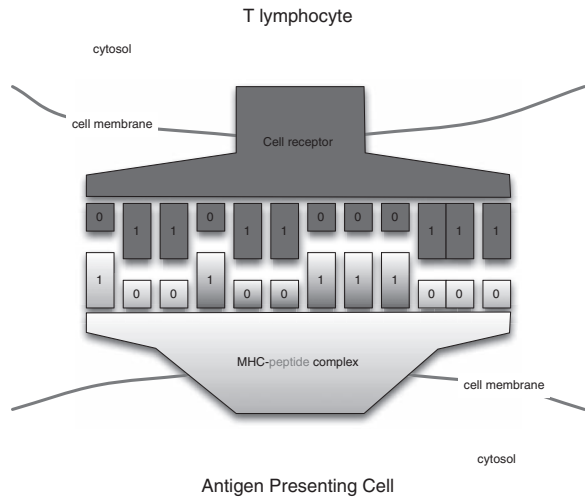


Fig. 1. Molecular affinity is calculated on the basis of the Hamming distance of the binary strings representing the binding sites of the interacting entities. In the figure a T lymphocyte receptor binds the MHC-peptide complex of an antigen presenting cell

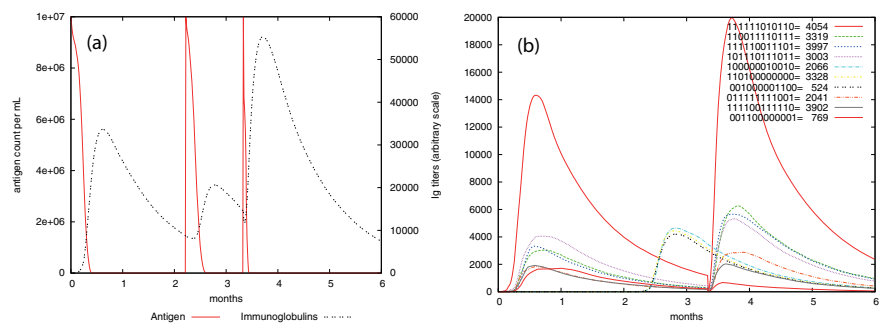


Fig. 2. Example of simulation results to show the selective ability of the model with respect to the injected binary string representing the antigen. In the panel (a) the first and the third injection are of the same compound. In that case the respective humoral immune response shown in panel (b) follows the classical primary / secondary response with immunological memory formation and affinity maturation. In contrast, the second injection of panel (a) does not cross react with the primary injection and in fact panel (b) shows antibody titers that are comparable to the first response but not to the third. Panel (b) also shows that the immunoglobulins are specific in that the clones of plasma B cells secreting antibodies have a binary string that is specific to the antigenic shape (that is also a binary string, not shown)

function $h(a, b)$ to give us the number of matching bits between two strings a and b (this corresponds to the Hamming distance in the space of the bit-strings). Then we define the function $\alpha(m)$ to give the affinity of an m -bit match. Typically we want to severely penalize mismatches by setting $\alpha(\ell)$ to a high value and $\alpha(m)$ ($m < \ell$) to much lower values, so as to ensure the perfect matches prevail over imperfect

ones. The $\alpha(\cdot)$ function is defined by a vector of length bits, called bit match vector with each component of the vector giving the affinity of a m -bit match where m ranges from ℓ down to 0. It completely specifies the affinity matrix. To specify the α vector, one method is to specify it directly by simply listing out the components of the vector. Another method uses the additional parameter arguments $\underline{m}$, that is the minimum match allowed, $\underline{a} = \alpha(\underline{m})$, that is the minimum level of affinity and δ_α a parameter specifying the gain in affinity proportional to a one bit more match, to calculate in the following way: i) using the parameter $\underline{m}$, set $\alpha(\underline{m}) = \underline{a}$ whereas for $m < \underline{m}$ set $\alpha(m)$ to 0 (this provides a level below which binding cannot occur); ii) the increase of strength on increasing a match by one bit is set to be the inverse of the ratio of number of clones with match $m + 1$ and m multiplied by the parameter δ_α . In formula,

$$\frac{\alpha(m+1)}{\alpha(m)} = \delta_\alpha \times \binom{\ell}{m} / \binom{\ell}{m+1}. \quad (3)$$

This allows to set the lower end value of $\alpha(m)$ and the steepness of its increase as the number of matching bits is incremented. It is usually more convenient than supplying the α vector directly. Generally it is advisable to set $\underline{m}$ somewhat close to ℓ bits in order to restrict the range of allowable matches to a few bits, so that the number of antibodies raised in response to a given antigen remains manageable. The affinity matrix thus generated has a regular structure.

MHC-peptide affinity. The affinity among MHC molecules and the antigen peptides is computed in a slightly different manner. Firstly, of all the match is computed over half bit string; secondly, there is no minimum match. The affinity value between two half strings whose match (i.e., Hamming distance) is m , $\forall m = 0, \dots, N/2$,

$$\alpha_{MHC}(m) = \left(\frac{1}{2}\right)^{\frac{N}{2}-m}. \quad (4)$$

This function is used in the process of antigen digestion and presentation.

Setting the parameters. The parameters of the model can be classified into three categories: (1) unknown values (free parameters) which are set after a tuning procedure that begins with an initial estimation of their values and iteratively improves the results of the simulations by small modifications of the parameters; (2) parameters that correspond to the initial conditions of the system and that determine the problem under investigation; (3) parameters whose value is well known and available from immunology literature.

Setting the free parameters requires running a high number of simulations for each parameter in order to get adequate statistical estimations. In practice, the parameters that are eventually selected are those enabling a good fit of the model dynamics to the available clinical (or experimental) data.

3.2 *Cancer Immunoprevention in Silico*

The efficacy of a treatment is strongly dependent on its dosage and schedule of administration. Many factors have to be considered in defining a new treatment schedule. Individual diversity and risk of side effects must be taken into account. If the former has effects on the minimal (lower bound) dosage of a treatment, the latter establishes an upper bound on the maximum allowed dosage to avoid side effects such as toxicity. To determine the schedule of a new treatment the common practice is to make use of the *medical consensus*, a typical policy where a representative group of experts in some medical areas commonly define the guidelines for the administration of a treatment, basing their decision on the state-of-the-art knowledge and past experimental evidence. The schedules that come out from consensus lie somewhere between minimal and maximal allowed dosages, usually far from the best (minimal) dosage needed to achieve at the same time efficacy among individuals and minimize the risk of side effects. Even the Triplex efficacy was related to the number and distribution of administrations along the mouse life. Complete protection at 1 year of age was achieved only using the chronic protocol ([3, 12] for the entire length of the experiment, whereas shorter protocols failed. However, the high number of vaccinations entitled with the use of the chronic schedule represents a limit for the use of the vaccine from a point of view of translational medicine. Further in vivo research focused on finding better vaccination protocols was discouraging, since it would require many sets of experiments in vivo, each lasting one year, with prohibitive costs.

To tackle the problem of determining if better vaccination protocols for the Triplex vaccine exist, we developed an in silico computational model (named SimTriplex) specifically designed to reproduce the effects of Triplex vaccine against the development of mammary carcinoma in HER-2/neu transgenic mice [23, 33].

3.2.1 A Modified Automaton

SimTriplex uses the same agent-based approach discussed in Sect. 3.1, and includes all the cellular and molecular entities considered essential for the reproduction of the in vivo experiment, such as tumor and vaccine cells, B and plasma cells, helper and cytotoxic T cells, macrophages, dendritic cells, antigens, antibodies, and various cytokines. Cell entities are distinguished on the basis of their position in the simulated volume, the age and an internal state (e.g., resting, activated, memory, antigen-presenting). Changes of this state (e.g., cell activation, cytotoxicity, cell death) are coded to mimic immunological rules like activation or duplication.

In order to model the continuous carcinogenic process of HER-2/neu transgenic mice, newborn tumor cells appear at each time step and are randomly placed on the lattice, whereas existing tumor cells duplicate. The simulation runs for a number of steps, typically equivalent to more than 1 year of real time. If the total number of tumor cells exceeds a given threshold, which indicates the formation of a palpable tumor mass, the simulation is stopped. Individual diversity observed in the experimental set-up is simulated through the use of pseudo-random number generators. Pseudo-random numbers affect the outcomes of various probabilistic events

at starting of the simulation (e.g., entities initial position in the lattice) as well as all the events that happen during the simulation, such as the order and outcomes of interactions. Each run of the simulator initialized with a given random number thus represents a *virtual mouse*. Experimental variability among mice is given by the use of different seeds for the pseudo-random number generator.

Once we validated the model against *in vivo* results, we had the opportunity to use the model as an *in silico* wet-lab to search for candidate protocols. To this end we needed to define the problem of finding optimal vaccination schedules in mathematical terms. The full story on how the biological concept of *vaccine effectiveness* has been translated in mathematical terms has been presented in [40]. We delineate here only the final results.

Let us consider a time interval $[0, T]$ in which the action of the vaccine is studied on a set of mice S . This represents, for example, the lifetime of the mouse or the time-length of the “*in vivo*” experiment. The time interval is divided in $N - 1$ equally spaced subintervals of width Δt , i.e. $\{t_1 = 0, t_2, \dots, t_i, \dots, t_N = T\}$. The time interval Δt represents a candidate time-window when one can administer the vaccine. In this case it has been set to 8 hours. Be $\mathbf{x} = \{x_1, x_2, \dots, x_i, \dots, x_N\}$ a binary vector representing a vaccine schedule where $x_i = 1/0$ represents if an administration is scheduled at time i or not. The label i indicates the time t_i of the vaccine administration and t_N the end of the vaccination period. The number of vaccine administrations is given by $n = \sum_{i=1}^N x_i$. Let's now call $\tau(\mathbf{x}, \lambda_j)$ the time of *carcinoma in situ* formation. This is a function of the vaccination schedule $\mathbf{x}$ administered to the mouse $j \in S$ and a parameter λ_j which describes the individual heterogeneity. The schedule is effective if $\tau \geq T$. Then the unconstrained optimization problem is to find $\bar{\mathbf{x}}$ such that

$$\begin{aligned} \tau(\bar{\mathbf{x}}, \lambda_j) &= \max\{\tau(\mathbf{x}, \lambda_j)\} \\ n(\bar{\mathbf{x}}) &= \min\{n(\mathbf{x})\} \end{aligned} \quad (5)$$

that is a multi-objective discrete optimization problem. The function τ depends on the action of the immune system stimulated by the vaccine against cancer cells and cannot be expressed in an analytical form. $\tau(\mathbf{x}, \lambda_j)$ is instead computed through the use of the SimTriplex simulator. We note here that the optimal search strategy has to be carefully biologically driven. Since the chronic schedule proved to be effective in tumor control, we would like to find an optimal protocol able to reproduce, *in silico*, the same time evolution of cancer cells given by the chronic schedule. This requirement leads now to place a two thresholds on the maximum number of cancer cells allowed,

$$\begin{cases} M_1(\bar{\mathbf{x}}) \leq \gamma_1, & t \in [0, T_{in}] \\ M_2(\bar{\mathbf{x}}) \leq \gamma_2, & t \in [T_{in}, T], \end{cases} \quad (6)$$

where $M_1(\mathbf{x})$ and $M_2(\mathbf{x})$ represent the maximum number of cancer cells in $[0, T_{in}]$ (cellular-mediated controlled phase) and in $[T_{in}, T]$ (humoral-mediated controlled phase) respectively, and $T_{in} \sim T/3$, while γ_1 and γ_2 represent cancer cells threshold in $[0, T_{in}]$ and in $[T_{in}, T]$, respectively. Moreover, biological variability has to be taken into account to achieve vaccine effectiveness for the population. Any candidate

optimal schedule must be tested on a representative sample and not only on a single virtual mouse. Bearing all this in mind, the optimization problem can be finally formulated in a following way: let $\{j_1, j_2, \dots, j_m\} \subset S$, (with $m = 8$), a random chosen subset of *in silico* mice,

$$\left\{ \begin{array}{l} \tau(\bar{\mathbf{x}}, \lambda_{j_1}) = \max(\tau(\mathbf{x}, \lambda_{j_1})) \\ \tau(\bar{\mathbf{x}}, \lambda_{j_2}) = \max(\tau(\mathbf{x}, \lambda_{j_2})) \\ \vdots \\ \tau(\bar{\mathbf{x}}, \lambda_{j_m}) = \max(\tau(\mathbf{x}, \lambda_{j_m})) \\ n(\bar{\mathbf{x}}) = \min(n(\mathbf{x})) \\ \text{subject to:} \\ M_1(\bar{\mathbf{x}}) \leq \gamma_1, \quad t \in [0, T_{in}] \\ M_2(\bar{\mathbf{x}}) \leq \gamma_2, \quad t \in [T_{in}, T]. \end{array} \right. \quad (7)$$

This is again a multi-objective optimization problem that can be reduced to a standard problem defining the new objective function as a linear combination of $f_{j_1}, \dots, f_{j_8}$, that are the objective functions relative to each of the eight *in silico* virtual mice, i.e.,

$$f(n, \tau_{j_1}, \tau_{j_2}, \dots, \tau_{j_8}) = \sum_{k=1}^8 \alpha_k f_{j_k}(n, \tau_{j_k}), \quad (8)$$

where all the $\alpha_k = 1$ as the mice are weighted equally.

In all the techniques adopted to solve the optimization problem (i.e., genetic algorithms first [22, 34] and simulated annealing afterwards [35–37]) we used the SimTriplex model to evaluate the objective function, in order to search for new vaccination schedules to prevent tumors in HER-2/neu transgenic mice.

The optimal protocol found *in silico* accounted for 32 injections and was not significantly different from the chronic vaccination in keeping all mice tumor free [32]. However, a fine grain analysis of the dynamics of tumor formation under the optimal therapeutic regimen and of the number of tumors subsequently appearing in each mouse (what is called *tumor multiplicity*) showed that the *in silico* vaccination schedule was more effective than other candidate protocols designed *in vivo* in reducing the number of tumors appearing over time. On the whole, mice vaccinated with the *in silico* schedule had a 60% reduction in the number of mammary tumors at 2 years, similar to that obtained with the Chronic protocol, even though the latter more efficiently controlled the kinetics of tumor development at earlier time points [32].

3.3 Simulating Immunotherapy of Lung Metastases

Metastatic dissemination represents a key step in malignant tumor progression after the formation of solid tumor masses. From the point of view of translational medicine, investigating the effect of the Triplex vaccine against lung metastases represents a crucial step, since clinical phase-I trials in humans are carried on advanced cancer patients rather than in patients at high risk of cancer.

As we have already mentioned above, the Triplex vaccine has been tested as a therapeutic approach against lung metastases derived by mammary carcinoma in BALBneuT mice [29]. Surprisingly the vaccine demonstrated able to show high levels of protection against the metastatic burden (see Sect. 2.4). We note here that induction of metastases in tumor-free mice can represent a typical scenario in human cancer, i.e., the scenario arising after the surgical removal of the primary tumor. Therefore the vaccine should in principle represent an effective treatment that can be used in early therapy of metastases, in substitution or in conjunction with other treatments such as chemotherapy.

In order to simulate the experimental outcome of immunotherapy experiments with Triplex against lung metastases we decided to develop a novel agent-based model (named *MetastaSim*) similarly to *SimTriplex* [38, 39]. The goal of this hybrid agent-based - ordinary differential equation model was to identify an optimal vaccine administration protocol in terms of maximization of prevented metastases and minimization of vaccine doses.

Since the biological and immunological background of the immune preventive and the immunotherapeutic *in vivo* experiments partially overlap, *SimTriplex* represented the unmistakable starting point for the development of *MetastaSim*. The models are similar, meaning that the latter can be considered an upgrade of the former. Differences worth to be noted are listed here. Biological, spatial and time-length differences are entitled with the two *in vivo* experiments and, therefore, with the two simulators. The therapeutic experiment lasts for 1 month whereas the immunoprevention experiment lasts for 1 year. Moreover *MetastaSim* uses a 128×128 hexagonal lattice to model the frontal ventral surface of the left lung of a mouse in order to mimic the spatial distribution and temporal evolution observed in the *in vivo* experiments, for an estimated volume which is approximately 64 times bigger than the spatial portion simulated by *SimTriplex*. The immune response of the two experiments is also different. Cancer immunoprevention is primarily driven by antibodies whereas anti-HER-2/neu antibodies, which are key effectors of the Triplex preventive ability, seem devoid of significant therapeutic activity. IFN- γ is the one of the major mediators both in cancer prevention and in the metastasis therapy whereas other vaccine-induced mechanisms in the therapeutic experiment are represented by T-helper activities at the systemic level and macrophages infiltration in the tumor cell nests.

The model is able to simulate multiple different metastatic nodules, each one with its own growth rate. The growth of the nodules is reproduced using the Gompertz law of growth in its differential form, and the initial growth rates are randomly generated using the inverse transform sampling method [14] in such a way that the *in silico* nodule measurements resemble the experimental measurements coming from the *in vivo* experiment. Moreover a first simplified chemotaxis and some of the effects of the immunosuppression induced by cancer cells have been introduced into the model as well, in order to improve the biological fidelity of the model.

After a tuning phase, necessary to determinate the free parameters, the model has been validated against existing *in vivo* experiments. Then we used the model to

find optimal protocols able to maximize the prevention of the metastatic burden and minimize the number of injections.

We note here that almost complete protection was achieved *in vivo* by using a vaccination protocol called Triplex+1 for which vaccine administration started one day after the intravenous injections of metastases and repeated two times a week up to the end of the experiment (32 days). The resulting protocol was then composed by 9 vaccine injections. Wet biology requirements (i.e., the time required for the preparation of the vaccine) as well as safeness for the mice (i.e., the need to avoid undesirable effects) entitle that no more than two vaccinations per week (in pre-established days) can be done. Shorter protocols should be therefore obtained by removing some injections from Triplex+1 protocol.

To this end we performed an exhaustive search over all the possible protocols lying in a time-window of one month. The results showed that it is possible a protocol with just 5 injections was able to entitle the same protection elicited by the 9-injections protocol tested *in vivo*. Moreover further analysis of best ranked protocols showed that in order to deal with an exponential-like growth of cancer cells (like a Gompertz growth), a significant vaccination should be administered in time, in order to establish a stronger immune response against cancer cells. Later vaccine recalls can indeed be useful to keep immunity at a high level in order to readily react against escaped or dormant cancer cells.

4 Discussion and Conclusion

We started from a real question: is it possible to reduce the number of Triplex vaccine administrations and maintain a high preventive efficacy? This is a typical clinical question that. The work we have described in this chapter showed that modeling is a concrete tool for study of cancer immunopreventive strategies and therefore can help in answering that question.

We learned that protective effectiveness of Triplex is provided once an immune plateau phase is reached and could be maintained with a reduced number of vaccinations. Moreover, vaccine efficacy was provided by a rapid increase of anti-HER-2/neu antibodies in the first phases of vaccination and the maintenance of an high level along life. In this respect the accuracy of mathematical modeling of early immune responses is critical. Finally, we have shown that elder mice reacted less efficiently to Triplex vaccine than younger ones so in elder mice more frequent boosts are required [32]. In other words, the design and modeling of cancer vaccines and vaccination protocols must take into account the progressive aging of the immune system, by striving to boost immune responses in elderly hosts.

It comes now very easy to answer the question of why the modeling is useful. Firstly, because mathematical and computational models are able to mimic high complex and dynamic systems that modify themselves. This is what biologists search because preclinical models are surely dynamic but often they offer only a piece of biological true. Secondly, because computational models can be modified in reasonable time, enabling a quick examination of hypothesis and model components.

Of course we can not forget what is the other side of the coin. Models are simplifications of reality and as such can leave apart important aspects of the phenomenon under study. A good model needs to be based on clinical or preclinical data and its improvement needs to make use of available clinical and preclinical testing of predicted data. Only through a carefully checked adherence with reality we can hope that our model produces meaningful biological knowledge. Perhaps not perfectly accurate but at least instructive.

5 The 21st Century Alliance

Years ago, when mathematical biology and bioinformatics started to be *trendy* disciplines, the general tendency was to have mathematicians looking for problems in hospitals and bio-labs. This was not always welcomed and sometimes the reaction was cold if not skeptic. Luckily there were exceptions.

Nowadays things have changed. Exact sciences like mathematics or physics are considered at the edge of research in biology thanks to the genome project that has brought into focus the urgency to looking into enormous amount of data. This has pushed biologists, to ask for help from computer scientists, statisticians and mathematicians to show them their data in the hope that the exact tools those people are used to handle could help disentangling the enormous complexity of biological systems at many space-temporal scales.

It is a novel alliance built on the idea that by addressing the challenges with a multidisciplinary “troop” armed with techniques and instruments coming from modern biotechnologies such as genomics, proteomics, bioinformatics and computational sciences, we will be able to unwind true sensible knowledge. The reward is highly priced as advances in clinical areas pave the way for improvement of human health through improved diagnostics and optimized treatment. Moreover, it promise to be tailored to both the individual and the disease, hence offering better targeted, optimized, and personalized treatments.

This 21st century alliance is stimulating for both counterparts. There is so much to understand, so much to learn about other specific scientific methods and viewpoint that the usual result is a lifelong involvement into this discipline.

Here we wish to share few words on experience matured during the few years of implementation of the work described in this chapter. This period has been extremely exciting and rewarding but not lacking in moments of discomfort. Here are our comments.

5.1 Computational Scientists in the Eyes of a Life Scientist

The interaction between biologists and mathematicians is based on the idea that each partner can contribute to increase the knowledge of the other one. This interaction can be describe by four verbs: to meet, to define, to reflect and to explore.

1. “To meet” is the first word to describe this interaction because only when these two worlds meet the collaboration can start. The Meeting is the time in which biologists communicate their problems and needs whereas mathematicians describe their approach and tools.
2. “To define” is the second action to build a strong and productive collaboration. Math and biology are two distinct disciplines that use completely different languages. In order to work proficiently together biologists and mathematicians have to share key concepts. This process is surely difficult and time-consuming but a necessary condition to build a robust interaction.
3. “To reflect” is the term to evaluate the interaction. Mathematicians have to ask themselves how much their models reflect the true biological reality. A mathematical model needs to be able to reproduce the behavior of the biological model with its elements and rules. Through careful comparison of the math model output with bio-models behavior the process of developing a useful model converges.
4. “To explore” is the final step of the collaboration. Biologists start to use the mathematical models as tools to explore their biological systems, addressing their studies to evidence new rules and events related to tumor onset. On the other hand mathematicians exploit the validation of their results with *in vivo* or *in vitro* experiments to refine their models.

5.2 Life Scientists in the Eyes of a Computational Scientist

The interdisciplinary nature of the experiences described here requires a strong effort of life scientists, which need to go beyond the only data supply, as it is extremely important in defining the biological scenario and ultimately construct a robust and validated mathematical or computational model. Only through a unceasing collaboration of life and computational scientists it is possible to turn software into a valuable tool in life sciences.

In our experience the first problem we had to face with has been to define a common language among life scientists and modelers. Communication is a fundamental step to start a fruitfully collaboration. We overcome this issue by using graphical and conceptual languages as those commonly used in computer science to describe complex and dynamical systems. Conceptual models are easily understood by life scientists and allow to transfer qualitative biological knowledge to modelers, which are then able to built mathematical and computational simulators.

The parameters of the computational models so constructed are adapted to reproduce the *in vivo* experimental data available. Also here the help of life scientists is very much required. They need to provide biological data first and interpretation (checking) of model results after, in order to move from qualitative to quantitative modeling.

After validation the model is ready to predict experimental outcomes, reproducing experimental setups that are either time or money consuming *in vivo*. Here is where computation turns out to be cheap. Predictions have to be carried by life scientists

for in vivo verifications. Model results are finally validated against in vivo or in vitro experiments.

Clearly, the involvement of both counterparts is important at every step of the modeling process. The vision must be the same. The interest high on both sides. Without a common effort, it would not be possible to reach any meaningful result.

References

1. Abbas, A.K., Lichtman, A.H., Pillai, S.: Cellular and Molecular Immunology, 6th edn. Elsevier, Philadelphia (2007)
2. Bernaschi, M., Castiglione, F.: Design and implementation of an immune system simulator. *Comp. Biol. Med.* **31**(5), 303–331 (2001)
3. Boggio, K., Nicoletti, G., Di Carlo, E., Cavallo, F., Landuzzi, L., Melani, C., Giovarelli, M., Rossi, I., Nanni, P., De Giovanni, C., Bouchard, P., Wolf, S., Modesti, A., Musiani, P., Lollini, P.L., Colombo, M.P., Forni, G.: Interleukin 12-mediated prevention of spontaneous mammary adenocarcinomas in two lines of Her-2/neu transgenic mice. *J. Exp. Med.* **188**, 589–596 (1998)
4. Brenner, S., Milstein, C.: Origin of antibody variation. *Nature* **211**, 242–243 (1966)
5. Burnet, F.M.: The Clonal Selection Theory of Acquired Immunity. Vanderbilt University, Nashville, TN (1959)
6. Castiglione, F.: Agent Based Modeling. *Scholarpedia* **1**(10), 1562 (2006)
7. Castiglione, F.: Agent Based Modeling and Simulation, Introduction to. In: Meyers, R. (ed.) *Encyclopedia of Complexity and Systems Science*, vol 1, pp. 197–200. Springer, New York (2009)
8. Castiglione, F., Duca, K.A., Jarrah, A., Laubenbacher, R., Luzuriaga, K., Hochberg, D., Thorley-Lawson, D.A.: Simulating Epstein Barr Virus Infection with C-ImmSim. *Bioinformatics* **23**, 1371–1377 (2007)
9. Castiglione, F., Poccia, F., D’Offizi, G., Bernaschi, M.: Mutation, fitness, viral diversity, and predictive markers of disease progression in a computational model of HIV type 1 infection. *AIDS Res. Hum. Retrovirus* **20**(12), 1314–1323 (2004)
10. Castiglione, F., Santoni, D., Rapin, N.: CTLs’ repertoire shaping in the thymus: a Montecarlo simulation. *Autoimmunity* **44**(4), 1–10 (2011)
11. Cavallo, F., Calogero, R.A., Forni, G.: Are oncoantigens suitable targets for anti-tumour therapy? *Nat. Rev. Cancer* **7**, 707–713 (2007)
12. Cavallo, F., De Giovanni, C., Nanni, P., Forni, G., Lollini, P.-L.: The immune hallmarks of cancer. *Cancer Immunol. Immunother.* **60**, 319–326 (2011)
13. De Giovanni, C., Nicoletti, G., Landuzzi, L., Astolfi, A., Croci, S., Comes, A., Ferrini, S., Meazza, R., Iezzi, M., Di Carlo, E., Musiani, P., Cavallo, F., Nanni, P., Lollini, P.L.: Immunoprevention of HER-2/neu transgenic mammary carcinoma through an interleukin 12-engineered allogeneic cell vaccine. *Cancer Res.* **64**, 4001–4009 (2004)
14. Devroye, L., Non-uniform random variate generation. Springer-Verlag, New York (1986)
15. Dunn, G.P., Old, L.J., Schreiber, R.D.: The immunobiology of cancer immunosurveillance and immunoediting. *Immunity* **21**, 137–148 (2004)
16. Finn, O.J.: Cancer immunology. *N. Engl. J. Med.* **358**, 2704–2715 (2008)
17. Francis, K., Palsson, B.O.: Effective intercellular communication distances are determined by the relative time constants for cyto/chemokine secretion and diffusion. *Proc. Natl. Acad. Sci. USA* **94**(23), 12258–12262 (1997)

18. Hayflick, L., Moorhead, P.S.: The serial cultivation of human diploid cell strains. *Exp. Cell. Res.* **25**, 585–621 (1961)
19. Jerne, N.K.: Towards a network theory of the immune system. *Ann. Immunol.* **125C**, 373–389 (1974)
20. Lederberg, J.: Genes and antibodies. *Science* **129**, 1649–1653 (1959)
21. Lollini, P.L., Cavallo, F., Nanni, P., Forni, G.: Vaccines for tumor prevention. *Nat. Rev. Cancer* **6**, 204–216 (2006)
22. Lollini, P.L., Motta, S., Pappalardo, F.: Discovery of cancer vaccination protocols with a genetic algorithm driving an agent based simulator. *BMC Bioinformatics* **7**, 352 (2006)
23. Lollini, P.L., Motta, S., Pappalardo, F.: Modeling tumor immunology, *Mathematical Models & Methods in Applied Sciences* **16**(7S), 1091–1124 (2006)
24. Lollini, P.L., Nicoletti, G., Landuzzi, L., Cavallo, F., Forni, G., De Giovanni, C., Nanni, P.: Vaccines and other immunological approaches for cancer immunoprevention. *Curr. Drug. Targets* **12**, 1957–1973 (2010)
25. Matzinger, P.: Tolerance, Danger, and the Extended Family. *Ann. Rev. Immunol.* **12**, 991–1045 (1994)
26. Murphy, K., Travers, P., Janeway, C., Walport, M.: *Janeway's Immunology*. Garland Science, Taylor and Francis, New York (2008)
27. Nanni, P., Landuzzi, L., Nicoletti, G., De Giovanni, C., Rossi, I., Croci, S., Astolfi, A., Iezzi, M., Di Carlo, E., Musiani, P., Forni, G., Lollini, P.L.: Immunoprevention of mammary carcinoma in HER-2/neu transgenic mice is IFN-gamma and B cell dependent. *J. Immunol.* **173**, 2288–2296 (2004)
28. Nanni, P., Nicoletti, G., De Giovanni, C., Landuzzi, L., Di Carlo, E., Cavallo, F., Pupa, S.M., Rossi, I., Colombo, M.P., Ricci, C., Astolfi, A., Musiani, P., Forni, G., Lollini, P.L.: Combined allogeneic tumor cell vaccination and systemic interleukin 12 prevents mammary carcinogenesis in HER-2/neu transgenic mice. *J. Exp. Med.* **194**, 1195–1205 (2001)
29. Nanni, P., Nicoletti, G., Palladini, A., Croci, S., Murgo, A., Antognoli, A., Landuzzi, L., Fabbi, M., Ferrini, S., Musiani, P., Iezzi, M., De Giovanni, C., Lollini, P.L.: Antimetastatic activity of a preventive cancer vaccine. *Cancer Res.* **67**, 11037–11044 (2007)
30. Novellino, L., Castelli, C., Parmiani, G.: A listing of human tumor antigens recognized by T cells. *Cancer Immunol. Immunother.* **54**, 187–207 (2005)
31. Nossal, G.J.V., Pike Beverley, L.: Clonal anergy: Persistence in tolerant mice of antigen-binding B lymphocytes incapable of responding to antigen or mitogen. *Proc. Natl. Acad. Sci. USA* **77**(3), 1602–1606 (1980)
32. Palladini, A., Nicoletti, G., Pappalardo, F., Murgo, A., Grosso, V., Stivani, V., Ianzano, M.L., Antognoli, A., Croci, S., Landuzzi, L., De Giovanni, C., Nanni, P., Motta, S., Lollini, P.L.: In silico modeling and in vivo efficacy of cancer-preventive vaccinations. *Cancer Res.* **70**, 7755–7763 (2010)
33. Pappalardo, F., Castiglione, F., Lollini, P.L., Motta, S.: Modelling and Simulation of Cancer Immunoprevention Vaccine. *Bioinformatics* **21**(12), 2891–2897 (2005)
34. Pappalardo, F., Mastriani, E., Lollini, P.L., Motta, S.: Genetic Algorithm against Cancer. *Lect. Notes Comp. Sci.* **3849**, 223–228 (2009)
35. Pappalardo, F., Pennisi, M., Castiglione, F., Motta, S.: Vaccine protocols optimization: in silico experiences. *Biotechnology Advances* **28**, 82–93 (2010)
36. Pennisi, M., Catanuto, R., Mastriani, E., Cincotti, A., Pappalardo, F., Motta, S.: Simulated Annealing And Optimal Protocols. *J. Circuits Systems and Computers* **18**(8), 1565–1579 (2009)
37. Pennisi, M., Catanuto, R., Pappalardo, F., Motta, S.: Optimal vaccination schedules using Simulated Annealing, *Bioinformatics* **24**(15), 1740–1742 (2008)

38. Pennisi, M., Pappalardo, F., Motta, S.: Agent based modeling of lung metastasis-immune system competition. *Lect. Notes Comp Sci* **5666**, 1–3 (2009)
39. Pennisi, M., Pappalardo, F., Palladini, A., Nicoletti, G., Nanni, P., Lollini, P.L., Motta, S.: Modeling the competition between lung metastases and the immune system using agents. *BMC Bioinformatics* **11**(S7), S13 (2010)
40. Pennisi, M., Pappalardo, F., Zhang, P., Motta, S.: Searching of optimal vaccination schedules: application of genetic algorithms to approach the problem in cancer immunoprevention. *IEEE Eng. Med. Biol. Magazine* **28**(4), 67–72 (2009)
41. Rapin, N., Lund, O., Bernaschi, M., Castiglione, F.: Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. *PLoS ONE* **5**(4), e9862 (2010)
42. Rosenberg, S.A., Yang, J.C., Restifo, N.P.: Cancer immunotherapy: moving beyond current vaccines. *Nat. Med.* **10**, 909–915 (2004)
43. Santoni, D., Pedicini, M., Castiglione, F.: Implementation of a regulatory gene network to simulate the TH1/2 differentiation in an agent-based model of hyper-sensitivity reactions. *Bioinformatics* **24**(11), 1374–1380 (2008)
44. Schwartz, R.H.: T cell anergy. *Ann. Rev. Immunol.* **21**, 305–334 (2003)
45. Segovia-Juarez, J.L., Ganguli, S., Kirschner, D.: Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection using an agent-based model. *J. Theo. Biol.* **231**(3), 357–376 (2004)
46. Slamon, D.J., Godolphin, W., Jones, L.A., Holt, J.A., Wong, S.G., Keith, D.E., Levin, W.J., Stuart, S.G., Udove, J., Ullrich, A.: Studies of the HER-2/neu protooncogene in human breast and ovarian cancer. *Science* **244**, 707–712 (1989)
47. Ursini-Siegel, J., Schade, B., Cardiff, R.D., Muller, W.J.: Insights from transgenic mouse models of ERBB2-induced breast cancer. *Nat. Rev. Cancer* **7**, 389–397 (2007)
48. Wolfram, S.: *A New Kind of Science*. Wolfram Media, Champaign, Illinois, USA (2002)
49. Zhang, X., Mosser, D.M.: Macrophage activation by endogenous danger signals. *J. Pathol.* **214**, 161–171 (2008)

On the Dynamics of Tumor-Immune System Interactions and Combined Chemo- and Immunotherapy

Alberto d’Onofrio, Urszula Ledzewicz, and Heinz Schättler

Abstract Tumor-immune system interplay is extremely complex, and, as such, it represents a big challenge for mathematical oncology. Here we investigate a simple general family of models for this important interplay by considering both the delivery of a cytotoxic chemotherapy and of immunotherapy. Then methods of geometrical optimal control are applied to a special case (the Stepanova model) in order to infer (under suitable constraints) the best combination of drugs scheduling to transfer – through therapy – the system from an initial condition in the malignant region of the state space into a benign region. Our findings suggest that chemotherapy is always needed first to reduce a large tumor volume before the immune system can become effective.

1 Introduction

Tumors are a family of high-mortality diseases, each different from the other, but all exhibiting a derangement of cellular proliferation that often leads to uncontrolled cell growth. Tumor cells (TCs) are characterized by a vast number of genetic and epigenetic events leading to the appearance of specific antigens (e.g., mutated proteins,

A. d’Onofrio

Department of Experimental Oncology, European Institute of Oncology, Via Ripamonti 435, 20141 Milan, Italy

e-mail: alberto.donofrio@ieo.eu

U. Ledzewicz (✉)

Department of Mathematics and Statistics, Southern Illinois University Edwardsville, Edwardsville, IL 62026-1653, USA

e-mail: uledzew@siue.edu

H. Schättler

Department of Electrical and Systems Engineering, Washington University, St. Louis, MO 63130-4899, USA

e-mail: hms@wustl.edu

A. d’Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_13, © Springer-Verlag Italia 2012

under- or over-expressed normal proteins and many others) triggering reactions by both the innate and adaptive immune system (IS) [17, 27, 36]. These observations have provided a theoretical basis to the old empirical hypothesis of *immune surveillance*, i.e., that the immune system may act to eliminate tumors, only recently experimentally and epidemiologically confirmed [11]. Of course, the competitive interaction between tumor cells and the immune system involves a considerable number of events and molecules, and as such is extremely complex and the kinetics of the interplay is strongly nonlinear. Moreover, to fully describe the immuno-oncologic dynamics, one has to take into account a range of spatial phenomena. Indeed, the tumor-immune system interplay is strongly shaped by the motility of both tumor cells and the effector cells of the immune system [25]. The outcome of this interplay is not only constituted by either tumor suppression or tumor outbreak, but there exist many intermediate scenarios. Indeed, the vast majority of theoretical immuno-oncology studies predict that a dynamic equilibrium may be established that allows the tumor to survive in a microscopic (and undetectable) dormant state [6, 8, 25]. This was largely inferred from clinical data. However, recently this theoretical prediction was directly confirmed by Koebel and coworkers [19], who were able to experimentally show, through an ad hoc mouse model, that adaptive immunity can maintain occult cancer in an equilibrium state. It is quite intuitive that this equilibrium can be disrupted by sudden events affecting the immune system. Indeed, if disease related impairments of innate and adaptive immune systems or immuno-suppressive treatments preceeding organ transplantations occur, then the tumor may restart developing. This has experimentally been shown both by mouse models and through epidemiologic studies [11, 35]. There also is a major class of causes of disruption of the equilibrium that are not related to immuno-suppression, but to adaptive processes called *immuno-editing* [11]. Over a long period of time, the neoplasm may develop multiple strategies to circumvent the action of the immune system [11, 27] that may allow it to re-commence growing [8, 11] into clinically apparent tumors [19], which theoretically can reach their carrying capacity [8].

Overall, tumor-immune system interactions exhibit a multitude of dynamic properties that include multi-stability, i.e., persistence of both benign and malignant scenarios. Unmodeled outside effects such as immuno-suppressive treatments or immuno-editing may move the current state (comprising cancer and immune cells) across the boundary between these regions of attraction and in effect turn a benign situation into a malignant one.

On the other hand, an interesting therapeutic approach is the immunotherapy [28], consisting in stimulating the immune system in order to better fight, and hopefully eradicate, a cancer. Here we will only be considering generic immunostimulations, for example via cytokines. The basic idea of immunotherapy is simple and promising, but the results obtained in medical investigations are globally controversial [1, 15], even if in recent years there has been evident progress. As far as the mathematical description of tumor and immune system interaction is concerned, many works have appeared using deterministic [5]–[9], [20, 21, 25, 34, 37] or stochastic [4, 10] models, as well as models introduced by N. Bellomo based on kinetic theories of nonlinear statistical mechanics [2].

In this paper, we describe a metamodel proposed by d’Onofrio for tumor-immune system interactions and then formulate treatment by cytotoxic agents and immunostimulations as an optimal control problem. Based on the biological background and the dynamic properties of the resulting system, we formulate a general mathematical objective function that expresses the underlying aim of moving initial conditions that lie in the malignant region of the state space (characterized through a macroscopic equilibrium or unbounded tumor growth) into its benign region (characterized by a microscopic equilibrium point). We then give a couple of numerical illustrations of optimal solutions for a modification of the classical model by Stepanova [34].

2 A Metamodel for Tumor-Immune System Interactions

In this section, a metamodel for tumor-immune interactions will be developed and an analysis of the tumor-free equilibrium will be given.

2.1 Classical Mathematical Models of Tumor-Immune System Interactions

The analysis of classical finite dimensional models for tumor-immune system interactions [18, 21, 34, 37], reveals the following dynamical features:

1. the existence of a tumor-free equilibrium;
2. the possibility that a tumor grows to a macroscopic size or unboundedly;
3. the possible coexistence of small or tumor-free and large tumor-size locally stable equilibria (bistability);
4. a constant influx of effector cells;
5. a varying proliferative profile of effector cells that depends on the tumor burden;
6. tumor-size dependent death rates of effectors;
7. the possible existence of limit cycles.

From this rough summary, one may understand that the puzzling results obtained up to now by immunotherapy [1] may be strictly linked to the complex dynamical properties of the immune system-tumor competition. In general, it is possible that a cancer-free or microscopic equilibrium coexists with other stable equilibria or with unbounded growth, so that the success of treatment or cure very much depends on the initial conditions.

Summarizing and generalizing the above cited biological features and models, in [6] d’Onofrio introduced the following family of models [A]:

$$\dot{x} = x(f(x) - \phi(x)y), \quad (1)$$

$$\dot{y} = \beta(x)y - \mu(x)y + \sigma q(x) + \theta(t), \quad (2)$$

with x and y denoting the “sizes” (given by the total number, cell volumes, densities or adimensional quantities) of, respectively, tumor cells and effector cells (ECs) of the immune system.

It is assumed that $0 < f(0) \leq +\infty$ and that f is nonincreasing, $f'(x) \leq 0$. In some relevant cases we shall suppose that $\lim_{x \rightarrow 0^+} xf'(x) = 0$ and that there exists a positive carrying capacity K , i.e., $f(K) = 0$. Equations f of this type incorporate widely used models of tumor growth rates such as exponential, logistic, generalized logistic or Gompertzian growth models and many more (e.g., see [13, 14, 22]). The functional parameters are given by differentiable functions that satisfy the following qualitative properties: ϕ is positive and decreasing, $\phi(0) = 1$, and $x\phi(x) \rightarrow \ell \leq +\infty$; β is nonnegative, monotonically nondecreasing with $\beta(0) = 0$. In the model, it is assumed that there is a tumor-induced loss of effector cells [40] which is described by a strictly increasing function μ . The function $q = q(x)$ satisfies $q(0) = 1$ and it may be nonincreasing or also initially increasing and then decreasing. Thus, we may assume that the growth of the tumor either decreases the influx of immune cells or that, on the contrary, it actually initially stimulates this influx before leading to an inhibitory pattern. In fact, for some cases of cancer it has been observed experimentally that progression may cause generalized immunosuppression which is reflected in an assumption of the type $q'(x) < 0$ for $x \gg 1$; for example, see [32] and the references therein. Finally, the term $\theta(t)$ is a positive function that describes the influx of effector cells from the primary organs and will also be used to model immunostimulation.

This metamodel includes, as particular cases, the ones by Stepanova [34], Kuznetsov, Makalkin, Taylor and Perelson [21] and de Vladar and Gonzalez [37]. For instance, for the Stepanova model [34] we have that tumor growth is exponential, $f(x) \equiv \alpha = \text{const}$ and

$$\phi(x) \equiv 1, \quad \beta(x) = \beta_1 x, \quad \mu(x) = \mu_0 + \mu_2 x^2 \quad \text{and} \quad q(x) \equiv 1. \quad (3)$$

The terms in the model by de Vladar and Gonzalez [37] are similar except that the tumor growth is Gompertzian, $f(x) = \alpha \log(K/x)$, [26]. In the model by Kuznetsov et al. [21], a logistic tumor growth model is used, $f(x) = \alpha(1 - x/K)$, and the functions are given by

$$\phi(x) \equiv 1, \quad \beta(x) = \frac{\beta_\infty x}{m + x}, \quad \mu(x) = \mu(0) + \mu_1 x, \quad \text{and} \quad q(x) \equiv 1. \quad (4)$$

2.2 A Generalised Model [B]

While the model [A] provides a general start, to make it biologically more accurate, it needs to be extended. A first modification involves the uptake rate of T-cells. We shall allow this function to be a nonmonotone function of the tumor burden and a nonlinear function of the number of effector cells. The latter feature is considered to take into account possible cooperative and/or competitive effects between effector cells. This leads to the following equation for the dynamics of tumor cells:

$$\dot{x} = x(f(x) - \phi(x)\pi(y)),$$

where now the predation function $\pi(y)$ no longer is linear, but is a generic strictly increasing growth function, $\pi'(y) > 0$. The overall functional response of tumor cells, $U(x) = x\phi(x)$, can be non-monotone. The other change concerns the proliferation of the cells of the immune system in response to tumors. It is reasonable to assume that the proliferation of effector cells depends on y . Apart from crowding effects, indeed, the proliferation might depend on the relative abundance of the two competing populations [12]. Taking this into account leads to equations of the form [B]:

$$\dot{x} = x(f(x) - \gamma\xi(t) - \phi(x)\pi(y)), \quad (5)$$

$$\dot{y} = (P(x, y) - \mu(x) - \eta\xi(t))y + \sigma q(x) + \theta(t). \quad (6)$$

Here we included the possibility of the delivery of cytotoxic chemotherapy with blood profile $\xi(t)$ that may also affect the immune system effector cells. For sake of simplicity, in this section we take $\xi(t)$ constant and such that $0 \leq \xi < f(0)/\gamma$. The proliferation rate $P = P(x, y)$ satisfies $P(0, y) = 0$, is strictly increasing in x and nonincreasing in y . For example, Forys et al. [12] propose the following functional relationship depending on the ratio x/y :

$$P(x, y) = \beta_\infty \frac{\left(\frac{x}{y}\right)^\alpha}{1 + \left(\frac{x}{y}\right)^\alpha} = \beta_\infty \frac{x^\alpha}{x^\alpha + y^\alpha}. \quad (7)$$

2.3 Equilibria and Null-Clines of Model [B]

The equilibria of (5) and (6) are determined by the equations $\dot{x} = 0$ and $\dot{y} = 0$. The equation $\dot{x} = 0$ yields $x = 0$ (which we will examine later), or

$$f(x) - \gamma\xi - \phi(x)\pi(y) = 0,$$

which implies that y is given in the form

$$y = Y_C(x; \xi) = \pi^{-1}(y_C(x; \xi)), \quad (8)$$

with

$$y_C(x; \xi) = (f(x) - \gamma\xi)/\phi(x). \quad (9)$$

We call $Y_C(p; \xi)$ the C null-cline and call $y_C(p; \xi)$ the baseline C-curve. This is the curve that coincides with the C null-cline in the case $\pi(y) = y$. Note that $\frac{\partial y_C}{\partial \xi}(x; \xi) < 0$, and that one can rewrite the tumor-dynamics equation in the following, technically useful form:

$$\dot{x} = U(x)(y_C(x; \xi) - \pi(y)). \quad (10)$$

Clearly, if

$$\pi(\infty) \geq \max_{x \in [0, K]} y_C(x; \xi), \quad (11)$$

then $Y_C(x; \xi)$ is bounded; otherwise it has vertical asymptotes where $\pi_\infty = y_C(x; \xi)$. For sake of simplicity, we assume that (11) holds.

The equation $\dot{y} = 0$ yields the I-nullcline $Y_I(x; \theta, \xi)$ determined by the equation

$$P(x, y) = \mu(x) + \eta\xi - \frac{\theta + \sigma q(x)}{y}. \quad (12)$$

If $\lim_{y \rightarrow \infty} P(x, y) = 0$, then Eq. (12) has a unique solution y in an interval $(y^*(x), \infty)$ and there exists no $\hat{x} < \infty$ such that

$$\lim_{x \rightarrow \hat{x}} Y_I(x; \theta, \xi) = \infty. \quad (13)$$

The boundedness of $Y_I(x)$ also holds if we relax the assumptions to the less restrictive condition

$$\lim_{y \rightarrow \infty} P(x, y) < \mu(x) \quad \text{for all } x > 0. \quad (14)$$

Thus, defining the function

$$\Psi_\infty(x) = \mu(x) - \lim_{y \rightarrow \infty} P(x, y), \quad (15)$$

if $\Psi_\infty(x)$ is strictly positive, then the I null-cline is bounded (biologically, the tumor may be classified as “lowly immunogenic” or “highly aggressive”). On the other hand, when $\Psi_\infty(x)$ has varying signs, then the I null-cline is unbounded (and the tumor may be classified as “highly immunogenic” or “lowly aggressive”). It is easy to show that the effect of a constant immunotherapy on the I null-cline is to increase it ($\frac{\partial Y_I}{\partial \theta}(x; \theta, \xi) > 0$), whereas the effect of a constant chemotherapy is to decrease it ($\frac{\partial Y_I}{\partial \xi}(x; \theta, \xi) < 0$).

Single or multiple equilibria may be determined by the intersections of the null-clines in $\mathbb{R}_+^2$. Stability properties of these equilibria can easily be related to the properties of the functions $Y_C(x; \xi)$ and $Y_I(x; \theta, \xi)$:

Proposition 1. *An equilibrium point $E = (x_e(\theta, \xi), y_e(\theta, \xi))$ of the system (5) and (6) is locally asymptotically stable if*

$$Y'_I(x_e; \theta, \xi) > Y'_C(x_e(\theta, \xi); \xi) \quad \text{and} \quad U(x_e)\pi'(y_e)Y'_C(x_e; \xi) < -\frac{\partial B}{\partial y}(x_e, y_e). \quad (16)$$

2.4 Tumor-Free Equilibrium

In absence of therapy, there is a disease-free equilibrium at

$$TF = \left(0, \frac{\sigma q(0)}{\mu(0)}\right), \quad (17)$$

which for $\sigma = 0$ degenerates to $(0, 0)$.

Proposition 2. *In the absence of immunotherapy (i.e., for $\theta = 0$), if $\sigma > 0$ and $Y_I(0) > Y_C(0; \xi)$, i.e., if*

$$\frac{\sigma q(0)}{\mu(0)} > \pi^{-1}(y_C(0; \xi)), \quad (18)$$

then the tumor-free equilibrium TF is locally asymptotically stable. If $\sigma = 0$, then the degenerate tumor-free equilibrium $TF = (0, 0)$ is always unstable with the y -axis the stable manifold.

There are examples of observations of persistent oscillations in oncological diseases in the medical literature [16, 18, 38]. In [6], d'Onofrio studied the limit cycle dynamics present in the case of $q(x) \geq 0$, and (in absence of chemotherapy) in [6, 7] it was shown that a necessary condition for the presence of periodic or other closed orbits is that $y'_C(x)$ has varying sign. These results can also be applied in the case of co-presence of constant chemo- and immunotherapies:

Proposition 3. *Suppose there exists exactly one equilibrium point $EQ = (x_e, y_e)$ other than TF . If $y_C(x; \xi)$ is decreasing, then there are no closed orbits. If EQ is unstable with*

$$U(x_e)\pi'(y_e)Y'_C(x_e; \xi) > -\frac{\partial B}{\partial y}(x_e, y_e; \theta, \xi),$$

then there exists at least one locally asymptotically stable limit cycle.

For the tumor-free equilibrium TF under constant therapy θ , there exists a value $\theta_{LAS}(\xi) \geq 0$ such that for $\theta > \theta_{LAS}(\xi)$ the tumor-free equilibrium TF is locally asymptotically stable. It is thus reasonable to investigate if there is a second value $\theta_{GAS}(\xi) \geq \theta_{LAS}(\xi) \geq 0$ such that for $\theta(\xi) > \theta_{GAS}(\xi)$, the disease-free equilibrium becomes globally asymptotically stable. Concerning this point, here we give the following proposition on the global eradicability of a tumor, under time-varying chemo- and immunotherapies, which extends the slightly less general result in Proposition 6.1 of [7]:

Proposition 4. *If $\theta(t) \geq 0$ and $\xi(t) \geq 0$ are such that*

$$Y_I(x; \theta_{\min}, \xi_{\min}) > Y_C(x; \xi_{\min}), \quad (19)$$

where

$$\theta_{\min} = \min_{t \geq 0} \theta(t) \quad \text{and} \quad \xi_{\min} = \min_{t \geq 0} \xi(t),$$

then TF is globally asymptotically stable for nonnegative initial conditions, i.e., the tumor is eradicated independently of the initial tumor burden and state of the immune system.

3 Optimal Control for Mathematical Models of Tumor-Immune System Interactions

In view of harmful side effects of both cytotoxic drugs and also of drugs that boost the immune system, it is not feasible to consider indefinite administrations of agents. Thus, and proceeding beyond models of constant blood profiles, the practically important question becomes how are therapeutic agents best administered. The goal is to transfer an initial condition (x_0, y_0) that lies in the region of malignant cancer growth for the uncontrolled system into the region of attraction of the stable, benign (microscopic) equilibrium point thus controlling the cancer volume. In an optimal control formulation, the aim is to achieve this goal in an efficient and effective way.

Intuitively, such a transfer requires to minimize the cancer cells x while not depleting the effector cell density y too strongly. The boundary between the benign and malignant regions consists of the stable manifolds of unstable equilibria. For most models, such as the classical version of Stepanova's model [34] or its extension by de Vladar and Gonzalez [37], there exists a unique saddle point whose stable manifold defines this boundary, the so-called separatrix. In general, it is rather difficult to give analytic descriptions of these manifolds, but their tangent spaces are readily computed and provide a first approximation for the separatrix. This motivates to tailor a term in the mathematical objective to be minimized that induces the system to move across this boundary. The underlying system is two-dimensional and thus the saddle has a unique stable eigenvector $v_s = (B, A)^T$. It follows from the underlying geometry that typically these coefficients are positive. Thus, by including a penalty term $Ax(T) - By(T)$ at the final time, this not only conforms with the heuristic notion of minimizing cancer cells and maximizing effector cells, but in a very precise mathematical way justifies the chosen weights by providing an incentive for the system to move across the separatrix into the benign region.

Furthermore, side effects of the treatment need to be taken into account. The model above, in a first approximation, does not incorporate healthy cells and tissue and thus side effects are only modeled indirectly. It is assumed that chemo- and immunotherapies have a proportional effect on healthy tissue and thus in our formulation we add weighted integral terms $\int_0^T u(t)dt$ and $\int_0^T v(t)dt$ that measure the total amounts given of a cytotoxic agent u and an immunostimulator v , respectively. Side effects of the immune boosts are less severe, but generally cannot be ignored. Thus both integrals are included in the objective with proper weights. Alternatively, based on medical expertise, these amounts could be limited a priori and then a minimization problem subject to isoperimetric constraints

$$\int_0^T u(t)dt \leq U_{\max} \quad \text{and} \quad \int_0^T v(t)dt \leq V_{\max}$$

might be considered.

Finally, we keep the terminal time T in our problem formulation free. However, because of the existence of a tumor-free or microscopic benign equilibrium, it is pos-

sible – and this indeed is observed in simple models [23, 24] – that trajectories use the zero controls over very long time horizons. These trajectories, which improve the value of the mathematical objective without incurring any cost, provide a “free pass,” but lead to a mathematically ill-posed problem formulation as no minimum may exist in this case. Indeed, the infimum arises as the control switches to follow $u = 0$ and $v = 0$ when the controlled trajectory intersects the separatrix, then follows the separatrix for an infinite time to the saddle and then again leaves this saddle point along the unstable manifold, once more taking an infinite time. This indeed would be the “optimal” solution for this problem formulation, but it is not an admissible trajectory in our system. In order to avoid this scenario, we include a penalty term on the final time as well. From a biological point of view, the addition of this term induces optimal solutions to give more drugs and thus reach the benign equilibrium point faster rather than take a very long time with a smaller amount of agents. Clearly, it is not desirable for the system to evolve along the border between benign and malignant tumor behavior and it is this term that forces the system into the safer benign region more quickly. In view of imprecise and mathematically unmodeled dynamics and other random perturbations, from a system theoretic perspective, the addition of this term provides desired robustness and stability properties for the underlying real system.

Summarizing, we formulate a mathematical *objective* of the following form that consists of a weighted average of

1. the penalty term $Ax(T) - By(T)$ which induces the system to move across the separatrix from the malignant into the benign region of the state space;
2. the cumulative side effects of the chemotherapeutic agent and the immune boost taken as an indirect measure of the side effects of treatment;
3. a small penalty term on the terminal time T that makes the problem mathematically well-posed:

$$J(u, v) = Ax(T) - By(T) + C \int_0^T u(t)dt + D \int_0^T v(t)dt + ST.$$

As explained earlier, A and B are positive coefficients determined by the stable eigenvector $v_s = (B, A)^T$ of the saddle and C , D and S are positive weights. We emphasize that the latter coefficients are variables of choice and may be chosen to calibrate the response of the system. The choice of the weights aims at striking a balance between the benefit at the terminal time T , $Ax(T) - By(T)$, and the overall side effects measured by the total amount of drugs given while it guarantees the existence of an optimal solution by also penalizing the free terminal time T . The integrals of the dose rates model the side effects of the therapies on the healthy tissue and if there exist clinical data as to the severity of the drugs, then this will be reflected in the choices for C and D . Naturally, also the specific type of tumor and even the stage of cancer the patient has, will enter into the calibration of these coefficients. For the case of a more advanced stage, higher side effects need to be tolerated and thus smaller values

of C and D will be taken. Overall, the coefficients C , D and S are variables of choice that can be fine tuned to calibrate the system's optimal response.

Thus, for a system with tumor-immune interactions, the following optimal control problem in Bolza form arises:

[OC] for a free terminal time T , minimize the objective

$$J(u, v) = Ax(T) - By(T) + \int_0^T (Cu(t) + Dv(t) + S)dt, \quad (20)$$

over all Lebesgue measurable functions

$$(u, v) : [0, T] \rightarrow [0, 1] \times [0, 1], \quad t \mapsto (u(t), v(t))$$

subject to the dynamics (1) and (2) respectively (5) and (6).

It is clear that for any correct modeling of the problem we must have that the states x and y remain positive for positive initial conditions x_0 and y_0 and arbitrary admissible controls u and v . Since $x = 0$ is an equilibrium solution of (1) and (5), the variable x cannot cross 0 and if $y = 0$, then because of the influx of effector cells that is part of the positive function $\theta(t)$, we always have that $\dot{y} > 0$. Thus the region $P = \{(x, y) : x > 0, y > 0\}$ is positively invariant and no state-space constraints need be imposed on the variables.

4 Optimal Controls for a Modified Stepanova Model

The structures of optimal controls no longer just depend on the qualitative features that are incorporated into the general models [A] and [B], but on the specific choices of the functional parameters such as the tumor growth model. In this section, we give two examples of the structure of optimal controls for a modified version of the classical model by Stepanova [34] when exponential tumor growth has been replaced with a Gompertzian model. Specifically, we consider the following equations:

$$\dot{x} = -\mu_C x \ln\left(\frac{x}{x_\infty}\right) - \gamma xy - \kappa_X x u, \quad x(0) = x_0, \quad (21)$$

$$\dot{y} = \mu_I (x - \beta x^2) y - \delta y + \alpha + \kappa_Y y v, \quad y(0) = y_0. \quad (22)$$

The controls u and v denote the blood profiles of a cytotoxic agent and a generic immune stimulation agent, respectively; α is a positive constant that models a constant influx of effector cells from the primary organs.

The uncontrolled system (i.e., for $u = 0$ and $v = 0$), has two locally asymptotically stable equilibria, one microscopic, the other macroscopic, with the regions of attraction separated by the stable manifold of an intermediate saddle point. These dynamic properties persist under full dose immunotherapy if no chemotherapy is

Table 1. Numerical values for the variables and parameters used in computations

<i>Variable/ parameters</i>	<i>Interpretation</i>	<i>Numerical value</i>	<i>Dimension</i>	<i>Ref.</i>
x	tumor volume		10^6 cells	[34]
x_0	initial value for x	600	10^6 cells	
y	immuno-competent cell density		non-dimensional	[34]
y_0	initial value for y	0.10	non-dimensional	
α	rate of influx	0.1181	1/day	[21]
β	inverse threshold for tumor suppression	0.00264	non-dimensional	[21]
γ	interaction rate	1	10^7 cells/day	[21]
δ	death rate	0.37451	1/day	[21]
μ_C	tumor growth parameter	0.5599	10^7 cells/day	
μ_I	tumor stimulated proliferation rate	0.00484	1/day	
x_∞	fixed carrying capacity	780	10^6 cells	
κ_x	chemotherapeutic killing parameter	1	10^7 cells/day	

used and thus, for some initial conditions, immunotherapy alone is not sufficient to eliminate the cancer. Fig. 1 shows the phase portraits for the parameter values given in Table 1 for a constant full dose control $v \equiv 1$ and $u \equiv 0$. The stable manifold of the saddle at $(x_s, y_s) = (555.1, 0.191)$ still separates a region where the immune system aided by the immune boost can eliminate the cancer (the y -values of the system approach $+\infty$ while x converges to 0 from the right) from a region where the cancer eventually will dominate and trajectories converge to the asymptotically stable malignant equilibrium point $(x_m, y_m) = (715.6, 0.048)$. Any initial condition that lies to the right of the stable manifold of the saddle cannot be transferred into the benign region with immunotherapy alone.

Necessary conditions for optimality of the controls u and v are given by the *Pontryagin maximum principle* (for some recent texts, see [3, 30]). For a 2-dimensional row-vector $\lambda = (\lambda_1, \lambda_2)$, define the Hamiltonian $H = H(\lambda, x, y, u, v)$ as

$$H = (Cu + Dv + S) + \lambda_1 \left(-\mu_C x \ln \left(\frac{x}{x_\infty} \right) - \gamma xy - \kappa_x xu \right) \\ + \lambda_2 (\mu_I (x - \beta x^2) y - \delta y + \alpha + \kappa_Y y v).$$

(23)

If (u_*, v_*) is an optimal control defined over an interval $[0, T]$ with corresponding trajectory $z_* = (x_*, y_*)^T$, then, by the Pontryagin maximum principle, there exists

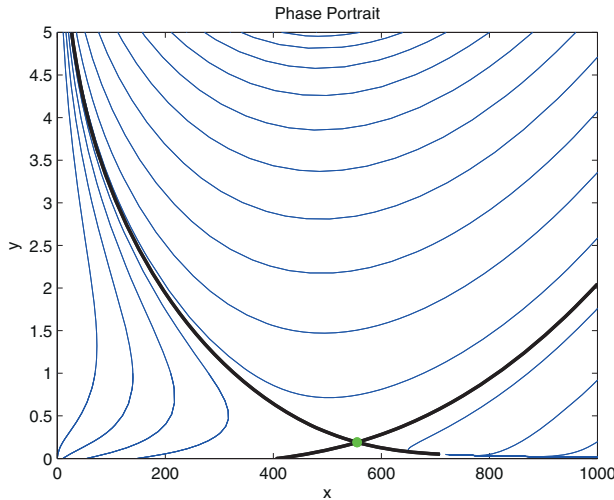


Fig. 1. Phase portrait of the controlled system (21) and (22) for $u = 0$ and $v = 1$

an absolutely continuous covector $\lambda = (\lambda_1, \lambda_2)$ defined on $[0, T]$, such that the following conditions hold:

(a) λ_1 and λ_2 satisfy the adjoint equations

$$\dot{\lambda}_1 = -\frac{\partial H}{\partial x} = \lambda_1 \left(\mu_C \left(1 + \ln \left(\frac{x}{x_\infty} \right) \right) + \gamma y + u \kappa_X \right) - \lambda_2 \mu_I (1 - 2\beta x) y, \quad (24)$$

$$\dot{\lambda}_2 = -\frac{\partial H}{\partial y} = \lambda_1 \gamma x - \lambda_2 (\mu_I (x - \beta x^2) - \delta + \kappa_Y v), \quad (25)$$

with terminal conditions $\lambda_1(T) = A$ and $\lambda_2(T) = -B$, and

(b) for almost every time $t \in [0, T]$, the optimal controls $(u_*(t), v_*(t))$ minimize the Hamiltonian H along $(\lambda(t), x_*(t), y_*(t))$ over the control set $[0, 1] \times [0, 1]$ with minimum value given by 0.

A *controlled trajectory* $((x, y), (u, v))$, consisting of admissible controls (u, v) and corresponding solution (x, y) of the initial value problem (21) and (22), for which there exists a multiplier λ such that the conditions of the maximum principle are satisfied is called an *extremal* (pair) and the triple $((x, y), (u, v), \lambda)$ is an extremal lift (to the cotangent bundle). More generally, in an optimal control problem there is another constant multiplier λ_0 associated with the objective, but in our case it can always be normalized to $\lambda_0 = 1$ and we already have anticipated this in our formulation.

By condition (b), the optimal controls $u_*(t)$ and $v_*(t)$ minimize the Hamiltonian H along the extremal $(\lambda(t), x_*(t), y_*(t))$ over the control set $[0, 1] \times [0, 1]$ a.e. on $[0, T]$. Since H is linear in the controls and the control set is a rectangle, these minimizations

decouple and can be carried out separately. Defining the *switching function* $\Phi_1 = \Phi_1(t)$ for u as

$$\Phi_1(t) = C + \langle \lambda(t), g_1(z_*(t)) \rangle = C - \lambda_1(t) \kappa_X x_*(t), \quad (26)$$

it follows that

$$u_*(t) = \begin{cases} 0 & \text{if } \Phi_1(t) > 0, \\ 1 & \text{if } \Phi_1(t) < 0, \end{cases} \quad (27)$$

and with $\Phi_2 = \Phi_2(t)$ the switching function for v ,

$$\Phi_2(t) = D + \langle \lambda(t), g_2(z_*(t)) \rangle = D + \lambda_2 \kappa_Y y_*(t), \quad (28)$$

the optimal control v_* satisfies

$$v_*(t) = \begin{cases} 0 & \text{if } \Phi_2(t) > 0, \\ 1 & \text{if } \Phi_2(t) < 0. \end{cases} \quad (29)$$

We refer to the constant controls given by the extreme values 0 and 1 as *bang* controls. The minimum conditions by themselves do not determine the controls at times when the switching functions vanish, $\Phi_i(\tau) = 0$. But if the time-derivative $\dot{\Phi}_i(\tau)$ does not vanish, then at such a time the control switches between its extreme values with the order depending on the sign of $\dot{\Phi}(\tau)$. Thus also the name of bang-bang controls. In the other extreme, if $\Phi(t) \equiv 0$ on an open interval I , then all derivatives of $\Phi(t)$ must vanish as well and typically this does allow to compute the control. Controls of the second kind are called *singular* [3, 30]. In principle, in a multi-input case, as it is considered here, it is even possible that both controls are singular at the same time and in this case the controls are called totally singular. Overall, optimal controls need to be synthesized from these classes of candidates.

In the papers [24] and [31], we have undertaken an extensive analysis of the local optimality of singular controls. For the monotherapy problem with chemotherapy as the only control, singular controls satisfy the Legendre-Clebsch condition, a second-order necessary condition for optimality, in its strengthened form away from the strip $\{(x, y) : \frac{1}{4\beta} \leq x \leq \frac{1}{2\beta}\}$ and indeed are locally optimal. Explicit formulas for the singular control and corresponding singular arcs are given in [24] and using these formulas, optimal controls have been computed numerically using GPOPS, a Gauss hp-adaptive pseudospectral method [29]. Below, in Fig. 2, we give a numerical illustration for this case.

Scenario 1: chemotherapy only. We used the numerical values given in Table 1 and the following coefficients in the objective:

$$A = 0.00192, \quad B = 1, \quad C = 0.08, \quad D = 0, \quad S = 0.05.$$

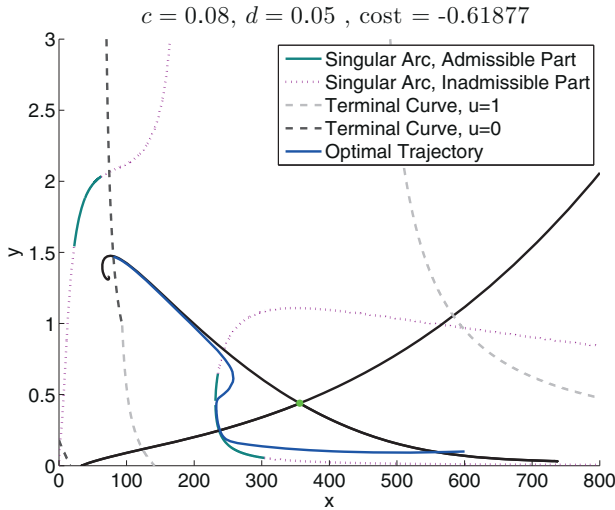


Fig. 2. Example of an optimal controlled trajectory of the form $1s0$ for the problem with chemotherapy only

Singular controls are only admissible along the specific curves in the state space shown as solid green segments in Fig. 2. The tumor volume is measured in multiples of 10^6 while the immunocompetent cell densities are on a scale relative to 1. The optimal trajectory starting at $(x_0, y_0) = (600, 0.1)$ initially administers cytotoxic agents at full dose until the corresponding trajectory (i.e., the response of the system) hits an admissible singular arc. At this time, the optimal control switches to a singular control and the trajectory follows the singular arc across the separatrix. Then, at a certain time τ the administration of chemotherapy ceases, i.e., the system follows the uncontrolled trajectory ($u = 0$) towards the benign equilibrium point. The optimal final time T is realized when the objective becomes minimized along this trajectory. Overall, a concatenation sequence for the control of the form $1s0$ results with the 1 indicating therapy at maximum dose, the s denoting a time-varying profile that follows a singular control and the 0 denoting no chemotherapy. Fig. 2 shows the resulting controlled trajectory along with the admissible and inadmissible portions of singular arcs identified.

Scenario 2: chemotherapy combined with immune boost. Fig. 3 gives the analogous phase portrait of an optimal controlled trajectory for a case when the immune boost v is included in the model. Again, the parameters used were the ones given in Table 1 and here the coefficients in the objective are

$$A = 0.00192, \quad B = 1, \quad C = 0.025, \quad D = 0.005, \quad S = 0.0005.$$

We have chosen the coefficients so that the side effects of chemotherapy are more severe than those of the immune boost. Thus, naturally one expects that the immune

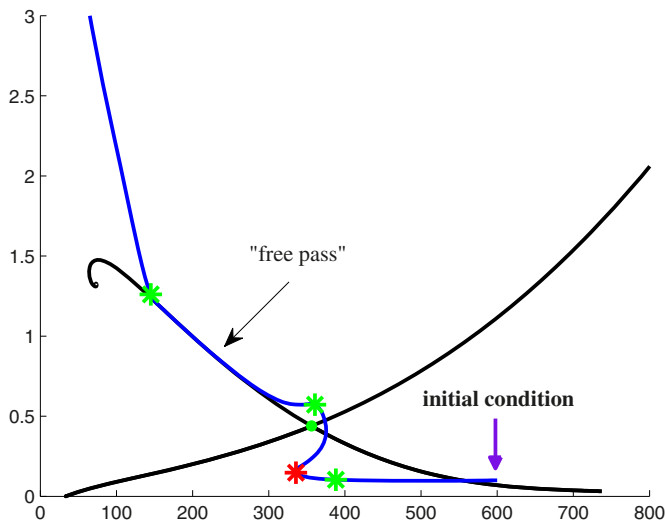


Fig. 3. Optimal controlled trajectory in Scenario 2. The asteriks indicate the points when switchings in the optimal controls occur (there is a switch in the control u for the chemotherapeutic agent at the second asterisk (from the initial condition) and there are switchings in the control v for the immune boost at all other asterisks). The curve gives the response of the system to the optimal controls

boost will play a significant role in this scenario. However, the initial condition still lies in the malignant region for the controls $u \equiv 0$ and $v \equiv 1$ and thus a constant immune boost in itself is not able to control the cancer volume. Thus, and although the cost is high, initially full dose chemotherapy must be given to move the state out of the region of attraction of this malignant equilibrium into the region where an immune boost can control the cancer volume. Once there, chemotherapy turns off (the red asterisk along the trajectory) and the immune boost (which starts at the first green asterisk counting from the initial condition) moves the state of the system into the benign region. Then, as in scenario 1, the uncontrolled trajectory (“free pass”) becomes the most cost-effective strategy (only paying a small penalty for the time spent along it) until, because of the low cost of the immune boost, it becomes once again beneficial to give an immune boost as the cancer volume becomes small.

Naturally, the structure of optimal controls depends on the coefficients selected for the objective. If the weights for the side effects of chemotherapy become less severe, then chemotherapy overall has the better effect and it remains the dominant therapy with the control structure similar to the one in the monotherapy case with chemotherapy only (e.g., see [31]). In fact, this is the typical scenario for a broad range of parameters and there exists only a rather small window of parameters when true interactions between the tumor immune system occur. Around this region, the underlying system is quite sensitive to parameters.

5 Conclusion

We outlined a metamodel for tumor-immune system interactions that incorporates various classical models from the literature (e.g., [21, 34, 37]) and becomes the basis for the problem of finding optimal procedures for administering cancer treatments. Mathematically, the problem to cure the patient becomes the problem to transfer the system from an initial condition in the malignant region of the state space through therapy into a benign region. In this paper, combinations of a chemotherapeutic agent and an immune boost are applied to achieve this goal in an optimal way in the sense of scheduling the dosages and their sequencing. Methods of optimal control are used to solve the problem and some examples of solutions are presented for a specific case of the metamodels presented, a modified version of the classical model by Stepanova. These solutions indicate chemotherapy will always be needed first to bring down a large tumor volume before the actions of the immune system can be become effective.

Acknowledgements This material is based upon research supported by the National Science Foundation under collaborative research grants DMS 1008209 and 1008221 (U.L. and H.S.), and by the EU project “p-Medicine: Personalized Medicine” (FP7-ICT-2009.5.3-270089) (A. d'O.). We also would like to thank our students Mohamad Naghnaeian and Mozhdeh Faraji for carrying out the numerical computations and making the figures used in the paper.

References

1. Agarwala, S.A. (Guest Editor): New Applications of Cancer Immunotherapy. *Sem. Oncol.* **29**(3), Special Issue, Suppl. 7 (2003)
2. Bellomo, N., Delitala, M.: From the mathematical kinetic, and stochastic game theory for active particles to modelling mutations, onset, progression and immune competition of cancer cells. *Phys. Life Rev.* **5**, 183–206 (2008)
3. Bonnard, B., Chyba, M.: Singular Trajectories and their Role in Control Theory. *Mathématiques & Applications*, vol. 40, Springer, Paris (2003)
4. Caravagna, G., d'Onofrio, A., Milazzo, P., Barbuti, R.: Antitumour Immune Surveillance Through Stochastic Oscillations. *J. Theor. Biol.* **265**, 336–345 (2010)
5. de Pillis, L.G., Radunskaya, A.E., Wiseman, C.L.: A validated mathematical model of cell-mediated immune response to tumor growth. *Cancer Res.* **65**, 7950–7958 (2005)
6. d'Onofrio, A.: A general framework for modelling tumor-immune system competition and immunotherapy: Mathematical analysis and biomedical inferences. *Physica D* **208**, 202–235 (2005)
7. d'Onofrio, A.: The role of the proliferation rate of effectors in the tumor-immune system competition. *Math. Mod. Meth. Appl. Sci.* **16**, 1375–1401 (2006)
8. d'Onofrio, A.: Tumor evasion from immune control: strategies of a MISS to become a MASS. *Chaos, Solitons and Fractals* **31**, 261–268 (2007)
9. d'Onofrio, A.: Metamodeling tumor-immune system interaction, tumor evasion and immunotherapy. *Math. Comput. Modelling* **47**, 614–637 (2008)
10. d'Onofrio, A.: Bounded-noise-induced transitions in a tumor-immune system interplay. *Phys. Rev. E* **81**, (2010), 021923 (2010)

11. Dunn, G.P., Old L.J., Schreiber, R.D.: The three ES of cancer immunoediting. *Ann. Rev. Immunol.* **22**, 322–360 (2004)
12. Forsys, U., Waniewski J., Zhivkov, P.: Anti-tumor immunity and tumor anti-immunity in a mathematical model of tumor immunotherapy. *J. Biol. Syst.* **14**, 13–30 (2006)
13. Guiot, C., Degiorgis, P.G., Delsanto, P.P., Gabriele, P., Deisboecke T.S.: Does tumor growth follow a “universal law”? *J. Theor. Biol.* **225**, 147–151 (2003)
14. Hart, D., Shochat, E., Agur, Z.: The growth law of primary breast cancer as inferred from mammography screening trials data. *Br. J. Cancer* **78**, 382–387 (1999)
15. Kaminski, J.M., Summers, J.B., Ward, M.B., Huber, M.R., Minev, B.: Immunotherapy and prostate cancer. *Canc. Treat. Rev.* **29**, (2004), 199–209 (2004)
16. Kennedy, B.J.: Cyclic leukocyte oscillations in chronic myelogenous leukemia during hydroxyurea therapy. *Blood* **35**, (1970), 751–760 (1970)
17. Kindt, T.J., Osborne, B.A., Goldsby, R.A.: *Kuby Immunology*. W.H. Freeman, New York (2006)
18. Kirschner, D., Panetta, J.C.: Modeling immunotherapy of the tumor-immune interaction. *J. Math. Biol.* **37**, 235–252 (1998)
19. Koebel, C.M., Vermi, W., Swann, J.B., Zerafa, N., Rodig, S.J., Old, L.J., Smyth, M.J., Schreiber, R.D.: Adaptive immunity maintains occult cancer in an equilibrium state. *Nature* **450**, 903–907 (2007)
20. Kogan, Y., Forsys, U., Shukron, O., Kronik, N., Agur, Z.: Cellular immunotherapy for high grade gliomas: mathematical analysis deriving efficacious infusion rates based on patient requirements. *SIAM J. Appl. Math.* **70**, (2010), 1953–1976 (2010)
21. Kuznetsov, V.A., Makalkin, I.A., Taylor, M.A., Perelson, A.S.: Nonlinear dynamics of immunogenic tumors: parameter estimation and global bifurcation analysis. *Bull. Math. Biol.* **56**, 295–321 (1994)
22. Ledzewicz, U., d’Onofrio, A., Schättler, H.: Tumor development under combination treatments with anti-angiogenic therapies. in: Ledzewicz, U., Schättler, H., Friedman, A., Kashdan, E. (eds.) *Mathematical Methods and Models in Biomedicine, Lecture Notes on Mathematical Modeling in the Life Sciences*, Vol. 1, pp. 301–327. Springer, Heidelberg (2012)
23. Ledzewicz, U., Naghnaeian, M., Schättler, H.: Dynamics of tumor-immune interactions under treatment as an optimal control problem. *Proc. of the 8th AIMS Conf.*, Dresden, Germany, pp. 971–980 (2010)
24. Ledzewicz, U., Naghnaeian, M., Schättler, H.: Optimal response to chemotherapy for a mathematical model of tumor-immune dynamics. *J. Math. Biol.* **64**, 557–577 (2012)
25. Matzavinos, A., Chaplain, M., Kuznetsov, V.A.: Mathematical modelling of the spatio-temporal response of cytotoxic T-lymphocytes to a solid tumour. *Math. Med. Biol.* **21**, (2004), 1–34 (2004)
26. Norton, L.: A Gompertzian model of human breast cancer growth. *Cancer Res.* **48**, (1988), 7067–7071 (1988)
27. Pardoll, D.: Does the immune system see tumors as foreign or self? *Ann. Rev. Immunol.* **21**, (2003), 807–839 (2003)
28. Peckham, M., Pinedo, H.M., Veronesi, U.: *The Oxford Textbook of Oncology*. Oxford University Press, Oxford (1995)
29. Rao, A.V., Benson, D.A., Huntington, G.T., Francolin, C., Darby, C.L., Patterson M.A.: User’s Manual for GPOPS: A MATLAB Package for Dynamic Optimization Using the Gauss Pseudospectral Method. University of Florida Report, <http://www.gpops.org> (2008)
30. Schättler, H., Ledzewicz, U.: *Geometric Optimal Control: Theory, Methods and Examples*. Springer, Heidelberg (2012)

31. Schättler, H., Ledzewicz, U., Faraji, M.: Optimal controls for a mathematical model of tumor-immune interactions under chemotherapy with immune boost. *Disc. Cont. Dyn. Syst. Ser. B*, (2013), to appear
32. Schmielau, J., Finn, O.J.: Activated granulocytes and granulocyte-derived hydrogen peroxide are the underlying mechanism of suppression of T-cell function in advanced cancer patients. *Cancer Res.* **61**, 4756–4760 (2001)
33. Skipper, H.E.: On mathematical modeling of critical variables in cancer treatment (goals: better understanding of the past and better planning in the future). *Bull. Math. Biol.* **48**, 253–278 (1986)
34. Stepanova, N.V.: Course of the immune reaction during the development of a malignant tumour. *Biophysics* **24**, 917–923 (1980)
35. Stewart, T.J., Abrams, S.L.: How tumours escape mass destruction. *Oncogene* **27**, 5894–5903 (2008)
36. Swann, J.B., Smyth, M.J.: Immune surveillance of tumors. *J. Clin. Inv.* **117**, 1137–1146 (2007)
37. de Vladar, H.P., González, J.A.: Dynamic response of cancer under the influence of immunological activity and therapy. *J. Theor. Biol.* **227**, 335–348 (2004)
38. Vodopick, H., Rupp, E.M., Edwards, C.L., Goswitz, F.A., Beauchamp, J.J.: Spontaneous cyclic leukocytosis and thrombocytosis in chronic granulocytic leukemia. *New Engl. J. Med.* **286**, (1972), 284–290 (1972)
39. Wheldon, T.E.: *Mathematical Models in Cancer Research*. Hilger Publishing, Boston-Philadelphia (1988)
40. Whiteside, T.L.: Tumor-induced death of immune cells: its mechanisms and consequences. *Sem. Canc. Biol.* **12**, 43–50 (2002)

Modeling the Kinetics of the Immune Response

Ami Radanskaya and Sarah Hook

Abstract In this chapter we develop a mathematical model of the immune response to a weak antigen, suitable for modeling a cancer vaccine. The parameters are calibrated to a murine model and the model is validated by comparing simulations to experimental results. The model is then used to develop a dosing strategy that optimizes the immune response.

1 Introduction

Vaccines have been the most successful preventative health care initiative of all time [8, 11]. However many diseases exist for which prophylactic vaccines are not available and with the increase in the numbers of multi drug resistant pathogens and the advent of extensively drug resistant bacteria, the use of vaccines for both disease prevention and treatment is becoming more important [1]. There is also the potential for vaccines to be used to treat and/or prevent non-communicable diseases such as cancer, autoimmunity and atopy [10].

Traditionally most immunization regimes have been optimized in order to stimulate high levels of antibody [13]. This is because for many diseases protection can be predominantly antibody mediated. However this is not the case for all diseases, for example cytotoxic immune responses are important for protection against intracellular bacteria and viruses [13]. Cellular and cytotoxic immune responses are also required for treating cancer. It is more difficult to optimized vaccination regimes aimed at stimulating cytotoxic immune responses as cytotoxic immune responses are much more difficult to measure than antibody titres [7]. Measuring cytotoxic immune responses may involve taking peripheral blood samples from subjects, puri-

A. Radanskaya (✉)

Pomona College, 333 North College Way, Claremont, CA 91711, USA

e-mail: aradanskaya@pomona.edu

S. Hook

Otago University, School of Pharmacy, 9054 Dunedin, NZ

e-mail: sarah.hook@stonebow.otago.ac.nz

fying lymphocytes and restimulating the cells in the laboratory to measure proliferation or cytokine production. The advent of MHC tetramer and pentamer technology has made the detection of cytotoxic lymphocytes directly *ex vivo* possible, however these cells are present only at very low levels and are often difficult to detect following immunisation [15]. Due to these problems, subjective markers for vaccine efficacy are often utilised [12]. In the case of therapeutic cancer vaccines, these markers may be such things as disease free survival or time to disease progression. However as these parameters do not provide an immediate read out of the efficacy of the immunization regime they provide little help for optimizing the therapy. Therefore often the only option available for researchers and clinicians is the “Goldilocks” approach to cancer vaccination: give not too much of the vaccine, or too little and give it not too often or too few times.

Therefore the aim of this work is to begin to establish a mathematical model of the T cell immune response to a vaccine with the goal that this can in the future be used to aid in the optimization of vaccine induced prophylactic and therapeutic cellular immune responses. The model will be established using data from mouse experiments where it is possible to directly measure cellular immune responses. Additionally in the mouse many parameters such as time to cell division, the rate of cell division and of cell death after exposure to antigen are known. Once the basic model is established the impact of the numerous factors that can affect the quality and quantity of an immune response generated by a vaccine will be simulated. This will include such factors as antigen dose, the duration of antigen persistence, the presence of immune potentiating agents and the number and timing of any booster immunizations.

2 The Mathematical Model

The kinetics of T-cell response to antigen determine the effectiveness of any vaccine therapy. In the case of cancer vaccines, when the immunogenicity of the tumor cells are typically low and memory cells as well as tumor-induced anti-immune response play crucial roles, understanding the details of these kinetics is especially important. A mathematical model can be a useful tool in predicting the response of sub-populations of T-cells to different vaccination strategies. In the spirit of parsimony, we begin by focusing on three populations: antigen presenting cells (APC), $CD4^+$ and $CD8^+$ T cells. In the course of an immune response, T-cell populations go through several phases: rapid expansion, rapid contraction, and differentiation into memory cell populations. These phases have been represented in different ways in previous models: either by changing the proliferation and death rates in the model at several “checkpoints” in time, as in [4], or by the addition of a time-delay in the activation term, as in [9]. We develop a model that combines these approaches.

We model two T-cell populations in the spleen: $CD8^+$ T cells and $CD4^+$ T cells. Each T cell population is further sub-divided into four sub-populations: naive T cells, N , active proliferating cells, P , highly apoptotic cells, A_p , active circulating cells, B , and memory cells, M . We also include two APC populations: the population in the

blood compartment, D_B , and the population in the spleen, D . The T-cell populations are all in the spleen compartment or exiting to circulate in the blood.

While the model described in [9] includes the kinetics of the distribution of APC's in the blood, spleen, and other organs following an intravenous injection of adoptively-transferred cells, we choose to focus for the present on the concentrations of cell types in the spleen. In other work we include a spatial component which will take into account the delivery of the vaccine as well as the migration of certain cell types and cytokines through tissue. This choice was motivated by the fact that, for the present, we are not tracking the antigen concentrations or the antigen-presenting cell populations, but rather we choose to focus on the dynamics of the immune response. We therefore use a very general model to capture the effect of antigen in the host: the amount of antigen in the blood at time t is a function with three parameters: π , roughly representing the duration of the administration of the vaccine, $\max_D$, representing the maximum level of antigen present in the blood, and μ_{DB} , representing the per-capita elimination rate of antigen from the bloodstream. The elimination rate is currently taken from the literature, [9], but will be measured experimentally for specific therapies in future studies. The parameters describing the maximum antigen level and the duration of the administration will be varied in order to predict the effect of maximum dosage and duration of vaccine administration on the kinetics of the immune response.

The model described in [9] for the evolution of the APCs and the T cells is:

$$\frac{dD}{dt} = \mu D_B(t) - \delta_D D - cAD \quad (1)$$

$$\frac{dA}{dt} = g(A) + b \frac{D(t-\tau)A(t-\tau)}{\theta + D(t-\tau)} - h(A)A + b_M DM. \quad (2)$$

In this formulation, $D(t)$ represents the number of APC's at time t , while $A(t)$ represents the number of activated effector cells at time t . $g(A)$ represents the rate of influx of the new activated effector cells into the population - explicitly given by $a_E(E^{naive} - A(t))$ in [9], δ_D represents the natural death rate of the APCs, c is the fraction of APC-T cell interactions that result in the destruction of a dendritic cell, and μ is the per-cell rate at which dendritic cells in the peripheral blood ($D_B(t)$) enter the spleen. The remaining terms involve the effect of the APCs on the number of activated effector cells: the second term in Eq. (2) gives the increased proliferation rate of the activated cells stimulated by the APCs. Note that this term has a saturation level: the parameter b gives the maximum per-cell proliferation rate; and it also includes a "steepness" coefficient, θ , which is the number of DC's at which the proliferation rate reaches half of its maximum value: $b/2$. This antigen-induced proliferation term also includes a time delay, τ , which represents the average amount of synaptic connection time necessary to provide an adequate co-stimulatory response. The third term, $h(A)$, is the rate at which activated cells leave the population. In [9], this term consists of a "quieting" rate, at which active cells become memory cells (no apoptosis is considered), and a term that describes the rate of recirculation of T cells into the blood. The last term represents activation of memory cells by dendritic cells.

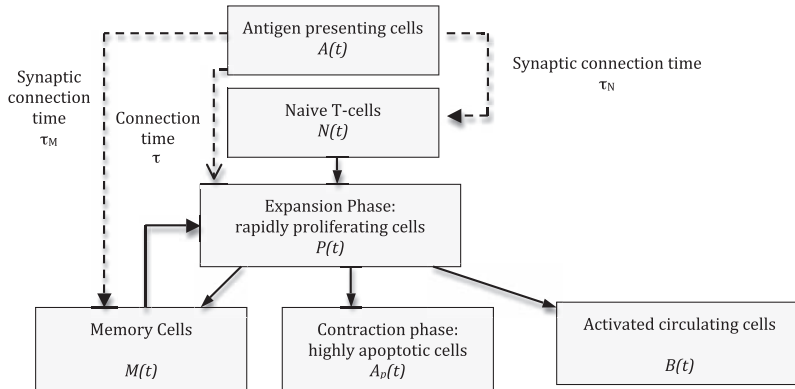


Fig. 1. Schematic of the flow from one immune cell sub-population to another

If we assume that resting, naive T cells are activated by APCs, and that only non-proliferating cells are responsible for APC inactivation, we can combine this model with the one given in [4] to explicitly model the T cell populations in the different phases. The transitions from one phase to another replaces the influx term, $g(A)$, in Eq. (2). The flow from compartment to compartment is shown schematically in Fig. 1. The homeostatic term, $h(A)$, is replaced by a death rate for cells in the rapidly proliferating, rapidly contracting or normally active phases. This death rate differs from phase to phase, reflecting the difference in apoptosis rates in the three phases. The memory cell sub-population contains a homeostatic term that reflects the observed long-term steady memory cell populations, despite relatively high cell turnover rates, [14]. While these steady states are thought to be controlled by cytokines produced by other cell types, the terms used here are the standard logistic growth functions, chosen for their simplicity and the fact that they imply self-regulating growth. We make one modification to the standard form of the logistic growth function, based on experimental observations that long-term memory cell populations depend on the size of the initial immune cell response, (e.g. [5, 6]). While the carrying capacity in the homeostatic term, k , is a constant, independent of immune cell response, the term is truncated so that it never becomes negative. This allows the initial growth in the memory cell population due to the primary immune response to affect the long-term population levels. Since our current experiments do not provide data on the evolution of the APC populations, we choose at this time to simplify the model, and omit the interaction term in Eq. (1). The model assumptions are summarized here:

- antigen in the blood stimulates the activation of immune cells through antigen presenting cells (APC's) that move from the blood to the spleen at a fractional rate of μ_{BS} . APC's are eliminated from the spleen at a per-cell rate of δ_D ;
- there is an average synaptic connection time between APC's and naive immune cells, τ_N , required to provide a co-stimulatory response;

- the proliferation rate depends on the concentration of APC's in the spleen, and saturates at some concentration. The maximum proliferation rate is given by ρ ;
- APC's also activate memory cells, with a different necessary synaptic connection time, τ_M . The average per-cell activation rate is denoted by λ ;
- activated effector cells proliferate rapidly for some finite time period. This proliferation window has an average duration of T . Proliferation requires a minimum period, τ , of continuous contact with APC's;
- after a phase of rapid proliferation, activated effector cells move into either a phase of rapid contraction, circulate in the blood, or become memory cells;
- the contraction phase has an average duration Δ . During this phase, the normal death rate, δ_T , is increased by an amount, α ;
- on average, a fraction r of the effector cells become memory cells after the proliferation phase;
- the rate of recirculation depends on the density of APC's in the spleen: a higher concentration of APC's slows down the rate of recirculation. The maximum rate of recirculation is denoted by $\mu_{SB}^* + \Delta$.

These assumptions yield the following system:

Model Equations

$$\begin{aligned}
 \frac{dA}{dt} &= \mu_{BS}D_B(t) - \delta_D A(t) \\
 \frac{dN}{dt} &= -\delta_N N(t) - gN(t - \tau_N)A(t - \tau_N) \\
 \frac{dP}{dt} &= gA(t - \tau_N)N(t - \tau_N) + \rho \frac{A(t - \tau)P(t - \tau)}{\theta + A(t - \tau)} A(t) \\
 &\quad + \lambda A(t - \tau_M)M(t - \tau_M) - \left(\delta_A + \frac{1}{T} \right) P(t) \\
 \frac{dA_p}{dt} &= \frac{1}{T} \left(1 - r - \mu_{SB}^* - \frac{\Delta \theta_{shut}}{\theta_{shut} + A(t)} \right) P(t) - (\alpha + \delta_T) A_p(t) \\
 \frac{dB}{dt} &= \frac{1}{T} \left(\mu_{SB}^* + \frac{\Delta \theta_{shut}}{\theta_{shut} + A} \right) P(t) - \delta_T B(t) \\
 \frac{dM}{dt} &= \frac{r}{T} P(t) - \lambda A(t - \tau_M)M(t - \tau_M)
 \end{aligned} \tag{3}$$

where we have denoted the naive, proliferating (rapidly expanding), apoptotic (rapidly contracting), circulating (normal death rate), and memory immune cell subpopulations by N , P , A_p , B and M , respectively. In the first equation, $D_B(t)$ gives the amount of circulating antigen, which depends on the vaccine regimen. For example, if a single dose of vaccine of size β with a pulse length of PL is given at time t_0 ,

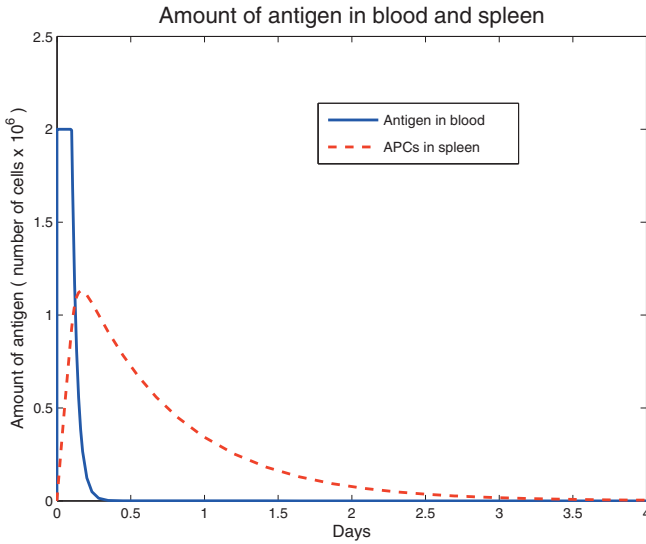


Fig. 2. Graph of the function $D_B(t)$ for one dose of vaccine. Parameters are: β , dose size, PL , pulse length, t_0 , time of administration, and μ_D , the decay rate of the antigen in the blood

$D_B(t)$ would be given by the following function (see Fig. 2):

$$D_B(t) = \begin{cases} 0 & t < t_0 \\ \beta & t_0 < t < t_0 + PL \\ \beta e^{-\mu_D(t-(t_0+PL))} & t \geq t_0 + PL. \end{cases}$$

Note that this set of equations describes antigen presenting cells and one immune cell population, with five subpopulations. Since we are modeling both $CD8^+$ and $CD4^+$ cells, we will have eleven equations in the model: one for the APC's, and five for each of the $CD8^+$ and $CD4^+$ populations.

3 Model Calibration and Validation

In this section we compare the calibrated model to experimental results. More details on the laboratory methods used in the experiment are given in the supplementary material at the end of this chapter.

The first step in making the model useful was to determine values for all of the parameters. We used our own experimental results in conjunction with values established in the literature.

- **Model calibration.** The model was calibrated using mean data collected from two separate experiments. The parameters in the model were determined by carrying out a least-squares minimization of the difference between the model simulation

Table 1. Model parameters

<i>Name</i>	<i>Value: CD4</i>	<i>Value: CD8</i>	<i>Units</i>	<i>Meaning</i>
α	2.5542	0.002	d^{-1}	activated T cell rapid apoptosis rate
δ_A	0.17	0.4	d^{-1}	proliferating T cell apoptosis rate (increase from normal)
ρ	4842	11431	d^{-1}	maximum T-cell proliferation rate
τ_N	0.45	0.47	d	co-stimulatory APC-Naive T-cell synaptic time
τ_M	0.01	0.0137	d	co-stimulatory APC-Memory cell synaptic time
θ	1300	2000	cells	number of APC's necessary to stimulate half the ma proliferation rate
λ	0.36	0.36	d^{-1}	activation rate of memory cells by APC's
T	1.6	1.05	d	average duration of rapid expansion phase
τ	.5	.25	d	co-stimulatory APC-proliferating T-cell synaptic time
Δ	10.00	3.00		maximum APC-induced <i>reduction</i> in rate of transfer of <i>T</i> -cells from spleen to blood
r	0.13	0.2		rate of production of new memory cells from activated <i>T</i> -cells
μ_{SB}^*	0.012	0.4055		APC-reduced rate of transfer of <i>T</i> -cells from spleen to blood
θ_{shut}	16.8859	13.0	cell	threshold in APC density in spleen for half-maximal rate of transfer of <i>T</i> cells to blood
δ_D	0.2370	0.2370	d^{-1}	per cell death rate of APCs
δ_N	1.3	.08	d^{-1}	per cell death rate of naive <i>T</i> cells
δ_T	.3	.6	d^{-1}	per cell death rate of activated <i>T</i> cells in the “basic” state
μ_{BS}	0.4907	.4907		transfer rate of <i>T</i> cells from blood to spleen

and experimental data. The minimization process was carried out by running simulations using initial cell values corresponding to the experiments. Some parameter values could be estimated from the literature. These sources are given in Table 1.

- Optimization method. The minimization was carried out using a simplex method built into the routine `fminsearch` in MATLAB, R2007b. Parameters estimated from the literature were checked against the data, and other parameters were fit directly to the data.

The results of the parameter estimation are summarized in Table 1. Model predictions are compared to experimental data in Fig. 3. A mathematical model can only be useful if it predicts the behavior of real systems. Any mathematical model, including the one discussed here, is a simplification of the actual biological process. However, since we are interested in the suggesting vaccine dosing strategies, we are mainly

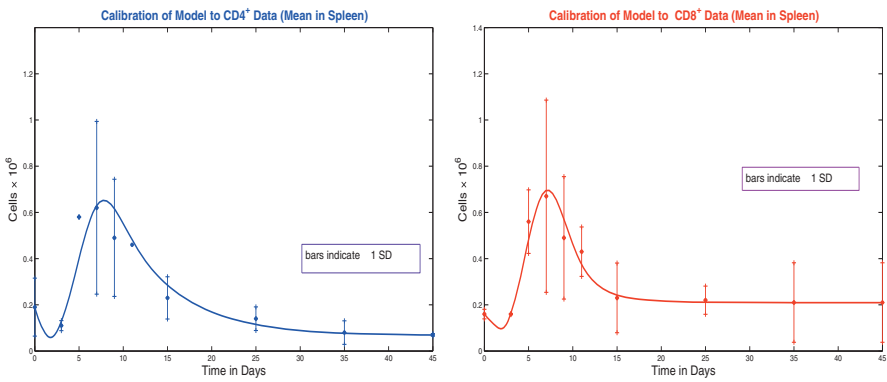


Fig. 3. Calibration of Model for Mean Data in Spleen for CD4⁺ Cells (left) and CD8⁺ Cells (right). The error bars represent one standard deviation. Note that the peak CD4⁺ levels are slightly lower and come slightly later than the peak CD8⁺ levels

concerned with the effect of varying the amount of antigen, as well as the initial immune levels.

The figures below show that the model simulations qualitatively reproduce results from the literature that study both the effects of initial levels of antigen-specific immune cells (Fig. 4) and the effect of varying the length of antigen exposure (Fig. 5).

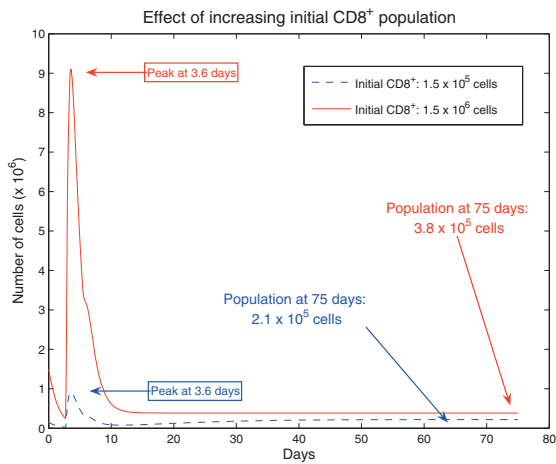


Fig. 4. Different initial CD8⁺ levels do not affect the timing of the peak, but they do affect the size of the peak as well as the long term population sizes [5]

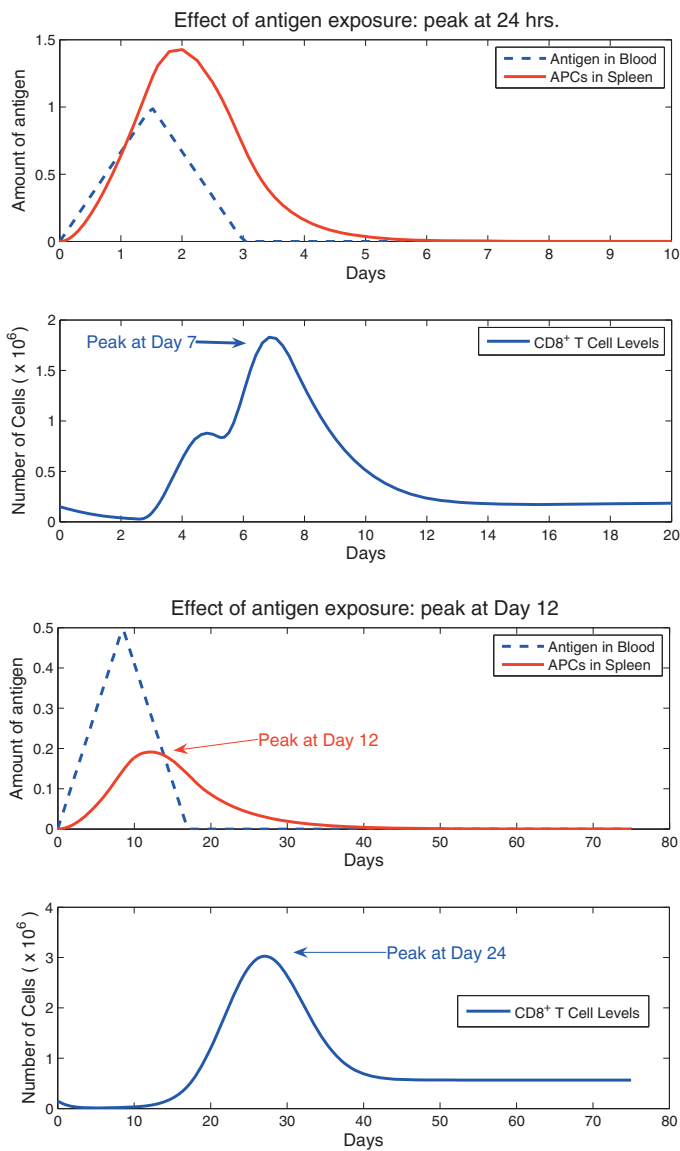


Fig. 5. Antigen exposure affects the timing of the peak response for CD8⁺ T Cells. These simulations reflect the results discussed in [5]

4 Optimization of Vaccine Protocols

The most urgent questions in the design of treatment protocols are “How much?” and “How often?”. Once the model has been validated, we can use it to suggest answers to these questions that can then be tested first in the laboratory and then in clinical trials.

In this section we address only the second question, since we did not perform experiments with varying dosages (see, however, Fig. 5 and the related experiments.) In seeking the best dosage strategy, we recall that the number of antigen presenting cells in the spleen, $A(t)$, is directly affected by an input function, $u(t)$, that represents a controlled dose of weak antigen, or vaccine, entering the blood stream. The optimization goal is to find this input, or *control* function, so that the immune response is maximized. Due to the interconnections between the components of the immune system, it is unclear what should be optimized. For example, we could choose to maximize the the number of effector T-cells circulating in the blood, or the number of memory cells, or some combination thereof.

Mathematically, the optimization problem takes the form:

Maximize

$$J(u) = k_1 \mathbf{v}^T \mathbf{x}(T_f) + k_2 \int_{T_0}^{T_f} \mathbf{w}^T \mathbf{x}(t) dt \quad (4)$$

subject to $\mathbf{x}(t)$ is a solution to the system (3) and $0 \leq u \leq u_{\max}$. (In practice, u will be a step function.) The parameters $k_1, k_2, \mathbf{v}, \mathbf{w}$ indicate relative weights.

We note that, due to the delays in the system of differential equations, this control problem is difficult to solve using analytical techniques. In what follows we use a heuristic method to suggest optimal dosing strategies.

4.1 Heuristic Optimization of Dose Times

We use a genetic algorithm (GA) approach to explore the space of dose strategies. As a first step, we assume that the dosing function, $u(t)$, is a step function taking only two values: 0 or $u_{\max}$. The method described here could also be applied to dosing functions taking on a finite set of values but, since our experiments involved bolus injections given over a short period of time, restricting the control functions to “on-off” dosing functions was not overly restrictive.

For more details on genetic algorithms and evolutionary algorithms as well as simulated annealing applied to dosing strategies for cancer treatments, see the paper by Villasana et al., [16]. Briefly, a GA searches through a space of solutions, or a “population”, where each solution is thought of as a “chromosome”. The algorithm successively selects subsets of the population that maximize a “fitness” function, in our case the objective function, J , given in Eq. (4). For the purpose of this discus-

sion, the search space, or populations, is the set of functions, $\{u(t)\}$, each of which describes a the timings and durations of a sequence of dosages.

4.2 GA Example

We represent each admissible function, $u(t)$ as a discrete sequence: $\{u(t_i)\}$, where we sample $u(t)$ at discrete points in time: $t_i, i = 0, 1, 2, \dots, n$. Since we are considering only dosing at full strength, we represent this sequence as a sequence of 1's, corresponding to $u(t_i) = u_{\max}$, and 0's, corresponding to $u(t_i) = 0$. Thus, each “chromosome” is a sequence of “on”s and “off”s, or a sequence of 1's and 0's, which we will denote by $\hat{u}_j$. We denote the set of *all* such sequences as U (see Fig. 6).

The GA algorithm consists of randomly selecting a subset of the population, and then allowing it to evolve, using the fitness function to drive the evolution. The algorithm stops when the subset reaches a quasi-steady state, i.e. the chromosomes change very little from one iteration to the next. We describe the steps of the algorithm very briefly.

At the initiation of the algorithm, a random set of size n of $\hat{u}_j \in U$ are selected. One iteration of the GA algorithm consists of the following three steps:

- 1. The fitness function, J , is evaluated at each $\hat{u}_j$, and the n_1 chromosomes with the largest fitness values are stored as part of the *current sub-population*.
- 2. These n_1 chromosomes are then allowed to “mutate”, creating another set of chromosomes. We use two types of mutations in this example: point mutations, where one element in the sequence switches between a ‘1’ and a ‘0’; and cross-over mutations, where two chromosomes exchange subsequences at a random point. The probability of each type of mutation are parameters of the GA algorithm.

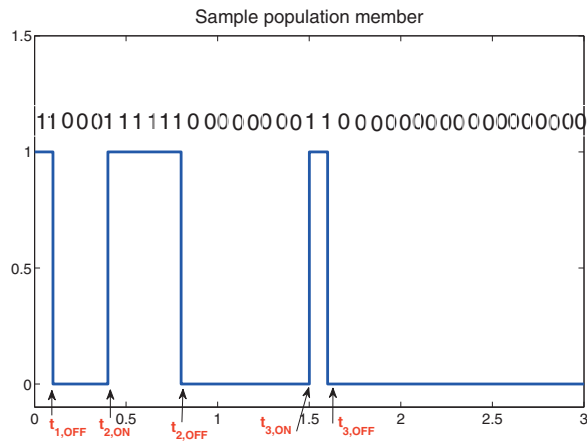


Fig. 6. Each chromosome in the population used in the GA represents an “on-off” dosing schedule given by a sequence of 1’s and 0’s

In this example, we set the probability of a point mutation to $p_m = .8$, and the probability of a cross-over mutation to $p_c = .2$.

3. The n_1 mutated chromosomes are added to the current sub-population, along with an additional set of n_2 chromosomes randomly chosen from the entire population. We require that the number of chromosomes in the current sub-set is always equal, so that $n_1 + n_1 + n_2 = n$.

The algorithm is iterated until it either passes a convergence test, which determines how many chromosomes have changed during an iteration, or until a maximum number of iterations have been executed. Since the GA algorithm searches a high-dimensional space, it can find local maxima. Therefore, the algorithm is repeated N times, and the results are compared, selecting the chromosome with the highest fitness as the “best”.

In the example shown here, we further restricted the admissible set, U to sequences that represent one vaccine with one boost, where the duration of each dose was fixed. This choice mimicked the laboratory setup, in which one initial vaccine and one boost were given to each animal, and the immune response was measured at several time points. Our goal was to find the optimal time at which to administer the second dose, or “boost”. The motivation behind this particular GA implementation was the low $CD8^+$ response to a preliminary boosting experiment, where the second dose was administered on Day 6, near the peak of the initial $CD8^+$ response. The GA algorithm suggests that the second dose be given on Day 3 (see Fig. 7). The model simulations (Fig. 8) and subsequent laboratory experiments confirmed that a boost on Day 3 does, in fact, yield a higher specific immune response (Fig. 9).

5 Conclusion and Challenges for the Future

In this chapter we have described a system of delay equations that models the specific immune response to a weak antigen, such as a cancer vaccine.

In particular, the model describes the activation of the $CD8^+$ and $CD4^+$ T-cells in the lymph organs. The model was calibrated to murine experiments where immune cells in the spleen were counted post-innoculation, with different dosing strategies. We found that a GA could inform the design of dosing timings, and we validated those results in the laboratory.

These results are very preliminary. A more detailed study of the optimal strategies suggested by either a GA approach, or by a more analytical approach, could yield insight into the kinetics of the immune response. It is important to get more refined experimental measurements of the time course of the immune response in mice. Since it is impossible to do these measurements *in vivo* in humans, a validated model can be useful in proposing more individualized therapies. In order for the models to be effective however, we need to be able to scale them from the murine system to the human model. The following challenges still need to be overcome:

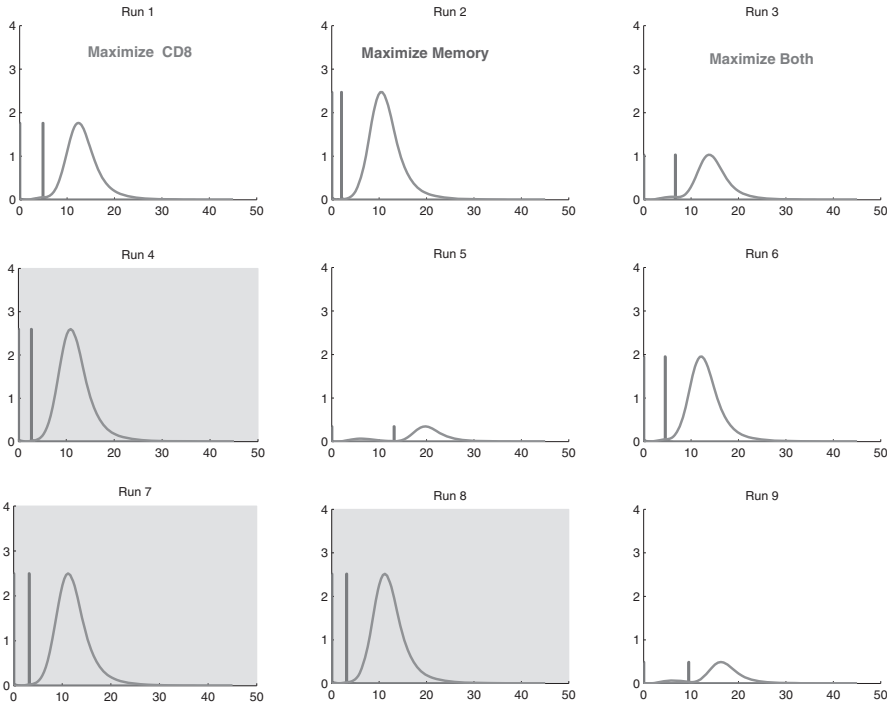


Fig. 7. The genetic algorithm is run N times, and then those results are compared. Each run pictured here is the end-point of one GA run, typically thousand of iterations. Vertical lines show the timing of the vaccines, and the curves depict the immune cell levels over time. The chromosomes, or treatment sequences, with the best fitness are then selected (highlighted). Three fitness functions were used: maximize the final number of $CD8^+$ cells in the blood ($k_2 = 0, k_1 = 1, v_5 = 1, v_i = 0$ for $i \neq 5$); maximize the final number of memory immune cells ($k_2 = 0, k_1 = 1, v_6 = v_{11} = 1, v_i = 0$ for $i \neq 6, 11$); maximize both the final number of active and memory immune cells ($k_2 = 0, k_1 = 1, v_5 = v_6 = v_{10} = v_{11} = 1$). For this numerical experiment, $N = 1000, n = 150, n_1 = n_2 = 50, p_m = .8$ and $p_c = .2$

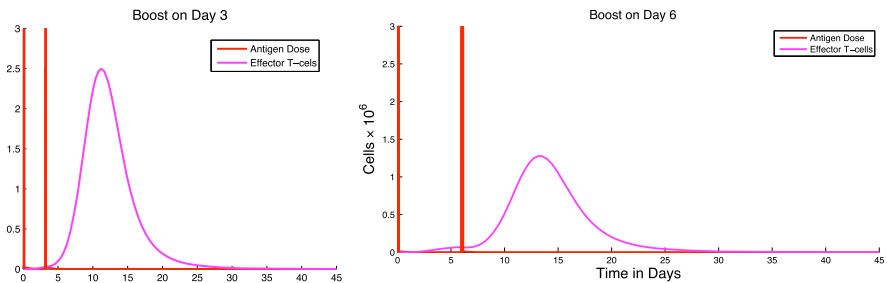


Fig. 8. Model simulations support the conclusions of the GA algorithm that boosting once on Day 3 results in a higher immune cell count that boosting on Day 6. In each case, the same amount of vaccine is given

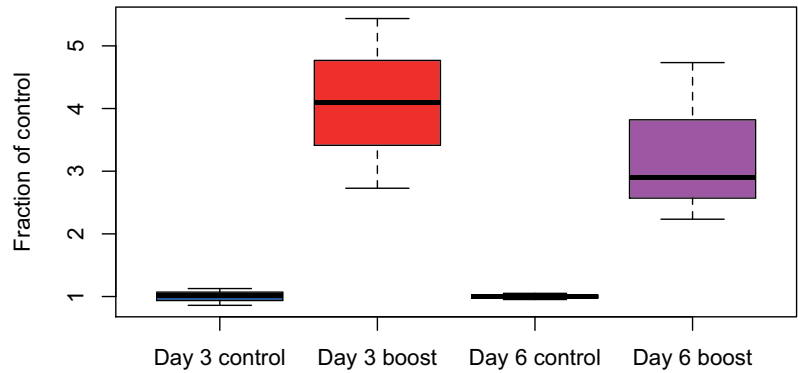


Fig. 9. Laboratory experiments are consistent with the model conclusions that a boost on Day 3 results in higher immune levels than a boost on Day 3. Two different experiments are compared: the first two boxplots are the result of one experiment, where mice were given a vaccine on Day 0, and then a subset were given a second vaccine on Day 3. The two boxplots on the right are the result of another experiment, in which mice were given a vaccine on Day 0, and then a subset were given a second vaccine on Day 6. In both cases, cells were counted 3 days after the vaccination, so we are not necessarily catching the peak of the response, but we do see a higher response, compared to the control group, when the vaccine boost is given on Day 3 instead of on Day 6. Since the experiments were done under different conditions with slightly different vaccine delivery, the boosted counts are compared to the cell counts of mice who received only the initial vaccine. For reference, the mean of the control group for the Day 3 boost experiment was 2.14×10^5 cells and the mean of the control group for the Day 6 boost experiment was 2.66×10^5 cells

- what aspect of the immune response is most important in predicting a successful attack against a tumor? The answer to this question is necessary in order to define the objective function, or fitness function, used in any optimization routine;
- we have only considered $CD8^+$ and $CD4^+$ cells in this model. Other immune cells, such as B cells and T-regs are important in mounting and maintaining an immune response. In order to quantitatively measure these other immune cell populations, suitable laboratory experiments must be designed and carried out;
- how should a mathematical model that is calibrated to mice be adapted so that it can simulate the immune response in humans? It will be necessary to find surrogate markers for immune cell counts, or non-invasive methods for quantitatively measuring the adaptive immune response in humans.

6 Supplementary Material

Here we collect together details on the laboratory experiments:

- Mice: C57Bl/6J, OT-I and OT-II mice were bred and maintained in specific-pathogen-free (SPF) conditions at the Hercus Tairei Resource Unit. All experiments were approved by the University of Otago Animal Ethics Committee.

- Reagents: Ovalbumin protein (Grade V, 98 % purity) was purchased from Sigma Chemicals, (St Louis, MO, USA). Anti-mouse 2.4G2 (Fc III/II Receptor), CD4-FITC, CD8-APC, V α 2-PE, V β 5.1/5.2-biotin and SA-PerCP were purchased from BD Pharmingen (San Diego, CA, USA) and were used according to the manufacturer's instructions. Alum (Alu-Gel-S) was purchased from Serva (Heidelberg, Germany).
- Analysis of expansion of transgenic T cells: The expansion of ovalbumin specific CD4 and CD8 T cells following immunisation was investigated by vaccinating 6-8 week old C57Bl/6J mice that had been given 4×10^6 OVA-specific CD4 and CD8 T cells by intravenous injection. Transgenic OT-I [3] and OT-II [2] mice were used as the source of the transgenic cells. The C57Bl/6J mice were then immunised by the subcutaneous route with 100 μ g of ovalbumin in alum. At various time points post immunisation the mice were euthanised using a lethal dose of anaesthetic and spleens, axillary and brachial lymph nodes were removed. Single cell suspensions were then prepared from all harvested lymphoid tissues and the cells were labelled with monoclonal antibodies directed against Fc III/II receptor, CD4, CD8, V α 2 and V β 5.1/5.2 according to the manufacturer's instructions. Expansion of the OVA specific T cell populations were determined by identifying OVA-specific V2^{+ve}/V5^{+VE} CD8 and CD4 T cells by flow cytometry. Results were analysed using CellQuest Pro (BD, USA).

References

1. Arnon, R., Ben-Yedidia, T.: Old and new vaccine approaches. *Int. Immunopharmacol.* **3**, 1195–1204 (2003)
2. Barnden, M., Allison, J., Heath, W., Carbone, F.: Defective tcr expression in transgenic mice constructed using cdna-based- α - and β -chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol.* **76**, 34–40 (1998)
3. Carbone, F., Bevan, M.: Induction of ovalbumin-specific cytotoxic t cells by in vivo peptide immunization. *J. Exp. Med.* **169**, 603–612 (1989)
4. De Boer, R.J., Homann, D., Perelson, A.S.: Different dynamics of CD4⁺ and CD8⁺ T cell responses during and after acute lymphocytic choriomeningitis virus. *J. Immunol.* **171**, 3928–3935 (2003)
5. van Faassen, H., Saldanha, M., Gilbertson, D., Dudani, R., Krishnan, L., Sad, S.: Reducing the stimulation of CD8⁺ T cells during infection with intracellular bacteria promotes differentiation primarily into a central (CD62L^{high} CD44^{high}) subset. *J. Immunol.* **174**, 5341–5350 (2005)
6. Homann, D., Teyton, L., Oldstone, M.: Differential regulation of anti-viral T cell immunity results in stable CD8⁺ but declining CD4⁺ T cell memory. *Nature Med.* **7**, 913–919 (2001)
7. Keilholz, U., Martus, P., Scheibenbogen, C.: Immune monitoring of t-cell responses in cancer vaccine development. *Clin. Cancer Res.* **12**, 2346–2352 (2006)
8. Liu, M.: Vaccine developments. *Nature Medicine* **4**, 515–519 (1998)
9. Ludweig, B., Krebs, P., Junt, T., Metters, H., Ford, N.J., Anderson, R.M., Bocharov, G.: Determining control parameters for dendritic cell-cytotoxic T lymphocyte interaction. *Eur. J. Immunol.* **34**, 2407–2418 (2004)
10. Makela, P.: Vaccines, coming of age after 200 years. *FEMS Microbiol. Rev.* **24**, 9–20 (2000)

11. Melief, C., Van Der Burg, S., Toes, R., Ossendorp, F., Offringa, R.: Effective therapeutic anticancer vaccines based on precision guiding of cytolytic t lymphocytes. *Immunol. Rev.* **188**, 177–182 (2002)
12. Rosenberg, S., Yang, J., Restifo, N.: Cancer immunotherapy: moving beyond current vaccines. *Nature Med.* **10**, 909–915 (2004)
13. Shiina, M., Rehermann, B.: Hepatitis c vaccines: Inducing and challenging memory t cells. *Hepatology* **43**, 1395–1398 (2006)
14. Sprent, J., Surh, C.D.: T cell memory. *Annu. Rev. Immunol.* **20**, 551–579 (2002)
15. Srivastava, P.: Therapeutic cancer vaccines. *Curr. Opin. Immunol.* **18**, 201–205 (2006)
16. Villasana, M., Ochoa, G.: Heuristic design of cancer chemotherapies. *IEEE Trans. Evolutionary Comput.* **8**, 513–521 (2004)

Optimizing Cancer Chemotherapy: from Mathematical Theories to Clinical Treatment

Zvia Agur and Yuri Kheifetz

Abstract Cancer chemotherapy is much influenced by the “dose dense” paradigm, advocating maximally possible dose density early in treatment. This paradigm is based on a controversial mathematical model, assuming Gompertz tumor growth law. Alternatively, it has been suggested that for maximizing efficacy/toxicity ratio in cancer chemotherapy, the inter-dosing intervals should be determined according to the Resonance theory. This theory asserts that cell growth is maximal when the periodicity of drug administration is an integer or fractional multiple of the characteristic periodicity of the cell population. Model analysis and *in vitro* and *in vivo* experiments, suggest that differences in cell-cycle distributions of host and cancer cells can be taken advantage of in chemotherapy by cell-cycle phase-specific drugs which use Resonance or Anti-Resonance (stochastic) drug pulsing. Mathematical proofs showing long-term prediction accuracy of cell population growth models under cell-cycle phase-specific drugs, enabled developing a heuristic optimization method for drug scheduling. Using this method in conjunction with personalized models of vascular tumor growth under chemotherapy by docetaxel and bevacizumab, an optimal combination regimen was tailored to a particular mesenchymal chondrosarcoma patient. The personalized regimen was prospectively validated, leading to increased longevity and quality of life of the patient. This patient’s model was further simulated, suggesting that the relative advantage of “Dose Dense” drug therapy depends on personal cytokinetic and angiogenic parameters.

Z. Agur (✉) · Y. Kheifetz

Institute for Medical Biomathematics, APOB 282, 60991 Bene Ataroth, Israel

Optimata Ltd., Abba Hillel St. 7, 52522 Ramat-Gan, Israel

e-mail: agur@imbm.org, yurih@imbm.org

A. d’Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_15, © Springer-Verlag Italia 2012

1 Introduction

Efficacy of cancer therapy is often unpredictable and significantly different among patients, as a result of differences in age, gender, diet, organ function, and genetic variability [13, 29]. Yet, under the current paradigm of clinical trials, new drug candidates are approved if they show statistically significant advantage, when applied by precisely the same protocol to a large patient population. Due to the high cost of such large scale population trials, only a few protocols are tested, inevitably implicating suboptimal treatment for the patient.

In trying to identify an efficacious general chemotherapy policy, Larry Norton and Richard Simon constructed a simple mathematical model for relating the efficacy of cytotoxics to the growth dynamics of the tumor. Mathematically, the Norton-Simon model can be written $N(t) = f(N(t))(1-d(t))$, where $N(t)$ denotes the number of tumor cells at time t , $N(t)$ is the growth rate of the tumor at time t , $d(t)$ reflects the rate of removal of cells as a result of treatment, and f is a function which describes the growth dynamics of the unperturbed tumor. Norton and Simon assumed that this growth function obeys Gompertz law, and that the rate of tumor regression induced by chemotherapy is proportional to the rate of unperturbed growth of tumor of that size. Consequently, they argued, to be more efficient, the dose rate of chemotherapy should be increased, for example by applying the maximum tolerated dose (MTD) and decreasing the inter-dosing intervals [18, 19].

The Norton-Simon hypothesis has been influential in oncology and much inspired the clinical investigation. It served as a basis and justification for the “dose-dense” approach to breast cancer chemotherapy, shown in many studies to achieve drops in cancer recurrence and death, as the model predicts [11].

The rationale of the “dose-dense” paradigm is rooted in the nature of the underlying Gompertz model. However, Norton and Simon never checked if the Gompertz model is valid, or if the paradigm is valid also for alternative models of tumor growth (see, for example [15, 20, 21, 28]. Can their model be validated? Can it be generalized, for example, to all vascular cancers, or all chemotherapeutic drugs? In this work we will sketch some prerequisites for the use of mathematical models of cell population dynamics for chemotherapy optimization. It will be shown, both theoretically and experimentally, that the relative advantage in dose dense therapy may depend on cytokinetic parameters of drug-susceptible host and cancer cells, and on the level of tumor angiogenesis.

We will summarize theoretical research, which shows that periodically applied chemotherapy creates a resonance effect by which cell population growth can be maximized or minimized, depending on the periodicity of the drug pulsing. Experiments supporting the validity of this *Resonance phenomenon* in real-life will be summarized and it will be shown that models of cell populations which are treated by schedules of cell cycle phase-specific (CCPS) drugs, yield reliable long range predictions. This opens the way to chemotherapy optimization by use of mathematical models of cancer progression, and, further, to the development of an optimization method by which the inter-dosing interval is timed, essentially, to minimize the toxicity/efficacy ratio of the drug.

Finally, we will briefly report a particular application of the optimization method for treatment personalization. In this case, chemotherapy of a cancer patient was successfully optimized using a personalized mathematical model of tumor growth. The combined method was prospectively validated and it was suggested that the relative advantage of dose-dense periodic chemotherapy mainly depends on personal angiogenic and cytokinetic parameters.

2 Universal Resonance Phenomenon Suggests a New Method for Cancer Chemotherapy

Models of population dynamics, under various distributions of environmentally-inflicted loss processes, suggest that population persistence depends on the level of synchronization of the environmental and population processes. Population growth is maximized when the periodicity of environmental disturbances is an integer or fractional multiple of the characteristic periodicity of the population. This *Resonance Phenomenon* is observed in as diverse models as those of mussel growth under harsh weather regimes in the intertidal zone of the sea, or in models of measles epidemics in human populations under vaccination. The internal periodicity caused by some population characteristics and the external periodicity inflicted by environmental forces may be unique to each case, but the resonances created by these relationships are a universal phenomenon [1, 2, 6, 25].

The universality of the Resonance Phenomenon offers an attractive possibility for the control of the cancer disease. Thus, it was shown, both theoretically and experimentally, that the period of drug dosing can be determined so as to create resonances which maximize the growth of host drug-susceptible cells and, at the same time, minimize the growth of the malignant cells.

First, the generality of the Resonance Phenomenon for different host cell populations was investigated in a much simplified model of a cancer patient. The patient's model assumed that both tumor and target host tissues are collections of cells that vary only in cell-cycle parameters; the spatial arrangement of these cells was ignored. In this way, the complex physiological and pathological dynamics of an oncology patient under chemotherapy was reduced into one essential variable, namely, the ratio between tumor growth and that of the host target tissues under different chemotherapy regimens of CCPS drugs.

Two types of treatments were considered : (i) drug efficacy is stepwise “infinite,” i.e. all cells in the vulnerable cycle-phases are instantly eliminated during the fixed period in which the drug is effective, while no cells are affected during the drug-free interval [5]; (ii) drug efficacy is “finite” and drug decay kinetics is considered [10]. The previous case, a limit of the latter one, can be considered as a reasonable approximation to reality. Assuming stochastic inter-dosing intervals, this model successfully shows a clear advantage for population growth when the period of the loss process coincides with an inherent population reproduction period.

In [10], the resonance effect was examined under improved realism of the drug's pharmacokinetics (PK). A probabilistic model was constructed, which assumes (i) a

constant cell-cycle duration, (ii) a constant drug-sensitive cycle-phase, (iii) S-phase drug susceptibility and (iv) negative exponential drug elimination. This is a major departure from the previous model, where a dichotomous, “all or none” drug effect was assumed. Analysis of this model provides explicit formulae for the growth or decay of the tumor cell population. Notably, it is shown that the resonance phenomenon is preserved even when model complexity is increased.

In the cases discussed above the models are probabilistic. In contrast, in [12], a systems of deterministic differential equations is used to analyze similar dynamics. The authors show that when the mean cell-cycle times of normal and tumor cells differ considerably, optimal drug schedules can significantly improve therapeutic efficacy, when drug periodicity is close to the mean cell-cycle time of the normal population and is practically independent of the tumor cell-cycle parameters. In this case, optimal treatment regimen heavily depends on precise estimation of the mean cell cycle duration of both cancer and host cells.

In [30], chronotherapy is discussed in structured models of tumor cell population growth, where drug treatment corresponds to a loss of cells in a periodic on-off schedule. Treatment is specific to the structure variable, so that only cells in a certain range of this variable are eliminated. The author demonstrates that resonances occur at a treatment period, p , being at, or near, integral or fractional multiples of the median age of cell division.

A rigorous mathematical analysis of the Resonance Phenomenon in periodic chemotherapy scheduling is presented in [31]. Here, realism is traded-off by rigor, and for model simplification it is assumed that cell cycle has a constant length and that treatment eliminates cells only at the time of cell division (to distinguish from the S-Phase killing, as in the model described in [5]). Moreover, differential equations model is used rather than probabilistic approach in [10].

2.1 Laboratory Experiments Verifying the Resonance Phenomenon

A series of laboratory experiments was conducted for testing the clinical validity of the Resonance phenomenon. To this end, the effect of drug pulsing rhythm on cell proliferation was examined both *in vitro*, *in vivo* and *ex vivo*. The *in vitro* experiments tested the dependence of cell population growth on the inter-dosing interval *per se*. Thus, short pulses of Cytosine Arabinoside (ara-C) – an S-Phase-specific cancer drug, were applied to several sets of cells. Experimental results show that, regardless of the total drug dose or the experimental procedures, drug schedules whose periodicity was an integer multiple of the average population cell-cycle time showed significantly higher growth rate, as evaluated by rates of DNA synthesis [5]. In the *in vivo* murine experiments ara-C was applied to lymphoma-bearing mice, treated by short drug pulses of different distributions of the inter-dosing intervals, including stochastically determined periodicity; the total dose and total treatment duration were kept constant. Toxicity was evaluated by spleen weight, by peripheral blood measurements and by the proportion of bone-marrow cells in the S-Phase gate of the DNA content distributions. Results support theory in showing that an Anti-

Resonance regimen, namely random duration of the inter-dosing intervals, inflicts higher cytotoxicity than a regimen of constant inter-dosing intervals. The results also show that minimal myelotoxicity is exerted when the inter-dosing interval is a multiple of the inter-mitotic time of bone-marrow stem and progenitor cells, thus supporting model predictions [4]. To verify that this approach is feasible using methods that are available in clinical practice, Ubezio and colleagues [27] determined the optimal intervals between treatments, by measuring the effects of single-dose ara-C on proliferation kinetics of murine bone-marrow cells. Results showed that ara-C is toxic to S-Phase cells, causing an arrest at the G1/S boundary of the cell-cycle for about 4 – 6hr, following which cells transit through the S-Phase in a nearly synchronized manner. Further results suggest that an optimal window for a second ara-C dose, designed to preserve the bone-marrow proliferating pool, would be optimally administered at 14–16hr (for 5mg/kg) and 12–14hr (for 1mg/kg) following the first ara-C dosing. This is so because in these time windows most of the surviving cells are found in the less susceptible cycle phases (G1 and G2/M). In contrast, a time periods of 7–11hr is expected to impose maximal toxicity to the bone-marrow, since at that time surviving cells will be transiting into the susceptible S-Phase of the cell-cycle [27]. In order to verify that, indeed, the optimal dosing period in murine chemotherapy is 14hr, it was first shown in bone-marrow aspirates of ara-C-injected mice, that cell kinetics following 2 or 3 ara-C dosings remained similar to that observed after a single dosing. It was further shown that, as predicted, cell-cycle kinetics was less affected when using the optimal dosing period. Indeed, when a second dose was administered after 7hr or 14hr period much less toxicity was inflicted than after a 10hr period, and a third dose of 14hr period had a considerable smaller effect on the percent of proliferating cells in the bone-marrow than any of the second dosings. Subsequently, toxicity to healthy mice by different ara-C protocols was evaluated by measuring mice survival. Thus, 4/5 mice died as a result of treatment with 4 ara-C doses given at 10hr intervals. Reducing the number of doses to 3 did not decrease mortality. In contrast, increasing dosing intervals to 14hr dramatically improved survival since 5/5 mice survived 3 ara-C doses, given at 14hr intervals. The next step was to examine the above scheme in tumor bearing mice. Mice were inoculated with 38C–13B lymphoma cells and treatment was initiated 3 days later. Results show that a regimen of $4 \times 14hr$ applications was not only less toxic, but it also delayed tumor progression, nearly doubling survival time, as compared to the untreated control. The protocol of 16hr dosing interval had similar results to the 14hr protocol, defining thus the width of the bone-marrow preserving time-window. In contrast, all untreated mice in the control group died within 14 days due to tumor progression, whereas the treatment with random drug application was highly toxic causing 100% death within 7 days [27].

2.2 Long-Range Predictability in Models of Cell Populations Subjected to Cell Cycle Phase-Specific Drugs

Can one *a priori* use mathematical modelling to predict whether a given cell population of distributed cell-cycle duration will be effectively eliminated by prescribed chemotherapy schedules administered according to the suggested method? In this section we show analytically that the answer is positive, thus paving the way for the development of heuristic optimization methods for chemotherapeutic drugs, which may be applicable for prolonged treatments.

In [16], numerical procedures are developed and estimated, for determining the growth rate of a cell population having distributed cell cycles under a prescribed chemotherapy protocol. It was shown that the cumulative error in estimating cells proliferation rate stays bounded and does not accumulate with time, if the considered treatment regimen is periodic. It was also proved that this rate is almost independent (with unity probability) from initial distribution of the ages in the cell population.

Let $n(a, t)$ denote the cell age density, a denote a cell age and t denote a time. It will be shown that the behavior of $r(t) = n(0, t)$ determines the behavior of $N(t)$, that is, the number of cells at time t , as t tends to infinity.

The model equations, describing cell age distribution dynamics are

$$n_t(a, t) + n_a(a, t) = -(\beta(a) + \eta(a, t))n(a, t), \quad a > 0, t > 0, \quad (1)$$

$$n(0, t) = 2 \int_{\tau_b}^{\tau_m} \beta(a)n(a, t)da, \quad t > 0, \quad (2)$$

$$n(a, 0) = n_0(a), \quad a > 0, \quad (3)$$

where n_a and n_t denote partial derivatives $\partial n/\partial a$ and $\partial n/\partial t$, respectively [12, 17, 26, 30, 31]. The age-specific division rate of cells is $\beta(a)$, the age-specific mortality rate (due to natural causes or to a treatment) of cells is $\eta(a, t)$, and the initial age distribution of cells is $n_0(a)$. The function $\beta(a)$ satisfies $\beta(a) = f(a)/\alpha(a)$, where $\int_{a_1}^{a_2} f(a)da$ is the probability that a cell divides between ages a_1 and a_2 and $\alpha(a) = \int_a^\infty f(\hat{a})d\hat{a}$ (the function $\alpha(a)$ gives the fraction of cells undivided by age a). In our case the support of β (the set of all of points on which β has nonzero values) is a subset of $[\tau_b, \tau_m]$.

In this model it was assumed that the division and mortality rates of the population are independent of the population density. Note, however, that in other cases, the division rate of cell population may depend on the total cell number. This can be so in noncancerous cells populations, where β and η depend also on the total number of the cells, $N(t)$ is given by integrating $n(a, t)$ from 0 to infinity. In such cases, being out of the scope of the present article, $\beta(N, a)$ is a decreasing and $\eta(N, a, t)$ is an increasing functions of N (negative feedback regulation of overall cells number).

The differential equation (1) is integrated, and the following formula for $r(t) = n(0, t)$ is obtained:

$$r(t) = 2 \int_{\tau_b}^{\tau_m} f(a) r(t-a) \Theta(a, t) da. \quad (4)$$

Here, $\Theta(a, t) = e^{-\int_0^a \widetilde{\eta}(\alpha, t-\alpha) d\alpha}$ stands for treatment activity, where $\widetilde{\eta}(a, t) = \eta(a, t+a)$. For a very aggressive treatment $\Theta(a, t)$ tends to 0, and in the absence of the drug and in the absence of a natural loss we have $\Theta(a, t) = 1$. $f(a)$ is a probability density function of cell division at the age a with support $[\tau_b, \tau_m]$, $\tau_b > 0$.

It is shown in [16] that in the generic case, the growth rate of the total cell population, $N(t)$, is similar to $r(t)$, for t large enough.

Let $r_n(s) = r(s + n\tau_b)$ for $s \in (0, \tau_b)$ and $n \in \mathbb{Z}$. The fact that $\tau_b > 0$ converts (4) to recursive formula:

$$r_n(s) = \sum_{j=0}^M \int_0^{\tau_b} E_{j,n}(u, s) r_{n-j-1}(\tau_b - u) du = \sum_{j=0}^M T_{j,n} r_{n-j-1}(s), \quad (5)$$

where $M = \lceil \frac{\tau_m - \tau_b}{\tau_b} \rceil + 1$, $E_{j,n}(v, s) = 2f(v + s + j\tau_b) \Theta(v + s + j\tau_b, s + n\tau_b)$ and $T_{j,n}g(s) = \int_0^{\tau_b} E_{j,n}(u, s) g(\tau_b - u) du$ for any $g \in L^2_{[0, \tau_b]}$. (5) define recursively $\{r_n(t)\}_{n=0}^\infty$, provided that $r_{-M}(t), \dots, r_0(t)$ are known.

The integration kernel $E_{j,n}(u, s)$ has been approximated in [16] by step functions, $E_{j,n}^u(u, s)$ and $E_{j,n}^l(u, s)$ bounding it from above (superscript “u”) and from below (superscript “l”), defining $\{r_n(t)^u\}$ and $\{r_n(t)^l\}$ as the upper and lower approximations for $\{r_n(t)\}$.

Let H be a set $(L^2_{[0, \tau_b]})^{M+2}$ with a following inner product: $\langle v, w \rangle = \sum_{j=1}^{M+2} \langle v_j, w_j \rangle$, where $\langle v_j, w_j \rangle$ is an inner product in L^2 . It is clear that H with this inner product constitutes Hilbert space. Thus Eq. (5) can be re-written in the following form:

$$\begin{pmatrix} r_{n-M-1}(t) \\ r_{n-M}(t) \\ \dots \\ r_n(t) \end{pmatrix} = \begin{pmatrix} 0 & Id & 0 & \dots & 0 \\ & & \dots & & \\ 0 & \dots & 0 & \dots & Id \\ 0 & T_{M,n} & T_{M-1,n} & \dots & T_{0,n} \end{pmatrix} \begin{pmatrix} r_{n-M-2}(t) \\ r_{n-M-1}(t) \\ \dots \\ r_{n-1}(t) \end{pmatrix}. \quad (6)$$

Let us denote

$$T_n = \begin{pmatrix} 0 & Id & 0 & \dots & 0 \\ & & \dots & & \\ 0 & \dots & 0 & \dots & Id \\ 0 & T_{M,n} & T_{M-1,n} & \dots & T_{0,n} \end{pmatrix}. \quad (7)$$

Clearly $\{T_n\}_{n=1}^\infty$ is a sequence of compact operators from H to H . In the case of cyclic treatment there is a minimal $\hat{p} > 0$, such that $\Theta(a, t) = \Theta(a, t + \hat{p})$ for any $a > 0$ and any $t > t_0$. It was easily shown in [16] that $\Theta(a, t) = \Theta(a, t + P\tau_b)$ for any $a > 0$, $t > t_0$ and P is the minimal natural number with this property. From this it easily

follows that $\{T_n\}_{n=0}^{\infty}$ is a periodic sequence of operators with the minimal period P . Then $T = T_1 T_2 \dots T_P$ is a compact operator on H and for every initial vector $v_0 \in H$ we have P sequences $\{T^m v_0\}_{m=0}^{\infty}$, $\{T^m T_1 v_0\}_{m=0}^{\infty}$, $\dots$, $\{T^m T_0 \dots T_{P-1} v_0\}_{m=0}^{\infty}$ and obtain sequence of vectors of functions $\{(r_n(t), r_{n+1}(t), \dots, r_{n+M}(t))\}_{n=1}^{\infty}$ in the following way: for any natural n $r_n(t) = T^m T_0 \dots T_i v_0$, where $n = mP + i$, $0 \leq i < P$. This means that behavior of $r(t)$ in infinity is determined by $\{T^m\}_{m=1}^{\infty}$ and in particular by $\lim_{n \rightarrow \infty} \sup_{g \in H} \frac{\|T^n g\|}{\|g\|} = \lim_{n \rightarrow \infty} \|T^n\|$. From elementary course in Banach spaces [22] follows that $\lim_{n \rightarrow \infty} \sqrt[n]{\|T^n\|} = \rho(T)$, $\rho(T)$ is a spectral radius of T . Thus the problem is reduced to the evaluation of $\rho(T)$.

If $\{T_n\}_{n=0}^{\infty}$ has a period P , then given N and equal partition of $[0, \tau_b]$, $\{T_n^l\}_{n=0}^{\infty}$ and $\{T_n^u\}_{n=0}^{\infty}$ have also period P . Let $T^l = T_1^l T_2^l \dots T_P^l$ and $T^u = T_1^u T_2^u \dots T_P^u$. These are obviously finite rank operators. Thus the behavior of $\hat{r}^l(t)$ and $\hat{r}^u(t)$ is determined by $\rho(T^l)$ and $\rho(T^u)$ respectively. $\rho(T^l)$ and $\rho(T^u)$ can be readily calculated. This is because these operators are of finite rank and can be identified with finite dimensional matrices by means of isometry of $\mathbb{C}^N$ with the space of step functions with equal partition of $[0, \tau_b]$ to N equal subintervals, as described earlier.

The following propositions have been established:

Proposition 1. *If $\rho(T) < 1$ then $\lim_{n \rightarrow \infty} \|T^n v\| = 0$.*

If $\rho(T) > 1$ then $\lim_{n \rightarrow \infty} \|T^n v\| = \infty$, $\forall v \in H$.

Proposition 2. *For any nonnegative compact operator K on H , the spectral radius $\rho(K)$ is an eigenvalue of K with at least one eigenfunction $x \geq 0$.*

Proposition 3. *For any nonnegative compact operator K on H and for almost every $v \in H_+$, $\lim_{n \rightarrow \infty} \|K^n v\|^{\frac{1}{n}} = \rho(K)$.*

where H_+ denotes the subspace of H , containing vectors with nonnegative elements.

We summarize the results as follows: The value of the spectral radius $\rho(T)$ fully characterizes population growth or decay. The population exponentially grows if $\rho(T) > 1$ and decays if $\rho(T) < 1$. These rates are bounded by $\rho(T)$ and tend to it for almost all initial conditions (which are represented by nonnegative functions).

The following proposition states that order is preserved by a map, which is defined on the set of all bounded nonnegative operators on H , and assigns its spectral radius for any such operator. This means that if $A_1 \leq T \leq A_2$, where A_1, A_2 and T are bounded operators on H , and if $\rho(A_1), \rho(A_2)$ are known, then $\rho(T) \in [\rho(A_1), \rho(A_2)]$. This fact in conjunction with results of Proposition 5 below gives a quantitative estimation of $\rho(T)$ through $\rho(A_1)$ and $\rho(A_2)$, in the case that $\rho(T)$ is compact and $\rho(A_1)$ with $\rho(A_2)$ are operators of finite rank.

Proposition 4. *Given S_1, S_2 being any bounded nonnegative operators on H , such that $S_1 \geq S_2$, their spectral radiuses satisfy similar inequality: $\rho(S_1) \geq \rho(S_2)$.*

The following proposition is significant as it directly estimates the radius of any given compact operator through the spectral radius of finite rank operator close to it.

Proposition 5. Assume that K is a compact operator on separable Hilbert space H , $\rho(K) > 0$ and A is a finite rank operator, such that $\|K - A\| < \varepsilon$ in operator norm for some $\varepsilon > 0$. If $\varepsilon < \rho(K)$, then

$$\rho(K) < \rho(A) + \varepsilon \left(3 + \frac{\|K\|}{\rho(K)} \right). \quad (8)$$

From Proposition 1 follows a property, which is associated with fully periodic treatments. When subjected to fully periodic drug treatments, the behavior of a cell population at large t is determined by the spectral radius of some operator T . This means that the inexactness in the evaluation of cell population dynamics is equal to the inexactness of the population growth rate estimation at any time. Thus, the error in the initial data will lead to a bounded error in the evaluation of the population growth rate. From Propositions 1, 4, 5 and from the approximation of compact operator by finite rank operators developed above, it follows that the spectral radius is a continuous function on the set of operators, whose spectrum is nontrivial. This guarantees the stability of the population growth rate under fully periodic treatment.

One of the problems any predictive mathematical model has to cope with is the inexactness of the biological data. In the present model, if we know f up to some accuracy, the following question arises: given a cyclic drug treatment, is the inexactness in the biological input data crucial for our prediction ability of population growth or extinction?

Let $f_1(a) < f(a) < f_2(a)$ for any $a \in [\tau_b, \tau_m]$. Let ${}_1T$ and ${}_2T$ be the respective operators on H . Then by Proposition 5 it follows that $\rho({}_1T) \leq \rho(T) \leq \rho({}_2T)$. Given any $\varepsilon > 0$, we can evaluate $\rho({}_1T)$ and $\rho({}_2T)$ with accuracy less than ε . In the generic case $\rho({}_1T) \neq 1$ and $\rho({}_2T) \neq 1$. Thus, for sufficiently small ε we will know whether both $\rho({}_1T)$ and $\rho({}_2T)$ are larger or less than 1, or $\rho({}_1T) < 1 < \rho({}_2T)$. In the first case the population increases, in the second it decreases and in the third case the provided data for f are not sufficient.

3 Optimizing Chemotherapy Regimens

3.1 Simple Optimization Methods for Minimizing Toxicity/Efficacy Ratio in Chemotherapy

Having guaranteed the long range predictability of models with periodic drug dosing, one could develop methods for predicting the optimality of prescribed schedules, when both efficacy to the cancer and toxicity to the patient are concurrently considered [5].

First, a simple optimization problem was defined, as follows: minimize the ratio between the average elimination time of the malignant population and that of the limiting host population, employing a newly defined efficacy coefficient, Z , such that

$$Z = 1 - T_M/T_H,$$

where T_M and T_H denote the mean elimination time of the malignant and the host cells, respectively. The efficacy coefficient, Z tends to unity for treatments that rapidly eliminate the cancer and bear little toxicity for host cells, it will be close to zero for indiscriminating treatments, and will be less than zero for treatments that cause severe damage to the host but have little effect on the cancer. The efficacy coefficient, Z was calculated over large schedule and parameter spaces, for a different formalism of the optimization problem.

In [7], the mathematical properties of the suggested method are analyzed by applying $\{\phi n\}$ series, where $0 < \phi < 1$ and $\{\phi n\}$ is the fraction part of ϕn . Using continued fractions a fast algorithm is put forward for computing the required treatment duration and the desired number of drug applications for eliminating the tumor under the suggested method. Note that this algorithm relies on a model that assumes deterministic temporal cell-cycle parameters and a “bang-bang” (all or none) drug effect.

3.1.1 A Heuristic Method for Personalizing Clinical Cancer Therapy – an Independent Support for the Resonance Phenomenon

For complying with the many constraints of modern clinical oncology strategies, a heuristic method for personalizing drug schedules was developed [8]. The method is based on simulations of the patient’s pharmacodynamics (PD) in models for cancer and host cell dynamics, so that the number of cells that are susceptible to drug at every moment of therapy is calculated. Local search heuristics are then employed for finding the optimal solution, as clinically prescribed. The suggested method does not depend on the exact assumptions of the model, thus enabling its use in complex mathematical descriptions of the biomedical scenario.

The main constituents of the method are two procedures that simulate the growth of cells during treatment and when no treatment is given. For simplicity, logistic growth is assumed for the host cells, while the growth of the cancer cells is taken as exponential.

This model has been employed for examining how to improve regimens of high-dose chemotherapy (HDC), currently believed to be one of the most effective strategies for obtaining higher anti-tumor efficacy for breast cancer. In particular, a high-dose ($210\text{mg}/\text{m}^2$) 3hr infusion of Paclitaxel (half-life: $10\text{--}12\text{hr}$) is routinely used in the adjuvant setting for high-risk patients after mastectomy. Under these regimens dose-limiting myelosuppression is common [23] so that autologous bone-marrow transplantation becomes mandatory.

Using the mathematical algorithms in [8], the killing effect on cancer cells and on host bone-marrow cells was calculated for several treatment schedules of a Paclitaxel-like HDC by a CCPS drug. Subsequently, a comparison has been made between schedules of a single dosing of the drug, and those of a fractionated dosing with the same or a smaller total drug dose. In these calculations it was assumed that cells in the $G1$ and $G2/M$ phases of the cell-cycle are more susceptible to the drug than S-phase cells [24], and that breast cancer cell cycle lasts 40hr , while that of bone-marrow cells is roughly 24hr . The computations show that by splitting the dose

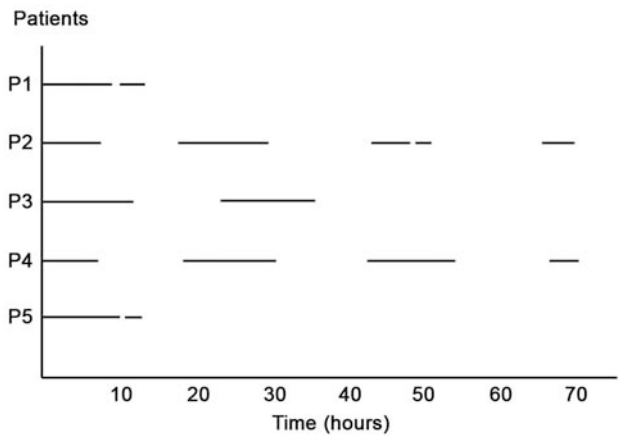


Fig. 1. Optimal treatment regimen for five patients, having different putative cancer indications, P1-P5. The patients are assumed to be similar in all; cell cycle and susceptible phase duration of the drug susceptible host cells being 24hr and 10hr, respectively. The patients vary with respect to the cancer cell cycle parameters: cell cycle duration (16hr (P1, P3), 28hr (P2, P4), 20hr (P5), and duration of the drug susceptible phase (10hr (P1, P2), 6hr (P3, P4), 14hr(P5))

in a specific manner some advantages are possible for increasing the survival/killing ratio of the host and breast cancer cells. It was shown that a HDC treatment regimen, in which a Paclitaxel-like drug exerts a very high cell kill (50%/hr) during 10hr, reduces tumor size to 0.148 of its initial size. However, this treatment causes more damage than cure, as the bone-marrow cell population is reduced even more significantly, to 0.137 of its initial size. This drug effect is reversed when the dosing is fractionated, so as to resonate with the susceptible bone marrow cell-cycle. Under these circumstances toxicity to normal cells is reduced while cancer cell kill is maintained, as expected by the Resonance Theory.

It is important to note that the heuristic optimization method put forward in [8] does not demand any *a priori* treatment periodicity. Nevertheless, its results show that, theoretically, a periodic treatment is optimal over a large range of biologically realistic system parameters. More specifically, two general classes of optimal cancer drug schedules are identified, depending on the temporal cycle parameters of the host and cancer cells. One class comprises one-time intensive treatments, while the other comprises treatments characterized by series of quasi-periodic non-intensive pulsing. We see, then, the Resonance Phenomenon emerging again in an independent model, this time under a more complex optimization method (Fig. 1).

4 From Theory to the Clinic

Mesenchymal Chondrosarcoma (MCS) is a rare malignant disease. One MCS patient was diagnosed with mediastinal located MCS at age 45. Shortly after the resec-

tion of the primary tumor, multiple bilateral pulmonary nodules were discovered. The patient underwent aggressive chemotherapy, but disease progression was not arrested, and additional liver and bone metastases appeared. The patient also developed severe myelosuppression with pancytopenia due to toxic side-effects of the prolonged chemotherapy.

4.1 Integrating in Silico and in Vivo Models for Treatment Personalization

To determine the best possible treatment for the MCS patient, tumor fragments, taken from his lung metastases, were implanted in mice (denoted tumorgrafts), and different pharmacotherapy regimens were applied to the animals.

A general mathematical model for angiogenesis-dependent solid tumor growth was used for model simulations, replicating the experiments performed in the tumorgrafts, for predicting the MCS dynamics in the control and treated animals [9]. PK/pharmacodynamics (PD) models of potential drugs were constructed, using publicly available data. In addition, qualitative chemosensitivity tests of several cytotoxic drugs were performed on tumor cells from the patient's biopsy. Incorporating the data of these chemosensitivity tests into the calculations allowed a certain level of personalization of the general PK/PD models.

The mathematical model of the MCS tumorgrafts was successfully validated in mice with the average accuracy of 87%. Subsequently, gene expression analysis of key proteins in the grafted tumors and in the MCS patient was performed in order to adjust the model to describe the tumor dynamics in the patient. The resulting personalized model of the patient's disease was then used to perform patient-specific predictions of various anti-cancer treatments.

Guided by the results of the personalized in silico/in vivo combined model, the clinicians administered the MCS patient once-weekly regimen of docetaxel (DOC). Previously, this regimen was found least toxic to hematopoiesis, e.g., [3]. Eventually, the patient had a dramatic response to therapy with an immediate substantial recovery of all 3 blood elements (hemoglobin, white blood cells and platelet count). Soft tissue disease in the lungs and liver remained stable and the patient enjoyed a period of good quality of life, ending only after pulmonary progression of his disease to which he finally succumbed [14].

4.2 The Rate of Angiogenesis Determines the Optimal Inter-Dose Interval of Chemotherapy

The simulations of the human MCS model for DOC delivery every 7, 14, 21 or 28 days (keeping the same average weekly dose), showed that once weekly regimen is more efficacious than all other tested alternatives. However, suggesting this DOC administration regimen may be problematic in some patients. Therefore, it would be useful to identify the patients that are more likely to benefit from the weekly chemotherapy schedules.

Cytotoxic agents, such as DOC, disturb the dynamic equilibrium between the growing tumor mass and the vessel bed that supports it, by direct killing of tumor cells. The mathematical model for angiogenesis-dependent tumor growth takes account of the cascade of compensating events which is triggered as a consequence of chemotherapy administration.

To assess tumor recovery from the cytotoxic drug shock, let us define tumor growth inhibition (TGI) in terms of the tumor volume before and after the treatment: TGI has a value of zero if the volume of the treated tumor equals to that in the control, untreated, tumor at the given time point. Larger TGI values indicate bigger inhibition, while negative TGI values indicate the situation where the treated tumor is bigger (in terms of living cells volume) than the untreated one: in other words, negative TGI values mean that the treatment was harmful, rather than beneficial.

In the human MCS model simulations [14], TGI values 7 days after a single DOC administration was 46%. As time goes on, this residual cytotoxic effect decreases to the value of 10% at day 21 – a difference of 36%. Thus, if the second DOC dose is delivered on day 21, after a substantial tumor recovery, the overall efficacy of the treatment would be small, compared to once-weekly regimens.

Moreover, if the human MCS model is simulated with VEGF secretion rate reduced by a factor of two, compared to its original value, the predicted decrease in TGI from day 7 to day 21 is only 29% (39% and 10%, respectively), indicating slower recovery from drug-induced tumor inhibition. In contrast, if the rate of new vessel formation is doubled, as compared to that calculated for the real MCS patient, the difference in the extent of tumor inhibition by DOC dose, between day 7 and day 21, increases to 92%; 69% TGI in day 7 versus –22% in day 21. This growth is made possible due to the extensive formation of blood vessels, triggered by the chemotherapy [14].

These results imply that for patients having less intensive angiogenesis, the less dense (for example tri-weekly) regimens would be approximately as efficacious as the weekly one, thus providing the clinicians more treatment alternatives [14].

5 Conclusion

The work reviewed in this chapter shows both theoretically and experimentally, that the advantage of dose-dense chemotherapy is not universal, but rather, depends on the patient's cytokinetic and angiogenic parameters and the length of the inter-dosing interval itself.

We believe that the above-reported treatment personalization study marks a transition point in the status of mathematical modelling in biomedicine. This is one of the first times that a mathematical model generates quantitative predictions that are prospectively validated in the clinic.

In general, the modelling procedures reviewed here and their experimental verification provide solid grounds for the use of rigorous biomathematical models in drug development and in the clinic. Clearly, this is an uphill struggle, which requires

patience and endurance. However, we believe that in the long run decision-making in medicine will be primarily based on biomathematical modelling.

Acknowledgements We thank the Chai Foundation for financial support.

References

1. Agur, Z.: Persistence in uncertain environments. In: Freedman, H.I., Strobeck, C. (eds.) *Population Biology*, pp. 125–132. Lecture Notes in Biomathematics 51. Springer, Heidelberg (1982)
2. Agur, Z.: Randomness, synchrony and population persistence. *J. Theor. Biol.* **112**, 677–693 (1985)
3. Agur, Z.: From the evolution of toxin resistance to virtual clinical trials: the role of mathematical models in oncology. *Future Oncol.* **6**, 917–927 (2010)
4. Agur, Z., Arnon, R., Sandak, B., Schechter, B.: The effect of the dosing interval on myelotoxicity and survival in mice treated by cytarabine. *Eur. Jour. Cancer.* **28A**, 1085–1090 (1992)
5. Agur, Z., Arnon, R., Schechter, B.: Reduction of cytotoxicity to normal tissues by new regimes of cell-cycle phase-specific drugs. *Math. Biosci.* **92**, 1–15 (1988)
6. Agur, Z., Cojocaru, L., Mazor, G., Anderson, R., Danon, Y.: Pulse mass measles vaccination across age cohorts. *Proc. Natl. Acad. Sci. USA* **90**, 11698–11702 (1993)
7. Agur, Z., Dvir, Y.: Use of knowledge on $\{\phi n\}$ series for predicting optimal chemotherapy treatment. *Random & Computational Dynamics* **2**, 279–286 (1996)
8. Agur, Z., Hassin, R., Levy, S.: Optimizing chemotherapy scheduling using local search heuristics. *Operations Research* **54**, 829–846 (2006)
9. Arakelyan, L., Vainstein, V., Agur, Z.: A computer algorithm describing the process of vessel formation and maturation, and its use for predicting the effects of anti-angiogenic and anti-maturation therapy on vascular tumor growth. *Angiogenesis* **5**, 203–214 (2002)
10. Cojocaru, L., Agur, Z.: A theoretical analysis of interval drug dosing for cell-cycle-phase-specific drugs. *Math. Biosci.* **109**, 85–97 (1992)
11. Chabner, B.A., Murphy, M.J. Jr.: Breast cancer: A tale of two centuries: With implications for understanding cancer metastasis and cancer stem cell biology. *Oncologist* **10**, 369 (2005)
12. Dibrov, B.F., Zhabotinsky, A.M., Neifakh, Yu.A., Orlova, M.P., Churikova, L.I.: Mathematical model of cancer chemotherapy. Periodic schedules of phase-specific cytotoxic-agent administration increasing the selectivity of therapy, *Math. biosci.* **73**, 1–31 (1985)
13. Evans, W.E., McLeod, H.L.: Pharmacogenomics—drug disposition, drug targets, and side effects. *N. Engl. J. Med.* **348**, 538–549 (2003)
14. Gorelik, B., Ziv, I., Shohat, R., Wick, M., Webb, C., Hankins, D., Sidransky, D., Agur, Z.: Efficacy of once weekly docetaxel combined with bevacizumab for patients with intense angiogenesis: validation of a new theranostic method in mesenchymal chondrosarcoma xenografts. *Cancer Research* **68**, 9033–9040 (2008)
15. Hart, D., Shochat, E., Agur, Z.: The growth law of primary breast cancer as inferred from mammography screening trials data. *Br J Cancer* **78**, 382–387 (1998)
16. Kheifetz, Y., Kogan, Y., Agur, Z.: Long-range predictability in models of cell populations subjected to phase-specific drugs: growth-rate approximation using positive compact operators properties. *Math. Mod. & Meth. Appl. Sci.* **16**, 1–18 (2006)
17. Murray, J.D.: *Mathematical Biology*, Chapter 1.7, pp. 29–33. Springer, Heidelberg (1993)
18. Norton, L., Simon, R.: Tumor size, sensitivity to therapy, and design of treatment schedules. *Cancer Treat. Rep.* **61**, 1307–1317 (1977)
19. Norton, L., Simon, R.: The Norton-Simon hypothesis revisited. *Cancer Treat. Rep.* **70**, 163–169 (1986)
20. Retsky, M.: Metronomic chemotherapy was originally designed and first used in 1994 for early stage cancer – why is it taking so long to proceed? *J. Bioequiv. Bioavailab.* **3**, 1195–1216 (2011)

21. Retsky, M.W., Swartzendruber, D.E., Wardwell, R.H., Bame, P.D.: Is Gompertzian or exponential kinetics a valid description of individual human cancer growth? *Medical Hypotheses* **33**, 95–106 (1990)
22. Riesz, F., Sz.-Nagy, B.: *Functional analysis*. Frederick Ungar Publishing, New York (1965)
23. Schiller, J.H., Storer, B., Tutsch, K., Arzooonian, R., Alberti, D., Feierabend, C., Spriggs, D.: Phase I trial of 3-hour infusion of paclitaxel with or without granulocyte colony-stimulating factor in patients with advanced cancer. *J. Clin. Oncol.* **12**, 241–248 (1994)
24. Sena, G., Onado, C., Cappella, P., Montalenti, F., Ubezio, P.: Measuring the complexity of cell cycle arrest and killing of drugs: kinetics of phase-specific effects induced by taxol. *Cytometry* **37**, 113–124 (1999)
25. Shulgin, B., Stone, L., Agur, Z.: Pulse vaccination strategy in the sir endemic model. *Bull. Math. Biol.* **60**, 1123–1148 (1998)
26. Thompson, J.R., Brown, B.W.: *Cancer Modelling*. Marcel Dekker, New York (1987)
27. Ubezio, P., Tagliabue, G., Schechter, B., Agur, Z.: Increasing 1- β -D-Arabinofuranosylcytosine efficacy by scheduled dosing interval based on direct measurement of bone marrow cell kinetics. *Cancer Res.* **54**, 6446–6451 (1994)
28. Vainstein, V., Kirnasovsky, O.U., Kogan, Y., Agur, Z.: Strategies for cancer stem cell elimination: Insights from mathematical modeling. *J. Theor. Biol.* **298**, 32–41 (2012)
29. Watters, J.W.: McLeod HL. Cancer pharmacogenomics: current and future applications. *Biochim Biophys Acta.* **1603**, 99–111 (2003)
30. Webb, G.F.: Resonance Phenomena in cell population chemotherapy models. *Rocky Mountain J. Math.* **20**, 1195–1216 (1990)
31. Webb, G.F., Johnson, M.E.: Resonances in aged structured cell population models of periodic chemotherapy. *Internat. J. Appl. Sci. Comp.* **3**, 57–67 (1996)

A Systems Biomedicine Approach for Chronotherapeutics Optimization: Focus on the Anticancer Drug Irinotecan

Annabelle Ballesta, Jean Clairambault, Sandrine Dulong, and Francis Levi

Abstract Most physiological functions in mammals display rhythms of period around 24 h, also called *circadian rhythms*. This temporal organization of the organism results in variations in the toxicity and efficacy of many antidrugs with respect to their circadian time of administration. Recent experimental and clinical results support the need of personalizing the chronomodulated administration pattern according to the patient genetic and circadian profile. We propose here a systems biomedicine approach for the optimization of the circadian delivery of irinotecan (CPT11), an anticancer drug approved for the treatment of colorectal cancer. First, CPT11 pharmacokinetics-pharmacodynamics (PK-PD) has been experimentally studied in Caco-2 colon cancer cell cultures. After cell synchronization, circadian rhythms with a period of 26 h 50 (SD 63 min) were observed as well as circadian variations in the protein amount of DNA-bound topoisomerase 1 in presence of CPT11, a marker of the drug PD. A mathematical model of CPT11 molecular PK-PD was then designed, fitted to experimental data and used in therapeutic optimization procedures. We adopted the therapeutics strategy of maximizing efficacy in non-synchronized cells, considered as cancer cells, under a constraint of

A. Ballesta (✉) · J. Clairambault

INRIA Paris-Rocquencourt, BANG project-team, Domaine de Voluceau, BP105, 78153 Le Chesnay Cedex, France

e-mail: annabelle.ballesta@inria.fr, jean.clairambault@inria.fr

S. Dulong

INSERM, U776 “Rythmes biologiques et cancers”, Hôpital Paul Brousse, F-94807 Villejuif, France and Université Paris-Sud, UMR-SO776, F-91405 Orsay, France

e-mail: sandrine.dulong@inserm.fr

F. Levi

INSERM, U776 “Rythmes biologiques et cancers”, Hôpital Paul Brousse, F-94807 Villejuif, France and Université Paris-Sud, UMR-SO776, F-91405 Orsay, France and Assistance Publique – Hôpitaux de Paris, Unité de Chronothérapie, Département d’oncologie médicale, Hôpital Paul Brousse, Villejuif, France

e-mail: francis.levi@inserm.fr

A. d’Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_16, © Springer-Verlag Italia 2012

maximum toxicity in synchronized cells, representing healthy ones. For any dose of CPT11, optimal exposure durations varied from 3h40 to 7h10. Optimal schemes started between CT2h10 and CT2h30, a time interval corresponding to 1 h 30 to 1 h 50 before the nadir of CPT11 bioactivation rhythm in healthy cells. The second step of our approach has consisted in optimizing CPT11 administration in mice. Within the European project TEMPO, CPT11 chronotoxicity has been studied in mice and three classes have been determined with regards to CPT11 best circadian time of administration (i.e. the time which induces the minimal toxicity). We have developed a whole-body physiologically-based model of CPT11 PK-PD based on the previous in vitro study, which aims at identifying molecular biomarkers which could discriminate between the mouse classes and at designing optimal chronomodulated infusion scheme for each of them. Parameters of the model have been estimated for B6D2F1 male mice (chronotoxicity class 2) by fitting available data on tissular PK for two different circadian times of administration and on circadian rhythms of relevant proteins. The same parameter estimation is in progress for the two other classes, which will allow the comparison of the three parameter sets in order to identify molecular differences between the classes. Moreover, optimization algorithms will be applied to the data-calibrated model to design theoretically optimal chronomodulated scheme of administration. In a clinical perspective, this whole-body model of CPT11 PK-PD designed for mice will be adapted to humans by keeping its main structure and resizing parameter values. This will provide clinicians with a new tool towards the personalization of CPT11 administration according to the patient genetic and circadian profile.

1 Chronotherapeutics of Cancer

Most physiological functions in mammals display rhythms of period around 24 h, also called *circadian rhythms* from the latin *circa diem*, around a day. Rest-activity rhythm, core temperature, cardiac rhythm, blood pressure, or intracellular concentrations of metabolic enzymes present variations over the 24-hour span [28]. This circadian regulation allows an anticipation upon the energetic supply of the body. It contributes to the optimal control of the energy needed for the smooth functioning of the organism.

1.1 The Circadian Timing System

1.1.1 The Central Pacemaker

The circadian timing system in mammals is controlled by a central pacemaker located in the hypothalamus in the brain: the suprachiasmatic nuclei (SCN). The SCN display endogenous circadian oscillations whose period depends on the genetic background of the individual and ranges from 23 to 25 h in humans [38]. Endogeneity of those rhythms has been demonstrated by their persistence when the individual is kept under constant darkness or constant light. The self-sustained oscillating SCN

are also entrained by environmental factors such as light, socio-professional activities or food intake which exactly calibrate their period each day to 24 h [28].

The central pacemaker exerts a control on peripheral clocks through different physiological signaling. Indeed, each nucleated cell of the organism is endowed with a molecular circadian clock which is constituted of clock genes interconnected in regulation feedback loops. The structure of this gene network intrinsically generates oscillations in clock genes mRNA and protein levels which in turn induce rhythms in the expression of other genes, in particular of those involved in drug metabolism, cell cycle, DNA repair, apoptosis and angiogenesis [29].

1.1.2 Molecular Description of the Cellular Clock

To be more specific, circadian oscillations are generated by transcriptional or post-transcriptional feedback loops. Clock genes can be categorized into two groups. On one hand, the positive category is composed of transcriptional factors such as BMAL1, CLOCK, or NPAS2. On the other hand, the negative group mainly consists of Cryptochrome (Cry) and Period (Per) genes. BMAL1-CLOCK or BMAL1-NPAS2 heterodimers activate Cry and Per transcription which leads to the subsequent accumulation of CRY and PER proteins in the cytosol. Those proteins then associate into heterotypic complexes which inhibit the action of BMAL1-CLOCK/NPAS2 dimers and thus repress their own expression. This inhibition decreases CRY and PER protein amounts which goes beneath the threshold concentration allowing autorepression, and triggers a new cycle of Cry and Per transcription [29].

In parallel, positive (BMAL1, CLOCK, NPAS2) and negative (CRY, PER) transcriptional regulators respectively activate and repress the transcription of the orphan nuclear receptor *Rev – Erba* α , and probably that of *Rev – Erba* β as well. *Rev – Erba* α/β proteins then strongly repress Bmal1 transcription and probably that of Clock but in a weaker manner [36]. This coupling is not mandatory to generate oscillations of the system but rather participates in the robustness of the molecular clock.

1.1.3 Experimental Assessment of the Circadian Timing System

In cell culture. Mammalian nucleated cells are endowed with a molecular circadian clock whose rhythms persist *in vitro* [4, 9, 43]. However, in the absence of an external synchronizer, the millions of cells which are contained in a single Petri dish do not oscillate neither with the same period nor with the same phase [9, 43]. Cell synchronization may be undertaken through a seric shock (exposure to a large amount of nutrient, [5]), through drug exposure [22] or temperature cycles [8]. The main effect of those synchronizers lies in the simultaneous reset of all cellular clocks which then oscillate in synchrony with a period close to 24 h. *In vitro* measurements of circadian rhythms of gene expression are then possible during several consecutive days. The time unit is then Circadian Time (CT), expressed in hours, where CT0 is the beginning of cell synchronization.

In rodents. Experimental demonstration of circadian rhythms can be undertaken in preclinical models such as the Rat or the Mouse. This involves the use of specific equipments dedicated to chronobiologic studies. They are composed of isolated shelves in which groups of animals are synchronized at different circadian stages using light control. Time is expressed in Zeitgeber Time (ZT) or in Hours After Light Onset (HALO). The typical synchronization pattern consists of an alternation of 12 hours of light and 12 hours of darkness (LD12:12). Animals can also be kept under constant darkness (DD) or constant light (LL). Animals may then be sacrificed at particular circadian stages to allow relevant biological measurements.

In humans. Minimally- or non-invasive procedures are nowadays available to provide high quality and reliable data about the patient circadian clocks and their coordination. Whenever circadian physiology is concerned, frequent sampling over several days has been advocated and used in order to provide an insight into the Circadian Timing System (CTS) of the patient. Chronobiology rhythms can be measured in patients in several ways. Rest-activity rhythms can be monitored through actimetry which has been considered as the method of choice regarding reliability, convenience and continuity in recordings [33]. Then temperature rhythms can be non-invasively assessed using different devices [33]. Moreover salivary samples can be collected in order to measure gene expression levels in the oral mucosa, those of cortisol and melatonin being considered as relevant circadian biomarkers in cancer patients [28, 33].

1.2 Current Knowledge on Chronotoxicity and Chronoefficacy of Anticancer Drugs

The circadian organization in mammals results in variations in the toxicity and efficacy of many drugs with respect to their circadian time of administration, named chronotoxicity and chronoefficacy. We focus here on the chronopharmacology of anticancer drugs.

Chronotoxicity in preclinical models. Tolerance to at least 40 anticancer agents is modified by their circadian time of administration in rats and mice [28]. Survival and maximal body weight loss after a drug administration of a potentially-lethal dose vary by a factor ranging from 2 to 10 according to the circadian time of injection. Those large differences are observed independently of the administration mode – oral, intra-peritoneal or intra-arterial – and of the number or frequency of administrations [29]. Moreover, circadian rhythms in the tolerability of anticancer drugs persist in rodents kept in constant darkness or in constant light, which demonstrates their endogeneity [39].

Optimal administration times of anticancer drugs are spread over the 24 h span and cannot be predicted by the drug pharmacological classification or main toxicity target organs [28]. The combination of chemotherapy drugs does not seem to affect their optimal administration times which remain the same observed when they are administered as single agents. In particular, this was demonstrated for the com-

bination doxorubicin-cisplatin in rats, for irinotecan-oxaliplatin and gemcitabine-cisplatin in B6D2F1 mice, and for docetaxel-doxorubicin in C3H/He mice [28]. The circadian control of the determinants of anticancer drug chronotoxicity seems to persist after the first drug exposure, at least when the latest is given at its best tolerability time [28].

Chronoefficacy in preclinical models. Circadian timing also affects the antitumor efficacy of at least 28 anticancer drugs in rodents with various kinds of malignancies [28]. Appropriately circadian-timed and dosed chemotherapy with one or several drugs at least halves tumor growth rate and/or significantly increases life span in tumor-bearing mice [28].

The circadian pattern of chronoefficacy usually coincides with that of chronotolerance. This is true for cytostatics, interferons, antiangiogenic agents, and cell cycle inhibitors, as well as for combination chemotherapy, such as irinotecan-oxaliplatin, gemcitabine-cisplatin, and docetaxel-doxorubicin, three widely used clinical regimens [28]. Experimental chronotherapeutics thus strongly supports circadian timing as a relevant method for improving anticancer treatments.

Chronotoxicity and chronoefficacy in patients with metastatic colorectal cancer. The conception of multichannel programmable delivery pumps has allowed the clinical development of cancer chronotherapeutics. Those pumps deliver a combination of anticancer drugs to non hospitalized patients, according to previously-implemented administration schemes (e.g. constant, semi-sinusoidal, linearly increasing or decreasing infusion) [28].

Over 100 phase I and II clinical trials of cancer chronotherapeutics have involved patients with advanced or metastatic cancer of almost all origins according to a PubMed search. Randomized phase III trials have compared chronotherapeutics administration schemes to their paired constant-rate infusion schedule lasting an integral multiple of 24 h and involving the same drug doses [28].

In particular, two international randomized phase III trials have compared the chronomodulated scheme ChronoFLO5 to an equivalent constant delivery in 278 patients with metastatic colorectal cancer. ChronoFLO5 combines the daily delivery of oxaliplatin over 11.5 h with peak flow rate at 4:00 p.m. and that of 5-FU-leucovorin over 11.5 h with peak flow rate at 4:00 a.m., for 5 consecutive days. The other cohort of patients received the same doses of the same three drugs, at a constant rate over the same 5-day span. In those trials, chronomodulated delivery reduced the incidence of grade 3–4 mucositis by fivefold and halved the incidence of peripheral sensory neuropathy [28].

A third randomized trial has compared the chronomodulated administration of the same three drugs over 4 days (ChronoFLO4) to a conventional constant-rate infusion over 2 days (FOLFOX2) in 564 patients with metastatic colorectal cancer. Overall survival, the main endpoint in this large international study, did not differ as a function of treatment schedule. However, the relative risk of an earlier death on ChronoFLO4 significantly increased by 38% in women and significantly decreased by 25% in men compared with conventional delivery [28]. A recent meta-analysis of these three randomized trials in 842 patients with metastatic colorectal cancer

confirms that the three-drug chronomodulated infusion achieves similar or worse efficacy compared with conventional delivery in women. In men, however, the same ChronoFLO treatment significantly increased tumor response and survival compared with conventional delivery, independently of all other prognostic factors. This result highlights the need for chronotherapeutic personalization in which chronomodulated administration schemes would be tailored according to the patient circadian and genetic profile.

2 Focus on the Anticancer Drug Irinotecan

We now focus on the anticancer drug irinotecan (CPT11) which was clinically approved for the treatment of colorectal cancer in 1994. CPT11 is part of the camptothecin family whose active principle comes from the Chinese tree *Camptotheca acuminata decne*. Other anticancer drugs are derived from the same molecule such as topotecan which is widely used for breast cancer treatment, or 9-aminocamptothecin. CPT11 administration may be responsible for severe toxicities in the bone marrow and in the intestine both in pre-clinical models and in cancer patients. The improvement of tolerability to CPT11 is thus a current clinical concern in particular in the case of digestive cancers. CPT11 toxicity and efficacy display rhythms both in mice and in cancer patients. Our approach aims at taking advantage of those circadian rhythms to reduce CPT11 toxicity and increase its efficacy.

2.1 Molecular Pharmacology of Irinotecan

Here we give molecular details about CPT11 PK-PD. CPT11 is an inhibitor of topoisomerase 1 (TOP1). The TOP1 enzyme is present in all normal nucleated cells. Its function is to relax supercoiled DNA [35]. Indeed transcription or replication mechanisms may increase the tension in DNA strands. The TOP1 enzyme wraps DNA and temporarily cuts one strand which rotates around the DNA molecule. Then TOP1 dissociates allowing the DNA strand to reconnect.

CPT11 attaches DNA/TOP1 complexes when the DNA strand is cut, thus creating CPT11/TOP1/DNA complexes whose lifetime is much longer than that of DNA/TOP1 ones. However those ternary complexes can still spontaneously dissociate. Collisions between those complexes and transcription or replication mechanisms induce irreversible single- or double-stranded DNA damage which triggers DNA repair, cell cycle arrest and may lead to cell apoptosis.

Concerning CPT11 pharmacokinetics, CPT11 is a prodrug which has to be activated into its metabolite SN38 under the activity of carboxylesterases (CES) enzymes. SN38 cytotoxic activity is a hundred to a thousand times higher than that of CPT11 [4]. SN38 is deactivated into SN38G by glucuronidation through the activity of UGT1A enzymes. The inverse reaction which consists in SN38G re-activation into SN38 occurs in several cancer cell lines and in the intestinal cells where the β -glucuronidase enzyme is expressed and catalyzes the reaction. CPT11 is converted in two other metabolites, namely APC and NPC, through the enzymatic activ-

ity of P450 3A4/3A5 cytochromes. Only NPC can be subsequently activated into SN38.

The uptake of CPT11 in the intracellular medium seems to occur in a passive manner involving a free diffusion through the cell membrane, as suggested by the absence of saturation in the uptake rate in the case of high drug concentrations [4]. Conversely, CPT11 and its metabolites are actively expelled outside of the cells by transporters of the ATP-Binding Cassette (ABC) family. Transporters involved in CPT11 efflux are mainly ABCB1 (P-gp), ABCC1 (MDR1), ABCC2 (MDR2) and ABCG2 (BCRP) [4].

2.2 Chronotoxicity and Chronoefficacy of Irinotecan

The expression of several genes involved in CPT11 PK-PD displays circadian variations in cell culture, in mice and in humans. This is the case for P450 cytochromes whose mRNA expression, protein amount and enzymatic activity vary according to the circadian time in mice and in humans [26, 34]. The activation enzymes CES1 and 2, the deactivation protein UGT1A1, the drug target TOP1 and the efflux transporters display circadian rhythms in their gene expression in cell culture and in mice [3, 4, 20, 23–25, 32, 41, 44]. Those molecular circadian rhythms induce different responses to CPT11 exposure depending on its administration time.

In mice. In B6D2F1 male mice, all toxicity criteria such as survival, body weight loss, hematological or intestinal lesions display circadian variations. CPT11 administration given as a single agent or combined with oxaliplatin is better tolerated at the end of the rest phase, at ZT11, both in GOS tumor-bearing and non-bearing mice [15, 19]. A similar chronotolerance has been observed in ICR male mice in which CPT11 intraperitoneal injection at ZT10 induce half leucopenia than the same injection performed at ZT22 [31].

Concerning CPT11 efficacy on tumor progression, CPT11 alone or in combination with oxaliplatin is more efficient when injected at ZT7 or 11 in GOS-tumor-bearing mice [19]. On the contrary, in ICR male mice bearing S-180 sarcoma, CPT11 antitumor activity is higher in mice treated at ZT10 compared to those treated at ZT22. This difference in CPT11 chronoefficacy could be explained by the fact that S-180 sarcoma is a low-growing tumor whereas GOS cancer cells proliferate much faster. Those different growth rates could be responsible for the differences in the chrono-chemosensitivity, as theoretically demonstrated in a recent study [6].

In patients with metastatic colorectal cancer. Experimental results on CPT11 chronoactivity in B6D2F1 male mice are in favor of an administration of the drug at the end of the rest phase. Thus, a chronomodulated administration scheme has been conceived for cancer patients and consists in a sinusoidal drug infusion from 2am to 8am, peaking at 5am. This delivery pattern was administered in monotherapy to patients suffering from metastatic colorectal cancer at the dose of 350 mg/m² and then in a randomized study [18, 27]. It was compared with a conventional administration of the same dose over 30 min at 10 am. Those two clinical trials involved

respectively 30 and 36 patients. They concluded to a slightly better drug tolerance in the chronomodulated cohort and to differences in CPT11 pharmacokinetics and biotransformation between the two groups of patients.

CPT11 chronotherapy has then been combined with the intravenously administered ChronoFLO4 scheme in two clinical trials. This ChronoIFLO protocol associated CPT11 with 5-FU-leucovorin and oxaliplatin in 70 colorectal cancer patients whose tumor had become resistant to conventional chemotherapies. This drug combination achieved an efficient and durable tumor control with a lower hematological toxicity than that observed when such quadritherapy was given in a constant conventional way [14, 17]. This scheme combined or not with cetuximab has been studied in several clinical trials which demonstrate an excellent antitumor activity and an acceptable toxicity, in particular concerning hematological lesions [16].

3 Optimization of Irinotecan Exposure in Cell Culture

We propose here a combined experimental and mathematical approach to optimize CPT11 circadian delivery [4]. The first step of our approach consisted in a proof of concept which involved *in vitro* experiments on human cell cultures and *in silico* mathematical modeling [4]. CPT11 molecular pharmacokinetics (PK) and pharmacodynamics (PD) were studied in human colorectal adenocarcinoma Caco-2 cells. An ODE-based mathematical model of CPT11 PK-PD was developed. It guided the design of experiments which were performed in order to estimate parameter values of the model. Optimization procedures were then applied to the data-calibrated model in order to compute theoretically optimal exposure schemes for the Caco-2 cell line.

3.1 A Mathematical Model of CPT11 PK-PD and Its Calibration to Caco-2 Cell Experimental Data

CPT11 PK-PD were experimentally studied in Caco-2 cells [4]. CPT11 accumulated in the intracellular medium where it was bioactivated into its metabolite SN38. The pre-incubation of cells with verapamil, a non-specific inhibitor of ABC transporters, increased CPT11 intracellular accumulation, thus demonstrating the involvement of those efflux pumps in CPT11 transport. After cell synchronization by a seric shock which defined the circadian time (CT) 0, circadian rhythms of period 26 h 50 (SD 63 min) were observed in the expression of the three clock genes REV-ERB α , PER2, and BMAL1; and of six metabolic genes: the drug target topoisomerase 1 (TOP1), the activation enzyme CES2, the deactivation enzyme UGT1A1, and the four ABC transporters ABCB1, ABCC1, ABCC2, ABCG2. On the contrary, TOP1 protein level and activity remained constant. The amount of DNA-bound TOP1 in the presence of CPT11 is a PD marker of the drug and it displayed circadian rhythms as it was equal to $47 \pm 5.2\%$ of the total amount of TOP1 protein after an exposure to CPT11 at CT14, as compared to $35.5 \pm 1.8\%$ after an exposure at CT28.

Molecular pathways of CPT11 PK-PD were modeled according to biological data published in the literature and experimental results obtained in Caco-2 cells. Briefly, the intracellular uptake of CPT11, SN38 and SN38G was assumed to be passive and modeled as a free diffusion across a membrane. CPT11 and SN38 efflux were mediated respectively by *ABC_CPT* (mainly standing for the sum of activities of ABCB1, ABCC1, ABCC2) and *ABC_SN* (for ABCC1, ABCC2, ABCG2). Efflux followed Michaelis-Menten kinetics as experimentally demonstrated in the literature [4]. Diffusion from inside to outside of the cells was neglected. CPT11 was bioactivated into SN38 through CES representing the sum of all carboxylesterases activity. This pathway was also modeled by Michaelis-Menten kinetics. SN38 was expected to be deactivated into SN38G which was modeled by Michaelis-Menten kinetics in which the mathematical variable UGT stands for the sum of all UGT1As enzymatic activities.

CPT11 ability to bind to TOP1 was neglected so that SN38 was the only molecule able to stabilize DNA/TOP1 complexes into DNA/TOP1/SN38 ones (denoted by *Compl* hereafter). Those ternary complexes were able to spontaneously dissociate or could be converted into irreversible complexes (*Icompl*) after collision with transcription or replication mechanisms [4]. Parameters of this CPT11 PK-PD model were estimated from experimental data in Caco-2 cells combined with information from literature using a bootstrap approach [4].

3.2 Theoretical Optimization of CPT11 Exposure in Caco-2 Cells

Optimization procedures were then applied to the data-calibrated model in order to compute theoretically optimal exposure schemes for Caco-2 cells. Synchronized cells were considered as healthy cells and non-synchronized cells as cancer ones as the circadian organization is often disrupted in tumor tissues. The adopted therapeutics strategy consisted in maximizing DNA damage in cancer cells under the constraint that DNA damage in the healthy population remained under a tolerability threshold. We considered administration schemes in the form of a cell exposure to an initial extracellular concentration of CPT11, over 1 to 27 h, starting at a particular CT.

For all considered doses, the optimal exposure scheme consisted in administering CPT11 over 3 h 40 to 7 h 10 starting between CT2h10 and CT2h30 which corresponded to 1 h 30 to 1 h 50 before the nadir of CES protein amount [4]. The optimal schemes were not centered on the nadir of rhythm but rather extended after it, when *UGT*, *ABC_CPT* and *ABC_SN* amounts were higher and therefore protected more efficiently healthy cells. For any maximum allowed toxicity, the optimal duration did not exceed 7 h 10, highlighting the need of short exposure durations to optimally exploit the temporal difference between healthy and cancer cells. Regarding efficacy, those optimal schemes induced twice more DNA damage in cancer cells than in healthy ones. A clinical interpretation can be obtained by rescaling to 24 h those results for Caco-2 cells which displayed a period of 26 h 50. Thus, an optimal administration of CPT11 to cancer patients should result in the presence of the drug

in the blood during 3 h 30 to 6 h 30, starting 1 h 30 to 1 h 40 before the minimum value of CES activity in the patient.

4 Optimization of Irinotecan Administration in Mice

CPT11 chronotoxicity was studied in three mice strains, in both male and female animals. From this study, three classes were characterized with respect to CPT11 chronotolerance. Those recent experimental studies demonstrated the existence of three classes of mice regarding CPT11 chronotoxicity which was assessed by survival, body weight loss, intestinal and hematological toxicities [1]. Female mice of the strain B6D2F1 represented the first class and showed worst tolerability after an injection of CPT11 at ZT3 and best tolerability at ZT15. Class 2 was constituted by B6D2F1 male mice and displayed worst toxicity at ZT23 and best toxicity at ZT11. Finally, class 3 was B6CBAF1 female mice and showed worst tolerability at ZT7 and best tolerability at ZT15.

Our combined *in vivo* and *in silico* approach aimed at characterizing the three chronotoxicity classes at the molecular level and at designing optimal administration schemes for each of them. In order to address this concern, the mathematical model which was built for Caco-2 cell culture was adapted to design a whole-body physiologically-based model of CPT11 PK-PD. Parameters were estimated for class 2 (B6D2F1 male mice) by fitting both blood and tissue pharmacokinetics data together with measurements of circadian rhythms of proteins involved in CPT11 pharmacology. Similar parameter estimations are ongoing for classes 1 and 3.

4.1 A Whole-Body Physiologically-Based Model of CPT11 PK-PD

We have built a physiologically-based whole-body model of CPT11 PK-PD. It is composed of compartments which represent the simulated organs. The liver is modeled for its major part in CPT11 metabolism. Then the two main toxicity targets of the drug which are the intestine and the bone marrow are modeled, as well as the blood, and the tumor in order to take into account the drug efficacy. Finally, the Non-Eliminating Tissue (NET) compartment stands for all other tissues such as muscles or skin (Fig. 1).

CPT11 and its metabolites circulate in and out of the tumor, the bone marrow, the NET and the liver compartments through the blood circulation. concerning the intestine, it is modeled by two compartments which represent the cells of the intestinal mucosa and the intestinal lumen. A bidirectional transport is assumed between the mucosa and the lumen. Moreover, the drug and its metabolites can be transported from the intestinal cells to the liver through the hepatic portal vein. The entero-hepatic circulation is modeled by a drug transport from the liver to the intestinal lumen which stands for biliary excretion [13].

Finally, renal clearance was modeled as degradation terms for CPT11, SN38 and SN38G in the blood compartment [42]. The intestinal lumen compartment

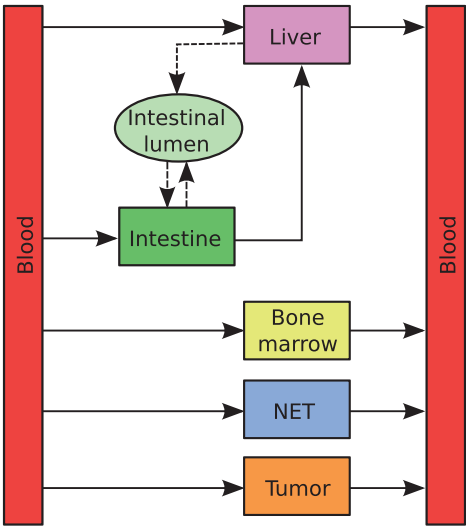


Fig. 1. Schematic view of the whole-body physiologically-based model of CPT11 PK-PD Detailed mathematical equations of the model are presented in the Appendix

presents degradation terms accounting for CPT11 and SN38 intestinal clearance, SN38G being exclusively eliminated through the kidneys [37]. Degradation speeds are assumed to be proportional to the drug concentrations. Each compartment contains an adaptation of the existing mathematical model of CPT11 tissue PK-PD as described in the following [4].

Modeling CPT11 pharmacokinetics. The uptake of CPT11 and its metabolites from the blood to the different organs is modeled as passive diffusion [4]. On the contrary, the drug efflux is assumed to be mediated by active transporters (such as ABC transporters) and is therefore modeled using Michaelis-Menten kinetics. CPT11 bioactivation into SN38 through CES activity occurs in every compartment but the NET one. SN38 glucuronidation into SN38G is assumed to be active in every compartment except in the blood and the NET compartments.

Bidirectional flux between the intestinal mucosa cells and the lumen involves active transporters in both directions [7, 40]. Intestinal cells receive the arterial blood and actively expel the drug into the portal vein, here modeled as a direct link from the intestinal compartment to the liver. The entero-hepatic circulation is represented by an active transport performed by liver cells towards the intestinal lumen through the biliary circulation. The re-activation of SN38G into SN38 under the β -glucuronidase activity only occurs in the intestinal lumen where this enzyme is produced by the bacteria of the intestinal flora. In the absence of concluding experimental data on this subject, we model this reaction by the law of mass action, in which the β -glucuronidase is in large excess compared to SN38G concentration. Differential equations of the model are presented in the Appendix.

Modeling CPT11 pharmacodynamics. The tumor, the intestine and the bone marrow being highly proliferating tissues, we have added the influence of the cell cycle on CPT11 PD to the model built in the *in vitro* study for quiescent cells. Indeed, the rate of conversion of reversible DNA/TOP1/irinotecan complexes into irreversible DNA damage depends on the cell cycle phase in which the cell stands. In the S-phase, many more replication forks run along the DNA which increases the probability of collisions with reversible complexes. This molecular mechanism may explain why CPT11 is preferentially toxic for cells in the S-phase [35]. Cyclin A ([CycA]) was used as a marker of the S-phase [30]. The speed of conversion of reversible into irreversible complexes which appears in Eqs. (3) and (4) is now assumed to be proportional to CycA concentration. Equations of CPT11 PD were adapted from the model at the tissue level as follows:

$$\begin{aligned} \frac{d[TOP1]}{dt} = & k_{f_{top}} - k_{d_{top}}[TOP1] - k_{f_1}[TOP1][DNA_{free}] \\ & + k_{r_1}[DNATOP1] + k_{r_2}[Compl], \end{aligned} \quad (1)$$

$$\begin{aligned} \frac{d[DNATOP1]}{dt} = & k_{f_1}[TOP1][DNA_{free}] - k_{f_2}[DNATOP1][SN] \\ & - k_{r_1}[DNATOP1], \end{aligned} \quad (2)$$

$$\begin{aligned} \frac{d[Compl]}{dt} = & k_{f_2}[DNATOP1][SN] - k_{r_2}[Compl] \\ & - k_{Irr}[CycA][Compl], \end{aligned} \quad (3)$$

$$\frac{d[Icompl]}{dt} = k_{Irr}[CycA][Compl]. \quad (4)$$

We assumed that CPT11 is active in mice only in the tumor, the intestine and the bone marrow as experimental results suggest [2]. Therefore CPT11 PD was modeled only in these three compartments. For the sake of simplicity, parameters k_{f_1} , k_{r_1} , k_{f_2} , k_{r_2} and k_{Irr} were assumed to be equal in all compartments and set to their values determined in the *in vitro* study [4].

Modeling pharmacological protein circadian rhythms. The protein amount of CES, UGT, CycA, TOP1 and efflux transporters of CPT11, SN38 and SN38G are assumed to display circadian rhythms in the liver, the bone marrow, the intestinal cells and lumen, and the tumor. For the sake of simplicity, we assumed that proteins of CPT11 pharmacology did not oscillate in the blood and the NET compartment. The period of oscillations was set to 24 h as mice were synchronized in LD12:12. For all proteins but TOP1, the following equation was used to model their circadian rhythms, as in the *in vitro* study [4]:

$$\frac{d[Protein]}{dt} = M + A \cos\left(\frac{2\pi}{T}(t - \phi)\right) - k_d[Protein]. \quad (5)$$

For TOP1, reaction terms presented in Eq. (1) were added to Eq. (5).

Initial conditions. At the initial instant, CPT11 was administered to the mice by intravenous injection. All variables representing the drug or its metabolites concentrations were set to zero except CPT11 blood concentration CPT_{blood} . Variables of the drug PD were set to their steady state values [4]. Concerning pharmacological protein concentrations, their numerical values were firstly computed over 24 h in the absence of CPT11, with an initial condition set to 1. Then, the computed values were used as initial condition for considered ZTs.

4.2 Model Calibration for B6D2F1 Male Mice

A set of parameters was estimated for B6D2F1 male mice (class 2 of the three chronotoxicity classes). Mainly two kinds of experimental data were used for parameter estimation. On one hand, protein amounts were assessed over the 24 h span for *ces1*, *ugt1a1* and *top1* in the liver; *ces1*, *ugt1a1*, *top1* and *cyclin A2* in the colon; and *abcc2* in the ileum. On the other hand, blood and tissue concentrations of CPT11 and SN38 were measured after a CPT11 injection at the worst and best circadian time regarding tolerability. Parameter estimation consisted in a least-squares approach in which the minimization task was performed by the CMAES algorithm [21]. Initial search values were set to parameter values determined in the *in vitro* study on Caco-2 cells.

Estimation of parameters of pharmacological protein circadian expression. Circadian variations of proteins of interest were measured in mice synchronized in LD12:12. Animals were sacrificed at indicated circadian times and their organs were collected. Protein quantities were then assessed by immunohistochemistry. Protein expression was determined for *ces1*, *ugt1a1* and *top1* in the liver; *ces1*, *ugt1a1*, *top1* and *cyclin A2* in the colon; and *abcc2* in the ileum [1, 2, 32]. Time series were normalized by dividing all values by the mean (Fig. 2).

Eq. (5) was used to simulate protein amounts. M and kd of Eq. (5) were set to 1, so that the mean value of simulated quantities is equal to 1. Amplitude A is searched between 0 and 1 and phase ϕ between 0 and 24 h. For each of the above-mentioned proteins in each compartment, Eq. (5) was calibrated to normalized experimental data. Estimated values of A and ϕ are shown in Table 1. In the whole-body mathematical model, parameter values of Table 1 were used to set amplitudes and phases of corresponding proteins in mentioned compartments. Circadian variations of *TOP1* and *CycA* in the bone marrow were assumed to be the same as those in the colon as no experimental data was available. Furthermore, they were set to constant values in the tumor compartment.

In the absence of relevant biological data on the circadian rhythms of ABC transporter activity at the tissue level, we assumed simplification hypotheses which may be improved in a future version of the model. Thus, efflux transporters from the intestine cells towards the intestinal lumen and the liver were assumed to display the same circadian variations. Moreover, we set circadian rhythms for the transport from the intestine cells to the lumen, but not for the opposite direction which was considered as constant. Finally, transporters in liver cells were assumed to display

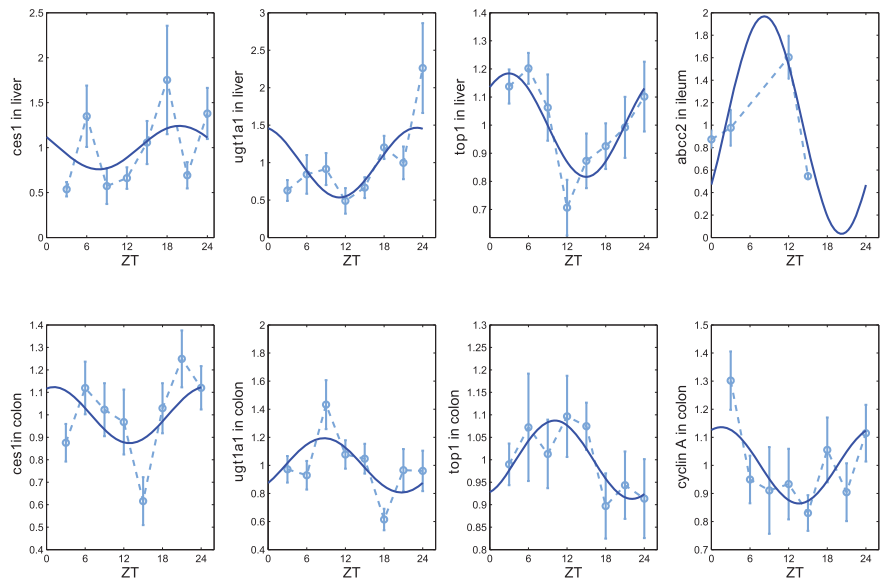


Fig. 2. Circadian variations of proteins involved in CPT11 pharmacology in B6D2F1 male mice synchronized in LD12:12 Data points represent the average of 10 to 15 animals ($\pm$ SEM). The solid curves show Eq. (5) which was calibrated to experimental data by a least square approach. Estimated parameter values are shown in Table 1

Table 1. Estimated parameter values for circadian rhythms of proteins of CPT11 pharmacology Values were computed by comparing Eq. (5) to experimental data using a least square approach. a.u. = arbitrary units

<i>Proteins</i>	<i>Amplitude A (a.u.h⁻¹)</i>	<i>Phase ϕ (h)</i>
LIVER		
ces1	0.25	18.85
top1	0.19	2.0
ugt1a1	0.48	22.09
COLON		
ces1	0.13	23.93
top1	0.09	9.1
ugt1a1	0.2	7.7
cyclin A2	0.14	0.63
ILEUM		
abcc2	1	7.25

Table 2. Parameter values for circadian rhythms of proteins of CPT11 pharmacology. Parameters were estimated using experimental results on blood and tissular pharmacokinetics of CPT11 and SN38

<i>Genes</i>	<i>A (a.u.)</i>	<i>ϕ (h)</i>
<i>ABC_CPT_{liver_blood}</i>	0.9	6.5
<i>ABC_SN_{liver_blood}</i>	0.75	16
<i>ABC_CPT_{int_lum}</i>	0.2	7.8
<i>ABC_SN_{int_lum}</i>	0.6	22.5
<i>ABC_CPT_{marrowle_blood}</i>	0.4	9.78
<i>ABC_SN_{marrowle_blood}</i>	0	-
<i>UGT_{marrow}</i>	0.5	7
<i>ABC_CPT_{tumor_blood}</i>	0.9	6.2
<i>ABC_SN_{tumor_blood}</i>	0	—
<i>CES_{tumor}</i>	0	—
<i>UGT_{tumor}</i>	0	—

the same rhythms whether they efflux the drug into the intestine or into the blood. Amplitudes and phases of all efflux transporters were inferred from ChronoPK data (Table 2). During parameter estimation, initial values of protein quantities were set as follows. Parameter search was run a first time in order to get optimal values of *A* and *ϕ* for each gene. then the value at *t* = 24 h of the best-fit curve was used as an initial condition during the second run of parameter estimation which provided us with the final optimal parameter values. Of note, the influence of initial condition on parameters *A* and *ϕ* slightly change them by approximately 1% of their value.

4.2.1 Estimation of PK-PD Parameters

Kinetic parameters of the whole-body model were evaluated using data on CPT11 and SN38 concentrations in the blood and the tissues of interest, after a CPT11 injection at best and worst circadian time regarding toxicity. Experimental study of CPT11 PK in chronotoxicity class 2 (B6D2F1 male mice) involved animals synchronized in LD12:12 which underwent an intravenous injection of CPT11 at the dose of 50 mg/kg, at ZT11 and ZT23, which are respectively the best and worst times for tolerability. Mice were sacrificed 30 min, 1, 2, 6 and 18 h after the drug injection and tissues of interest were collected. CPT11 and SN38 concentrations were determined in the blood, the liver, the intestine cells, the bone marrow and the tumor by High Performance Liquid Chromatography (HPLC, [1]).

In the mathematical model, initial drug concentrations were all equal to zero except that of CPT11 in the blood compartment which was set to 860 μM. This cor-

responds to a dose of 50 mg/kg for a B6D2F1 male mouse whose average weight was estimated at $w = 26$ g and blood volume at $V_{blood} = 1.7\text{mL}$ [1, 12]. Values of other organ volumes were inferred from literature: $V_{liver} = 1.3\text{mL}$, $V_{intestine} = 1.5\text{mL}$, $V_{marrow} = 1.8\text{mL}$, $V_{NET} = 16\text{mL}$, $V_{tumor} = 1\text{mL}$ [12].

For the sake of simplicity, uptake and efflux parameters for SN38G were assumed to be equal to those for SN38, except for those of the entero-hepatic circulation. Moreover we have set all Km parameters of Michaelis-Menten kinetics to the values determined in the *in vitro* study. Thus, Km for CPT11 bioactivation was set to $K_{CES} = 146\mu\text{M}$, that for SN38 deactivation to $K_{UGT} = 2.85\mu\text{M}$, that for CPT11 transport to $K_{CPT} = 47.8\mu\text{M}$ and that for SN38 and SN38G transport to $K_{SN} = 1.6\mu\text{M}$. Parameters of CPT11 PD were set to their values determined in the *in vitro* study [4].

Remaining parameters were estimated by confronting the model to biological data on CPT11 and SN38 concentrations in the blood, the liver, the intestinal cells, the bone marrow and the tumor, after an injection of CPT11 at ZT11 or at ZT23. For each compartment, experimentally-observed CPT11 concentration was compared with the corresponding mathematical variable, namely CPT_{organ} . Measured concentrations of SN38 were compared to the sum of $SN_{organ} + COMPL + Icompl$ as the experimental method assessed the totality of intracellular SN38, including molecules linked to DNA. Estimated parameter values are showed in the Appendix. The data-calibrated whole-body model mimics the main features observed in the biological data of chronotoxicity class 2 (Figs. 3 and 4).

Interpretation of estimated parameter values. Concerning CPT11 uptake, the drug seemed to enter preferentially in the liver and in the NET compartments (see parameter values in the appendix). A high SN38 uptake rate was estimated in the target organs (intestine cells, bone marrow) and in the tumor which may explain the observed toxicities and efficacy in those compartments [1].

CPT11 bioactivation into SN38 was mainly present in the liver compartment and weakly active in the blood and in the bone marrow. CES activity was estimated to values close to zero in the intestinal cells where SN38 experimentally appeared 2 h after CPT11 injection, despite the presence of CPT11 in the same organ only 30 min after its injection. CPT11 metabolism in the tumor was also estimated to quasi-zero.

SN38 glucuronidation appeared to be high in the bone marrow, and less active in the liver, the intestine and the tumor.

5 Discussion and Perspectives

Current clinical knowledge supports a personalization of chronomodulated delivery regimens according to the patient circadian and genetic profile. We propose here a combined biological and mathematical approach for the optimization of CPT11 circadian delivery. CPT11 chronomodulated exposure was optimized in cultured Caco-2 cells. This *in vitro* study provided the basis for the design of a whole-body

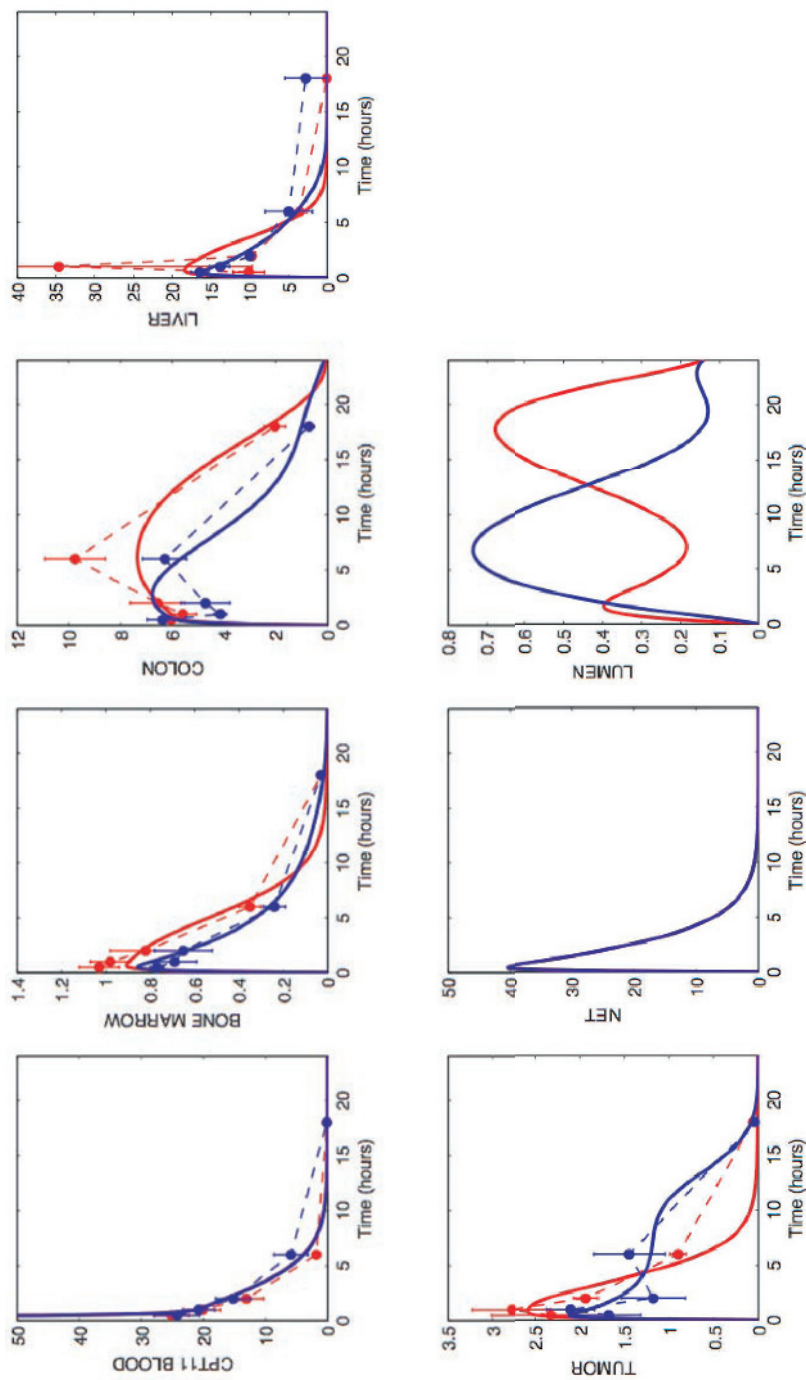


Fig. 3. Blood and tissue CPT11 PK in class 2, after injection at best (ZT11, blue) and worst (ZT23, red) time regarding toxicity. Data points stand for the mean value ($\pm SEM$) of 10 to 15 mice. Solid curves represent the best-fit simulations of the whole-body model

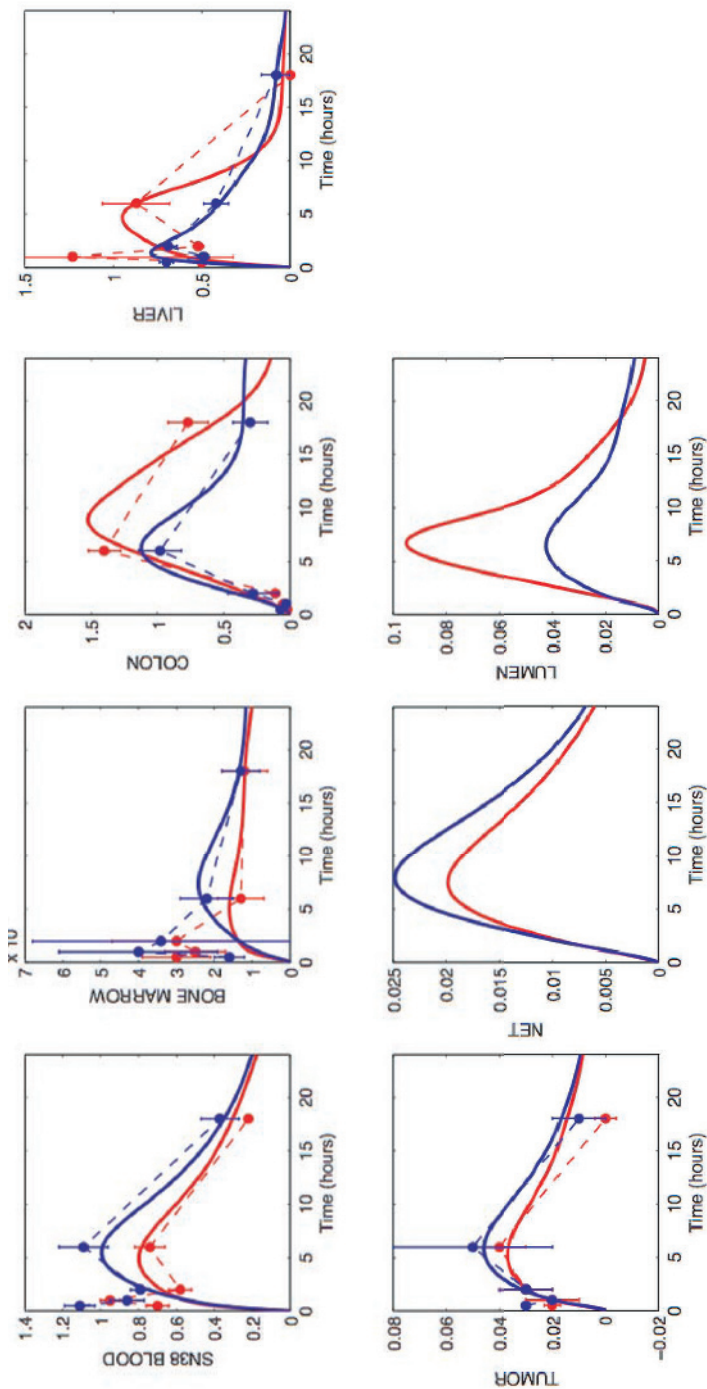


Fig. 4. Blood and tissue SN38 PK in class 2, after injection at best (ZT11, blue) and worst (ZT23, red) time regarding toxicity. Data points stand for the mean value ($\pm$ SEM) of 10 to 15 mice. Solid curves represent the best-fit simulations of the whole-body model

physiologically-based model of CPT11 PK-PD which aims at optimizing CPT11 administration in three chronotoxicity classes in mice.

The whole-body model was calibrated to experimental data obtained for class 2 (B6D2F1 male mice). It provided simulations which were consistent with biological results. It also suggested molecular explanations for CPT11 and SN38 chronoPK which could lead to predictive biomarkers of CPT11 chronotoxicity in this chronotoxicity class.

The next step will consist in applying optimization procedures to the calibrated whole-body model in order to design theoretically optimal administration schemes for this class. The same study is in progress for classes 1 (B6D2F1 female mice) and 3 (B6CBAF1 female mice). Once the three parameter sets have been estimated, their comparison may allow the molecular characterization of the three classes by highlighting differences in protein activities.

The data-calibrated whole body model simulated the main features observed in the biological data of class 2 (Figs. 3 and 4). However, the rebound in CPT11 intestine concentration which occurred 2 h after the injection at the best time of tolerability (ZT11) was not captured by the model. SN38 blood concentrations were underestimated in the simulations during the first hours after CPT11 injection, which probably induced the under-estimation in SN38 bone marrow concentration during the same period of time.

In the tumor compartment, parameter estimation led us to conclude to an absence of circadian rhythms of all protein activities, except in CPT11 efflux. This suggests an absence of circadian organization in the cancer tissues; the rhythm found in CPT11 efflux may be interpreted as a circadian variation of the permeability of healthy cells in the tumor environment. Indeed, drug uptake from the blood to the tumor was assumed to be constant in the current mathematical model. This should be modified in a future version of the model.

Experimental data are available on intestinal and hematological chronotoxicities, together with chronoefficacy results on tumor growth for the three mouse classes. They are currently being integrated into the parameter estimation and will especially allow a more precise evaluation of circadian rhythms of TOP1 and CycA in the intestine, the bone marrow and the tumor. Parameters of CPT11 PD will also be adjusted.

In vitro experiments on human blood sample have demonstrated that CPT11 and SN38 associate with blood components such as plasmatic proteins, platelets or red blood cells [10, 11]. 60 to 66% of CPT11 molecules were linked to blood components when investigating CPT11 concentrations from 100 to 4000 ng/mL, whereas 94 to 96% of SN38 was linked with blood components for investigated SN38 concentrations from 50 to 200 ng/mL. These associations may interfere with the drug PK-PD and will therefore be integrated into a future version of the model. Moreover, CPT11 metabolism through P450 cytochromes should also be added to the model as it may also display circadian rhythms [26].

In a clinical perspective, this whole-body model of CPT11 PK-PD designed for mice will be adapted to humans by keeping its main structure and resizing parameter values. This will provide clinicians with a new tool towards the personalization of CPT11 administration according to the patient genetic and circadian profile.

Appendix

Equations of the Whole-Body Model of CPT11 PK-PD

Equations of the whole body physiologically-based model of CPT11 PK-PD. *int* stands for intestine, *lum* for intestinal lumen, *marrow* for bone marrow.

Liver:

$$\begin{aligned} \frac{d[CPT_{liver}]}{dt} = & C_{blood\_liver} - C_{liver\_blood} + C_{int\_liver} - \frac{V_{lum}}{V_{liver}} C_{liver\_bile} \\ & - C_{CES\_liver}; \end{aligned} \quad (6)$$

$$\begin{aligned} \frac{d[SN_{liver}]}{dt} = & S_{blood\_liver} - S_{liver\_blood} + S_{int\_liver} - \frac{V_{lum}}{V_{liver}} S_{liver\_bile} \\ & + C_{CES\_liver} - S_{UGT\_liver} + S_{RC\_liver} - S_{C\_liver}; \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{d[SNG_{liver}]}{dt} = & G_{blood\_liver} - G_{liver\_blood} + G_{int\_liver} - \frac{V_{lum}}{V_{liver}} G_{liver\_bile} \\ & + S_{UGT\_liver}. \end{aligned} \quad (8)$$

Intestine:

$$\frac{d[CPT_{int}]}{dt} = C_{blood\_int} - \frac{V_{liver}}{V_{int}} C_{int\_liver} + C_{lum\_int} - C_{int\_lum} - C_{CES\_int}; \quad (9)$$

$$\begin{aligned} \frac{d[SN_{int}]}{dt} = & S_{blood\_int} - \frac{V_{liver}}{V_{int}} S_{int\_liver} + S_{lum\_int} - S_{int\_lum} + C_{CES\_int} \\ & - S_{UGT\_int} + S_{RC\_int} - S_{C\_int}; \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{d[SNG_{int}]}{dt} = & G_{blood\_int} - \frac{V_{liver}}{V_{int}} G_{int\_liver} + G_{lum\_int} - G_{int\_lum} \\ & + S_{UGT\_liver}. \end{aligned} \quad (11)$$

Intestinal lumen:

$$\frac{d[CPT_{lum}]}{dt} = \frac{V_{int}}{V_{lum}} (C_{int\_lum} - C_{lum\_int}) + C_{liver\_bile} - C_{clear\_int}; \quad (12)$$

$$\frac{d[SN_{lum}]}{dt} = \frac{V_{int}}{V_{lum}} (S_{int\_lum} - S_{lum\_int}) + S_{liver\_bile} - S_{clear\_int} + G_{beta}; \quad (13)$$

$$\frac{d[SNG_{lum}]}{dt} = \frac{V_{int}}{V_{lum}} (G_{int\_lum} - G_{lum\_int}) + G_{liver\_bile} - G_{beta}. \quad (14)$$

Bone marrow:

$$\frac{d[CPT_{marrow}]}{dt} = C_{blood\_marrow} - C_{marrow\_blood} - C_{CES\_marrow}; \quad (15)$$

$$\begin{aligned} \frac{d[SN_{marrow}]}{dt} = & S_{blood\_marrow} - S_{marrow\_blood} + C_{CES\_marrow} \\ & - S_{UGT\_marrow} + S_{RC\_marrow} - S_{C\_marrow}; \end{aligned} \quad (16)$$

$$\frac{d[SNG_{marrow}]}{dt} = G_{blood\_marrow} - G_{marrow\_blood} + S_{UGT\_liver}. \quad (17)$$

Net:

$$\frac{d[CPT_{NET}]}{dt} = C_{blood\_NET} - C_{NET\_blood}; \quad (18)$$

$$\frac{d[SN_{NET}]}{dt} = S_{blood\_NET} - S_{NET\_blood} + S_{RC\_NET} - S_{C\_NET}; \quad (19)$$

$$\frac{d[SNG_{NET}]}{dt} = G_{blood\_NET} - G_{NET\_blood}. \quad (20)$$

Tumor:

$$\frac{d[CPT_{tumor}]}{dt} = C_{blood\_tumor} - C_{tumor\_blood} - C_{CES\_tumor}; \quad (21)$$

$$\begin{aligned} \frac{d[SN_{tumor}]}{dt} = & S_{blood\_tumor} - S_{tumor\_blood} + C_{CES\_tumor} - S_{UGT\_tumor} \\ & + S_{RC\_tumor} - S_{C\_tumor}; \end{aligned} \quad (22)$$

$$\frac{d[SNG_{tumor}]}{dt} = G_{blood\_tumor} - G_{tumor\_blood} + S_{UGT\_tumor}. \quad (23)$$

Blood:

$$\begin{aligned} \frac{d[CPT_{blood}]}{dt} = & \frac{V_{liver}}{V_{blood}}(C_{liver\_blood} - C_{blood\_liver}) - \frac{V_{int}}{V_{blood}}C_{blood\_int} \\ & + \frac{V_{marrow}}{V_{blood}}(C_{marrow\_blood} - C_{blood\_marrow}) \\ & + \frac{V_{NET}}{V_{blood}}(C_{NET\_blood} - C_{blood\_NET}) \\ & + \frac{V_{tumor}}{V_{blood}}(C_{tumor\_blood} - C_{blood\_tumor}) \\ & - C_{CES\_blood} - C_{clear\_ren}; \end{aligned} \quad (24)$$

$$\begin{aligned}
\frac{d[SN_{blood}]}{dt} = & \frac{V_{liver}}{V_{blood}}(S_{liver\_blood} - S_{blood\_liver}) - \frac{V_{int}}{V_{blood}}S_{blood\_int} \\
& + \frac{V_{marrow}}{V_{blood}}(S_{marrow\_blood} - S_{blood\_marrow}) \\
& + \frac{V_{NET}}{V_{blood}}(S_{NET\_blood} - S_{blood\_NET}) \\
& + \frac{V_{tumor}}{V_{blood}}(S_{tumor\_blood} - S_{blood\_tumor}) + C_{CES\_blood} \\
& - S_{clear\_ren}; \tag{25}
\end{aligned}$$

$$\begin{aligned}
\frac{d[SNG_{blood}]}{dt} = & \frac{V_{liver}}{V_{blood}}(G_{liver\_blood} - G_{blood\_liver}) - \frac{V_{int}}{V_{blood}}G_{blood\_int} \\
& + \frac{V_{marrow}}{V_{blood}}(G_{marrow\_blood} - G_{blood\_marrow}) \\
& + \frac{V_{NET}}{V_{blood}}(G_{NET\_blood} - G_{blood\_NET}) \\
& + \frac{V_{tumor}}{V_{blood}}(G_{tumor\_blood} - G_{blood\_tumor}) - G_{clear\_ren}. \tag{26}
\end{aligned}$$

Fluxes of Equations of the Whole Body Physiologically-Based Model of CPT11 PK-PD (Organ Designates Liver, Intestine, Bone Marrow, Tumor or NET)

Cellular uptake:

$$C_{blood\_organ} = k_{upCPT\_organ}[CPT_{blood}]; \tag{27}$$

$$S_{blood\_organ} = k_{upSN\_organ}[SN_{blood}]; \tag{28}$$

$$G_{blood\_organ} = k_{upSN\_organ}[SNG_{blood}]. \tag{29}$$

Active transport:

$$C_{organ\_blood} = \frac{V_{effCPT}[ABC\_CPT][CPT_{organ}]}{K_{effCPT} + [CPT_{organ}]}; \tag{30}$$

$$S_{organ\_blood} = \frac{V_{effSN}[ABC\_SN][SN_{organ}]}{K_{effSN} + [SN_{organ}]}; \tag{31}$$

$$G_{organ\_blood} = \frac{V_{effSN}[ABC\_SN][SNG_{organ}]}{K_{effSN} + [SNG_{organ}]}; \tag{32}$$

$$C_{int\_lum} = \frac{V_{int\_lumCPT}[ABC\_CPT\_int\_lum][CPT_{int}]}{K_{int\_lumCPT} + [CPT_{int}]}; \quad (33)$$

$$S_{int\_lum} = \frac{V_{int\_lumSN}[ABC\_SN\_int\_lum][SN_{int}]}{K_{int\_lumSN} + [SN_{int}]}; \quad (34)$$

$$G_{int\_lum} = \frac{V_{int\_lumSN}[ABC\_SN\_int\_lum][SNG_{int}]}{K_{int\_lumSN} + [SNG_{int}]}; \quad (35)$$

$$C_{lum\_int} = \frac{V_{lum\_intCPT}[ABC\_CPT\_lum\_int][CPT_{lum}]}{K_{lum\_intCPT} + [CPT_{lum}]}; \quad (36)$$

$$S_{lum\_int} = \frac{V_{lum\_intSN}[ABC\_SN\_lum\_int][SN_{lum}]}{K_{lum\_intSN} + [SN_{lum}]}; \quad (37)$$

$$G_{lum\_int} = \frac{V_{lum\_intSN}[ABC\_SN\_lum\_int][SNG_{lum}]}{K_{lum\_intSN} + [SNG_{lum}]}. \quad (38)$$

PT11 bioactivation into SN38:

$$C_{CES\_organ} = \frac{V_{CES\_organ}[CES][CPT_{organ}]}{K_{CES} + [CPT_{organ}]}. \quad (39)$$

SN38 glucuronidation into SN38G

$$S_{UGT\_organ} = \frac{V_{UGT\_organ}[UGT][SN_{organ}]}{K_{UGT} + [SN_{organ}]}. \quad (40)$$

Entero-hepatic circulation

$$C_{liver\_int} = \frac{V_{bileCPT}[ABC\_CPT\_bile][CPT_{liver}]}{K_{bileCPT} + [CPT_{liver}]}; \quad (41)$$

$$S_{liver\_int} = \frac{V_{bileSN}[ABC\_SN\_bile][SN_{organ}]}{K_{bileSN} + [SN_{liver}]}; \quad (42)$$

$$G_{liver\_int} = \frac{V_{bileSN}[ABC\_SN\_bile][SNG_{organ}]}{K_{bileSN} + [SNG_{liver}]}. \quad (43)$$

β -glucuronidase:

$$G_{beta} = V_{beta}[SNG_{lum}]K_{beta} + [SNG_{lum}]. \quad (44)$$

Renal and intestinal clearance:

$$C_{clear\_ren} = k_{renCPT}[CPT_{blood}]; \quad (45)$$

$$S_{clear\_ren} = k_{renSN}[SN_{blood}]; \quad (46)$$

$$G_{clear\_ren} = k_{renSNG}[SNG_{blood}]; \quad (47)$$

$$C_{clear\_int} = k_{intCPT}[CPT_{lum}]; \quad (48)$$

$$S_{clear\_int} = k_{intSN}[SN_{lum}]. \quad (49)$$

PD:

$$S_{C\_organ} = k_{r2}[DNATOP1_{organ}][SN_{organ}]; \quad (50)$$

$$S_{RC\_organ} = k_{r2}[Compl_{organ}]. \quad (51)$$

Parameter Values

Parameter values of the whole-body model of CPT11 PK-PD, for class2 (B6D2F1 male mice). Parameters were estimated by a least square approach using the CMAES algorithm for minimization tasks.

Cellular uptake (h^{-1})

$$k_{upCPT\_liver} = 0.16, k_{upSN\_liver} = 0.05;$$

$$k_{upCPT\_int} = 0.048, k_{upSN\_int} = 0.48;$$

$$k_{upCPT\_marrow} = 0.008, k_{upSN\_marrow} = 0.9;$$

$$k_{upCPT\_NET} = 0.4, k_{upSN\_NET} = 0.23;$$

$$k_{upCPT\_tumor} = 0.021, k_{upSN\_tumor} = 0.45.$$

Active transport (μMh^{-1})

$$V_{liver\_bloodCPT} = 14, V_{liver\_bloodSN} = 28;$$

$$V_{marrow\_bloodCPT} = 17, V_{marrow\_bloodSN} = 6000;$$

$$V_{NET\_bloodCPT} = 60, V_{NET\_bloodSN} = 8;$$

$$V_{tumor\_bloodCPT} = 17, V_{tumor\_bloodSN} = 30;$$

$$V_{int\_lumCPT} = 60, V_{int\_lumSN} = 0.002, V_{int\_lumSNG} = 15;$$

$$V_{lum\_intCPT} = 0.001, V_{lum\_intSN} = 15, V_{lum\_intSNG} = 0.001.$$

CPT11 bioactivation into SN38 (μMh^{-1})

$$V_{CES\_blood} = 10.5, \quad V_{CES\_liver} = 35, \quad V_{CES\_int} = 0.02, \\ V_{CES\_marrow} = 10, \quad V_{CES\_tumor} = 0.001.$$

SN38 glucuronidation into SN38G (μMh^{-1})

$$V_{UGT\_liver} = 80, \quad V_{UGT\_int} = 20, \quad V_{UGT\_marrow} = 5000, \\ V_{UGT\_tumor} = 300.$$

Entero-hepatic circulation (μMh^{-1})

$$V_{bileCPT} = 0.01, \quad V_{bileSN} = 0.4\mu M.$$

 β -glucuronidase

$$V_{beta} = 20\mu Mh^{-1}, \quad K_{beta} = 10\mu M.$$

Renal and intestinal clearance (h^{-1})

$$k_{renCPT} = 4.2, \quad k_{renSN} = 0.3, \quad k_{renSNG} = 0.5, \\ k_{intCPT} = 1, \quad k_{intSN} = 8.$$

References

1. Ahowesso, C.: Approche expérimentale de la personnalisation de la chronothérapie par irinotecan. Thèse Université Paris-Sud 11 (2011)
2. Ahowesso, C., Piccolo, E., Li, X.M., Dulong, S., Hossard, V., La Sorda, R., Filipski, E., Tinari, N., Delaunay, F., Iacobelli, S., Levi, F.: Relations between strain and gender dependencies of irinotecan toxicity and UGT1A1, CES2 and TOP1 expressions in mice. *Toxicol. Lett.* **192**, 395–401 (2010)
3. Ando, H., Yanagihara, H., Sugimoto, K., Hayashi, Y., Tsuruoka, S., Takamura, T., Kaneko, S., Fujimura, A.: Daily rhythms of P-glycoprotein expression in mice. *Chronobiol. Int.* **22**, 655–665 (2005)
4. Ballesta, A., Dulong, S., Abbara, C., Cohen, B., Okyar, A., Clairambault, J., Levi, F.: A combined experimental and mathematical approach for molecular-based optimization of irinotecan circadian delivery. *PLoS Comput. Biol.* **7**, e1002143 (2011)
5. Balsalobre, A., Damiola, F., Schibler, U.: A serum shock induces circadian gene expression in mammalian tissue culture cells. *Cell* **93**, 929–937 (1998)
6. Bernard, S., Cajavec-Bernard, B., Levi, F., Herzog, H.: Tumour growth rate determines the timing of optimal chronomodulated treatment schedules. *PLoS Comput. Biol.* **3**, e1000712 (2010)
7. Brand, W., Schutte, M.E., Williamson, G., van Zanden, J.J., Cnubben, N.H., Groten, J.P., van Bladeren, P.J., Rietjens, I.M.: Flavonoid-mediated inhibition of intestinal ABC transporters may affect the oral bioavailability of drugs, food-borne toxic compounds and bioactive ingredients. *Biomed. Pharmacother.* **60**, 508–519 (2006)
8. Brown, S.A., Zimbrun, G., Fleury-Olela, F., Preitner, N., Schibler, U.: Rhythms of mammalian body temperature can sustain peripheral circadian clocks. *Curr. Biol.* **12**, 1574–1583 (2002)
9. Carr, A.J., Whitmore, D.: Imaging of single light-responsive clock cells reveals fluctuating free-running periods. *Nat. Cell. Biol.* **7**, 319–321 (2005)

10. Chabot, G.G., Robert, J., Lokiec, F., Canal, P.: Irinotecan pharmacokinetics. *Bull. Cancer. Spec No*:11–20 (1998)
11. Combes, O., Barré, J., Duché, J.C., Vernillet, L., Archimbaud, Y., Marietta, M.P., Tillement, J.P., Urien, S.: In vitro binding and partitioning of irinotecan (cpt-11) and its metabolite, sn-38, in human blood. *Invest. New Drugs*. **18**, 1–5 (2000)
12. Davies, B., Morris, T.: Physiological parameters in laboratory animals and humans. *Pharm. Res.* **10**, 1093–1095 (1993)
13. De Buck, S.S., Mackie, C.E.: Physiologically based approaches towards the prediction of pharmacokinetics: in vitro-in vivo extrapolation. *Expert Opin. Drug. Metab. Toxicol.* **3**, 865–878 (2007)
14. Falcone, A., Ricci, S., Brunetti, I., Pfanner, E., Allegrini, G., et al.: Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. *J. Clin. Oncol.* **1**;25, 1670–1676 (2007)
15. Filipski, E., Lemaigre, G., Liu, X.H., Mery-Mignard, D., Mahjoubi, M., Levi, F.: Circadian rhythm of irinotecan tolerability in mice. *Chronobiol. Int.* **21**, 613–630 (2004)
16. Garufi, C., Vanni, B., Aschelter, A.M., Zappala, A.R., Bria, E., Nisticó, C., Sperduti, I., Cognetti, F., Terzoli, E.: Randomised phase II study of standard versus chronomodulated CPT11 plus chronomodulated 5-fluorouracil and folinic acid in advanced colorectal cancer patients. *Eur. J. Cancer.* **42**, 608–616 (2006)
17. Gholam, D., Giacchetti, S., Brezault-Bonnet, C., Bouchahda, M., Hauteville, D., Adam, R., Ducot, B., Ghémard, O., Kustlinger, F., Jasmin, C., Lévi, F.: Chronomodulated irinotecan, oxaliplatin, and leucovorin-modulated 5-Fluorouracil as ambulatory salvage therapy in patients with irinotecan- and oxaliplatin-resistant metastatic colorectal cancer. *Oncologist.* **11**, 1072–1080 (2006)
18. Giacchetti, S., Itzhaki, M., Gruia, G., Adam, R., Zidani, R., Kunstlinger, F., Brienza, S., Alafaci, E., Bertheault-Cvitkovic, F., Jasmin, C., Reynes, M., Bismuth, H., Misset, J.L., Levi, F.: Longterm survival of patients with unresectable colorectal cancer liver metastases following infusional chemotherapy with 5-fluorouracil, leucovorin, oxaliplatin and surgery. *Ann. Oncol.* **6**, 663–669 (1999)
19. Granda, T.G., D'Attino, R.M., Filipski, E., Vrignaud, P., Garufi, C., Terzoli, E., Bissery, M.C., Levi, F.: Circadian optimisation of irinotecan and oxaliplatin efficacy in mice with Glasgow osteosarcoma. *Br. J. Cancer* **86**, 999–1005 (2002)
20. Guo, Y., Lu, P., Farrell, E., Zhang, X., Weller, P.: In silico and in vitro pharmacogenetic analysis in mice. *Proc. Natl. Acad. Sci. USA* **104**, 17735–17740 (2007)
21. Hansen, N.: The CMA evolution strategy: a comparing review. In: *Towards a new evolutionary computation. Advances on estimation of distribution algorithms*, pp. 75–102. Springer, Heidelberg (2006)
22. Izumo, M., Sato, T.R., Straume, M., Johnson, C.H.: Quantitative analyses of circadian gene expression in mammalian cell cultures. *PLoS Comput. Biol.* **2**, e136 (2006)
23. Kanno, Y., Otsuka, S., Hiromasa, T., Nakahama, T., Inouye, Y.: Diurnal difference in CAR mRNA expression. *Nucl. Recept.* **2**, 6 (2004)
24. Kornmann, B., Preitner, N., Rifat, D., Fleury-Olela, F., Schibler, U.: Analysis of circadian liver gene expression by ADDER, a highly sensitive method for the display of differentially expressed mRNAs. *Nucleic Acids Res.* **29**, E51–1 (2001)
25. Kuramoto, Y., Hata, K., Koyanagi, S., Ohdo, S., Shimeno, H., Soeda, S.: Circadian regulation of mouse topoisomerase I gene expression by glucocorticoid hormones. *Biochem. Pharmacol.* **71**, 1155–1161 (2006)
26. Lavery, D.J., Lopez-Molina, L., Margueron, R., Fleury-Olela, F., Conquet, F.: Circadian expression of the steroid 15 alpha-hydroxylase (Cyp2a4) and coumarin 7-hydroxylase (Cyp2a5) genes in mouse liver is regulated by the PAR leucine zipper transcription factor DBP. *Mol. Cell. Biol.* **19**, 6488–6499 (1999)
27. Levi, F., Giacchetti, S., Zidani, R., Brezault-Bonnet, C., Tigaud, J., Goldwasser, F., Misset, J.L.: Chronotherapy of colorectal cancer metastases. *Hepatogastroenterology.* **38**, 320–322 (2001)

28. Levi, F., Okyar, A., Dulong, S., Innominato, P.F., Clairambault, J.: Circadian timing in cancer treatments. *Annu. Rev. Pharmacol. Toxicol.* **50**, 377–421 (2010)
29. Levi, F., Schibler, U.: Circadian rhythms: mechanism and therapeutics implications. *Annu. Rev. Pharmacol. Toxicol.* **47**, 593–628 (2007)
30. Morgan, D.: *The Cell Cycle, Principles of Control*. In: *Primers in Biology*. New Science Press Ltd and Oxford University Press, Oxford (2007)
31. Ohdo, S., Makinosumi, T., Ishizaki, T., Yukawa, E., Higuchi, S., Nakano, S., Ogawa, N.: Cell cycle-dependent chronotoxicity of irinotecan hydrochloride in mice. *J. Pharmacol. Exp. Ther.* **283**, 1383–1388 (1997)
32. Okyar, A., Piccolo, E., Ahowesso, C., Filipski, E., Hossard, V., Guettier, C., La Sorda, R., Tinari, N., Iacobelli, S., Levi, F.: Strain- and sex-dependent circadian changes in *abcc2* transporter expression: implications for irinotecan chronotolerance in mouse ileum. *PLoS One* **6**, e20393 (2011)
33. Ortiz-Tudela, E., Mteyrek, A., Ballesta, A., Innominato, P.F., Levi, F.: Cancer chronotherapeutics. Experimental, theoretical and clinical aspects. In: Kramer, A., Mewow, M. (eds) *Circadian clocks: overview and basic principles*, Springer, Heidelberg, in press
34. Panda, S., Antoch, M., Miller, B., Su, A., Schook, A.: Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* **109**, 307–320 (2002)
35. Pommier, Y.: Topoisomerase I inhibitors: camptothecins and beyond. *Nat. Rev. Cancer* **6**, 789–802 (2006)
36. Preitner, N., Damiola, F., Luis Lopez, M., Zakany, J., Duboule, D., Albrecht, U., Schibler, U.: The orphan nuclear receptor REV-ERB α controls circadian transcription within the positive limb of the mammalian circadian oscillator. *Cell* **110**, 251–260 (2002)
37. Slatter, J.G., Schaaf, L.J., Sams, J.P., Feenstra, K.L., Johnson, M.G., Bombardt, P.A.: Pharmacokinetics, metabolism, and excretion of irinotecan (CPT-11) following I.V. infusion of [(14)C]CPT-11 in cancer patients. *Drug. Metab. Dispos.* **28**, 423–433 (2000)
38. Smith, M.R., Burgess, H.J., Fogg, L.F., Eastman, C.I.: Racial differences in the human endogenous circadian period. *PLoS One* **4**, e6014 (2009)
39. Sothorn, R., Halberg, F., Hrushesky, W.: Circadian stage not time of day characterizes doxorubicin susceptibility rhythm of mice in continuous light. *Annu. Rev. Chronopharmacol.* **5**, 385–388 (1988)
40. Stanton, B.A.: Summary of the speaker presentations: ABC transporters in liver, kidney, and intestine. *Kidney International*, **62**, 1520–1521 (2002)
41. Storch, K.F., Lipan, O., Leykin, I., Viswanathan, N., Davis, F.C.: Extensive and divergent circadian gene expression in liver and heart. *Nature* **417**, 78–83 (2002)
42. Tsukamoto, Y., Kato, Y., Ura, M., Horii, I., Ishitsuka, H., Kusuhara, H., Sugiyama, Y.: A physiologically based pharmacokinetic analysis of capecitabine, a triple prodrug of 5-FU, in humans: the mechanism for tumor-selective accumulation of 5-FU. *Pharm. Res.* **18**, 1190–1202 (2001)
43. Welsh, D.K., Yoo, S.H., Liu, A.C., Takahashi, J.S., Kay, S.A.: Bioluminescence imaging of individual fibroblasts reveals persistent, independently phased circadian rhythms of clock gene expression. *Curr. Biol.* **14**, 2289–2295 (2004)
44. Zhang, Y., Yeager, R., Klaassen, C.: Circadian expression profiles of drug processing genes and transcription factors in mouse liver. *Drug Metab. Dispos.* **37**, 106–115 (2009)

Modeling the Dynamics of HCV Infected Cells to Tailor Antiviral Therapy in Clinical Practice: Can This Approach Fit for Neoplastic Cells?

Piero Colombatto, Filippo Oliveri, Ferruccio Bonino, and Maurizia R. Brunetto

Abstract The always present need to predict therapy outcomes and to understand complex phenomena in medicine can be addressed by the use of mathematical models that simulate which relevant events may be expected at the single patient level. In chronic hepatitis C patients undergoing effective antiviral treatment, we found that a model specifically developed to compute the infected cells surviving at the end of the treatment, was able to predict the final outcome of the therapy better than the single measures of the viral load decline after 4 and 12 weeks. This approach, that showed in a prospective trial to improve the cost efficacy of standard peg-interferon/ribavirin treatment, could become very useful after the first month of treatment to select patients in whom adding a more potent antiviral drug is actually required, as standard therapy outcome is unlikely favorable or its predicted duration is too long. Similarly, in patients with hepatocellular carcinoma where the efficacy of the novel biological treatments combined to standard ablation approaches is difficult to assess, modeling the dynamics of the neoplastic cells using experimental data from imaging techniques integrated with measures of liver cell function and proliferation, should become available to help clinical management of these patients.

1 Introduction

The simulation of complex phenomena in medicine by mathematical equations defined on the basis of biological assumptions has become a frequent approach

P. Colombatto · F. Oliveri · M.R. Brunetto (✉)

Hepatology Unit, University Hospital, Via Paradisa 2, 56124 Cisanello, Pisa, Italy
e-mail: colombatto@unipi.it, f.oliveri@do-pisa.toscana.it, brunetto@med-club.com

F. Bonino

Chair of Gastroenterology, Liver and Digestive – General Medicine 2 Unit, University Hospital, Via Roma 67, 56126 Pisa, Italy
e-mail: bonino@med-club.com

to investigate difficult to understand processes, including infections and cancer. In chronic viral hepatitis the persistence of the infection with long lasting liver damage may lead to the development of the hepatocellular carcinoma (HCC), therefore, early elimination of the virus represent the best treatment approach. Insights in the dynamics of chronic viral infection were obtained in the 90s using models that consider two distinct but related compartments: the blood where free virions circulate and the cells where the viruses replicate. This approach was originally applied by Perelson et al [6] in patients with Human Immunodeficiency Virus (HIV) infection treated with protease inhibitors and by Nowak et al. [5] in Hepatitis B Virus (HBV) infected patients during polymerase inhibitors therapy in their first models of viral dynamics. By the same model Neumann et al. [4] showed that the treatment with interferon (IFN) in patients with chronic Hepatitis C Virus (HCV) infection was associated with a more than 10-fold decrease of serum HCV RNA within the first 48 hours of treatment, explained by a reduction of the number of virions produced and released daily from infected cells. The slower further decline was attributed to the immune mediated clearance of the infected hepatocytes, occurring after the drastic reduction of the viral load that lowers “de novo” infection of target cells. This model contributed to the knowledge of HCV dynamics with the first reliable estimates of the very short half-life of the HCV viral particles in serum (about 2.7 hours), which implies an extremely high replication rate. It also enlightened the mechanisms of action of IFN revealing the initial dose dependent block on viral production followed by the slower decrease of the infected cell number killed by the immune system and supports the hypothesis that the final outcome of the therapy is mainly related to the infected hepatocytes clearance rate [4].

2 New Insights on HCV Infected Cell Dynamics

Using the standard biphasic model in IFN treated patients Neumann et al estimated the HCV infected cell half-life ranging from 1.7 to 70 days [4]. Part of such considerable variability may be related to the fact that the immune system activity that causes the clearance of the infected hepatocytes is not measured directly, but estimated by the 2nd phase decline of the viral load. According to the model, serum HCV RNA in this phase was assumed to decline exclusively because of the decline of the infected cells. Serum alanine aminotransferases (ALT), which are regarded as surrogate parameters of hepatocyte damage and turnover [3] can be used to better investigate the infected cell loss in this phase. When IFN treatment is effective, ‘de novo’ infection of target cells becomes irrelevant, therefore, the logarithmic decrease of viral load and the logarithmic decrease of infected cells load, after 48 hours of therapy, should parallel each other; that is, the decay time constant of these two quantities have to be equal. On the contrary, experimental data showed that the logarithmic decrease of the viral load $V(t)$ is greater than the logarithmic decrease of ALT subtracted by their normal value $ALT(t) - 20$ (Fig. 1) [2]. In addition, the biphasic model is meant to describe the viral load decline during the first 2–4 weeks of ther-

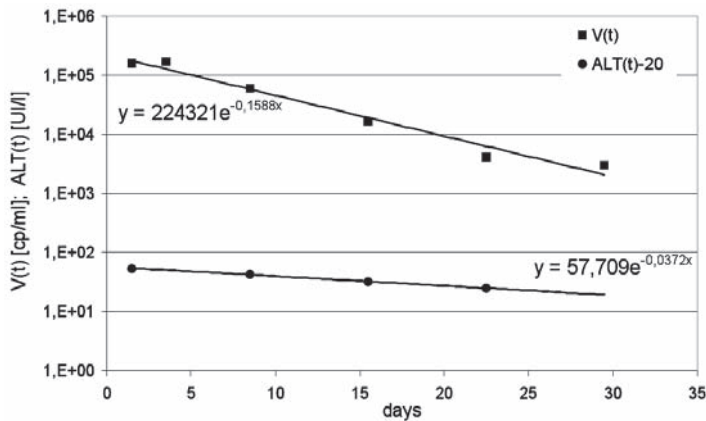


Fig. 1. Logarithmic decrease of the viral load $V(t)$ and of Alanine-Amino-Transferases (ALT) subtracted by their normal value ($ALT(t)-20$). From [2]: Colombatto et al. *Antivir. Ther.* **8**, 519–530 (2003) © International Medical Press

apy but not to simulate properly the dynamics viral infection throughout the whole treatment course. In fact the assumption of the constant clearance rate of the infected cells leads to an unrealistic estimation of the time elapsing before viral clearance [2]. To achieve a simulation throughout the treatment course we hypothesized a negative feedback on the immune system activity during therapy mediated by the reduction of the infected cells. As the infected cells trigger the immune response, being at the same time the target of the immune response, we hypothesize that the activity of the immune system responsible for the clearance of the infected cells could decrease during therapy according to the reduction of the infected cells. Therefore, we introduced a new parameter (k) that links the effectiveness of the immune clearance of infected cells to the number of infected cells (Fig. 2). Integrating the analysis of both ALT and HCV-RNA kinetics during the first month of treatment allowed to calculate with higher accuracy the infected cell clearance rate, then, under the feedback assumption, to simulate the dynamics of the infected cells until the end of therapy (I_{eot}) to predict the outcomes of the treatment [1, 2].

3 Mathematical Description of the Model

The dynamics of target (T) and infected cells (I), and that of the viral load (V) are described by the following equations:

$$\frac{dT}{dt} = \xi(T, I) - \Theta T(t) - (1 - \eta)\beta T(t)V(t); \quad (1)$$

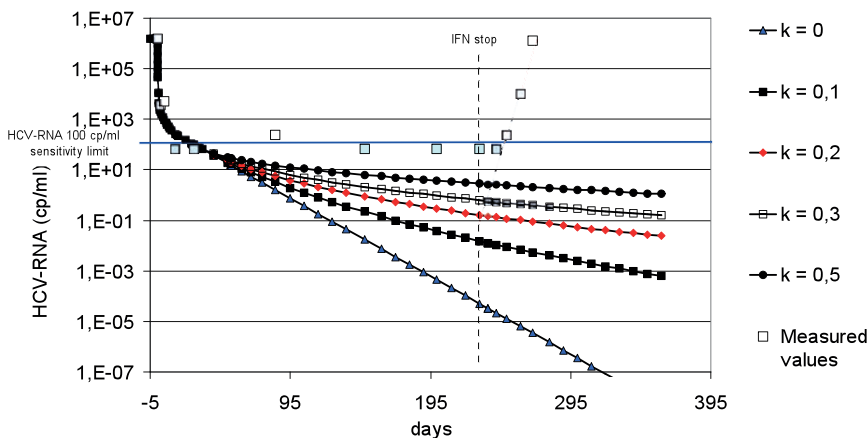


Fig. 2. Simulation of the viral dynamics throughout the whole treatment course under different assumptions on the immune system activity, feedback related to the reduction of the infected cells by the parameter k . From [2]: Colombatto et al. *Antivir. Ther.* **8**, 519–530 (2003) © International Medical Press

$$\frac{dI}{dt} = (1 - \eta)\beta T(t)V(t) - \left[\delta_0 \left(\frac{I(t - \sigma)}{I_0} \right)^k + \Theta + \alpha \right] I(t); \quad (2)$$

$$\frac{dV}{dt} = (1 - \epsilon)\Psi(t)I(t) - \lambda V(t). \quad (3)$$

In Eq. 1, $\xi(T, I)$ gives the number of target cells produced daily, $\Theta T(t)$ the number of target cells dying daily by natural death and $(1 - \eta)\beta T(t)V(t)$ the number of target cells infected daily. The term β represents the susceptibility of target cells to become infected and the term $(1 - \eta)$ the potential reduction of susceptibility due to IFN. In Eq. 2, $[\delta_0(I(t - \sigma)/I_0)^k + \Theta + \alpha]I(t)$ describes the number of infected hepatocytes lost every day, where δ_0 is the rate constant of the immune mediated clearance of the infected cells computed by the ALT decline according to the equation: $ALT(t) = \delta_0 I(t) + 20$ ($20 = ALT$ normal value). During therapy, after a delay time ($\sigma = 4$ weeks), the infected cell clearance rate is modulated by the exponent k (assumed value: 0.2) according to the reduction of the infected cell number: $(I(t - \sigma)/I_0)^k$; Θ is the hepatocyte natural turnover and α is the asymptotic immune system clearance activity for $I(t)$ approaching to zero. In Eq. 3, $\Psi(t)I(t)$ and $\lambda V(t)$ describe the number of virions produced and removed daily, where λ (assumed value: 6) is the clearance rate constant of the circulating virions. During therapy the viral production coefficient becomes $(1 - \epsilon)\Psi(t)$, where ϵ indicates the effectiveness of IFN in blocking viral production. As in most cases the 2nd phase HCV-RNA decay time constant ($1/\pi$) is shorter than that of the infected cells ($1/\delta$), the difference of the decay time constants ($1/\phi = 1/\pi - 1/\delta$) describes

the further decrease of the viral production to its asymptotic value γ in the equation:
 $\Psi(t) = (1 - \gamma)\Psi_0\exp(-\phi t) + \gamma\Psi_0$.

4 Clinical Application of the Model

By fitting ALT and HCV RNA in the first month of therapy into this model it has been possible to simulate the dynamics of the infected cells until the end of treatment, showing a strong association between the numbers of the infected hepatocytes at the end of therapy (I_{eot}) and the outcomes of standard IFN plus RBV treatments [2]. When the model was applied to compute I_{eot} in a larger cohort of patients treated with pegIFN α -2a/2b plus Ribavirin [1], the index $I_{eot} < 250$ was found to have 93% Positive Predictive Value (PPV) of sustained virological response (SVR) with a Diagnostic Accuracy (DA) of 81%, higher than that of a rapid or an early virological response at week 4 or 12 during treatment (RVR: 74%; EVR: 77%). Thus, I_{eot} appears to well represents the complexity of the viral-host interactions during therapy and provides an important tool to tailor therapy duration in the single patient according to her/his effective capacity of clearance of infected cells. The predicted duration of the treatment in the cohort where the clinical applicability of the model was tested pointed out that 60 patients could have reached $I_{eot} < 250$ between 3 and 18 months of therapy (Fig. 3), whereas 18 should have stopped treatment at week 4 because not effective. Overall, the therapy duration in the 78 modelled patients would have been of 49 months shorter than that based on the most recent algorithms (RVR plus week 12 stopping rule), thus providing a useful tool for more cost-effective Peg-IFN / RBV treatments in chronic hepatitis C patients.

Given these premises a prospective randomized trial was started to compare the cost-effectiveness of model tailored (MT) Peg-IFN / RBV treatment management as compared to Guide Lines (GL). The prospective application of our model confirmed

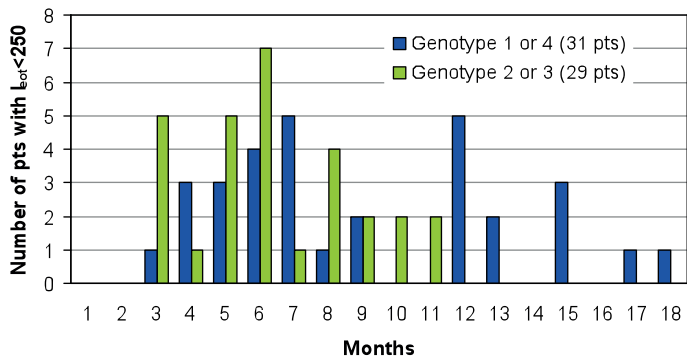


Fig. 3. Distribution according to viral genotype of the treatment duration predicted by the model to reach 250 end of treatment infected cells (I_{eot}) in 60 patients with effective antiviral therapy. From [1]: Colombatto et al. Clin. Pharmacol. Ther. **84**, 212–215 (2008) © Nature Publishing Group

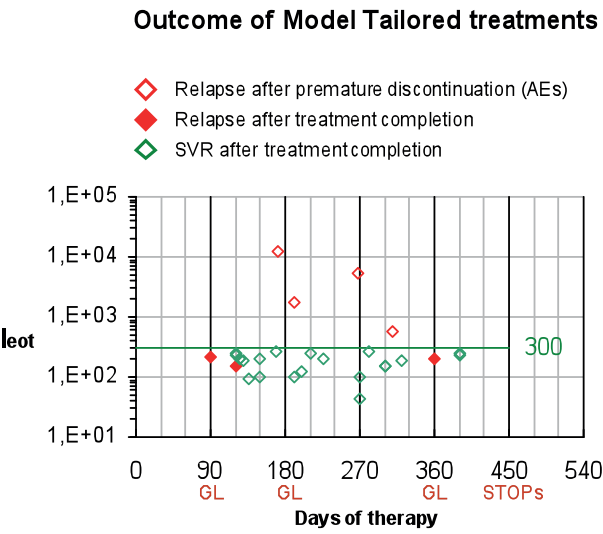


Fig. 4. Personalization of peginterferon/ribavirin treatment by model tailored duration to reach $I_{eot} < 250$ in comparison to the fixed schedules proposed by international guidelines (GL). Black empty diamonds represents patients (21) who completed treatment without recurrence, black full diamonds represents patients (3) who completed treatment but relapsed, grey empty diamonds represents patients (4) who prematurely withdrawn treatment because of adverse events (AEs) and relapsed

the wide diversification of the treatment duration required for SVR, as predicted by our previous retrospective study, and allowed in clinical practice a fine personalization of the antiviral treatment at the single patient level (Fig. 4). Tailoring treatment to $I_{eot} < 250$ showed SVR rates comparable to those of the standard schedules but with a significant reduction of futile treatments.

5 Potential for Modeling Hepatocellular Carcinoma Cell Dynamics

In patients for whom the preventive approach of the antiviral therapy fails or the liver disease is already too advanced at the time of viral clearance, HCC may occur at different hazard rates depending on the number and strength of several promoting factors and cofactors. Therapeutic options for HCC are now increasing with the introduction of a new systemic agent (sorafenib) for patients in the advanced stages, when the ablation techniques did not control the tumor or could not be applied. However, many open issues in the assessment of the response to the different treatments remain. Indeed, while the efficacy of a surgical or radiological ablation of a well defined lesion in the liver is easy to assess by imaging techniques, when Trans-Arterial Chemotherapy and Embolization (TACE) or sorafenib are used for local or

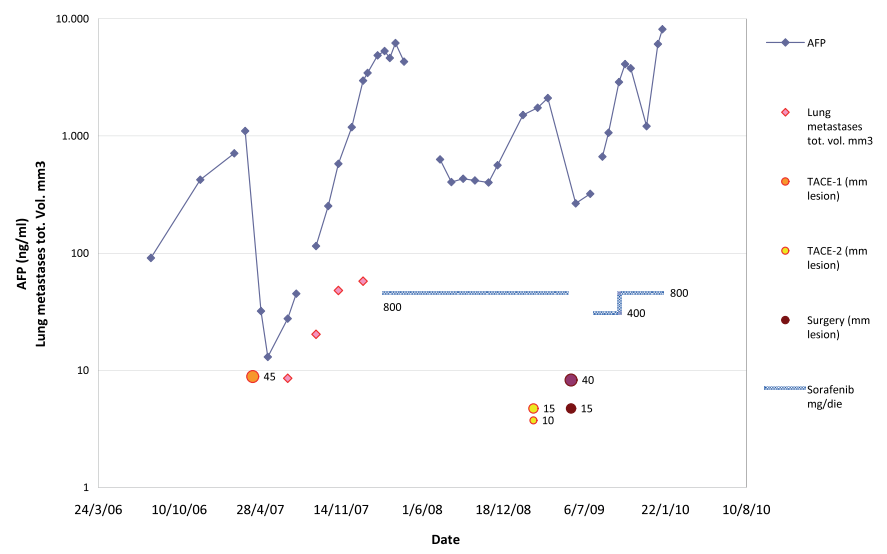


Fig. 5. Neoplastic liver proliferation biomarker (Alphafetoprotein, AFP) in response to multimodal treatment of Hepatocellular Carcinoma (HCC) including Trans-Arterial-Chemo-Embolization (TACE), liver resection and systemic chemotherapy with Sorafenib. The treated lesions are shown in the graph with their dimensions as detected by imaging techniques

systemic approaches, imaging of the treated lesions are not enough informative. In this context, the efficacy of the treatment could be related to the number of surviving neoplastic cells. However, such relevant prognostic variable is not directly assessable and only in a minority of cases it is possible to measure in the blood neoplastic markers, like the Alpha-Feto-Protein (AFP) levels in the case of the patient reported in Fig. 5, that provide information on the dynamics of cancer cells during therapy. In this field, the development of a mathematical model oriented to the clinical practice should predict the dynamics of the neoplastic cells in the individual patient after treatment, integrating experimental data obtained by imaging techniques with measures of liver function, liver cell proliferation and neoplastic markers.

6 Conclusion and Perspectives

Interesting suggestions with important therapeutic implications were derived from the analysis of HCV dynamics during therapy. Further implementations of current models will allow better simulations of HCV dynamics during the different antiviral treatments and predictions of the outcome of the infection after treatment discontinuation. This goal appears to be now achieved in patients in whom the number of HCV infected cells at the end of therapy (I_{eot}) can be calculated by fitting both ALT and HCV-RNA declines during the 1st month of therapy, thus providing a bet-

ter prediction of the treatment outcome than step-wise algorithms derived by time point measures of HCV-RNA declines. In addition, the use of a model computed I_{eot} threshold with high chance of SVR to predict treatment duration might be very helpful after a lead in phase of Peg-IFN+RBV therapy to better support the ‘add on’ of the direct antiviral agents, thus optimizing cost-effectiveness of the new antiviral therapies. In the next studies further efforts should be made to reduce the need for multiple measures of the experimental variables and to simplify the computational algorithms which make this personalized approach to antiviral therapy not suitable for universal use in clinical practice. Finally, it would be of great interest to investigate new hypothesis on the immune control of the residual infected cells in viral hepatitis models as well as to develop a liver cancer model where the effects of the atineoplastic treatments could be analysed by computing the number of neoplastic cells.

References

1. Colombatto, P., Ciccorossi, P., Maina, A.M., Civitano, L., Oliveri, F., Coco, B., Romagnoli, V., Bonino, F., Brunetto, M.R.: Early and accurate prediction of Peg-IFNs/ribavirin therapy outcome in the individual patient with chronic hepatitis C by modeling the dynamics of the infected cells. *Clin. Pharmacol. Ther.* **84**, 212–215 (2008)
2. Colombatto, P., Civitano, L., Oliveri, F., Coco, B., Ciccorossi, P., Flichman, D., Campa, M., Bonino, F., Brunetto, M.R.: Sustained response to interferon-ribavirin combination therapy predicted by a model of hepatitis C virus dynamics using both HCV RNA and alanine aminotransferase. *Antivir. Ther.* **8**, 519–530 (2003)
3. Kronenberger, B., Ruster, B., Lee, J.H., Sarrazin, C., Roth, W.K., Herrmann, G., Zeuzem, S.: Hepatocellular proliferation in patients with chronic hepatitis C and persistently normal or abnormal aminotransferase levels. *J. Hepatol.* **33**, 640–647 (2000)
4. Neumann, A.U., Lam, N.P., Dahari, H., Gretch, D.R., Wiley, T.E., Layden, T.J., Perelson, A.S.: Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* **282**, 103–107 (1998)
5. Novak, M.A., Bonhoeffer, S., Hill, A.M., Boeheme, R., Thomas, H.C., McDade, H.: Viral dynamics in hepatitis B infection. *Proc. Natl. Acad. Sci. USA* **93**, 4398–4402 (1996)
6. Perelson, A.S., Neumann, A.U., Markowitz, M., Leonard, J.M., Ho, D.D.: HIV-1 dynamics in vivo: virion clearance rate, infected cell life-span, and viral generation time. *Science* **271**, 1582–1586 (1996)

Introducing Drug Transport Early in the Design of Hypoxia Selective Anticancer Agents Using a Mathematical Modelling Approach

Kevin Hicks

Abstract Hypoxia is more severe in tumors than normal tissues, and contributes to tumor progression and therapeutic resistance. It is therefore a potentially important therapeutic target, which has led to development of hypoxia-activated prodrugs such as tirapazamine.. However extensive metabolism limits penetration of these prodrugs leading to reduced exposure in hypoxic cells which are generally most distant from the capillaries. To overcome this problem in drug design we combined in vitro drug transport and potency measurements and in vivo pharmacokinetics with spatially resolved pharmacokinetic-pharmacodynamic modelling, to successfully screen drug candidates and select a second generation tirapazamine analogue. The lead compound, CEN-209 (SN30000), is highly active in a range of tumor xenografts, and is now entering clinical evaluation. The combination of experiment and mathematical modelling in a multidisciplinary drug development setting improved screening efficiency and added valuable information for drug design.

1 Introduction

1.1 Tumor Hypoxia

Hypoxia is a common feature of solid tumors [1, 2, 37] and occurs as a result of the inefficient blood supply and rapid metabolism of oxygen in tumor tissue [5]. Hypoxia contributes to tumor progression and therapeutic resistance, and correlates with poor prognosis through a variety of mechanisms such as slowing of proliferation leading to resistance to many anticancer drugs. In particular hypoxic cells are less sensitive to killing by ionizing radiation [3, 40] and generally require approximately 3-fold higher dose to achieve the same cell killing as well oxygenated cells. This results

K. Hicks (✉)

Auckland Cancer Society Research Centre, The University of Auckland, Auckland, New Zealand
e-mail: k.hicks@auckland.ac.nz

in a large decrease in cell kill of hypoxic cells compared to well oxygenated cells given the same radiation dose. Consequently the extent of hypoxia can profoundly influence response to radiotherapy. Since the 1960s substantial research effort has been devoted to methods to improve radiation sensitivity of hypoxic cells, initially in the development of hypoxic radiosensitizers. By the 1990s accumulating knowledge from research on both hypoxia and radiosensitizers resulted in a paradigm shift – that it is possible to turn tumor hypoxia to advantage; extensive hypoxia is generally confined to tumors making it a target for cytotoxic prodrugs which are activated selectively in hypoxic cells.

1.2 Tirapazamine

The most intensively studied hypoxia-activated prodrug (HAP) is tirapazamine [40, 42], a benzotriazine di-N-oxide (TPZ, Fig. 1) which has completed extensive phase III trials [25, 31] but is unlikely to proceed to registration. The intracellular enzymatic reduction of TPZ is suppressed at high oxygen concentrations but proceeds under low oxygen concentrations via short-lived DNA damaging radicals to produce the non-toxic 1-oxide. TPZ is activated at higher oxygen concentrations than other hypoxia activated prodrugs [20, 22] ensuring its complementary action with radiation despite having no bystander effect [40]. TPZ is well suited to fractionated radiation protocols where cells at intermediate oxygen concentrations were shown to be critical in determining tumor response in a mathematical modelling study [41]. TPZ demonstrated promising activity in early clinical trials. Notably, in a small study where patient's head and neck squamous cell cancers (HNSCC) were classified using PET imaging the presence of hypoxia was shown to be a negative prognostic factor but local regional failure was substantially reduced when TPZ was added to the chemo-radiation treatment [30]. Despite this TPZ showed no overall benefit in large Phase III clinical trials for advanced HNSCC combined with chemo-radiation [31], although these trials were compromised by substandard radiotherapy [26] and lack of selection of patients with hypoxic tumors, which remains a major issue for the development of hypoxia activated prodrugs).

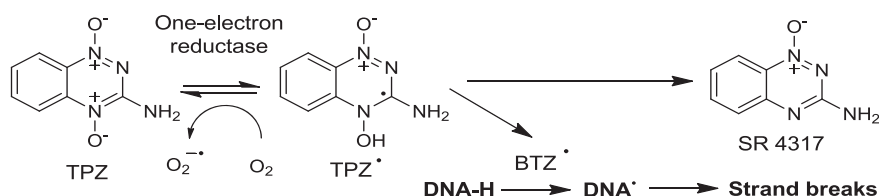


Fig. 1. Simplified scheme of TPZ metabolism. TPZ is metabolised to an initial radical which is rapidly back oxidised in the presence of oxygen. In the absence of oxygen the TPZ radical is further metabolised producing DNA damaging radicals (e.g. the benzotriazinyl radical BTZ. [35]) and the 2-electron reduction product (SR 4317)

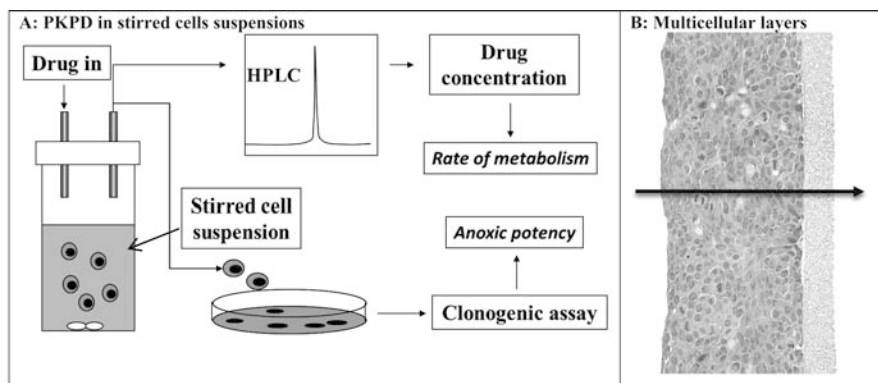


Fig. 2. (a) Schematic of experimental set up for metabolism and cytotoxicity experiments in single cell suspensions. After introduction of drug samples of medium containing cells are taken and analysed by HPLC for concentration and by clonogenic assay for cell survival, from which the parameters for metabolism and cytotoxicity were derived. (b) Multicellular layer (MCL) ca. 150 μm in thickness grown from HT29 colon cancer cells on a permeable support membrane. In transport studies in diffusion chambers drug is introduced to culture medium on one side of the MCL and samples taken from medium on both sides for HPLC analysis to determine concentrations. The arrow indicates drug transport from the donor to receiver compartments. For more details see [17, 18, 20]. Modified from [19], Hicks et al. *Clin. Cancer Res.* **16**, 4946–4957 (2010) © American Association for Cancer Research

1.3 The Extravascular Transport Problem

A major cause of sub-optimal activity of TPZ appears to be poor extravascular transport due to rapid drug metabolism in tumor tissue leading to reduced exposure in hypoxic cells which are most distant from the capillaries. It was shown early that selectivity for hypoxic cells is substantially reduced when multicellular spheroids [8] or tumor xenografts [9] are exposed to TPZ, relative to low cell density cultures, and studies with multicellular layers demonstrated TPZ transport was hindered by its high rate of bioreductive metabolism and relatively low diffusion coefficient [14, 23]. The transport problem has now been recognised for many small molecule anti-cancer agents [24] but efficient extravascular transport is particularly critical for HAP because the drugs must reach hypoxic cells which are generally the most distant from capillaries. Multicellular layers (MCL, Fig. 2b) are an *in vitro* model of the extravascular compartment of tumors, developed in our laboratory [4] where cells are grown on a semipermeable support membrane to a thickness of ca. 100–300 μm (approximately the distance between capillaries in many tumors). Due to their planar structure MCL may be used in diffusion chambers to study the rate of drug transport by measuring drug concentrations in culture medium using compound specific analytical techniques such as high-performance liquid chromatography (HPLC) or liquid chromatography/mass spectrometry (LCMS). In order to generalise the results from these transport studies with high spatial resolution it was recognised that mathematical modelling was necessary to extract the parameters governing transport in

3D and to combine these with the results from single cell metabolism and uptake studies [16].

2 Development of a Transport Model for Tirapazamine

It was clear that improving the transport properties of this class of drugs had potential to improve hypoxia selectivity in vivo and hence the therapeutic ratio of a second generation analogue. However the extensive drug metabolism (by reductases in tumor cells) which can lead to limited drug penetration cannot be eliminated entirely as it is required for cytotoxic activity. Consequently the problem of developing an improved analogue is a problem in optimising the rate of metabolism, not minimising it. To achieve this we first developed a detailed spatially resolved (SR) PKPD model for TPZ transport, metabolism and cytotoxicity [17, 18, 20]. 2.1 Drug metabolism and cytotoxicity TPZ metabolism and cytotoxicity was studied in vitro in stirred hypoxic suspensions of HT29 colon carcinoma cells where samples are taken at hourly intervals after drug was introduced into the vials [18]. Drug concentrations were measure by HPLC on the extracellular medium and cell survival (clonogenicity) was measured by plating cells from the same sample (Fig. 2a). The resulting concentration-time data was fitted with a mixed order kinetic model:

$$\frac{dC}{dt} = -\phi_c M(C), \quad (1)$$

where

$$M(C) = \left(k_{met} + \frac{V_{max}}{k_m + C} \right) C \quad (2)$$

is the rate of intracellular drug metabolism, C is the drug concentration at time t , k_{met} , V_{max} , k_m are the parameters of the (intracellular) metabolism model, $C(0) = C_0$, the initial drug concentration at time zero and ϕ_c is the cell volume fraction in the experiment

$$\phi_c = n \frac{V_c}{V} \quad (3)$$

where n is the number of cells in volume V and V_c is the volume of a single cell. The rate of cell killing was related to the rate of intracellular metabolism and the concentration of parent drug:

$$\frac{d \log SF}{dt} = -\gamma C M(C) \quad (4)$$

where SF is the cell surviving fraction, and γ is a proportionality constant.

2.1 MCL Flux

This homogenous cellular pharmacokinetic-pharmacodynamic (PKPD) model was then extended to tumor tissue by measuring transport of TPZ in HT29 MCL (Fig. 2b)

under aerobic conditions to suppress bio-reductive metabolism and under hypoxic conditions to stimulate drug metabolism. Drug was introduced into one side of diffusion chambers and samples taken from both the donor and receiver compartments at hourly intervals. Concentration-time profiles in the 2 compartments of the diffusion chamber were fitted using Fick's second law to describe transport in the MCL:

$$\frac{\partial C}{\partial t} = D \frac{\partial^2 C}{\partial x^2} - \phi_c M(C). \quad (5)$$

Drug was assumed to be homogeneously distributed in the intracellular and extracellular phase. D , the diffusion coefficient of drug, was estimated using aerobic MCL (k_{met} and $V_{max} = 0$) and ϕ , the scaling factor between the single cell metabolism model and that at MCL cell density, was estimated using hypoxic MCL. The thickness of the MCL was estimated by simultaneously measuring the transport of stable ^{14}C -urea and using the previously determined urea diffusion coefficient to fit the urea data with MCL thickness as fitted parameter. This transport model was then used to predict the gradients of concentration and cell killing at each point in hypoxic MCL. To test the model MCL were exposed to TPZ from both sides for 1 or 2 hours with concentrations determined by HPLC, then MCL were disaggregated and clonogenicity determined. Good agreement between measured and predicted average cell killing was observed [18].

3 Extension of the Model to Tumors

Prediction of TPZ distribution and cell killing in HT29 tumor xenografts, where there is a gradient of oxygen concentration, presented a greater challenge. This required additional parameters to be determined: (1) the oxygen dependence of metabolism and cytotoxicity of TPZ; (2) the plasma pharmacokinetics of TPZ in nude mice; (3) the plasma protein binding of TPZ in order to determine free drug fraction and (4) a model for the tumor microvascular network in xenografts.

3.1 Oxygen Dependence of Metabolism and Cytotoxicity

As solution oxygen concentration increases cell killing by TPZ is suppressed. The oxygen dependence of cell killing by TPZ was determined in metabolism and cytotoxicity experiments as described above with gassing at a range of oxygen partial pressures [20]. Solution oxygen concentrations were measured using with an OxyLite optical oxygen sensor. Fitting the metabolism model of (1) showed that oxygen suppressed the rate of reductive metabolism by a factor of

$$f(O_2) = \frac{K_{O_2}}{K_{O_2} + O_2}, \quad (6)$$

where K_{O_2} is the oxygen concentration for half-maximal suppression ($1.2\mu\text{M } O_2$).

Log cell kill was suppressed by the same amount, implying that the effect of oxygen is entirely mediated through the effect on suppressing the rate of metabolism.

3.2 Plasma Pharmacokinetics

Plasma PK of TPZ was determined after intraperitoneal (i.p.) injection at the maximum tolerated dose (MTD) in nude mice. Absorption from the peritoneum was rapid and plasma concentrations were well fitted by a simple first order one-compartment model

$$C_p(t) = C_0 e^{-kt}, \quad (7)$$

where C_0 is the maximum plasma concentration, and $k (= \ln(2)/t_{1/2})$ is the elimination rate constant.

3.3 Plasma Protein Binding

Since free drug concentration drives diffusion into tissues and cells the plasma protein binding of TPZ was determined (in vitro) by equilibrium dialysis and confirmed to be negligible as previously reported [38].

3.4 Microvascular Network

A model for the tumor microvascular network in xenografts was required to simultaneously model the spatial distribution of oxygen and drug. This presented the greatest challenge as no microvascular network was available for HT29 xenografts. However, a microvascular network from a rat R3230Ac window chamber tumor had been mapped in detail with measured blood flows [33]. In addition, the microvascular network was digitised (shown in Fig. 3) and a computational model of oxygen tissue distribution produced [33] using measured rates of oxygen metabolism [7]. This model reproduced many features of the dysfunctional vasculature in tumors that leads to hypoxia [10, 27, 28, 32], namely highly variable capillary diameters, blood flows and intercapillary distances together with extraction of oxygen along the length of the capillaries, all leading to highly heterogeneous distributions of oxygen typically seen in xenografts models [6, 39] and human tumors [36].

3.5 Application to TPZ PKPD

We collaborated with these groups to develop an in silico SR-PKPD model for oxygen and TPZ using the Greens function method which is described in detail elsewhere [17, 21, 33, 34]. Briefly the tissue is assumed to be homogeneous medium with a steady input concentration of oxygen and drug in the inflow microvessels. By conservation of mass

$$D_{O_2} \alpha \nabla^2 P = M(P), \quad (8)$$

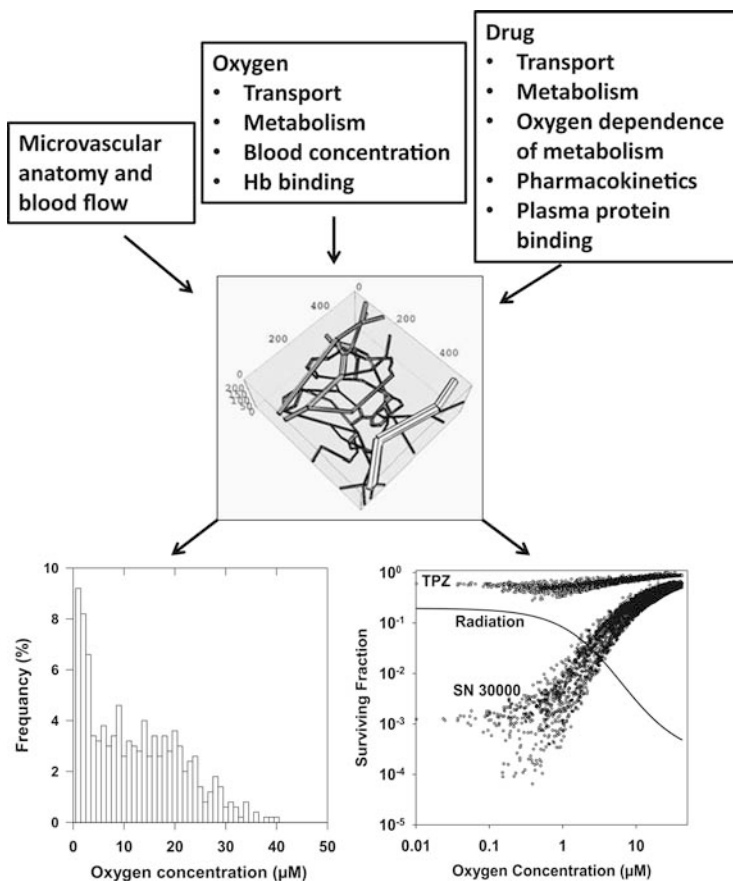


Fig. 3. Inputs to, and outputs from, the 3D spatially resolved pharmacokinetic-pharmacodynamic model. Measured microvascular anatomy and blood flows from a rat window chamber tumor were used to produce a digitized microvascular network (central figure) used for in silico modelling of oxygen and drug distribution based on their transport parameters shown as inputs. The lower graphs show distribution of oxygen in the region and the distribution of TPZ killing as a function of oxygen. Also shown is the distribution of cell killing predicted for CEN-209 which indicates the gain made for this compound with superior transport and potency properties. The complimentary cell killing by radiation as a function of oxygen concentration (which is used to calculate overall cell kill when combined with drug) is also shown. Lower left: modified from [17], Hicks, K.O., et al. *J. Natl. Cancer Inst.* **98**, 1118–1128 (2006) © Oxford University Press; lower right: modified from [19], Hicks et al. *Clin. Cancer Res.* **16**, 4946–4957 (2010) © American Association for Cancer Research

where P is the tissue oxygen partial pressure, D_{O_2} the is the diffusion coefficient, α is the oxygen solubility and $M(P)$ is the oxygen metabolism rate,

$$M(P) = M_0 \frac{P}{P_0 + P}, \quad (9)$$

where M_0 is the maximal oxygen consumption, P_0 is pO_2 when O_2 metabolism is half-maximal. For drug,

$$D\nabla^2 C = M(C), \quad (10)$$

where

$$M(C, O_2) = f(O_2)M(C), \quad (11)$$

with $M(C)$ and $f(O_2)$ as defined above. The rate of convective oxygen transport along each vessel segment is

$$f(P_b) = Q(H_D C_0 S(P_b) + \alpha_{\text{eff}} P_b), \quad (12)$$

where Q is the flow rate of blood, P_b is the blood pO_2 , H_D is the haematocrit, S is the haemoglobin saturation, C_0 is the maximum concentration of haemoglobin-bound oxygen and α_{eff} is the effective solubility of oxygen in blood. Haemoglobin saturation is:

$$S(P_b) = \frac{P_b^n}{P_b^n + P_{50}^n}, \quad (13)$$

where P_{50} is the pO_2 at 50% saturation and n is the Hill constant. For numerical implementation using the Green's function method blood vessels are represented as discrete oxygen sources, and the tissue region is represented as a set of discrete oxygen sinks. Vessels are represented as short segments and the tissue is discretised into small volumes approximately cubic in shape. The resulting system of nonlinear equations for the unknown source strengths is solved iteratively. Further details and a C++ implementation of the method (which now includes multiple diffusing solutes) is available for download at <http://www.physiology.arizona.edu/people/secomb/greens>. Parameters to describe oxygen transport in the microregion were set to reported values. This resulted in the calculated oxygen concentration distribution shown in Fig. 3 with hypoxic fraction (proportion below $4\mu M O_2$) of ca. 30%. The distribution of TPZ in the microvascular network was predicted using the transport (PK) parameters (diffusion, and metabolism as a function of oxygen concentration) and cytotoxicity parameters (potency) with free drug concentration from mouse plasma PK as the input [17]. Surviving fractions at each point were then calculated from the model-predicted tissue drug concentrations by integrating (3) to infinite time since the clonogenic assays were performed 18 hours (at least 20 drug plasma half-lives) after drug administration (see below). The model clearly predicted less cell killing in the microregion than the theoretical maximum attainable if TPZ was homogeneously distributed at the same concentration as in plasma (see [17] for more details). In addition, the model predicted heterogeneous distribution of TPZ in the tumor tissue as a function of distance from the capillary as expected in a complex 3D structure with a range of intercapillary distances and oxygen concentrations. The resultant heterogeneous distribution of cell killing is shown in Fig. 3.

3.6 Validation of the *in Silico* model

By using tumor cell survival in the Greens function model as the efficacy endpoint, the predictions of the model could be directly compared to the results of the clonogenic assays performed after tumors were excised from mice treated with TPZ (and its analogues) combined with radiation. Briefly, mice were treated with drug 5 minutes after 20 Gy tumor irradiation (which sterilises well oxygenated cells). Tumors were excised 18 hours after treatment, cells plated and observed cell killing was compared to that predicted by the *in silico* tumor network model. Radiation induced cell kill was predicted at each point in the *in silico* model by the linear quadratic model fitted to HT29 cells irradiated under hypoxia (parameters $\alpha_H = 0.0473 \pm 0.0158 \text{ Gy}^{-1}$ and $\beta_H = 0.0017 \pm 0.0004 \text{ Gy}^{-2}$) using literature parameters for the oxygen enhancement ratio and the K_{ms} value ($OER_{\alpha_m} = 2.5$ and $OER_{\beta_m} = 3.0$, and $K_{ms} = 4.3 \mu M$) [17, 41]:

$$-\ln SF = \alpha_H OER_{\alpha} D_r + \beta_H [OER_{\beta} D_r]^2, \quad (14)$$

$$OER_i = \frac{OER_{i,max} O_2 K_{ms}}{O_2 + K_{ms}}. \quad (15)$$

For each case (radiation alone and radiation with drug), predicted surviving fraction at each point was averaged over the whole tumor region. Hypoxic cell killing was then defined as the difference between log cell kill with radiation and drug and log cell kill with radiation alone

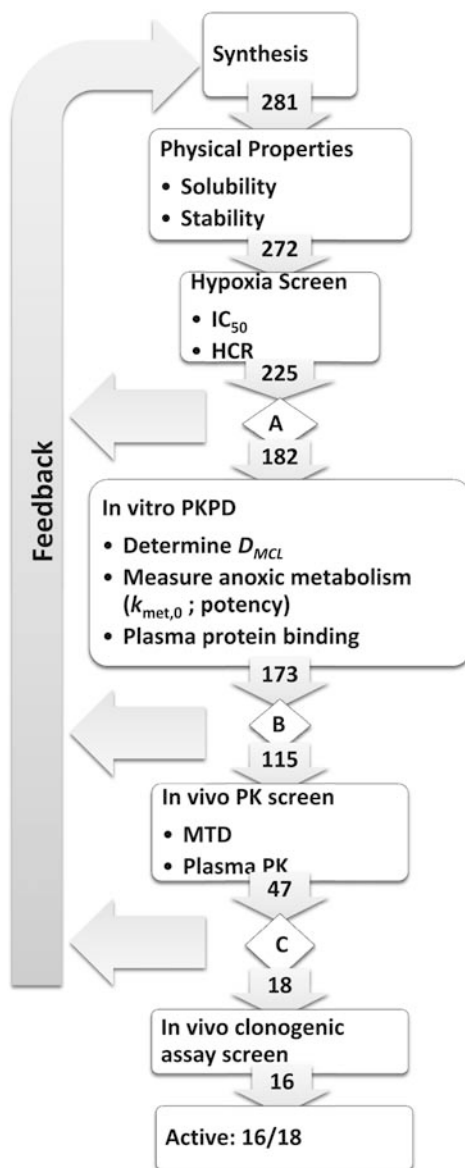
$$\log(\text{drug induced hypoxic cell kill}) = \log \left(\frac{\text{Average SF, radiation alone}}{\text{Average SF, radiation + drug}} \right). \quad (16)$$

Good agreement was seen between predicted and observed hypoxic cell killing for TPZ in HT29 xenografts and we subsequently showed that this *in silico* model predicted the *in vivo* activity of 14 other TPZ analogues in the initial screening program [17]. We also validated the model for TPZ in a second tumor (SiHa) xenograft [15]

4 Tirapazamine Analogue Screening

In order to introduce drug transport into the early stages of TPZA design we required a simple, repeatable model for screening large numbers of drugs, with easily accessible endpoints which are directly comparable *in vitro* and *in vivo*. To screen analogues and prioritize them for advanced *in vivo* testing we introduced additional steps in the traditional screening process by using a simplified form of the detailed SR-PKPD model developed for TPZ [11–13, 19]. The screening algorithm (shown in Fig. 4) essentially repeated the experiments for development of the TPZ PKPD model (3.1–3.6 above) with the following simplifications to enable rapid throughput: 1. Prodrug metabolism was not tested over a wide range of concentrations and was assumed to be first order. 2. Since the compounds are closely related they were assumed to have the same oxygen dependence of metabolism as TPZ (i.e. $K_{O_2} = 1.2 \mu M$, with no

Fig. 4. Flowchart of technique used for selecting tirapazamine analogues with improved extravascular transport properties. The critical decision points shown are A, after initial hypoxic selectivity screen; B, after using the in silico PKPD model to predict hypoxic selectivity in tumors without in vivo PK data; C, after using the in silico PKPD model to predict hypoxic cell killing when combined with radiation using in vivo PK data as input. The feedback loop emphasises that the high quality pharmacokinetic data obtained in this scheme could be used to continuously inform drug design. Modified from [19]: Hicks et al. Clin. Cancer Res. **16**, 4946–4957 (2010) © American Association for Cancer Research



oxygen-independent metabolism). 3. Compounds which failed at a particular step were dropped from PKPD model development and further testing. Analogues were initially screened by measurement of stability, solubility, hypoxic potency (determined by IC_{50} for inhibition of growth after exposure to the drugs in an anoxic chamber) and hypoxic selectivity determined by calculating the hypoxic cytotoxic ratio: $HCR = IC_{50}(20\% O_2)/IC_{50}(0\% O_2)$. For those compounds passing this ini-

tial screen of sufficient potency and $HCR > 20$ (decision point A, Fig. 4) we then measured metabolism and cytotoxic potency in hypoxic cell suspensions and measured or calculated [29] the diffusion coefficients in well oxygenated MCL for each analogue. These parameters were used in the in silico model to investigate predicted drug distribution and cytotoxicity in the microvascular network. At this stage (before plasma PK was measured) an arbitrary input plasma PK was used and the output from the model was the distribution of drug exposure and cytotoxicity. A simple parameter, the hypoxic cytotoxic differential was used to assess the predicted in vivo hypoxic selectivity of drugs at tumor tissue cell density:

$$HCD = \frac{LCK_{(<4\mu M)}}{LCK_{(>30\mu M)}}, \quad (17)$$

where $LCK_{(<4\mu M)}$ is the average log cell kill in cells at oxygen concentration $< 4\mu M$ and $LCK_{(>30\mu M)}$ is the average LCK in cells at $[O_2] > 30\mu M$. Only compounds penetrating the tumor tissue efficiently and thus demonstrating hypoxia selectivity in the 3D microvascular network model (i.e. $HCD > 1$) passed decision point B and were progressed to in vivo testing¹. For these compounds, after determination of MTD and plasma PK, the in silico model was re-run to predict hypoxic cell killing at the MTD, this time using the measured free drug C_{max} and AUC_{inf} as input. Those compounds where plasma PK was sufficient to demonstrate model predicted selective hypoxic cell killing in addition to radiation alone were then advanced to testing in the tumor clonogenic assay (Decision point C). Over the whole program 18 compounds were progressed to testing in xenografts by clonogenic assays with dosing at 75% of the MTD and 16/18 compounds predicted to be active were active in vivo while all (12/12) compounds predicted to be inactive were inactive. If PKPD screening had not been used at least 47 compounds would have been tested in vivo, while if a traditional screening approach had been used (potency and HCR both $> TPZ$) only 6 active compounds would have been identified. Now that the screening program is complete we have performed a retrospective analysis of the ability of this in silico model to predict antitumor activity in this series by regression of predicted vs observed hypoxic cell killing for all compounds (Fig. 5). There is a close correspondence between predicted hypoxic cell killing in the microvascular network model and that observed in tumor xenografts ($slope = 0.99$, $p < 0.001$, $R^2 = 0.85$; Fig. 5A). However, if the tumor compartment was assumed to be continuously stirred with respect to drug transport, as often assumed in PK analyses, the correlation between predicted and observed activity is poor ($slope = 0.24$, $p = 0.013$, $R^2 = 0.22$; Fig. 5B). This updates the previous validation of this approach which used 7 active and 9 inactive compounds [17]. The lead compound from this analogue program, CEN-209 (Fig. 6), was predicted to be highly active against hypoxic tumor cells by the in silico model (Fig. 4), due to improved potency and extravascu-

¹ Some compounds with predicted $HCD < 1$ were advanced to further testing to validate the in silico model.

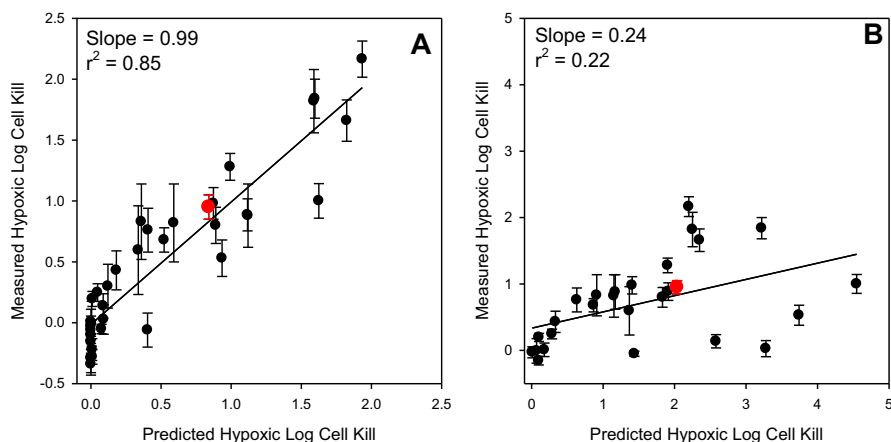


Fig. 5. (a) Correlation between hypoxic cell killing predicted by the spatially resolved PKPD model in the tumor microvascular network and cell killing observed in HT29 tumor xenografts for all analogues tested. (b) A similar correlation with cell kill predicted using the microvascular network with the same oxygen gradient but assuming the extravascular compartment is continuously stirred with respect to drug transport (tissue drug concentrations equal to those in plasma). Hypoxic log cell kill is defined as in the text. This updates the previously reported correlation in [17] with the addition of 25 further compounds

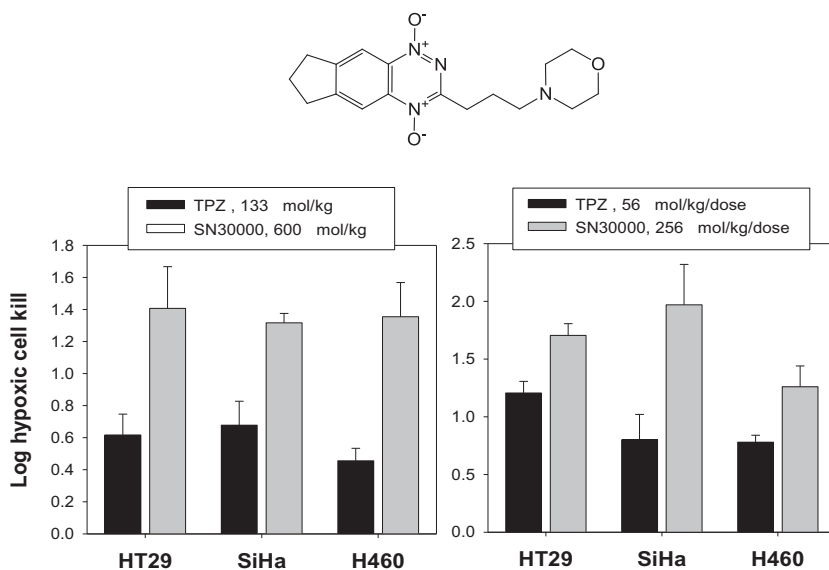


Fig. 6. Improved activity of CEN-209, the lead TPZ analogue selected from the PKPD guided screening program compared to TPZ in clonogenic assays 18 hours after administration radiation and drug at a single dose (left panel) or twice daily doses for 4 days. Activity is presented as logs of hypoxic cell kill (additional logs of cell kill in clonogenic assays when drug is combined with radiation, compared to radiation alone). Redrawn from [19]: Hicks et al. Clin. Cancer Res. **16**, 4946–4957 (2010) © American Association for Cancer Research

lar transport. It has been shown to be substantially more active than TPZ in excision assays across a range of tumor xenografts with both single dose (Fig. 6, Left panel) and fractionated (Fig. 6, right panel) radiation.

5 Discussion

In vitro drug transport and potency measurements and plasma PK, combined with in silico spatially resolved PKPD modelling, has been used to successfully screen drug candidates and select a second generation TPZ analogue (CEN-209) with predicted improved activity against hypoxic cells in tumor xenografts. CEN-209 also has improved solubility, anoxic cytotoxic potency and hypoxic selectivity and has been shown to be substantially more active than TPZ across a range of tumor xenograft types with both single dose and fractionated radiation (Fig. 6), and is now entering clinical evaluation. This program would not have been possible without a multidisciplinary team, which included synthetic chemists, cancer biologists, pharmacologists and mathematical modellers. The program was highly integrated with experimentalist and modellers giving rapid feedback to chemists with detailed information across the cellular, tissue, and whole animal scale being used to inform drug design. For example, detailed PKPD analysis of each compound enabled calculation of parameters useful for drug design such as the penetration half distance [11–13]

$$X_{1/2} = \ln(2) \sqrt{\frac{D}{\phi k_{met}}}, \quad (18)$$

which combines the diffusion coefficient and rate constant for metabolism in a simple measure of the ability of the drug to penetrate. Reviewing the screening program there is little correlation between IC₅₀ or HCR and in vivo activity [19]. Consequently, progressing drugs from the primary in vitro screen into in vivo testing would have been inefficient, whereas the in silico model successfully predicted the magnitude of in vivo activity in 38/40 compounds with a high correlation between predicted and observed cell killing. We can now use this data set to test other in silico models for predictive ability. In particular we have now shown that the in silico predictions are not highly sensitive to tumor geometry with good correlations between predicted and observed activity in simpler spheroid and cylindrical models, although the microvascular network model performed the best overall (paper in preparation). Our studies have also shown that it is important not to separate PK and PD. For TPZ analogues the same metabolism that limits transport also contributes to cell killing and hence use of a linked PKPD model is essential for prediction of the consequences of differences in drug transport (that is, in vivo activity). The SR-PKPD model has many potential applications to other classes of compounds. In addition, the in silico model will now be used to give an early indication of antitumor activity as human PK data becomes available in clinical trials of CEN-209.

6 Challenges for Systems Biomedicine

We are now extending this model to other classes of anticancer agents, for example to hypoxia selective prodrugs that selectively release molecularly targeted drugs and cytotoxins which can diffuse to and kill surrounding well oxygenated cancer cells (bystander effect). While the model worked well for a drug series such as TPZ analogues, with a simple mechanism of action (no bystander effect, non-toxic diffusible metabolic product), when there is metabolism to a diffusible cytotoxic species this simple approach fails [15]. If cell uptake is also involved and there are multiple metabolites the complexity increases substantially. The speed of drug screening may also be an issue as this method can appear slow compared to traditional mass screening techniques. However this must be balanced against the increase in high quality data this approach provides that can be used to provide feedback to medicinal chemistry. The generality of the microvascular network in being typical of all tumors is an issue. Because of the spatial and temporal heterogeneity of tumors, detailed analysis of their architecture and blood flow is cost and labor intensive. New tumor networks, based on microvessel structural adaptation models [27, 28], may allow further exploration of the effect of tumor structure and geometry on extravascular transport of anticancer drugs. In addition the lack of time dependence in the current Greens function model is a limitation and being addressed to allow time varying inputs. Finally, relating short term endpoints such as cell killing to other common endpoints (e.g. growth delay) is a challenge for this model. Modelling growth delay in a spatially resolved manner requires tumor angiogenesis and growth models which tend to be computationally intensive. The design of programs which are user friendly and have good databasing capability will be essential for the practical use of these techniques in drug design

Acknowledgements I wish to thank the team at the Auckland Cancer Society Research Centre (especially William Wilson, Michael Hay, Frederik Pruijn, William Denny, Bronwyn Siim), Timothy Secomb (University of Arizona) for 3D greens function modelling, Mark Dewhirst (Duke University) for making the experimental microvascular network model available and the program director of NCI Grant, J. Martin Brown (Stanford University). This work was supported by the U.S. National Cancer Institute under grant CA82566, the Health Research Council of New Zealand, and the Auckland Division of the Cancer Society of New Zealand.

References

1. Brown, J.M.: Tumor hypoxia in cancer therapy. *Methods Enzymol.* **435**, 297–321 (2007)
2. Brown, J.M.: Henry S. Kaplan Distinguished Scientist Award Lecture 2007. The remarkable yin and yang of tumour hypoxia. *Int. J. Radiat. Biol.* **86**, 907–917 (2010)
3. Brown, J.M., Wilson, W.R.: Exploiting tumor hypoxia in cancer treatment. *Nat. Rev. Cancer* **4**, 437–447 (2004)
4. Cowan, D.S., Hicks, K.O., Wilson, W.R.: Multicellular membranes as an in vitro model for extravascular diffusion in tumours. *Br. J. Cancer Suppl.* **27**, S28–S31 (1996)
5. Dewhirst, M.W.: Concepts of oxygen transport at the microcirculatory level. *Semin. Radiat. Oncol.* **8**, 143–150 (1998)

6. Dewhirst, M.W., Ong, E.T., Klitzman, B., Secomb, T.W., Vinuya, R.Z., Dodge, R., Brizel, D., Gross, J.F.: Perivascular oxygen tensions in a transplantable mammary tumor growing in a dorsal flap window chamber. *Radiat. Res.* **130**, 171–182 (1992)
7. Dewhirst, M.W., Secomb, T.W., Ong, E.T., Hsu, R., Gross, J.F.: Determination of local oxygen consumption rates in tumors. *Cancer Res.* **54**, 3333–3336 (1994)
8. Durand, R.E., Olive, P.L.: Evaluation of bioreductive drugs in multicell spheroids. *Int. J. Radiat. Oncol. Biol. Phys.* **22**, 689–692 (1992)
9. Durand, R.E., Olive, P.L.: Physiologic and cytotoxic effects of tirapazamine in tumor-bearing mice. *Radiat. Oncol. Investig.* **5**, 213–219 (1997)
10. Gillies, R.J., Schornack, P.A., Secomb, T.W., Raghunand, N.: Causes and effects of heterogeneous perfusion in tumors. *Neoplasia* **1**, 197–207 (1999)
11. Hay, M.P., Hicks, K.O., Pchalek, K., Lee, H.H., Blaser, A., Pruijn, F.B., Anderson, R.F., Shinde, S.S., Wilson, W.R., Denny, W.A.: Tricyclic [1,2,4]triazine 1,4-dioxides as hypoxia selective cytotoxins. *J. Med. Chem.* **51**, 6853–6865 (2008)
12. Hay, M.P., Hicks, K.O., Pruijn, F.B., Pchalek, K., Siim, B.G., Wilson, W.R., Denny, W.A.: Pharmacokinetic/pharmacodynamic model-guided identification of hypoxia-selective 1,2,4-benzotriazine 1,4-dioxides with antitumor activity: the role of extravascular transport. *J. Med. Chem.* **50**, 6392–6404 (2007)
13. Hay, M.P., Pchalek, K., Pruijn, F.B., Hicks, K.O., Siim, B.G., Anderson, R.F., Shinde, S.S., Phillips, V., Denny, W.A., Wilson, W.R.: Hypoxia-selective 3-alkyl 1,2,4-benzotriazine 1,4-dioxides: the influence of hydrogen bond donors on extravascular transport and antitumor activity. *J. Med. Chem.* **50**, 6654–6664 (2007)
14. Hicks, K.O., Fleming, Y., Siim, B.G., Koch, C.J., Wilson, W.R.: Extravascular diffusion of tirapazamine: effect of metabolic consumption assessed using the multicellular layer model. *Int. J. Radiat. Oncol. Biol. Phys.* **42**, 641–649 (1998)
15. Hicks, K.O., Myint, H., Patterson, A.V., Pruijn, F.B., Siim, B.G., Patel, K., Wilson, W.R.: Oxygen dependence and extravascular transport of hypoxia-activated prodrugs: comparison of the dinitrobenzamide mustard PR-104A and tirapazamine. *Int. J. Radiat. Oncol. Biol. Phys.* **69**, 560–571 (2007)
16. Hicks, K.O., Ohms, S.J., van Zijl, P.L., Denny, W.A., Hunter, P.J., Wilson, W.R.: An experimental and mathematical model for the extravascular transport of a DNA intercalator in tumours. *Br. J. Cancer* **76**, 894–903 (1997)
17. Hicks, K.O., Pruijn, F.B., Secomb, T.W., Hay, M.P., Hsu, R., Brown, J.M., Denny, W.A., Dewhirst, M.W., Wilson W.R.: Use of three-dimensional tissue cultures to model extravascular transport and predict in vivo activity of hypoxia-targeted anticancer drugs. *J. Natl. Cancer Inst.* **98**, 1118–1128 (2006)
18. Hicks, K.O., Pruijn, F.B., Sturman, J.R., Denny, W.A., Wilson, W.R.: Multicellular resistance to tirapazamine is due to restricted extravascular transport: a pharmacokinetic/pharmacodynamic study in HT29 multicellular layer cultures. *Cancer Res.* **63**, 5970–5977 (2003)
19. Hicks, K.O., Siim, B.G., Jaiswal, J.K., Pruijn, F.B., Fraser, A.M., Patel, R., Hogg, A., Liyanage, H.D.S., Dorie, M.J., Brown, J.M., Denny, W.A., Hay, M.P., Wilson, W.R.: Pharmacokinetic/pharmacodynamic modeling identifies SN30000 and SN29751 as tirapazamine analogues with improved tissue penetration and hypoxic cell killing in tumors. *Clin. Cancer Res.* **16**, 4946–4957 (2010)
20. Hicks, K.O., Siim, B.G., Pruijn, F.B., Wilson, W.R.: Oxygen dependence of the metabolic activation and cytotoxicity of tirapazamine: implications for extravascular transport and activity in tumors. *Radiat. Res.* **161**, 656–666 (2004)
21. Hsu, R., Secomb, T.W.: A Green's function method for analysis of oxygen delivery to tissue by microvascular networks. *Math. Biosci.* **96**, 61–78 (1989)

22. Koch, C.J.: Unusual oxygen concentration dependence of toxicity of SR-4233, a hypoxic cell toxin. *Cancer Res.* **53**, 3992–3997 (1993)
23. Kyle, A.H., Minchinton, A.I.: Measurement of delivery and metabolism of tirapazamine to tumour tissue using the multilayered cell culture model. *Cancer Chemother. Pharmacol.* **43**, 213–220 (1999)
24. Minchinton, A.I., Tannock, I.F.: Drug penetration in solid tumours. *Nat. Rev. Cancer* **6**, 583–592 (2006)
25. von Pawel, J., von Roemeling, R., Gatzemeier, U., Boyer, M., Elisson, L.O., Clark, P., Talbot, D., Rey, A., Butler, T.W., Hirsh, V., Olver, I., Bergman, B., Ayoub, J., Richardson, G., Dunlop, D., Arcenas, A., Vescio, R., Viallet, J., Treat, J.: Tirapazamine plus cisplatin versus cisplatin in advanced non-small-cell lung cancer: A report of the international CATAPULT I study group. Cisplatin and tirapazamine in subjects with advanced previously untreated non-small-cell lung tumors. *J. Clin. Oncol.* **18**, 1351–1359 (2000)
26. Peters, L.J., O’Sullivan, B., Giralt, J., Fitzgerald, T.J., Trotti, A., Bernier, J., Bourhis, J., Yuen, K., Fisher, R., Rischin D.: Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. *J. Clin. Oncol.* **28**, 2996–3001 (2010)
27. Pries, A.R., Cornelissen, A.J., Sloot, A.A., Hinkeldey, M., Dreher, M.R., Hoepfner, M., Dewhirst, M.W., Secomb, T.W.: Structural adaptation and heterogeneity of normal and tumor microvascular networks. *PLoS Comput. Biol.* **5**, e1000394 (2009)
28. Pries, A.R., Hopfner, M., le Noble, F., Dewhirst, M.W., Secomb, T.W.: The shunt problem: control of functional shunting in normal and tumour vasculature. *Nat. Rev. Cancer* **10**, 587–593 (2010)
29. Pruijn, F.B., Patel, K., Hay, M.P., Wilson, W.R., Hicks, K.O.: Prediction of tumour tissue diffusion coefficients of hypoxia-activated prodrugs from physicochemical parameters. *Aust. J. Chem.* **61**, 687–693 (2008)
30. Rischin, D., Fisher, R., Peters, L., Corry, J., Hicks, R.: Hypoxia in head and neck cancer: Studies with hypoxic positron emission tomography and hypoxic cytotoxins. *Int. J. Radiat. Oncol. Biol. Phys.* **69** (Suppl 2), S61–S63 (2007)
31. Rischin, D., Peters, L.J., O’Sullivan, B., Giralt, J., Fisher, R., Yuen, K., Trotti, A., Bernier, J., Bourhis, J., Ringash, J., Henke, M., Kenny, L.: Tirapazamine, cisplatin, and radiation versus cisplatin and radiation for advanced squamous cell carcinoma of the head and neck (TROG 02.02, HeadSTART): A phase III trial of the Trans-Tasman Radiation Oncology Group. *J. Clin. Oncol.* **28**, 2989–2995 (2010)
32. Secomb, T.W., Dewhirst, M.W., Pries, A.R.: Structural adaptation of normal and tumour vascular networks. *Basic Clin. Pharmacol. Toxicol.* **110**, 63–69 (2012)
33. Secomb, T.W., Hsu, R., Braun, R.D., Ross, J.R., Gross, J.F., Dewhirst, M.W.: Theoretical simulation of oxygen transport to tumors by three-dimensional networks of microvessels. *Adv. Exp. Med. Biol.* **454**, 629–634 (1998)
34. Secomb, T.W., Hsu, R., Park, E.Y., Dewhirst, M.W.: Green’s function methods for analysis of oxygen delivery to tissue by microvascular networks. *Ann. Biomed. Eng.* **32**, 1519–1529 (2004)
35. Shinde, S.S., Hay, M.P., Patterson, A.V., Denny, W.A., Anderson, R.F.: Spin trapping of radicals other than the $\cdot\text{OH}$ radical upon reduction of the anticancer agent tirapazamine by cytochrome P450 reductase. *J. Am. Chem. Soc.* **131**, 14220–14221 (2009)
36. Vaupel, P., Hockel, M., Mayer, A.: Detection and characterization of tumor hypoxia using pO₂ histography. *Antioxid Redox Signal* **9**, 1221–1235 (2007)
37. Vaupel P., Mayer, A.: Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev.* **26**, 225–239 (2007)

38. Walton, M.I., Workman, P.: Pharmacokinetics and bioreductive metabolism of the novel benzotriazine di-N-oxide hypoxic cell cytotoxin tirapazamine (WIN 59075; SR 4233; NSC 130181) in mice. *J. Pharmacol. Exp. Ther.* **265**, 938–947 (1993)
39. Wilson, D.F., Evans, S.M., Jenkins, W.T., Vinogradov, S.A., Ong, E., Dewhirst, M.W.: Oxygen distributions within R3230Ac tumors growing in dorsal flap window chambers in rats. *Adv. Exp. Med. Biol.* **454**, 603–609 (1998)
40. Wilson, W.R., Hay, M.P.: Targeting hypoxia in cancer therapy. *Nat. Rev. Cancer* **11**, 393–410 (2011)
41. Wouters, B.G., Brown, J.M.: Cells at intermediate oxygen levels can be more important than the “hypoxic fraction” in determining tumor response to fractionated radiotherapy. *Radiat. Res.* **147**, 541–550 (1997)
42. Zeman, E.M., Brown, J.M., Lemmon, M.J., Hirst, V.K., Lee, W.W.: SR-4233: a new bioreductive agent with high selective toxicity for hypoxic mammalian cells. *Int. J. Radiat. Oncol. Biol. Phys.* **12**, 1239–1242 (1986)

Top-Down Multiscale Simulation of Tumor Response to Treatment in the Context of In Silico Oncology.

The Notion of Oncosimulator

Georgios Stamatakos

Abstract The aim of this chapter is to provide a brief introduction into the basics of a *top-down* multilevel tumor dynamics modeling method primarily based on discrete entity consideration and manipulation. The method is clinically oriented, one of its major goals being to support patient individualized treatment optimization through experimentation *in silico* (= on the computer). Therefore, modeling of the treatment response of clinical tumors lies at the epicenter of the approach. Macroscopic data, including i.a. anatomic and metabolic tomographic images of the tumor, provide the framework for the integration of data and mechanisms pertaining to lower and lower biocomplexity levels such as clinically approved cellular and molecular biomarkers. The method also provides a powerful framework for the investigation of multilevel (multiscale) tumor biology in the generic investigational context. The *Oncosimulator*, a multiscale physics and biomedical engineering concept and construct tightly associated with the method and currently undergoing clinical adaptation, optimization and validation, is also sketched. A brief outline of the approach is provided in natural language. Two specific models of tumor response to chemotherapeutic and radiotherapeutic schemes are briefly outlined and indicative results are presented in order to exemplify the application potential of the method. The chapter concludes with a discussion of several important aspects of the method including i.a. numerical analysis aspects, technological issues, model extensions and validation within the framework of actual running clinico-genomic trials. Future perspectives and challenges are also addressed.

G. Stamatakos (✉)

In Silico Oncology Group. Laboratory of Microwaves and Fiber Optics. Institute of Communication and Computer Systems. National Technical University of Athens. Iroon Polytechniou 9, Zografos 157 80, Greece

e-mail: gestam@central.ntua.gr

Web: www.in-silico-oncology.iccs.ntua.gr

A. d'Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_19, © Springer-Verlag Italia 2012

1 Introduction

The majority of cancer modeling techniques developed up to now adopt the straightforward *bottom-up* approach focusing on the better understanding and quantification of rather microscopic tumor dynamics mechanisms and the investigation of crucial biological entity interdependences including i.a. tumor response to treatment in the generic investigational context¹. To this end several combinations of mathematical concepts, entities and techniques have been developed and/or recruited and appropriately adapted. They include i.a. population dynamics models [27], cellular automata and hybrid techniques [16, 17, 22, 30, 53, 54, 59], agent based techniques [9], diffusion related continuous and finite mathematics treatments [5, 8, 18, 19, 34, 38, 57] etc. In addition, a number of large clinical tumor models focusing mainly on invasion and tumor growth morphology rather than on tumor response to concrete therapeutic schemes as administered in the clinical setting have appeared. Finite difference and finite element based solutions of the diffusion and classical mechanics equations constitute the core working tools of the corresponding techniques [6, 34, 57].

Nevertheless, a number of concrete and pragmatic clinical questions of importance cannot be dealt with neither by the bottom-up approach nor by the morphologically oriented large tumor growth models in a direct and efficient way. Two examples of such questions are the following [25]: “Can the response of the local tumor and the metastases to a given treatment be predicted in size and shape over time? What is the best treatment schedule for a patient regarding drugs, surgery, irradiation and their combination, dosage, time schedule and duration?” A promising modeling method designed with the primary aim of answering such questions is the *top-down* method developed by the In Silico Oncology Group (ISOG) [2, 3, 10–14, 29, 43–52]. Macroscopic data, including i.a. anatomic and metabolic tomographic images of the tumor, provide the framework for the integration of *available* and *clinically trusted* biological information pertaining to lower and lower biocomplexity levels such as *clinically approved* histological and molecular markers. However, the method does also provide a powerful framework for the investigation of multiscale tumor biology in the *generic investigational* context.

From the mathematical point of view the *top-down* simulation method presented is primarily a *discrete* mathematics method, although *continuous* mathematics (continuous functions, differential equations) are used in order to tackle specific aspects of the models such as pharmacokinetics and cell survival probabilities based on pharmacodynamical and radiobiological models. Adoption of the discrete approach as the core mathematical strategy of the method has been dictated by the obvious fact

¹ List of Abbreviations: *A* = apoptosis; *ACGT* = Advancing Clinico-Genomic Trials on Cancer; *AHF* = Accelerated Hyper-Fractionation; *BIR* = British Institute of Radiology; *C_{pav}* = Average Plasma Concentration; *CT* = Computerized Tomography; *EC* = European Commission, GBM = Glioblastoma Multiforme; *G₀* = dormant phase; *G₁* = Gap 1 phase of the cell cycle; *G₂* = Gap 2 phase; *HF* = Hyper-Fractionation, *IEEE* = Institute of Electrical and Electronics Engineers; *M* = Mitosis; *MRI* = Magnetic Resonance Imaging; *N* = necrosis; *PET* = Positron Emission Tomography; *R&D* = Research and Development; *RTOG* = Radiation Therapy Oncology Group; *S* = DNA synthesis phase; *TDS* = Time Delay in the S phase compartment; *TMZ* = Temozolomide.

that from the cancer treatment perspective it is the discrete i.e. the integer number of the usually few biological cells surviving treatment and their discrete mitotic potential categorization (stem cells, progenitor cells of various mitotic potential levels and differentiated cells) that really matter. These discrete entities and quantities in conjunction with their complex interdependences may give rise to tumor relapse or to ensure tumor control over a given time interval following completion of the treatment course. Cell cycle phases have a clearly discrete character too. Moreover, the properties of the different cell phases may vary immensely from the clinical significance perspective. A classical example is the lack of effect of cell cycle specific drugs on living tumor cells residing in the resting G0 phase.

It is noted that complex interdependencies of microscopic factors in the surroundings of the cells such as oxygenation, nutrient supply and molecular signals emitted by other cells play a critical role in the mitotic fate of tumor cells. Their effect is taken into account in the method through the local mean values of the corresponding model parameters. To this end imaging, histological and molecular data is exploited as will be described further down.

Due to the numerical character of the method a careful and realistically thorough numerical analysis concerning consistency, convergence and sensitivity/stability issues is absolutely necessary before any application is envisaged. A discussion of this critical issue is included in Sect. 5.

Tumor neovascularisation is taken into account in an indirect yet pragmatic way by exploiting grey level and/or color information contained within slices of tomographic imaging modalities sensitive to blood perfusion and/or the metabolic status of the tumor [12, 14, 33, 45, 50, 53, 54]. The reason for adopting the above mentioned strategy rather than developing or integrating detailed tumor angiogenesis models is that no microscopic information regarding the exact mesh of the neovascularization *capillaries* throughout the tumor can be currently extracted from clinically utilized imaging modalities. Nevertheless, the microscopic functional capillary *density distribution* over the tumor can be grossly estimated based on various imaging modalities such as T1 gadolinium enhanced MRI in the case of glioblastoma multiforme (GBM) and arterial spin labelling (ASL) MRI.

Precursors of the method can be traced in the well established and clinically applicable disciplines of pharmacology and radiobiology. Integration of molecular biology in the *top-down* method may be viewed as the introduction of a perturbator or adaptor of the cellular and higher biocomplexity level parameters. In such a way *in vivo measurable* clinical manifestation of tumor dynamics is placed the foreground. This is one of the reasons why the method is gaining wider and wider acceptance within the clinical and the industrial environment including the emergent domain of *in silico* oncology [24–26, 41, 42, 49, 50]. Both the large scale European Commission (EC) and Japan funded ACGT and p-medicine research and development (R&D) projects [1, 39] and the EC funded ContraCancrum R&D project [7] have adopted the *top-down* method as their core cancer simulation method. It is worth noting that in both projects the role of clinicians is prominent. A biomedical engineering concept and construct tightly associated with the method, the *Oncosimulator*, which is currently under clinical adaptation, optimization and validation is also sketched.

In order to convey the core philosophy of the method to the reader in a concise way a symbolic mathematical formulation of the *top-down* method in terms of a hypermatrix and discrete operators is introduced. Two specific models of tumor response to chemotherapeutic and radiotherapeutic schemes/schedules are briefly outlined so as to exemplify the method's application potential. The chapter concludes with a discussion of several critical aspects including numerical analysis, massive parallel code execution, associated technologies, extensions and validation within the framework of clinico-genomic trials and future challenges and perspectives.

A rather encouraging fact as far as industrial and eventually clinical translation of the method is concerned is that both the *top-down* method outlined and the *Oncosimulator* have been selected and endorsed by a worldwide leading medical technology company and now constitute modules of their research and development line [7]. One of the envisaged final products of this endeavor is a radiotherapy treatment planning system based on both physical and multiscale biological optimization of the spatiotemporal dose administration scheme. A clinical trial based validation process for the system is currently at the final stage of its detailed formulation.

2 The *Oncosimulator*

The *Oncosimulator* can be defined as a concept of multilevel integrative cancer biology, a complex algorithmic construct, a biomedical engineering system and (eventually in the future) a clinical tool which primarily aims at supporting the clinician in the process of optimizing cancer treatment in the patient individualized context through conducting experiments *in silico*. Additionally it is a platform for simulating, investigating, better understanding and exploring the natural phenomenon of cancer, supporting the design and interpretation of clinicogenomic trials and finally training doctors, researchers and interested patients alike [26, 47, 52]

A synoptic outline of the clinical utilization of a specific version of the *Oncosimulator*, as envisaged to take place following an eventually successful completion of its clinical adaptation, optimization and validation process, is provided in the form of the following seven steps (Fig. 1).

1st step: Obtain patient's individual multiscale and inhomogeneous data. Data sets to be collected for each patient include: clinical data (age, sex, weight etc.), eventual previous anti-tumor treatment history, imaging data (e.g. MRI, CT, PET etc images), histopathological data (e.g. detailed identification of the tumor type, grade and stage, histopathology slide images whenever biopsy is allowed and feasible etc.), molecular data (DNA array data, selected molecular marker values or statuses, serum markers etc.). It is noted that the last two data categories are extracted from biopsy material and/or body fluids.

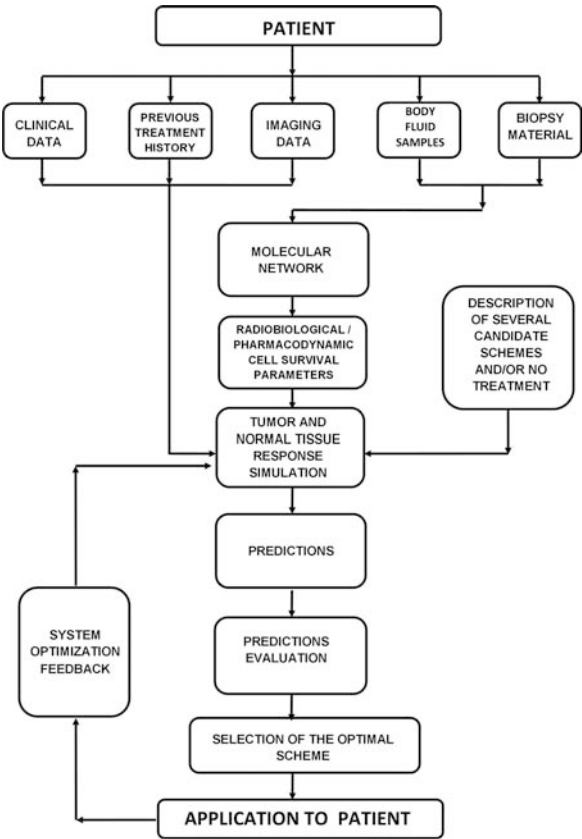


Fig. 1. The *Oncosimulator*: a gross workflow diagram

2nd step: Preprocess patient’s data. The data collected are preprocessed in order to take an adequate form allowing its introduction into the “Tumor and Normal Tissue Response Simulation” module of the *Oncosimulator*. For example the imaging data are segmented, interpolated, eventually fused and subsequently the anatomic entity/-ies of interest is/are three dimensionally reconstructed. This reconstruction will provide the framework for the integration of the rest of data and the execution of the simulation. In parallel the molecular data is processed via molecular interaction networks so as to perturb and individualize the average pharmacodynamic or radiobiological cell survival parameters. 3rd step: Describe one or more candidate therapeutic scheme(s) and/or schedule(s). The clinician describes a number of candidate therapeutic schemes and/or schedules and/or no treatment, obviously leading to free tumor growth, to be simulated *in silico* i.e. on the computer.

3rd step: Describe one or more candidate therapeutic scheme(s) and/or schedule(s). The clinician describes a number of candidate therapeutic schemes and/or schedules and/or no treatment, obviously leading to free tumor growth, to be simulated in silico i.e. on the computer.

4th step: Run the simulation. The computer code of tumor growth and treatment response is massively executed on distributed grid or cluster computing resources so that several candidate treatment schemes and/or schedules are simulated for numerous combinations of possible tumor parameter values in parallel (see Sect. 5 for detailed justification). Predictions concerning the toxicological compatibility of each candidate treatment scheme are also produced.

5th step: Visualize the predictions. The expected reaction of the tumor as well as toxicologically relevant side effect estimates for all scenarios simulated are visualized using several techniques ranging from simple graph plotting to four dimensional virtual reality rendering.

6th step: Evaluate the predictions and decide on the optimal scheme or schedule to be administered to the patient. The *Oncosimulator*'s predictions are carefully evaluated by the clinician by making use of their logic, medical education and even qualitative experience. If no serious discrepancies are detected, the predictions support the clinician in taking their final and expectedly optimal decision regarding the actual treatment to be administered to the patient.

7th step: Apply the theoretically optimal therapeutic scheme or schedule and further optimize the *Oncosimulator*. The expectedly optimal therapeutic scheme or schedule is administered to the patient. Subsequently, the predictions regarding the finally adopted and applied scheme or schedule are compared with the actual tumor course and a negative feedback signal is generated and used in order to optimize the *Oncosimulator*.

3 A Brief Outline of the Basics of the ISOG *Top-Down* Method

3.1 The Multilevel Matrix of the Anatomical Region of Interest

The anatomical region of interest, primarily including the tumor and possibly adjacent normal tissues and edema, in conjunction with its biological, physical and chemical dynamics is represented by a multilevel matrix i.e. a:

Matrix of (Matrices of (Matrices ... of (Scalars or Vectors or Matrices) ...)).

The multilevel matrix is created by a cubic discretization mesh which is virtually superimposed upon the anatomical region of interest. Biological cells residing within each geometrical cell of the mesh are conceptually clustered into mathematical equivalence classes. Equivalence classes primarily correspond to the various phases within or out of the cell cycle that a biological cell of the tumor can be found.

Since a tumor cell at any given instant also belongs to a mitotic potential category (stem, progenitor, terminally differentiated) the latter acts as a further partitioner of the biological cells into equivalence classes. One of the reasons, though not the single most important, for clustering biological cells into equivalence classes within each geometrical cell of the discretization mesh is computing resource limitations. Complex computational treatment of each single cell of a large clinical tumor undergoing therapeutic treatment as a *separate entity* is still not achievable within acceptable resource and time limits.

Discrete time is used. An important discretization aspect of the method is the mean time spent in the phase of an equivalence class by the biological cells belonging to the equivalence class [10, 13, 14, 44–46, 48, 50]. In order to allow for spatiotemporal perturbations of critical parameter values throughout the tumor and also avoid artificial cell synchronizations due to quantization, use of pseudo-random numbers is extensively made (Monte Carlo technique).

3.2 Practical Considerations Regarding the Construction of the Discretization Mesh

Collection of the appropriate mono-modality or far better multi-modality tomographic data of the patient such as slices of T1 weighted contrast enhanced MRI, T2 weighted MRI, CT, PET or other modalities, image segmentation, slice interpolation, three dimensional reconstruction of the anatomical entities of interest centered at the tumor, and eventually fusion of more than one modality images constitute the initial steps for the creation of the discretization mesh covering and discretizing the anatomical region of interest. Processed microscopic data (histological, molecular) are then utilized in order to enhance the patient individualization of the hypermatrix.

3.3 The Basics of the Top-Down Method

The multilevel matrix corresponding to the anatomical region of interest describes explicitly or implicitly the biological, physical and chemical dynamics of the region. The following parameters are used in order to identify a cluster of biological cells belonging to a given equivalence class within a geometrical cell of the mesh at a given time point:

- the spatial coordinates of the discrete points of the discretization mesh with spatial indices i, j, k , respectively. Each discrete spatial point lies at the center of a geometrical cell of the discretization mesh;
- the temporal coordinate of the discrete time point with temporal index l ;
- the mitotic potential category (i.e. stem or progenitor or terminally differentiated) of the biological cells with mitotic potential category index m ;
- the cell phase (within or out of the cell cycle) of the biological cells with cell phase index n . The following phases are considered: $\{G_1, S, G_2, M, G_0, A, N, D\}$ where G_1 denotes the G_1 cell cycle phase; S denotes the DNA synthesis phase; G_2 denotes the G_2 cell cycle phase; M denotes mitosis; G_0 denotes the quiescent

(dormant) G_0 phase; A denotes the apoptotic phase; N denotes the necrotic phase and D denotes the remnants of dead cells.

For the biological cells belonging to a given mitotic potential category AND residing in a given cell phase AND being accommodated within the geometrical cell whose center lies at a given spatial point AND being considered at a given time point – in other words for the biological cells clustered in the same equivalence class denoted by the index combination (i, j, k, l, m, n) – the following state parameters are provided:

- local oxygen and nutrient provision level. The following possible binary values of this parameter were initially considered: “oxygen and nutrient provision level *sufficient* for tumor cell proliferation” and “oxygen and nutrient provision level *insufficient* for tumor cell proliferation”. Obviously the binary character of the oxygen and nutrient provision level is to be considered only a first simplifying approximation. More elaborate descriptions have been proposed and applied [3, 13, 14, 43, 45, 50];
- number of biological cells;
- average time spent by the biological cells in the given phase;
- number of biological cells hit by treatment;
- number of biological cells not hit by treatment.

The initial constitution of the tumor i.e. its biological, physical and chemical state has to be estimated based on the *available* medical data through the application of pertinent algorithms [21, 32]. This state corresponds to the instant just before the start of the treatment course to be simulated.

The entire simulation can be viewed as the periodic and sequential application of a number of sets of algorithms on the multilevel matrix of the anatomical region of interest. Thus a stepwise multilevel matrix updating is achieved. Each *algorithm set application period* is equal to the time separating two consecutive complete scans of the discretization mesh. A complete scan includes mesh scans performed by all *algorithm sets* for any given time point. The *algorithm set application period* is usually taken 1 *h* since this is approximately the duration of mitosis, the shortest of the cell cycle phases. It should be noted that although the parameter values *exported* by the simulation execution at any desired instant for visualization and analysis purposes have a discrete character, certain parameters are handled by the computer *internally* and *temporarily* as real numbers (even with enhanced precision) in order to minimize quantization error propagation, in particular when dealing with small numbers of discrete entities in the stochastic context. By no means, however, does this technicality affect the fundamentally discrete character of the *top-down* method described.

The application of the algorithm sets on the multilevel matrix of the anatomic region of interest takes place in the following order:

1. time updating i.e. increasing time by a time unit (e.g. 1h);
2. estimation of the local oxygen and nutrient provision level;
3. estimation of the effect of treatment (therapy) referring mainly to cell hitting by treatment, cell killing and cell survival;

4. application of cell cycling, possibly perturbed by treatment. Transition between mitotic potential cell categories such as transition of the offspring of a terminally divided progenitor cell into the terminally differentiated cell category is also tackled by this algorithm set;
5. differential expansion or shrinkage or more generally geometry and mechanics handling;
6. updating the local oxygen and nutrient provision level following application of the rest of algorithm sets at each time step.

It is noted that the outcome of appropriate processing of the molecular and/or histopathological data via e.g. molecular networks and signaling pathways is used as a perturbator of the cell survival probabilities included in algorithm set “C” so as to considerably enhance patient individualization of the simulation. A realistic estimate of the extent of such perturbations for a given tumor type subclass in the framework of a clinico-genomic trial is achieved in a stepwise way. Initial rough modifications of the cell survival probabilities based on the baseline-pretreatment data, pertinent literature information and logic are subsequently corrected through utilization of the corresponding post treatment data via a process of parameter fitting. Obviously the above mentioned concepts and briefly outlined steps cannot convey all the details needed for the simulation to run. Their role is rather to identify and decompose the major conceptual mathematical and computational steps than to list all modeling details. The interested reader is referred to the website of the In Silico Oncology Group where they may find lists of pertinent publications providing detailed descriptions of several *top-down* multilevel cancer models including i.a. assumptions, mathematical treatment, numerical aspects such as convergence and quantization error minimization, sensitivity analysis, validation, applications and suggested extensions.

It is worth noting that discrete simulation under certain constraints can efficiently replace analytical solutions to a wide range of mathematical problems which, although being formulated in terms of continuous mathematics – usually including symbolically formulated differential equations – refer in fact to discrete physical quantities such as biological cells and cell state transition rates. Moreover, in many cases the *continuous* symbolic formulation of mathematical operators, such as the well known differential operator, when acting on discrete physical quantities can be readily replaced by a conceptually more straightforward *algorithmic* formulation. Several techniques leading to the minimization of error propagation for those cases where small numbers of discrete entities are dealt with by stochastic processes are available. The above generic policy has been extensively adopted in *top-down* multiscale models. For extensions of the method currently under implementation see Sect. 5.

4 Examples of Applications of the ISOG *Top-Down* Methods and Results

In this section two indicative models denoted by Model A and Model B are briefly outlined so as to exemplify the application potential of the method in the clinical context.

4.1 Model A: Tumor Response to Chemotherapeutic Schedules

Model A is a four dimensional, patient specific *top-down* simulation model of solid tumor response to chemotherapeutic treatment in vivo. The special case of imageable glioblastoma multiforme (GBM) treated by temozolomide (TMZ) has been addressed as a simulation paradigm. However, a considerable number of the involved algorithms are quite generic. The model is based on the patient's imaging, histopathologic and genetic data. For a given drug administration schedule lying within acceptable toxicity boundaries, the concentration of the prodrug and its metabolites within the tumor is calculated as a function of time based on the drug pharmacokinetics. A discretization mesh is superimposed upon the anatomical region of interest and within each geometrical cell of the mesh the basic *biological, physical* and *chemical laws* such as the rules concerning oxygen and nutrient provision, cell cycling [40], mechanical deformation etc. are applied at each discrete time point. The biological cell fates are predicted based on the drug pharmacodynamics [4, 20, 31, 35, 37, 56]. The outcome of the simulation is a prediction of the spatiotemporal activity of the entire tumor and is i.a. virtual reality visualized. A good qualitative agreement of the model's predictions with clinical experience [43, 44] supports the applicability of the approach. Model A has provided a basic platform for performing patient individualized *in silico* experiments as a means of chemotherapeutic treatment optimization in the *theoretical* context. A few indicative aspects of the model are outlined below. Since the complexity of the analysis is high the interested reader is referred to [43, 44] for a detailed description of the model. The work has also provided the basis for the development of the ACGT [1], p-medicine [39] and ContraCancrum [7] chemotherapy treatment response models. Fig. 2 depicts the simplified cytokinetic model of a tumor cell that has been proposed and adopted in Model A. The cytotoxicity produced by TMZ is primarily modeled by a delay in the S phase compartment (TDS) which is denoted by "Delay due to the effect of chemotherapy" in the diagram of Fig. 2 and by subsequent apoptosis. Further details are provided in the caption of Fig. 2. Fig. 3 provides a three dimensional visualization of the simulated response of a clinical GBM tumor to one cycle of the TMZ chemotherapeutic scheme: $150\text{mg}/\text{m}^2$ orally once daily for 5 consecutive days per 28-day treatment cycle [43]. Panel (a) shows the external surface of the tumor before the beginning of chemotherapy. Panel (b) shows the internal structure of the tumor before the beginning of chemotherapy. Panel (c) shows the predicted external surface of the tumor 20 days after the beginning of chemotherapy. Panel (d) shows the predicted internal structure of the tumor 20 days after the beginning of chemother-

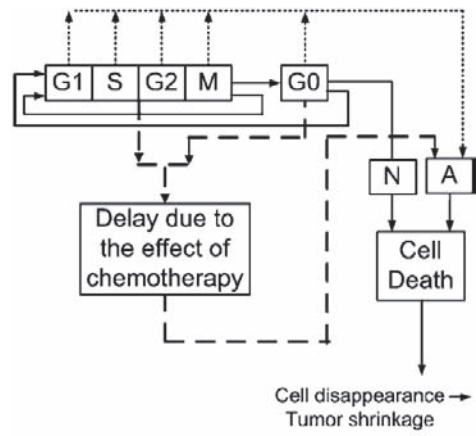


Fig. 2. Simplified cytokinetic model of a tumor cell proposed and adopted in model A. Symbols: G_1 : G_1 phase; S : DNA synthesis phase; G_2 : G_2 phase; G_0 : G_0 phase; N : necrosis; A : apoptosis. The cytotoxicity produced by TMZ is primarily modeled by a delay in the S phase compartment (TDS) (“Delay due to the effect of chemotherapy” in the diagram) and subsequent apoptosis. The delay box simply represents the time corresponding to at most two cell divisions being required before the emergence of temozolomide cytotoxicity. It is not a time interval additional to the times represented by the cell cycle phase boxes. See also [43]

apy. The pseudo-coloring criterion proposed and utilized is described in the caption of Fig. 3.

4.2 Model B: Tumor Response to Radiotherapeutic Schedules

Model B is a spatiotemporal simulation model of in vivo tumor growth and response to radiotherapy exemplified by the special case of imageable GBM treated by the treatment modality under consideration. The main constitutive processes of the model can be summarized as follows. A discretizing cubic mesh is superimposed upon a three dimensional virtual reconstruction of the tumor including its necrotic region and the surrounding anatomical features based on imaging data. In a way analogous to Model A, within each geometrical cell of the mesh a number of biological cell equivalence classes are defined based i.a. on the biological cell distribution over the various phases within or out of the cell cycle for the various mitotic potential categories. Sufficient registers are used in order to store the current state of each equivalence class such as the average time spent by clustered biological cells in phase G_1 etc. The mesh is scanned every one hour. The basic biological, physical and chemical “laws” or more precisely rules including the metabolic activity dynamics, cell cycling, mechanical and geometrical aspects, cell survival probability following irradiation with dose D [36, 55] are applied on each geometrical cell at each complete scan. A spatial and functional restructuring of the tumor takes place during each discrete time point since new biological cells are eventually produced, leading to differential tumor growth, or existing cells eventually die and subsequently

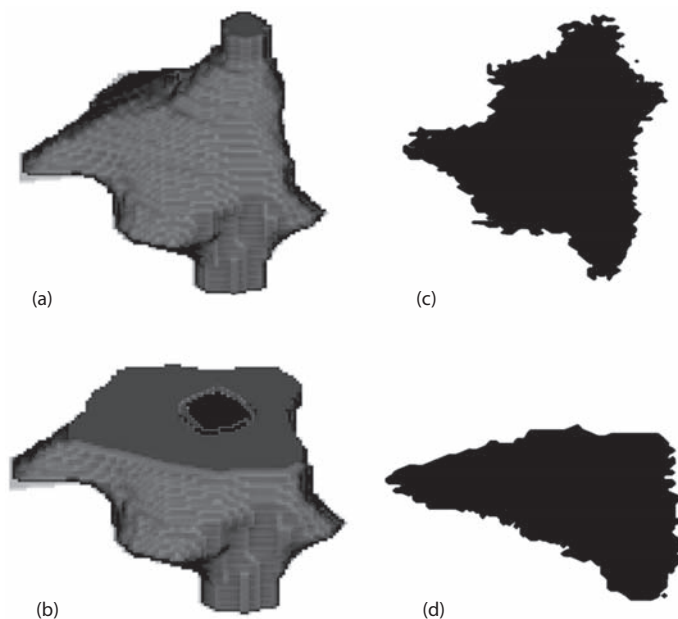


Fig. 3. A three dimensional visualization of the simulated response of a clinical GBM tumor to one cycle of the TMZ chemotherapeutic scheme: $150\text{mg}/\text{m}^2$ orally once daily for 5 consecutive days per 28-day treatment cycle. (a) external surface of the tumor before the beginning of chemotherapy, (b) internal structure of the tumor before the beginning of chemotherapy, (c) predicted external surface of the tumor 20 days after the beginning of chemotherapy (d) predicted internal structure of the tumor 20 days after the beginning of chemotherapy. The following pseudocolor code has been applied: light grey: outer tumor layer in (a) and (b), proliferating cell layer; dark grey: intermediate thin tumor layer in (a) and (b), dormant cell layer (G_0); black: dead cell layer. The “99.8%” pseudo-coloring criterion has been devised and applied as follows. “For a geometrical cell of the discretizing mesh, if the percentage of dead cells within it is lower than 99.8% then if *percentage of proliferating cells* > *percentage of G_0 cells*, then paint the geometrical cell red (proliferating cell layer), else paint the geometrical cell green (G_0 cell layer) else paint the geometrical cell blue (dead cell layer).” See also [43]

disappear through specific molecular and cellular event cascades, thus leading to differential tumor shrinkage. Simulation predictions can be two or three dimensionally visualized at any simulated instant of interest. In the particular model special attention has been paid to the influence of oxygenation on radiosensitivity in conjunction with the introduction of a refined imaging based description of the neovasculture density distribution. In order to validate the model two identical – except for the p53 gene status – virtual GMB tumors of large size, complex shape and complex internal necrotic region geometry were considered. The first one possessed a wild type p53 gene whereas the second one was characterized by a mutated p53 [14, 45]. The values of the α and β parameters of the standard linear quadratic radiobiological model for cell survival [36, 55] have been determined experimentally for the two cell lines considered [28]. Simulation predictions agree at least semi-

quantitatively with clinical experience and in particular with the outcome of the Radiation Therapy Oncology Group RTOG Study 83 02 [58]. The model allows for a quantitative study of the inter-relationship between the competing influences in a complex, dynamic tumor environment. Therefore, the model is already useful as an educational tool with which to theoretically study, understand and demonstrate the role of various parameters on tumor growth and response to irradiation. A long term quantitative clinical adaptation and validation of a considerably extended version of the model is in progress within the framework of the ContraCancrum project [7]. The long term goal is obviously integration into the clinical treatment planning procedure.

Fig. 4 shows simulation predictions corresponding to several branches of the RTOG Study 83 02. Upper panel provides the total number of proliferating and dormant tumor cells as a function of time for the hyperfractionated (1.2 Gy twice daily, 5 days per week to the dose of 72 Gy, “HF-72”) and accelerated hyperfractionated (1.6 Gy twice daily, 5 days per week to the dose of 48 Gy, “AHF-48”) radiotherapy schedules. All schemes start on the first day of the radiotherapy course. HF-72 is completed on day 40 after initiation of treatment whereas AHF-48 is completed on day 19. Lower panel of Fig. 4 shows the total number of proliferating and dormant tumor cells as a function of time for the hyperfractionated (1.2 Gy twice daily, 5 days per week to the dose of 64.8 Gy, “HF-64.8”) and accelerated hyperfractionated (1.6 Gy twice daily, 5 days per week to the dose of 48 Gy, “AHF-48”) radiotherapy schedules. Both irradiation schedules start on the first day of the first week of treatment. Both irradiation schedules start on the first day of the first week of treatment. HF-64.8 is completed on day 37 after initiation of treatment whereas AHF-48 is completed on day 19. According to the graphs, before completion of the AHF course cell kill due to AHF irradiation is more pronounced than cell kill induced by the HF scheme. This can be explained by the fact that a higher total dose has been administered to the tumor by the AHF scheme whereas for the period under consideration both schemes are characterized by the same time intervals between consecutive sessions. In case that not all living cells have been killed by AHF irradiation, tumor repopulation is considerable so that by the time the HF scheme is completed living tumor cells and their progeny which have escaped AHF irradiation outnumber tumor cells which have escaped HF irradiation. Improved tumor control following HF irradiation in comparison with the AHF scheme is in agreement with the conclusions of the clinical trial RTOG-83-02.

5 Discussion

Since the *top-down* multilevel method presented is a numerical method, a thorough convergence and sensitivity/stability analysis that includes the study of multiple parameter interdependences is necessary before any application is envisaged. Numerical analysis should satisfactorily cover at least those regions of the abstract parameter space that correspond to the envisaged applications. It is noted that of

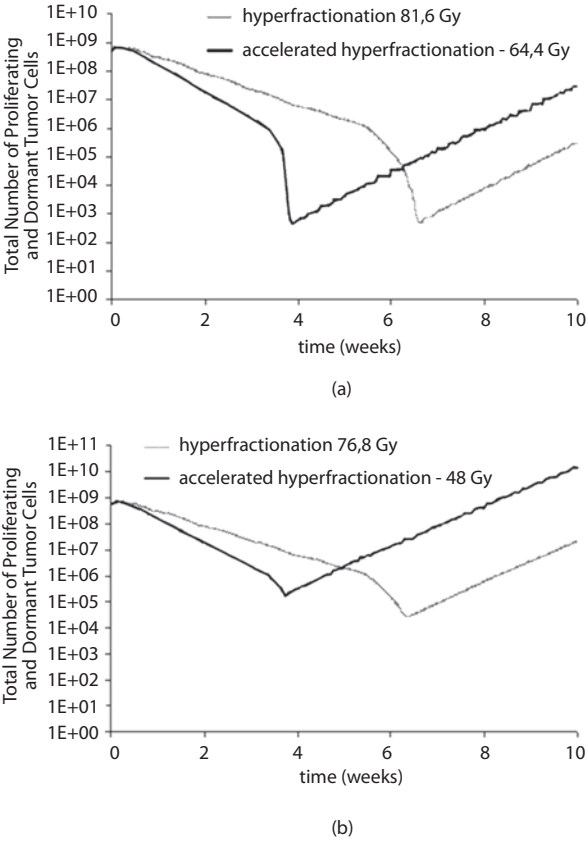


Fig. 4. Simulation predictions corresponding to several branches of the RTOG Study 83 02. (a) total number of proliferating and dormant tumour cells as a function of time for the hyperfractionated (1.2 Gy twice daily, 5 days per week to the dose of 72 Gy, “HF-72”) and accelerated hyperfractionated (1.6 Gy twice daily, 5 days per week to the dose of 48 Gy, “AHF-48”) radiotherapy schedules. HF-72 is completed on day 40 after initiation of treatment whereas AHF-48 is completed on day 19. (b) total number of proliferating and dormant tumour cells as a function of time for the hyperfractionated (1.2 Gy twice daily, 5 days per week to the dose of 64.8 Gy, “HF-64.8”) and accelerated hyperfractionated (1.6 Gy twice daily, 5 days per week to the dose of 48 Gy, “AHF-48”) radiotherapy schedules. HF-64.8 is completed on day 37 after initiation of treatment whereas AHF-48 is completed on day 19. Both irradiation schedules start on the first day of the first week of treatment. See also [45]

particular importance is the creation of the baseline tumor constitution by exploiting the relevant multilevel data available. Convergence of the tumor initialization has to be ensured. All of the above issues have been successfully addressed for specific tumor treatment cases such as breast cancer treated with epirubicin and nephroblastoma treated with vincristine and dactinomycin. The numerical behavior of the corresponding models has been checked through massive numerical experimenta-

tion. Concrete applicability intervals, restrictions and limitations have been identified (see Internal ACGT project reports and deliverables, and [21, 32]). Since the entire parameter space of the *top-down* models is rather large, numerical studies covering regions that correspond to further applications are in progress. Special attention is paid to the inherent relative biological instability of the cancer system itself when the model's stability is investigated.

It is well known that the values of critical parameters determining treatment outcome can vary considerably around what is assumed to be their population based average values. *Even after* incorporation of patient specific multiscale data into the simulation model no accurate evaluation of several critical model parameters is expected to be achieved. Moreover, as already mentioned, a tumor may behave as a relatively unstable system. Therefore, in order to compare candidate treatment schemes and/or schedules *in silico*, several possible combinations of parameter values lying around their apparently most probable estimates have to be constructed so as to cover the abstract parameter space as best as possible. Code executions have to be performed for all these selected parameter combinations. If for example the clinical question addressed is "Which out of the two candidate treatment schedules dented by I and II is the most promising for a given patient?", simulations have to run for both schedules I and II and for all parameter value combinations selected in the way briefly delineated above. If based on the simulation predictions schedule I outperforms schedule II for a sufficiently large percentage of the total parameter combinations considered, say 90%, then there is ground to suggest adoption of schedule I. Candidate scheme/schedule selection criteria are currently under formulation in tight collaboration with specialist clinicians within the framework of the ACGT [1], p-medicine [39] and ContraCancrum [7] *R&D* projects. Obviously the above sketched treatment optimization strategy dictates the need for a large number of parallel code executions on either cluster of grid platforms. This necessity has been addressed by specific actions of the previously mentioned projects.

Critical constraints imposed by toxicological limits of the treatment affected normal tissues should also be taken into account in order to judge whether or not a candidate scheme could be toxicologically acceptable. This issue may be addressed by exploiting the outcome of eventually relevant clinical trials and in particular of their phase I results. Ideally, direct multiscale spatiotemporal simulation of the effects of a given candidate scheme on specific normal tissues would provide quantitatively refined predictions. Nevertheless, due to the extremely high complexity of the homeostatic mechanisms governing normal tissue dynamics, the large number of normal tissue functional aspects and the potential induction of serious late effects by treatment such as radiotherapy, clinical translation of the second scenario seems to be a long term enterprise [2].

Since many solid tumors are microscopically inhomogeneous in space, the applications presented so far essentially make use of the *mean values* of certain biological parameters over each imaging based segmented sub-region of the tumor [2, 3, 11–14, 43–45, 49, 50, 52]. Small perturbations around these values are nevertheless implemented across each region through Monte Carlo simulation by the *top-down*

method presented. In the paradigmatic case of MRI T1 gadolinium enhanced imaging modality, strong grey level fluctuations over a tomographic slice can lead to an approximate delineation of the internal necrotic and the well neovascularized region of the tumor. Despite the fact that different values of certain parameters may be assigned to these two regions, sub-imaging scale inhomogeneities may still create spatial fluctuations of certain parameter values. In order to theoretically investigate the role of such biological inhomogeneities, pertaining for example to the genotypic and/or phenotypic tumor constitution, as well as the role of biochemical inhomogeneities of the extra tumoral environment such as acidity, necrosis exudate concentration etc the *top-down* basic platform can be still used provided that specific adaptations have taken place. Furthermore, tumor cell–tumor cell, tumor cell–host cell and tumor cell–local environment interactions in the microscopic setting can in principle be studied. In order to implement the above scenarios, the density of the discretization mesh should considerably increase, deeper level partitioning into more equivalence (sub-)classes has to be introduced into the multilevel matrix of the anatomical region of interest and the algorithm sets should be extended accordingly. However, such an approach dictates a sharp increase in computing memory and time demands and therefore tumor size must be kept small if restrictions in these resources apply, as is usually the case.

Following appropriate adaptation of specific modeling modules or equivalently algorithm sets such as the “C” set referring to cell killing etc. the *top-down* method outlined is in principle able to simulate tumors of any shape, size, geometry, macroscopic distribution of the metabolic or neovascularisation field, differentiation grade, spatial inhomogeneities, molecular profile and treatment scheme/schedule such as radiotherapeutic, chemotherapeutic, combined, new treatment modality etc. However, great care should be taken so that the model parameter values are estimated as best as possible based on real multilevel data. If such data is not available use of at least population based average parameter value estimates or qualitative experience based plausible values may be utilized only for generic exploratory reasons.

Hybridization of the *top-down* method presented with continuous and finite mathematics approaches such as diffusion based tumor growth modeling and detailed biomechanics is currently under implementation [7]. The aim of the task is to integrate into a *top-down* GBM model the microscopic tumor invasion process. The detailed biomechanics of the system calculated via a finite element module is also being integrated. Such a hybrid model is expected to be able to reproduce in relative detail both physical and biological aspects of tumor dynamics within the generic investigational framework. It should be noted that the non imageable diffusive component of GBM does play an important role in the development of the disease and therefore merits an in depth theoretical investigation. However, since the non imageable boundaries of GBM cannot be defined and monitored in a sufficiently objective way i.e. based on observational data such as clinically obtainable tomographic images, direct handling of the non imageable component by treatment planning systems in the patient individualized treatment context seems not to be a fully mature scenario as yet. Furthermore, by focusing on the imageable component within the treatment optimization context, one may argue that if for the imageable component

a candidate treatment scheme denoted by scheme I outperforms another candidate scheme denoted by scheme II *in silico*, the same would be true for the non imageable component of the tumor as well. From the *treatment perspective* again the main advantage of focusing on the imageable component, although this may represent even less than half of the total number of all viable tumor cells, is that this very component is amenable to relatively objective measurement *in vivo* and not only *post mortem*. Therefore, glioma dynamics models based on the imageable component are amenable to validation, at least in part, *in vivo*. Besides, incorporation of the immune system response to tumor [15] and simulation of the effects of antiangiogenic drugs on the tumor are two further scenarios, currently under investigation in the p-medicine *Oncosimulator* extension context [39].

Referring to the molecular level from the generic investigational standpoint, a large number of mechanisms, such as pathways leading to apoptosis or survival, that can be informed by available molecular data can be readily integrated into *top-down* models by applying the summarize and jump strategy of bio-data and bio-knowledge integration across bio-complexity scales [51]. This is in fact one of the actions currently taking place within the framework of the ContraCancrum project [7]. However, if the same biocomplexity level is viewed from the clinical perspective, care has to be taken so that only those characteristics and /or mechanisms whose predictive potential has been proved and established in the clinical setting – normally through clinical trials – may be incorporated into the models.

Regarding the envisaged clinical translation of *top-down* based models and systems, including the *Oncosimulator*, a sine qua non prerequisite is a systematic, formal and strict clinical validation. Designing the models so as to mimick actual clinical or far better clinico-genomic trials seems to be the optimal way to achieve this goal [23–26, 45, 47]. Therefore, involvement of clinicians in the model and system design and validation process should start at the *very beginning* of the endeavor [23, 24, 26]. Real clinicogenomic trials can provide invaluable multiscale data (imaging, histological, molecular, clinical, treatment) before, during and after a treatment course so as to best adapt and optimize the models and subsequently validate them. This was and is one of the core tasks of some international projects: ACGT [1], ContraCancrum *R&D* [7] and p-medicine [39]. Nephroblastoma and breast cancer are the solid tumor types addressed by ACGT and p-medicine, whereas gliomas and lung cancer are the ones addressed by ContraCancrum.

A further important challenge is to develop reliable, efficient, highly versatile and user friendly technological platforms which, following clinical adaptation, optimization and validation of the models would facilitate translation of *Oncosimulators* into the clinical practice so as to efficiently support, enhance and accelerate patient individualized treatment optimization. Advanced image processing, visualization and parallel code execution modules are but a few of the components necessary to achieve this goal. The ACGT, ContraCancrum and p-medicine *R&D* projects constitute exemplary initiatives towards this direction.

In summary both the *top-down* multilevel cancer simulation method briefly outlined above and the *Oncosimulator* have been designed so as to be readily optimizable, extensible and adaptable to changing clinical, biological, and research envi-

onments. Thus both entities, being primarily *multiscale physics* and *biomedical engineering* geared, have a *pragmatic* and *evolutionary* character.

Acknowledgements This work has been supported in part by the European Commission under the projects “ACGT: Advancing Clinicogenomic Trials on Cancer” (FP6-2005-IST-026996), ContraCancrum: Clinically Oriented Translational Cancer Multilevel Modelling” (FP7-ICT-2007-2-223979), TUMOR: Transatlantic Tumor Model Repositories (FP7-ICT-2009.5.4-247754) and p-Medicine: Personalized Medicine (FP7-ICT-2009.5.3-270089).

References

1. ACGT: Advancing Clinicogenomic Trials on Cancer: Open Grid Services for Improving Medical Knowledge Discovery. EC and Japan funded R&D project. (FP6-2005-IST-026996) <http://eu-acgt.org/acgt-for-you/researchers/in-silico-oncology/oncosimulator.html> and <http://www.eu-acgt.org/>
2. Antipas, V.P., Stamatakos, G.S., Uzunoglu, N.K.: A patient-specific in vivo tumor and normal tissue model for prediction of the response to radiotherapy: a computer simulation approach. *Meth. Inf. Med.* **46**, 367–375 (2007)
3. Antipas, V., Stamatakos, G. S., Uzunoglu, N.K., Dionysiou, D., Dale, R.: A spatiotemporal simulation model of the response of solid tumors to radiotherapy in vivo: parametric validation concerning oxygen enhancement ratio and cell cycle duration. *Phys. Med. Biol.* **49**, 1–20 (2004)
4. Bobola, M.S., Tseng, S.H., Blank, A., Berger, M.S., Silber, J.R.: Role of O6-methylguanine-DNA methyltransferase in resistance of human brain tumor cell lines to the clinically relevant methylating agents temozolomide and streptozotocin. *Clin. Cancer Res.* **2**, 735–741 (1996)
5. Breward, C.J., Byrne, H.M., Lewis, C.E.: A multiphase model describing vascular tumour growth. *Bull. Math. Biol.* **65**, 609–640 (2003)
6. Clatz, O., Sermesant, M., Bondiau, P.Y., Delingette, H., Warfield, S.K., Malandain, G., Ayache, N.: Realistic simulation of the 3-D growth of brain tumors in MR images coupling diffusion with biomechanical deformation. *IEEE Trans. Med. Imaging* **24**, 1334–1346 (2005)
7. ContraCancrum: Clinically Oriented Translational Cancer Multilevel Modelling. EC funded R&D project (FP7-ICT-2007-2- 223979). www.contracancrum.eu
8. Cristini, V., Frieboes, H.B., Gatenby, R., Caserta, S., Ferrari, M., Sinek, J.P.: Morphological instability and cancer invasion. *Clin. Cancer Res.* **11**, 6772–6779 (2005)
9. Deisboeck, T.S., Berens, M.E., Kansal, A.R., Torquato, S., Stemmer-Rachamimov, A.O., Chiocca, E.A.: Pattern of self-organization in tumour systems: Complex growth dynamics in a novel brain tumour spheroid model. *Cell Prolif.* **34**, 115–134 (2001)
10. Dionysiou, D.D., Stamatakos, G.S.: Applying a 4D multiscale in vivo tumor growth model to the exploration of radiotherapy scheduling: the effects of weekend treatment gaps and p53 gene status on the response of fast growing solid tumors. *Cancer Inform.* **2**, 113–121 (2006)
11. Dionysiou, D.D., Stamatakos, G.S., Gintides, D., Uzunoglu, N., Kyriaki K.: Critical parameters determining standard radiotherapy treatment outcome for glioblastoma multiforme: a computer simulation. *Open Biomed. Eng. J.* **2**, 43–51 (2008)
12. Dionysiou, D.D., Stamatakos, G.S., Marias, K.: Simulating cancer radiotherapy on a multi-level basis: biology, oncology and image processing. *Lect. Notes Comp. Sci.* **4561**, 569–575 (2007)
13. Dionysiou, D.D., Stamatakos, G.S., Uzunoglu, N.K., Nikita K.S.: A computer simulation of in vivo tumour growth and response to radiotherapy: New algorithms and parametric results. *Comp. Biol. Med.* **36**, 448–464 (2006)

14. Dionysiou, D.D., Stamatakis, G.S., Uzunoglu, N.K., Nikita, K.S., Marioli, A.: A four-dimensional simulation model of tumour response to radiotherapy in vivo: parametric validation considering radiosensitivity, genetic profile and fractionation. *J. Theor. Biol.* **230**, 1–20. (2004)
15. d’Onofrio, A.: A general framework for modeling tumor-immune system competition and immunotherapy: analysis and medical inferences. *Physica D* **208**, 220–235 (2005)
16. Duechting, W., Ulmer, W., Lehrig, R., Ginsberg, T., Dedeleit, E.: Computer simulation and modeling of tumor spheroid growth and their relevance for optimization of fractionated radiotherapy. *Strahlenther Onkol* **168**, 354–360 (1992)
17. Duechting, W., Vogelsaenger, T.: Three-dimensional pattern generation applied to spheroidal tumor growth in a nutrient medium. *Int. J. Biomed. Comput.* **12**, 377–392 (1981)
18. Enderling, H., Chaplain, M.A.J., Anderson, A.R.A., Vaidya, J.S.: A mathematical model of breast cancer development, local treatment and recurrence. *J. Theor. Biol.* **246**, 245–259 (2007)
19. Frieboes, H.B., Zheng, X., Sun, C.H., Tromberg, B., Gatenby, R., Cristini, V.: An integrated computational/experimental model of tumor invasion. *Cancer Res.* **66**, 1597–604 (2006)
20. FDA, Center for Drug Evaluation and Research. Application Number: 21029, Drug Name: Temozolomide (Temodal), New Drug Application, Clinical Pharmacology and Biopharmaceutics Reviews, (Revision date: 2/2/1999), pp. 19–20 (1999)
21. Georgiadi, E.C., Stamatakis, G.S., Graf, N.M., Kolokotroni, E.A., Dionysiou, D.D., Hoppe, A., Uzunoglu, N.K.: Multilevel cancer modeling in the clinical environment: simulating the behavior of Wilms tumor in the context of the SIOP 2001/GPOH clinical trial and the ACGT Project. In: *Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008)*, Athens, Greece, October 8–10 2008, Paper No. BE-2.1.2. (2008)
22. Ginsberg, T., Ulmer, W., Duechting, W.: Computer simulation of fractionated radiotherapy: further results and their relevance to percutaneous irradiation and brachytherapy. *Strahlenther Onkol* **169**, 304–310 (1993)
23. Graf, N., Desmedt, C., Buffa, F., Kafetzopoulos, D., Forgo, N., Kollek, R., Hoppe, A., Stamatakis, G., Tsiknakis, M.: Post-genomic clinical trials – the perspective of ACGT. *E-cancer Medical Science* **2**, (2008)
24. Graf, N., Desmedt, C., Hoppe, A., Tsiknakis, M., Dionysiou, D., Stamatakis, G.: Clinical requirements of “in silico oncology” as part of the integrated project ACGT (Advancing Clinico-Genomic Trials on Cancer). *Eur. J. Cancer Suppl.* **5**, 83 (2007)
25. Graf, N., Hoppe, A.: What are the expectations of a clinician from in silico oncology? In: Marias, K., Stamatakis, G.S. (eds.) *Proc. 2nd International Advanced Research Workshop on In Silico Oncology*, Kolymbari – Chania, Greece, September 25–26, 2006, pp. 36–38 (2006)
26. Graf, N., Hoppe, A., Georgiadi, E., Belleman, R., Desmedt, C., Dionysiou, D., Erdt, M., Jacques, J., Kolokotroni, E., Lunzer, A., Tsiknakis, M., Stamatakis, G.: “In silico oncology” for clinical decision making in the context of nephroblastoma. *Klin. Paediatr.* **221**, 141–149 (2009)
27. Guiot, C., Delsanto, P.P., Carpinteri, A., Pugno, N., Mansury, Y., Deisboeck, T.S.: The dynamic evolution of the power exponent in a universal growth model of tumors. *J. Theor. Biol.* **240**, 459–463 (2006)
28. Haas-Kogan, D.A., Yount, G., Haas, M., Levi, D., Kogan, S.S., Hu, L., Vidair, C., Deen, D.F., Dewey, W.C., Israel, M.A.: p53-dependent G1 arrest and p53 independent apoptosis influence the radiobiologic response of glioblastoma. *Int. J. Radiat. Oncol. Biol. Phys.* **36**, 95–103 (1995)
29. In Silico Oncology Group, Institute of Communications and Computer Systems, National Technical University of Athens. www.in-silico-oncology.iccs.ntua.gr
30. Kansal, A.R., Torquato, S., Harsh, G.R., Chiocca, E.A., Deisboeck, T.S.: Simulated brain tumour growth dynamics using a three-dimensional cellular automaton. *J. Theor. Biol.* **203**, 367–382 (2000)

31. Katzung, B.G. (ed.): Basic and Clinical Pharmacology, 8th Ed. Lange Medical Books, McGraw-Hill, New York (2001)
32. Kolokotroni, E.A., Stamatakos, G.S., Dionysiou, D.D., Georgiadi, E.C., Desmedt, C., Graf N.M.: Translating multiscale cancer models into clinical trials: simulating breast cancer tumor dynamics within the framework of the “Trial of Principle” clinical trial and the ACGT Project. In: Proc. 8th IEEE International Conference on Bioinformatics and Bioengineering (BIBE 2008), Athens, Greece, 8-10 October 2008, Paper No. BE-2.1.1. (2008)
33. Marias, K., Dionysiou, D., Stamatakos, G.S., Zacharopoulou, F., Georgiadi, E., Maris, T.G., Tollis I.: Multi-level analysis and information extraction considerations for validating 4D models of human function. *Lect. Notes Comput. Sci.* **4561**, 703–709 (2007)
34. Murray, J.D.: *Mathematical Biology II. Spatial Models and Biomedical Applications*. 3rd Edition, pp. 543–546, Springer, Heidelberg, (2003)
35. Newlands, E.S., Blackledge, G.R., Slack, J.A., Rustin, G.J., Smith, D.B., Stuart, N.S., Quartermann, C.P., Hoffman, R., Stevens, M.F., Brampton, M.H.: Phase I trial of temozolomide (CCRG 81045 *M&B* 39831 NSC 362856), *Br. J. Cancer* **65**, 287–291 (1992)
36. Perez, C., Brady, L.: *Principles and Practice of Radiation Oncology*, 3rd Ed.. Lippincott-Raven, Philadelphia. (1998)
37. Perry, M.C. (ed.): *The Chemotherapy Source Book*, 3rd Ed.. Lippincott Williams and Wilkins, Philadelphia (2001)
38. Ramis-Conde, I., Chaplain, M.A.J., Anderson, A.R.A.: Mathematical modelling of cancer cell invasion of tissue. *Math. Comput. Model.* **47**, 533–545 (2008)
39. P-medicine – From data sharing and integration via VPH models to personalized medicine. EC and Japan funded *R&D* project (FP6-2005-IST-026996). www.p-medicine.eu/
40. Salmon, S.E., Sartorelli, A.C.: Cancer chemotherapy. In: Katzung, B.G. (ed.), *Basic & Clinical Pharmacology*, pp. 923–1044. Lange Medical Books/McGraw-Hill, International Edition (2001)
41. Stamatakos, G.S.: Spotlight on cancer informatics. *Cancer Inform.* **2**, 83–86 (2006)
42. Stamatakos, G.S.: In silico oncology: a paradigm for clinically oriented living matter engineering. In: Stamatakos, G.S., Dionysiou, D. (eds.) *Proc. 3rd International Advanced Research Workshop on In Silico Oncology*, Istanbul, Turkey, September 23-24 2008, pp. 7–9. www.3rd-iarwiso.iccs.ntua.gr/procs.pdf (2008)
43. Stamatakos, G.S., Antipas, V.P., Uzunoglu, N.K.: A spatiotemporal, patient individualized simulation model of solid tumor response to chemotherapy in vivo: the paradigm of glioblastoma multiforme treated by temozolomide. *IEEE Trans. Biomed. Eng.* **53**, 1467–1477 (2006)
44. Stamatakos, G.S., Antipas, V.P., Uzunoglu, N.K.: Simulating chemotherapeutic schemes in the individualized treatment context: the paradigm of glioblastoma multiforme treated by temozolomide in vivo. *Comput. Biol. Med.* **36**, 1216–34 (2006)
45. Stamatakos, G.S., Antipas, V.P., Uzunoglu, N.K., Dale, R.G.: A four dimensional computer simulation model of the in vivo response to radiotherapy of glioblastoma multiforme: studies on the effect of clonogenic cell density. *Br. J. Radiol.* **79**, 389–400 (2006)
46. Stamatakos, G.S., Dionysiou, D.: Introduction of hypermatrix and operator notation into a discrete mathematics simulation model of malignant tumour response to therapeutic schemes in vivo. Some operator properties. *Cancer Inform.* **7**, 239–251 (2009)
47. Stamatakos, G.S., Dionysiou, D.D., Graf, N.M., Sofra, N.A., Desmedt, C., Hoppe, A., Uzunoglu, N., Tsiknakis, M. The Oncosimulator: a multilevel, clinically oriented simulation system of tumor growth and organism response to therapeutic schemes. Towards the clinical evaluation of in silico oncology. In: *Proc. 29th Annual Intern Conf IEEE EMBS*. Lyon, France, August 23-26 2007, *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, 6629–6632 (2007)

48. Stamatakos, G.S., Dionysiou, D.D., Nikita, K., Zamboglou, N., Baltas, D., Pissakas, G., Uzunoglu, N.: In vivo tumour growth and response to radiation therapy: a novel algorithmic description, *Int. J. Radiat. Oncol. Biol. Phys* **51**, 240 (2001)
49. Stamatakos, G.S., Dionysiou, D.D., Uzunoglu N.K.: In silico radiation oncology: a platform for understanding cancer behavior and optimizing radiation therapy treatment. In: Akay, M., (ed.) *Genomics and Proteomics Engineering in Medicine and Biology*, pp. 131–156, Wiley-IEEE Press, Hoboken (2007)
50. Stamatakos, G.S., Dionysiou, D.D., Zacharaki, E.I., Mouravliansky, N.A., Nikita, K.S., Uzunoglu N.K.: In silico radiation oncology: combining novel simulation algorithms with current visualization techniques. *Proc. IEEE. Special Issue on Bioinformatics: Advances and Challenges* **90**, 1764–1777 (2002)
51. Stamatakos, G.S., Kolokotroni, E., Dionysiou, D.D., Georgiadi, Giatili, S.: In silico oncology: a top-down multiscale simulator of cancer dynamics. Studying the effect of symmetric stem cell division on the cellular constitution of a tumour. In: Doessel, O., Schlegel, W.E. (eds.) *Proc. World Congress Medical Physics and Biomedical Engineering*, September 7–12 2012, Munich, Germany, pp. 1830–1833. Springer, Heidelberg (2009)
52. Stamatakos, G.S., Uzunoglu, N.K.: Computer simulation of tumour response to therapy. In: Nagl, S. (ed.) *Cancer Bioinformatics: from therapy design to treatment*, pp. 109–125. John Wiley & Sons Ltd, Chichester (2006)
53. Stamatakos, G.S., Zacharaki, E.I., Makropoulou, M.I., Mouravliansky, N.A., Marsh, A., Nikita, K.S., Uzunoglu, N.K.: Modeling tumor growth and irradiation response in vitro- a combination of high-performance computing and web based technologies including VRML visualization. *IEEE Trans. Inform. Technol. Biomed.* **5**, 279–289 (2001)
54. Stamatakos, G.S., Zacharaki, E.I., Makropoulou, M.I., Mouravliansky, N.A., Marsh, A., Nikita, K.S., Uzunoglu, N.K.: Tumor growth and response to irradiation in vitro: a technologically advanced simulation model. *Int. J. Radiat. Oncol. Biol. Phys.* **51**, Suppl. 1, 240–241 (2001)
55. Steel, G. (ed.): *Basic Clinical Radiobiology*, 3rd Ed.. Oxford University Press, Oxford (2002)
56. Stupp, R., Gander, M., Leyvraz, S., Newlands, E.: Current and future developments in the use of temozolomide for the treatment of brain tumours. *Lancet Oncol. Rev.* **2**, 552–560 (2001)
57. Swanson, K.R., Alvord, E.C., Murray, J.D.: Virtual brain tumours (gliomas) enhance the reality of medical imaging and highlight inadequacies of current therapy. *Br. J. Cancer* **86**, 14–18 (2002)
58. Werner-Wasik, M., Scott, C.B., Nelson, D.F., Gaspar, L.E., Murray, K.J., Fischbach, J.A., Nelson, J.S., Weinstein, A.S., Curran, W.J. Jr.: Curran. Final report of a phase I/II trial of hyperfractionated and accelerated hyperfractionated radiation therapy with carmustine for adults with supratentorial malignant gliomas. Radiation Oncology Therapy Group Study 83-02. *Cancer* **77**, 1535–43 (1996)
59. Zacharaki, E.I., Stamatakos, G.S., Nikita, K.S., Uzunoglu, N.K.: Simulating growth dynamics and radiation response of avascular tumour spheroid model validation in the case of an EMT6/Ro multicellular spheroid. *Comput. Meth. Progr. Biomed.* **76**, 193–206 (2004)

Challenges in the Integration of Flow Cytometry and Time-Lapse Live Cell Imaging Data Using a Cell Proliferation Model

Paolo Ubezio, Francesca Falcetta, and Monica Lupi

Abstract Multicellular systems are currently studied both *in vitro* and *in vivo* using different platforms, providing high throughput data of different types. Mathematical modelling is now called to interpret this reality and has to face more and more with quantitative data. This requires a connection between the basic theoretical model and the data structures, taking account of the processes of measure.

Working on the response to anticancer treatment, we considered the data provided by flow cytometry (FC) and time-lapse live cell imaging (TL) in time-course experiments *in vitro* with untreated and treated cell populations. We created a flexible cell cycle simulator including subsequent cell generations to achieve a full reconstruction *in silico* of the cell cycle progression under a variety of treatment effects. Unperturbed growth was modelled taking into account intercellular variability of G_1 , S and G_2M transit times, quiescent cells and natural cell loss. The effect of treatment was modelled by “perturbation modules” associated to each cell cycle phase and cell generation, containing a submodel of the checkpoint activity in that phase. Upon input of a set of parameters associated to unperturbed growth and perturbation modules, the program reproduced the time course of cell cycling through subsequent generations, providing outputs comparable with both TL and FC measures. The challenges to fit the data of specific experiments were discussed, indicating a feasible procedure for model building and identification. This lead to a dynamic rendering of proliferation midway between the macroscopic data level and the underlying molecular processes.

P. Ubezio (✉) · F. Falcetta · M. Lupi

Biophysics Unit, Department of Oncology, Institute for Pharmacological Research Mario Negri,
Via La Masa 19, 20156 Milan, Italy

e-mail: paolo.ubezio@marionegri.it, francesca.falcetta@marionegri.it, monica.lupi@marionegri.it

A. d’Onofrio, P. Cerrai, A. Gandolfi (Eds.), *New Challenges for Cancer Systems Biomedicine*
SIMAI Springer Series

DOI 10.1007/978-88-470-2571-4_20, © Springer-Verlag Italia 2012

1 Introduction

Cancer cell populations, according to the present biological understanding of cell proliferation, include cells in different cell cycle phases, cycling or quiescent [13, 14, 16]. Each cell experiences a lifespan of variable length, traversing in sequence G_1 , S and G_2M cell cycle phases, then divides and the subsequent generation of cells behaves in a similar way. At any time some death events naturally occur in a cell population. After an anticancer treatment, a series of events strongly perturbs the normal cell cycle flow. In these cases a plethora of possible responses can be played out by the individual cells of the population, with some cells dying, other being arrested or just delayed for a while. This usually occurs at specific points of the cell cycle, called checkpoints, where cell integrity is checked by internal molecular engines and defence reactions are activated to repair damages induced by treatments [5]. No direct way exists to measure the dynamics of the activity of each individual checkpoint, but we can observe their “macroscopic” consequences on the distribution of cells in the cell cycle phases by FC [7] or changes in division frequencies and distribution of intermitotic times by TL [6]. Time-course experiments usually show complex variations of the features observable by FC and TL, which cannot be interpreted only by the researcher’s intuition. In fact, the two platforms, considered singularly, convey a piece of the information, but are unable to suggest unequivocally the dynamics of the underlying cell cycle progression.

Up to now, no attempts have been made to integrate the two points of view. Our approach to the problem make use of a cell cycle simulator acting at the cell population level to reproduce in detail the flow of the cells under unperturbed and perturbed situations. The model is based on an age- and phase-structure (G_1 , S and G_2M phases) [2, 3] with drug effects superimposed, and uses only biologically sound parameters, like duration of cell cycle phases for unperturbed proliferation or modules miming the perturbations induced by each checkpoint, in terms of blocking, repair or death probability. Upon input of a set of parameters associated to unperturbed growth and perturbation modules, the program simulates the time course of cell cycling through subsequent generations, giving as output simulated data of both TL and FC experiments, which can be directly compared with the experimental databases. On the other side, we can tackle the inverse problem, to obtain the best fit scenario for the databases, avoiding over-parameterization by using biologically-consistent regularity assumptions on the time and dose dependence of the parameters. In the next sections we will show three successful examples of full rendering of cell proliferation obtained by contemporaneous fit of FC and TL data during exponential growth, approaching confluence and after an anticancer treatment.

2 Experiments and Data

2.1 Cell Culture and Treatment

Cancer cells grew in culture-wells with optimal nutrients' conditions. Experiments usually started one or two days after seeding, when cells have reached the exponential growth phase. A typical experiment [9, 10, 12, 18] involved the use of several wells, randomly assigned to each treatment group. The dose-range considered for treatments with drugs or other physical agents was based on the results of preliminary experiments, and it spanned from low to high efficacy. Then, we performed FC and TL observations on replicate samples at various times after treatment as shown in the next sections.

2.2 Flow Cytometry

We used different flow cytometric techniques to retrieve information on the distribution of the cells in the cycle (Fig. 1). Monoparametric flow cytometric analysis

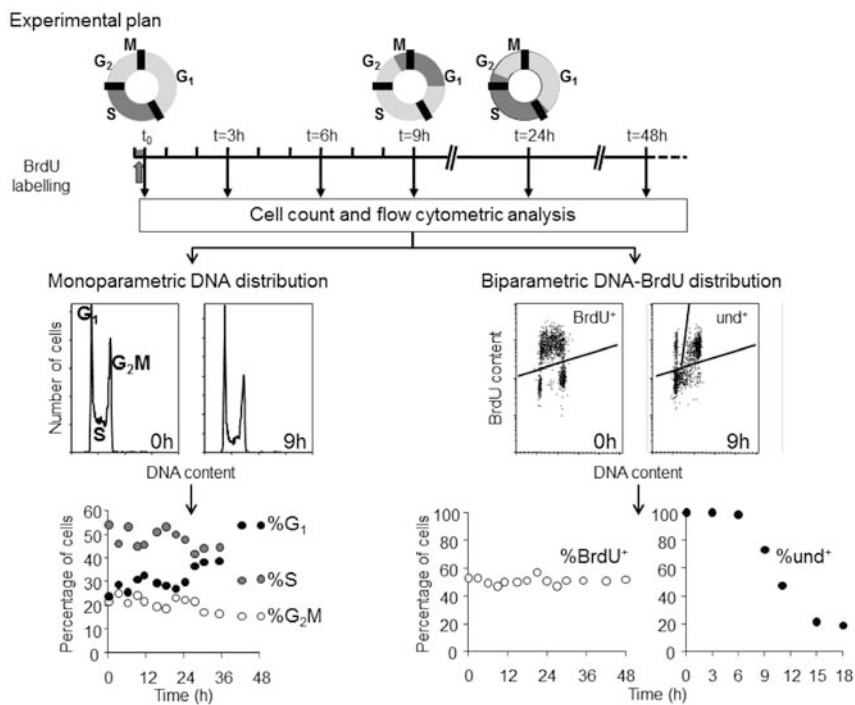


Fig. 1. Scheme of a typical FC experiment. It includes time-course measures of both DNA content (monoparametric histograms, left side) and DNA-BrdU (biparametric dot plots, right side). Percentages of cells in cell cycle phases and within BrdU subsets are calculated from mono- and biparametric histograms using conventional methods

of DNA content was performed staining a suspension of cells with a fluorescent probe, which binds stoichiometrically to DNA [17]. The fluorescence signal of individual cells was analysed by a flow cytometer and converted in an electronic pulse with amplitude proportional to the total fluorescent emission and thus to the cellular DNA content. At least 10000 cells for each sample were measured, obtaining the frequency distribution of DNA content. We calculated the percentage of cells within the major phases of cell cycle ($\%G_1$, $\%S$, $\%G_2M$), on the basis of the differences in DNA content between the cells in pre-replicative phase (G_1), those which have started DNA replication (S phase) and cells in the post-replicative plus mitotic phase (G_2M , with DNA content double than G_1 cells).

Short-term perturbations were additionally investigated by bromodeoxyuridine (BrdU) pulse-chase analysis [7]. In this technique BrdU was added to the cell culture for few minutes, then it was washed out and the cells were incubated in BrdU-free medium and analysed by FC. Incorporation of BrdU instead of thymidine during DNA synthesis enabled us to separate S-phase cells (“BrdU⁺” cells), from G_1/G_2M cells (“BrdU⁻”) at the time of labelling and follow them while exiting the original phases. Biparametric DNA-BrdU plots were then analysed to calculate the percentages of BrdU⁺ cells ($\%BrdU^+$) and undivided BrdU⁺ cells ($\%und^+$) as well as $\%G_1$, $\%S$, $\%G_2M$. Moreover, the percentages of BrdU⁺ cells within G_1 , S (or also SW, a given window within S phase) or G_2M could be calculated. With another BrdU protocol, cells were exposed to BrdU for a time long enough to assure that all cycling cells entered phase S and incorporated BrdU. In these “continuous BrdU exposure” experiments, $\%BrdU^-$ estimated the percentage of quiescent cells.

2.3 Time-Lapse Live Cell Imaging

TL enabled us to capture images of the culture-wells with a high frequency up to the end of the experiments (Fig. 2). Time-lapse microscopy movies were then analysed, tracking all cells in the fields of view (e.g. 50–100 cells per field) and their descendants [8]. Cells that were in the field at time zero (t_0) were defined as “generation 0” (gen0 or precursors), their descendants constitute gen1, gen2, etc. Hundreds of cell lineages were analysed for controls and each treatment group, recording the time and the kind of outcome of each cell (e.g. mitosis, death, cell loss due to exit from the field of view). These data were then mined, retrieving for each treatment group: i) the frequency distribution of intermitotic times (i.e. the cell cycle time (T_c) between subsequent mitosis); ii) the time course of the number of cells in each generation, relative to initial cell number and corrected to take account of cells that exit the field of view; iii) the frequency of the death events in each generation.

2.4 Data Inspection

Qualitative visual inspection of FC and TL data already allowed to draw a preliminary view of the effects of treatment. TL data indicated whether cell cycle delays and cell killing were detectable, at which doses and in which generation (but without

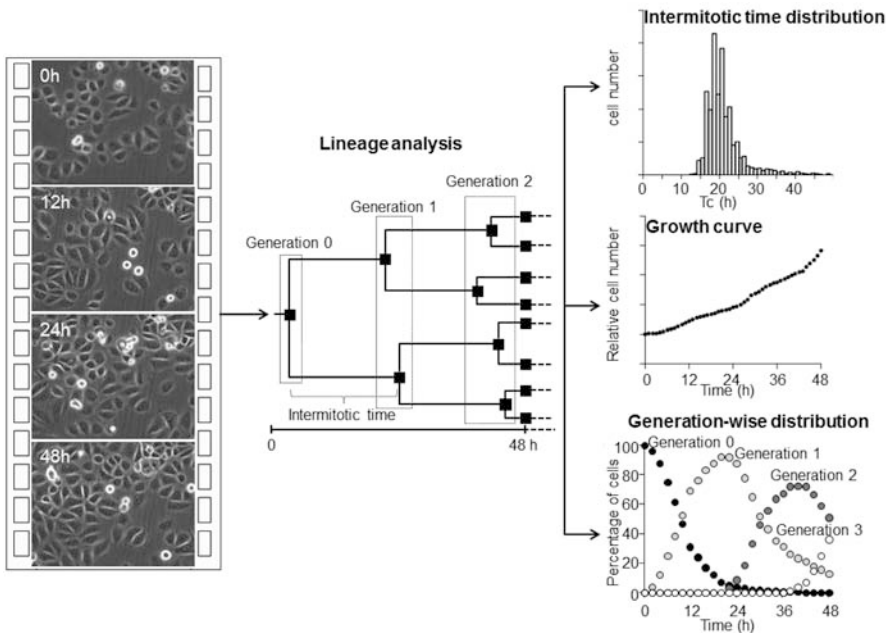


Fig. 2. Scheme of a typical TL experiment. From left to right: 1) acquisition (only representative images were shown, images were snapped every 20 minutes); 2) lineage analysis, tracking every single cell and its descendants through the whole movie; 3) final data, including distribution of intermitotic times and time courses of percentages of cells in each generation and of the overall cell number (growth curve relative to t0)

distinguishing cell cycle phases), and whether some sibling cells eventually refused, producing a subset of polyploid cells. FC data provided complementary information, suggesting the activation of blocks/delays in some phase(s), but without distinguishing different generations.

3 Modelling

Experimental evidences suggest that each cell population is characterised by high intercellular variability, even when growing in vitro. Cycling cells in different cell cycle phases or quiescent cells are contemporaneously present and some death events occur at any time [16]. Upon treatment, cell cycle arrest, repair or/and death were observed not only in the cells that were exposed to treatment (gen0), but also in their descendants. In our modelling approach, the rendering of the spontaneous proliferation of the cell population represents the first step. Then, cell cycling after treatment is rendered applying “perturbation modules” to the spontaneous proliferation in any phase and generation, miming either quiescence, cell cycle arrest or death phenomena at various levels of complexity.

3.1 Modelling Unperturbed Cell Cycling

Our modelling approach grounds on the theory of age-structured cell population models [2, 3, 20]. The basic single generation model is a simple discrete time and age-structured model [12, 17], where the time step Δt is sufficiently small compared to the time scale of variations of the available data (for our purposes, Δt was set to 30min). Cell cycle is divided in phases (G_1 , S and G_2M), each one divided in a suitable number of age-compartments of length Δt , the compartment “ k ” collecting cells with “age” (i.e. the time a cell has already spent in that phase) between $(k-1)\Delta t$ and $k\Delta t$. State variables are $G_1(k, t)$, $S(k, t)$, $G_2M(k, t)$, where $G_1(k, t)$ is the number of cells in the k^{th} compartment of G_1 , at time t , and similar definitions hold for $S(k, t)$ and $G_2M(k, t)$. The cells in each compartment represent a cohort of cells that have the same age (i.e. are at the same point of the cycle) at the same time.

The dynamics of cell cycling can be reproduced at different complexity levels. In the simplest model, with negligible intercellular variability, the cells of a cohort move together through the cycle, and when they arrive at the end of a phase they enter together at the beginning of the following one, until they reach the end of the cycle and divide. Then a new cohort with a double number of cells enters at the beginning of G_1 . In this case the dynamics depends only on three parameters: the durations of the phases $\{T_{G_1}, T_S, T_{G_2M}\}$.

3.1.1 Modelling Unperturbed Cell Cycling: Variable Phase Durations

At a higher complexity level we take into consideration that each cell spends a different time in a given phase, introducing a probability distribution $F_{ph}(k)$, defined as the probability that a cell completes a phase “ ph ” (G_1 , S or G_2M) at the age compartment k . For calculus purposes, the distribution is discretized, truncated at an appropriately long age $K_{ph}\Delta t$ (i.e. K_{ph} is the last compartment in a phase) and re-normalised. Fig. 3 shows the movement of the cells through the cycle with this model. Introducing the $\beta_{ph}(k)$ factor, defined as:

$$\beta_{ph}(k) = \frac{F_{ph}(k)}{1 - \sum_{j=1}^{k-1} F_{ph}(j)} \quad (1)$$

and giving the fraction of cells completing a phase at age k , among the cells that reach age $k-1$, the dynamics is described by the following balance equations:

$$G_1(1_{G_1}, t) = 2 \sum_{k=1}^{K_{G_2M}} G_2M(k, t - \Delta t) \beta_{G_2M}(k), \quad (2)$$

$$G_1(k+1, t) = G_1(k, t - \Delta t) [1 - \beta_{G_1}(k)], \quad \text{for } 1 \leq k \leq K_{G_1} - 1$$

$$S(1_S, t) = \sum_{k=1}^{K_{G_1}} G_1(k, t - \Delta t) \beta_{G_1}(k),$$

$$S(k+1, t) = S(k, t - \Delta t) [1 - \beta_S(k)], \quad \text{for } 1 \leq k \leq K_S - 1$$

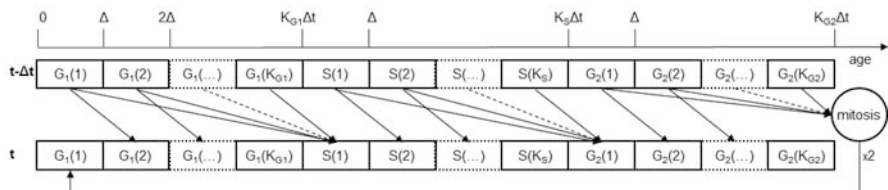


Fig. 3. Model of the flow through the cell cycle, with variable duration of phases. The arrows indicate the possible movements of cells in the cell cycle compartments in a step Δt . Cells may either progress to the subsequent age-compartment in a given phase or complete a phase entering the first compartment of the next phase

$$G_2M(1_{G2M}, t) = \sum_{k=1}^{K_S} S(k, t - \Delta t) \beta_S(k),$$

$$G_2M(k + 1, t) = G_2M(k, t - \Delta t) [1 - \beta_{G2M}(k)] \quad \text{for } 1 \leq k \leq K_{G2M} - 1.$$

The dynamics is defined upon choice of the three $F_{ph}(k)$ distributions, for which we normally adopt two parameters distributions, with a convenient choice of the six variable parameters such as $\{\bar{T}_{G1}, CV_{G1}, \bar{T}_S, CV_S, \bar{T}_{G2M}, CV_{G2M}\}$, i.e. the average and coefficient of variation of the phase durations. More complex distributions could be in principle adopted, but the data available are usually poorly sensitive to the shape of the distribution [19] and it would be impossible to identify such models.

3.1.2 Modelling Unperturbed Cell Cycling: Quiescence and Spontaneous Cell Death

Eqs. (2) do not consider quiescence and spontaneous cell death yet, the two other phenomena that are crucial for cell proliferation in both normal tissues *in vivo* (balancing cell cycling to maintain cellular homeostasis) and in tumours [15]. These are often not negligible even in untreated cell populations growing *in vitro* in optimal environmental situations. In our model, quiescent cells are localized in a distinct Q compartment in each phase, and the numbers of quiescent cells $G_1Q(t), SQ(t)$ and $G_2MQ(t)$ add to the list of the state variables. The probability that a cell entering G_1 , S or G_2M phase becomes quiescent is represented by p_{G1Q}, p_{SQ} and p_{G2MQ} respectively. Cell death among quiescent cells due to spontaneous death processes is included in the model by parameters $DRQ_{G1}, DRQ_S, DRQ_{G2M}$. We neglect cell death among cycling cells, assuming that cells stop cycling before dying. Eqs. (2) are accordingly modified:

$$G_1Q(t) = G_1Q(t - \Delta t)(1 - DRQ_{G1}) \quad (3)$$

$$+ p_{G1Q} 2 \sum_{k=1}^{K_{G2M}} G_2M(k, t - \Delta t) \beta_{G2M}(k)$$

$$G_1(1_{G_1}, t) = (1 - p_{G_1}Q) 2 \sum_{k=1}^{K_{G_2M}} G_2M(k, t - \Delta t) \beta_{G_2M}(k)$$

$$G_1(k + 1, t) = G_1(k, t - \Delta t) [1 - \beta_{G_1}(k)] \quad \text{for } 1 \leq k \leq K_{G_1} - 1$$

where similar equations for S and G_2M phases are omitted. The model includes now 12 independent parameters $\{\bar{T}_{G_1}, CV_{G_1}, \bar{T}_S, CV_S, \bar{T}_{G_2M}, CV_{G_2M}, p_{G_1}Q, p_SQ, p_{G_2M}Q, DRQ_{G_1}, DRQ_S, DRQ_{G_2M}\}$ and the system converges to an exponential growth with a steady-state percentage of cells in each compartment (including quiescence compartments). This asynchronous growth is univocally associated to a set of constant parameter values, as expected [2–4, 15].

3.1.3 Modelling Unperturbed Cell Cycling: A Generation-Wise Cell-Cycle Model

In the previous models new-born cells re-enter in the G_1 of the same cycle, eventually mixing with those of the previous generations. In order to include generation detail, required for connection with TL data, we consider a different cell cycle for each generation so that cells exiting G_2M of a generation enter G_1 of the following. The state variables of these “generation-wise” models are $G_1(k, t, gen_i)$, $S(k, t, gen_i)$ and $G_2M(k, t, gen_i)$, representing the number of cycling cells in the age compartment k of G_1 , S or G_2M of generation “ i ” at the time “ t ”, and $G_1Q(t, gen_i)$, $SQ(t, gen_i)$ and $G_2MQ(t, gen_i)$, representing the number of quiescent cells in each phase and generation. Eqs. (3) become:

$$G_1Q(t, gen_i) = G_1Q(t - \Delta t, gen_i)(1 - DRQ_{G_1}) \quad (4)$$

$$+ p_{G_1}Q 2 \sum_{k=1}^{K_{G_2M}} G_2M(k, t - \Delta t, gen_{i-1}) \beta_{G_2M}(k)$$

$$G_1(1_{G_1}, t, gen_i) = (1 - p_{G_1}Q) 2 \sum_{k=1}^{K_{G_2M}} G_2M(k, t - \Delta t, gen_{i-1}) \beta_{G_2M}(k)$$

$$G_1(k + 1, t, gen_i) = G_1(k, t - \Delta t, gen_i) [1 - \beta_{G_1}(k)]$$

$$\text{for } 1 \leq k \leq K_{G_1} - 1$$

$$SQ(t, gen_i) = SQ(t - \Delta t, gen_i)(1 - DRQ_S)$$

$$+ p_SQ \sum_{k=1}^{K_{G_1}} G_1(k, t - \Delta t, gen_i) \beta_{G_1}(k)$$

$$S(1_S, t, gen_i) = (1 - p_SQ) \sum_{k=1}^{K_{G_1}} G_1(k, t - \Delta t, gen_i) \beta_{G_1}(k)$$

$$\begin{aligned}
S(k+1, t, \text{gen}_i) &= S(k, t - \Delta t, \text{gen}_i) [1 - \beta_S(k)] \quad \text{for } 1 \leq k \leq K_S \\
G_2MQ(t, \text{gen}_i) &= G_2MQ(t - \Delta t, \text{gen}_i)(1 - DRQ_{G_2M}) \\
&\quad + p_{G_2M}Q \sum_{k=1}^{K_S} S(k, t - \Delta t, \text{gen}_i) \beta_S(k) \\
G_2M(1_{G_2M}, t, \text{gen}_i) &= (1 - p_{G_2M}Q) \sum_{k=1}^{K_S} S(k, t - \Delta t, \text{gen}_i) \beta_S(k) \\
G_2M(k+1, t, \text{gen}_i) &= G_2M(k, t - \Delta t, \text{gen}_i) [1 - \beta_{G_2M}(k)] \\
&\quad \text{for } 1 \leq k \leq K_{G_2M} - 1
\end{aligned}$$

where the terms containing gen_{i-1} should be omitted in the gen_0 equations.

3.1.4 Modelling Unperturbed Cell Cycling: Connecting the Model to Experimental Data

We developed a computer program with a friendly spreadsheet interface enabling the user to perform simulations and to compare the results with experimental data. Simulation intends to start at the laboratory time t_0 . In order to simulate the time evolution, according to the equations 4, the model requires the input of parameter values and the age distributions at t_0 , then cell cycling is rendered up to the desired end time. The values of the state variables at each time step allow calculating several kinds of simulated data. Simulated $\%G_1(t)$, $\%S(t)$ and $\%G_2M(t)$ are trivially calculated summing the cells of all compartments of the respective phase, including all represented generations.

By summing up the cells of all compartments of all phases in a given generation the program calculates $N_{\text{gen}_i}(t)$, the number of cells in generation i at time t , equivalent to the corresponding TL data. The time course of overall cell number ($N(t)$) is also calculated, summing the number of cells in all generations, and can be compared with experimental measures obtained by TL or with other techniques (e.g. Coulter counter).

Any starting distribution can be provided by the user, but the asynchronous distribution can be conveniently chosen whenever artificial “synchronization” methods are not applied in the experimental setting [9]. Actually any cell population, under constant environmental conditions, reaches an asynchronous exponential growth, and a good laboratory practice for data reproducibility requires that the cells are in such condition (or as near as possible) before initiating a treatment. In this condition the frequency distribution of cells in the cell cycle phases and ages is time-independent. A desynchronization routine of our program automatically calculates the asynchronous distribution with the desired precision, by running the cell cycling procedure from an approximated arbitrary initial distribution until the cell cycle percentages vary less than 0.1%. Then, the G_1 , S and G_2M age distributions are normalized in order to start the cell cycle simulation with a given number of cells at t_0 (e.g. 1000 cells), providing in this way $G_1(k, t_0, \text{gen}_0)$, $S(k, t_0, \text{gen}_0)$ and $G_2M(k, t_0, \text{gen}_0)$.

In order to simulate pulse-chase BrdU experiments, two sub-populations of cells are considered (i.e. BrdU⁻ cells and BrdU⁺ cells). The starting distribution is split assigning G₁ and G₂M cells to the BrdU⁻ cycle and S-phase cells to the BrdU⁺ cycle. Thereafter two cell cycle simulations proceed independently according to equations 4, and simulated percentages of cells in the BrdU⁻ and BrdU⁺ subsets are calculated and plotted with experimental counterparts. In particular, the time course of the percentage of undivided BrdU⁺ cells (%und⁺(*t*)) is obtained summing all the numbers of cells in all gen0 compartments of the BrdU⁺ cell cycle and dividing by the overall cell number.

Another valuable data type is the frequency distribution of intermitotic times ($F(Tc)$), measurable by TL. While the simulated average $\overline{Tc}$ is directly easily calculated in our model as $\overline{T}_{G1} + \overline{T}_S + \overline{T}_{G2M}$ and compared with the corresponding datum, a specific routine is designed to calculate the whole $F(Tc)$, on the basis of $F_{G1}(k)$, $F_S(k)$ and $F_{G2M}(k)$. Considering two phases, e.g. S and G₂M, the joint probability that a cells traverses S and G₂M with times k_S and k_{G2M} is $F_S(k_S) \times F_{G2M}(k_{G2M})$. Thus the probability that the duration of $S + G_2M$ is k_{SG2M} is given by the formula: $F(k_{SG2M}) = \sum F_S(k_S) F_{G2M}(k_{G2M})$ where the summation includes all combinations of k_S and k_{G2M} such that $k_S + k_{G2M} = k_{SG2M}$.

Then, $F(k_{G1SG2M})$ is calculated as $F(k_{G1SG2M}) = \sum F_{G1}(k_{G1}) F(k_{SG2M})$ where the summation includes all combinations of k_{G1} and k_{SG2M} such that $k_{G1} + k_{SG2M} = k_{G1SG2M}$.

Summarizing the simulation procedure, upon input of the values of the parameters, the program:

1. calculates $F_{ph}(k)$ and $\beta_{ph}(k)$;
2. performs the subroutine to produce the asynchronous starting cell distribution;
3. simulates the movement of the cells through the cell cycle in subsequent generations, calculating the time course of the state variables up to the desired end-time;
4. calculates the time course of all derived quantities (such as %G₁(*t*), %S(*t*) and %G₂M(*t*), %BrdU⁺(*t*), %und⁺(*t*), $\overline{Tc}$, $F(Tc)$, $Ngen_i(t)$ etc.) and updates their plots with the corresponding experimental data.

3.2 Rendering Unperturbed Cell Cycling: Practical Examples

We tackle in this section the inverse problem, attempting to find the values of the parameters that provide the best fit of data collected in an “unperturbed” (i.e. untreated) cell population. We consider first the case of an asynchronous exponential growth and then a non exponential growth. A model including all twelve available parameters may appear at first undue, but FC and TL experiments produced a large amount of information and their contemporary use imposes very stringent requirements, so that a simple model usually does not work. Thus we look for a compromise, attempting to find a model detailed enough to explain contemporaneously FC and TL data and to avoid overparametrization.

3.2.1 Rendering Unperturbed Cell Cycling: Exponential Growth

The first example concerns BxPC-3 cells, a human pancreatic adenocarcinoma. After data collection, the *in silico* rendering of the proliferation process was achieved in three steps.

Step 1. Model choice. By data inspection and trial simulations, we decided which were the main variable parameters to be optimized by fitting, and whether some phenomena had to be considered secondary, with the corresponding parameters fixed to reasonable values.

TL data indicated that about 10% of the cells either died or never divided in 72h. Because the average cell cycle time was 27h, it was reasonable to identify such cells as quiescent. FC experiments with continuous BrdU exposure were consistent with that figure, suggesting in addition that most quiescent cells were in G_1 phase. Because there was no experimental evidence suggesting a preferential death in a particular phase, or a different probability to become quiescent in S respect to G_2M , a first choice was to use just a single death parameter $DR (= DRQ_{G_1} = DRQ_S = DRQ_{G_2M})$ and two quiescence parameters $p_{G_1}Q$ and $p_{SG_2}Q (= p_SQ = p_{G_2}Q)$ and to fix them to values producing the observed order of magnitude of quiescent and dead cells. For what concerns CVs, trial simulations indicated that data were poorly sensitive to CV_S values distinct from CV_{G_2M} , thus we adopted a common variable parameter $CV_{SG_2M} (= CV_S = CV_{G_2M})$. In this way we selected a model with five variables $\{\bar{T}_{G_1}, CV_{G_1}, \bar{T}_S, \bar{T}_{G_2M}, CV_{SG_2M}\}$ and three fixed $\{p_{G_1}Q, p_{SG_2}Q, DR\}$ parameters.

Step 2. Fitting. The five variables were optimised using a nonlinear fitting program with random multistart, minimizing a negative log-likelihood function $(-ln(Lik))$, globally taking into account: i) steady state % G_1 , %S and % G_2M ; ii) time course of the overall cell number $N(t)$; iii) time course of the percentage of cells in each generation, and iv) the distribution of intermitotic times. A Gaussian error was considered with variance given by the experimental error of each procedure. We ran 50 fitting procedures from different starting values of the parameters, most of them converged in two minima with very close $-ln(Lik)$ values, with similar $\bar{T}_{G_1}, \bar{T}_S, \bar{T}_{G_2M}$ and different CVs (scenario A: $CV_{G_1} = 17\%$ and $CV_{SG_2M} = 62\%$; scenario B: $CV_{G_1} = 86\%$ and $CV_{SG_2M} = 30\%$). Both scenarios could fit the data in a very similar way, as shown in Fig. 4.

Step 3. Challenge concurrent solutions with additional experiments. Further experiments were planned to discriminate between the two concurrent scenarios emerging from the previous step. We made BrdU pulse-chase experiments with the samples harvested at specific times, suggested by the simulation to enable best discrimination between the concurrent scenarios. FC data and the simulated predictions with scenarios A and B are shown in Fig. 5, demonstrating that scenario B provided the best fit of the additional data.

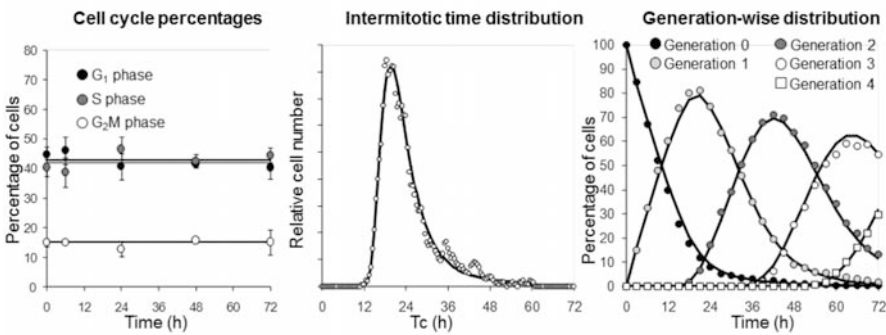


Fig. 4. Modelling exponential growth. BxPC-3 cells were analysed by FC and TL during exponential growth. Panels show data (symbols) and predictions from best fit scenarios (scenario B continuous line). Predictions from scenarios A and B were almost superimposable

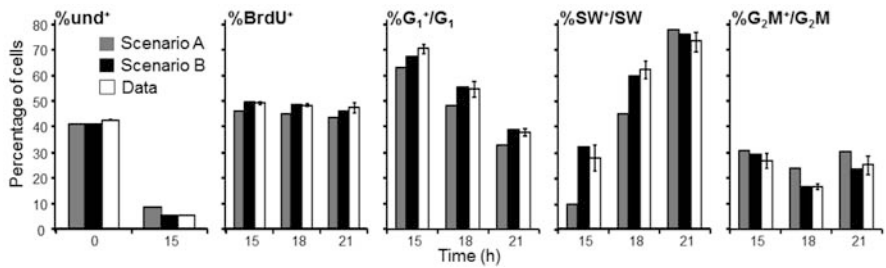


Fig. 5. Challenging concurrent scenarios in BxPC-3 cells. Panels show data of pulse-chase BrdU experiments and predictions from scenarios A and B. Only scenario B was consistent with all additional data

3.2.2 Rendering Unperturbed Cell Cycling: Non-Exponential Growth

Although the steady-state exponential growth was a reasonable approximation of the t_0 age distribution in most *in vitro* experiments, small departures from the steady-state were often observed in experiments where a treatment was applied, due to handling of the cells, e.g. when adding a drug at room temperature or taking it off by changing culture medium. These small perturbations resulted in a temporary delay of cell cycle progression detectable also in control cells, which as a rule were subjected to the same handlings of treated samples. We were able to model these disturbances – if necessary, e.g. to avoid misinterpretation of these delays as due to treatment – refining the fit of controls introducing gen0 delay parameters (see next section) and keeping fixed their values for treated samples.

More importantly, when observation extended over days, control cell populations approached confluence, as nutrients were consumed and/or space became crowded. During this process, the fraction of quiescent cells and death events increased, progressively shifting the equilibrium distribution reducing the fraction of cycling cells, so that the doubling time increased as well. Similar behaviour is usually observed also *in vivo*, although with time scales completely different, as the whole story of a

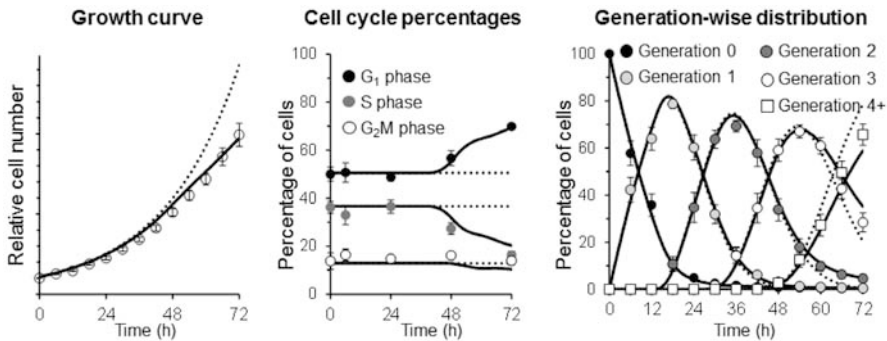


Fig. 6. Fitting the approach to confluence in Igrov-1 cells. Cell cycle percentages remained constant up to about 48h, then %G₁ increased and the doubling time lengthened. The data were fitted refining the exponential growth model (dotted line) introducing an increase of quiescence probabilities linked to the overall cell number (continuous line)

tumour in experimental mice typically lasts weeks and that of a human tumour years. In the framework of our modelling approach, it is possible to explain the decline of the growth rate in terms of the underlying changes of quiescence and death rates.

The experiment in Fig. 6 exemplifies the fitting procedure showing the approach to confluence in Igrov-1 cells. We started from a previously determined model of exponential growth, and we aimed to estimate the time variations of quiescence and death rates, best fitting FC and TL data. On a first instance, we tested whether this behaviour could be modelled increasing only the probability of quiescence in phase G₁ (p_{G1Q}). For this purpose, we connected the variation of p_{G1Q} with the crowding of cell population, assuming that p_{G1Q} departed from the steady state $p_{G1Q_{exp}}$ when the overall number of cells reached a threshold N_{thG1Q} according to the formula $p_{G1Q}(t) = p_{G1Q_{exp}} + (N(t) - N_{thG1Q})/N_{thG1Q}$, up to a maximum $p_{G1Q_{max}}$. We first optimized the values of N_{thG1Q} and $p_{G1Q_{max}}$, and then we refined the model introducing a smaller increase of p_{G2MQ} with the same procedure. The final model was able to fit reasonably well the approach to confluence of these samples (Fig. 6).

It should be noted that while the steady state exponential growth is well reproducible and the associated parameters can be considered cell line “characteristic”, the timing to confluence and the other small perturbations are more variable, depending on contingent, less controllable, environmental conditions, and the values of the parameters describing the deviation from the exponential growth should be refined for each experiment.

3.3 Modelling Treatment

Once a suitable model of proliferation for untreated samples was obtained, the effects of a treatment were modelled using parameters that “perturbed” the proliferation process. In order to limit the number of parameters, simple but biologically sound

perturbation modules were adopted, that can in principle be applied to any phase and generation to describe the action of the underlying cell cycle controls specific for that phase. Biological research is now disclosing the complex molecular networks that regulate the cell cycle and the response to a treatment insult [5], but there is no way to measure how much a specific network performs in his task to affect the cell cycle. In the model, the perturbing modules mime the overall activity (not the molecular interactions) of these networks and thus provide a kind of measure of their strength.

3.3.1 Delay Modules

The simplest effect of a treatment is a delay of the transit through a phase. To model this effect we used a module with a single parameter (Del_{ph}) affecting the progression through the age compartments in a given phase ph , ranging from 0 (no effect) to 1 (complete arrest). Del_{ph} was the fraction of cells that were not able to mature to the next age compartment, so that the equation for the number of cells in a generic compartment of a phase ph was modified as follows:

$$ph(k + 1, t, gen_i) = ph(k + 1, t - \Delta t, gen_i) Del_{ph} + (1 - Del_{ph}) ph(k, t - \Delta t, gen_i) [1 - \beta_{ph}(k)] \quad (5)$$

with straightforward modifications for the first compartment.

In the case of phase S, Del_S was equivalent to the fractional reduction of DNA synthesis rate, while Del_{G1} and Del_{G2M} represented a reduction of a generic “maturation rate” through the respective phases. As a consequence of the presence of delay, the $\bar{T}_{ph}$ became longer and a simple relationship was found between Del_{ph} and $\bar{T}_{ph}$: $\bar{T}_{ph} = \bar{T}_{phBasal} / (1 - Del_{ph})$, where $\bar{T}_{phBasal}$ was the average phase duration in the absence of delay.

By summing the average phase durations of all phases (with or without the presence of delay), it was possible to estimate the $\bar{T}_c$ for cells belonging to a given generation and to compare it with the direct measure obtainable by TL. Moreover, pulse-chase experiments provide estimates of T_S in BrdU⁺ cells in generation 0, enabling to calculate directly Del_S in this subset.

3.3.2 G₁ and G₂M Checkpoint Modules

Another typical response to treatment is given by G₁ and G₂M checkpoint dynamics. The current biological paradigm affirms that these checkpoints act arresting damaged cells in these phases to avoid progression in S and M where genomic errors could amplify and propagate to the next generations, and the molecular players within the respective networks have been identified. The module we designed reproduced this checkpoint dynamics, deviating part of the cells on exit G₁ and G₂M into specific “block” compartments. The fate of blocked cells was then either re-entering the cycle or die. The three parameters pBL_{ph} , Rec_{ph} and $DRBL_{ph}$ were respectively the fraction of cells that remained blocked and the rates of recycling and death within the subset of blocked cells. The corresponding equation for the compartment of blocked cells

was the following:

$$\begin{aligned} Bph(t, gen_i) = & (1 - Rec_{ph})(1 - DRBL_{ph})Bph(t - \Delta t, gen_i) \\ & + pBL_{ph} \sum_{k=1}^{K_{ph}} ph(k, t - \Delta t, gen_i) \beta_{ph}(k) \end{aligned} \quad (6)$$

while the equations of the quiescence and first compartments of each phase were accordingly modified, as exemplified for phase S:

$$\begin{aligned} SQ(t, gen_i) = & (1 - DRQ_S)SQ(t - \Delta t, gen_i) \\ & + p_S Q(1 - pBL_{G1}) \sum_{k=1}^{K_{G1}} G_1(k, t - \Delta t, gen_i) \beta_{G1}(k) \\ & + p_S Q Rec_{ph} (1 - DRBL_{G1}) B_{G1}(t - \Delta t, gen_i) \\ S(1_S, t, gen_i) = & (1 - p_S Q)(1 - pBL_{G1}) \sum_{k=1}^{K_{G1}} G_1(k, t - \Delta t, gen_i) \beta_{G1}(k) \\ & + (1 - p_S Q) Rec_{G1} (1 - DRBL_{G1}) B_{G1}(t - \Delta t, gen_i). \end{aligned} \quad (7)$$

A time-dependence of the parameters of these modules was optionally included, to render a delay between block and subsequent recycling or death rate, modelling the time dependence with a Hill function.

3.3.3 Cycling-Cells Death Module

An additional module was required to model immediate cell death possibly occurring at high doses without detectable previous cell cycle arrest. This was trivially obtained with a single death rate parameter (DR_{ph}), acting on all non-blocked cells in a phase. By definition, the module was used only in generation 0.

3.3.4 Polyploidization Modules

The polyploidization phenomenon was modelled introducing the parameter p_{pol} , representing the probability that new-born cells refuse together, and the corresponding death rate (DR_{pol}). Polyploid cells were localized in a new compartment and the corresponding equation for the number of polyploidy cells ($Pol(t, gen_i)$, for gen_1 and higher generations) was:

$$\begin{aligned} Pol(t, gen_i) = & Pol(t - \Delta t, gen_i)(1 - DR_{pol}) \\ & + p_{pol} Rec_{ph} (1 - DRBL_{G2M}) B_{G2M}(t - \Delta t, gen_{i-1}) \\ & + p_{pol} (1 - pBL_{G2M}) \sum_{k=1}^{K_{G2M}} G_2M(k, t - \Delta t, gen_{i-1}) \beta_{G2M}(k) \end{aligned} \quad (8)$$

with consequent changes of equations for $G_1 Q(t, gen_i)$ and $G_1(1_{G1}, t, gen_i)$.

3.4 *Building the Proliferation Model of a Perturbed Cell Population*

Within our model-building environment, the researcher selects the modules he wants to apply (if any) to each phase and each generation, possibly considering distinct values for BrdU⁺ and BrdU⁻ subsets. This opens an enormous number of possibilities that in principle can be explored in simulation studies, based on a given (theoretical or established) model of unperturbed proliferation of a cell population. On the other hand, wishing to select a suitable model to fit specific data sets, a judicious procedure of model building should be devised to avoid over-parameterization. The main steps of the procedure adopted in our lab are presented in the next subsections. They are conceptually similar to those shown in Sect. 3.2 for unperturbed proliferation, based on the principle of parsimony and going through subsequent refinements of the models when justified by the available data.

3.4.1 The Starting Model A

The first step was a thorough inspection of the available data, to establish a list of the phenomena for which there was objective direct evidence, possibly estimating some of the parameters, and those that could be excluded. Short term pulse-chase BrdU provided direct evidence of the phases affected by delays (or blocks) in generation 0. TL data provided direct measures of cell death and the average duration of the cell cycle in each generation. If the cell cycle duration and death events were equal to control values in a given generation, it would be possible to exclude any perturbation in that generation. In phases where an effect of treatment could not be excluded in this way, we applied a perturbation module at the lower complexity level, adopting only the delay parameter when cell death could be excluded, and three parameters otherwise (delay, block and death), using the block compartment only for definitively blocked cells. The dynamics of recycling was not taken into consideration at this first level, and cells that eventually divided were considered as delayed. In this way, it was possible to calculate the average cell cycle duration in each generation, and during the optimization procedure, this value was constrained to be close to the corresponding experimental $\bar{T}_c$ with a given tolerance.

However, the stronger limitation to the available model options was provided by experiments where the effects of different doses of treatment were measured. In these experimental plans, different treatments were made in parallel in independent samples and the same kind of FC and TL measures were retrieved in each dose group. This strongly improved the possibility to estimate the parameters, not only because of the increase of the experimental data but also because of constraints that could be applied to guarantee biologically feasible dose-dependence of the phenomena. For instance, if the presence of a dose-increasing effect is excluded in a high-dose treatment, it could be excluded also in the lower dose treatments.

At the end of this step, a tentative starting model (Model A) was established with the minimum number of modules required by the data of each dose.

3.4.2 Data Fitting with the Starting Model A

Model A was fitted to the data of each individual dose “ d ” minimizing a suitable objective function. We evaluated different candidate objective functions suitable for the experimental setting including time course of TL and FC measures, and then we adopted the negative log-likelihood given by the formula:

$$- \ln(Lik_d) = -\ln(Lik_{NTL,d}) - \ln(Lik_{\%G1SG2,d}) - \ln(Lik_{\%und+,d}) - \ln(Lik_{death,d}) \quad (9)$$

obtained summing the contributes of the time courses of the number of cells in each generation, of $\%G_1$, $\%S$ and $\%G_2M$, of $\%und^+$ and of the number of cell death events observed.

In multi-dose experiments, the presence of a limited number of minima of the objective function for each single-dose fits can be acceptable at this stage, otherwise a simpler model A should be considered. These alternative minima provided for each parameter one or more alternative “best fit” values, referred to as “Model A parameter values”.

3.4.3 Analysis of Dose Dependence of Single Parameters

The possible not uniqueness of the single dose fits was faced imposing dose-response regularities and eventually building a multi-dose model. With such model, we performed a multi-dose fit, constraining parameter values to follow specific dose-response functions and thus strongly reducing the alternatives for individual doses. This is like cross-validating the data obtained with one dose using the data obtained with all the other doses.

For this purpose, we first studied the dose-dependence of each parameter, based on Model A parameter values. Each parameter was then classified in one of the three following types of dose-dependence (excluding for biological regularity the existence of multiple maxima or minima):

1. constant with a dose threshold, chosen when the best fit parameter values were zero at the lower doses and the first not zero value did not vary at higher doses;
2. monotone. Parameters expected to vary monotonically with the dose (e.g. a block or a kill parameter expected to increase with the dose) were fitted with a Hill function:

$$Par(d) = Par_{min} + (Par_{max} - Par_{min}) \frac{d^{Par_m}}{Par_{D_{1/2}}^{Par_m} + d^{Par_m}} \quad (10)$$

where Par_{max} , Par_{min} , $Par_{D_{1/2}}$, Par_m coefficients represented the maximum, the minimum, the dose of half maximum effect and the sigmoidicity respectively;

3. non-monotone, when a parameter value increased at lower doses and diminished at higher ones. Typically this behaviour was expected for delay parameters that decreased at higher doses due to the competition of an irreversible block in the same phase. In order to provide flexibility to the fit of the dose-response with no

more than three coefficients, the following function (a scaled probability density function of a shifted gamma distribution with shape parameter equal to two) was found suitable:

$$Par(d) = Par_{max} \frac{d - Par_{Dmin}}{Par_{\vartheta}} e^{(1 - \frac{d - Par_{Dmin}}{Par_{\vartheta}})} \quad (11)$$

where $Par(d) = 0$ when $d \leq Par_{Dmin}$, Par_{max} was the maximum value of the function, Par_{Dmin} the threshold dose and Par_{ϑ} was proportional to both the mean and the width of the curve.

For each parameter, tentative values of dose-response coefficients were found by fitting with the appropriate functions the Model A parameter values, including all possible options emerging from the different minima. Thus, we obtained for each parameter one or more dose-response functions differing by one or more coefficients. Coefficients that did not change -with a given tolerance – when considering all alternative values, were taken as fixed in the following steps, while the other were left variable.

3.4.4 Multi-Dose Model A and Fitting

The dose-response functions (either constant, monotone or non-monotone) obtained in the previous step were embedded in a version of our simulation program that runs all doses contemporaneously using the whole TL and FC databases, constraining the parameters to the given dose-response functions. In this multi-dose model, the variable coefficients of the dose-response function of each parameter became the new variables to be fitted, minimizing a global merit function which was the sum of the negative log-likelihood of the fits of each dose. Non-linear fit of all the data of all doses together with such “multi-dose Model A”, taking into account all constraints, was expected to provide a unique multi-dose best fit scenario A of parameters values. In case of multiple solutions, the procedure should be repeated using simpler models, until an identifiable alternative to Model A is found.

3.4.5 Model Refinements

Starting from multi-dose scenario A, perturbation modules were progressively substituted, starting from gen0 modules, with more complex module types and tested whether the substitution improved the fit. For this purpose, we adopted a recursive two-step procedure. In the first step all the coefficients of the new module under test were fitted, keeping fixed all coefficients of the other modules. In the second step only the coefficients selected as variable were optimized together with the variables of all modules of the model. The procedure was repeated to obtain the best fit for the new model and decision on its acceptance was taken. Upon acceptance of the new model, it became the “running best fit model” and the reference for subsequent refinements. Eventually all modules were refined and the final model was obtained. Best fit with the final model provided the optimal dynamic rendering of cell cycling consistent with the available data.

3.5 Rendering Perturbed Cell Cycling: A Practical Example

Fig. 7 shows an example of a fitting with the strategy described in the previous section. Igrov-1 cells were treated while in exponential growth phase with low (L), intermediate (I) or high (H) doses of an anticancer agent. FC and TL data were obtained up to 72h after treatment. Controls of this experiment have been shown in Fig. 6. Following the first data analysis we initially worked with a model including perturbation modules in all cell cycle phases up to gen2, a unique polyploid model for gen1 and gen2 cells and a cycling-cell death module in G_1 in gen0. A higher number of death events than in control samples were observed only with the doses I and H, so that single-dose model A for dose L included delay parameters only, while I and H models included also pBL and $DRBL$ parameters in G_1 and G_2M . Following the fits with single-dose model A, a single minimum of the objective function was found with L and H while a few minima were found with I, providing alternative “Model A parameters values” for some parameters. For each parameter, the dose-response was studied, taking into account all possible alternative values, the type of dose-dependence was selected and tentative dose-response functions were provided, defining fixed (e.g. a dose threshold) and variable coefficients. A multi-dose model A was obtained with 40 variable coefficients and fitted to the whole data set, includ-

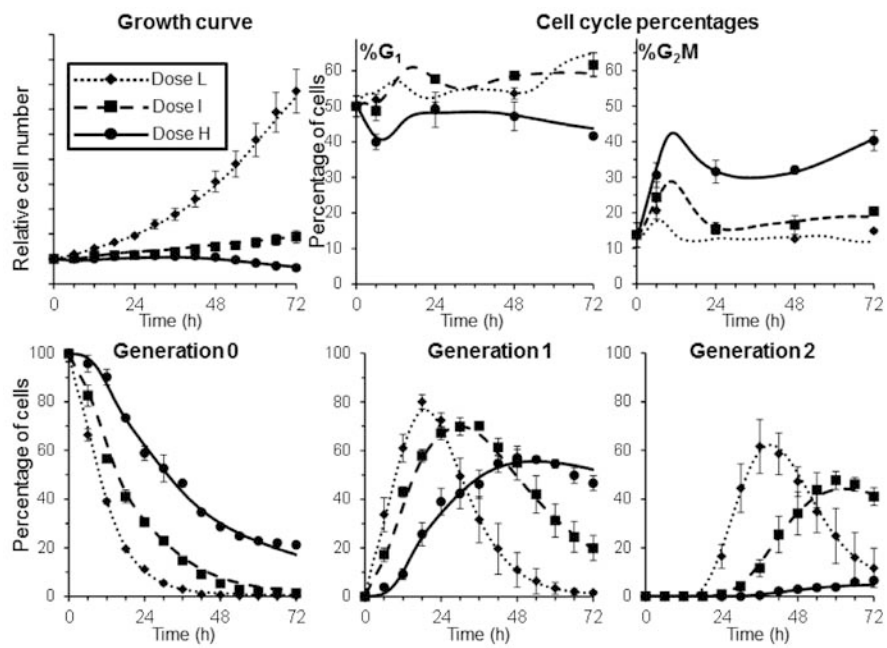


Fig. 7. Fitting the dose-response of an anticancer treatment. The panels show data (symbols) and model fit (line) for overall cell number, cell cycle percentages (%G₁ and G₂M, with %S omitted for clarity) and percentages of cells in gen0, gen1 and gen2, with three representative doses of treatment: L (low dose, diamond and dotted line), I (intermediate dose, squares and dashed line); H (high dose, circles and continuous line)

ing all doses, obtaining a unique best fit, catching the trend of the data of all doses. Then the single modules were refined, including proper checkpoint modules in G_1 and G_2M phases, with recycling instead of delay parameters and eventually a model with 33 variable coefficients was selected. The model converged to a unique best fit, in more than one hundred optimization procedures from random starting values, which well fitted the data as shown in Fig. 7.

4 Discussion

Mathematical modelling of biological or biomedical systems is not only directed to theoretical studies, but it is now facing the challenge of a quantitative rendering of the biological phenomena, taking into consideration the amount of data nowadays produced by experimental biomedicine. This should be particularly desirable to biologists, as they could pursue a quantitative comprehension of biological phenomena and a correct and more formal interpretation of the results of their measures, and to mathematicians, to build realistic models with appreciation of the processes of measure and their limitations.

New challenges emerge from a deep interaction of experimental and modelling approaches [1, 11], as the studies shown in this chapter exemplified. In particular, an equilibrated choice of the complexity level of the model becomes crucial, trying to find a rational compromise between the need to take into account all the phenomena demonstrated in play by the data and the necessity to avoid over-parameterization.

Working in the field of cancer research, we showed in this chapter the way we are currently tackling these problems to model proliferation in cell populations and we suggested a feasible general strategy possibly broadly adaptable to other fields where the dynamics of proliferation is investigated, e.g. to study differentiation chains or stem cell-based proliferative structures. The approach grounds on both experimental expertise (in our case in FC and TL measures) and availability of a flexible computational model of the process (in our case a very detailed simulation program of cell cycle proliferation). Our program provides a model-building environment, offering the user a wide choice of options to render the proliferation process at the scale of cell populations, in keeping with present day biological/pharmacological knowledge and reproducing the outputs of several experimental platforms (FC, TL, but also Coulter counter absolute cell counts and other growth inhibition assays). The model has a basic structure, able to render the dynamics of the “unperturbed” process (in our case: basic cell proliferation), on which perturbation modules apply to reproduce an external intervention (in our case: anticancer treatment). Once the model of unperturbed proliferation is established, it is possible to study how the system responds to a hypothetical treatment. Different types of modules have been implemented in the model, miming the activity of cell cycle controls at increasing complexity levels, and the user can select a specific effect and look at its consequences in simulated experiments. On the other hand one may wish to identify a model for a real treatment. This is the most challenging task, because the set of appropriate modules is not in general known in advance and a rational procedure for model building is required. In the past,

using only FC data combined with absolute cell counts, we set up reasonable models for a number of anticancer treatments with a trial-and-error procedure (databases available in [18]), describing a checkpoint activity as function of time and dose. By introducing TL data, the generation detail was considered, entailing a different view of the dynamics of checkpoint activities, distinguishing the effects directly observed in treated cells from those in their descendants. We integrated the cell-cycle view of FC data with the generation-wise perspective of TL using a computer reconstruction of the common underlying proliferation process fitting the two different kinds of data at the same time. In the attempt to achieve robust parameter estimation, we devised to use data from different treatment doses contemporaneously, performing multi-dose fitting.

This combination of data from different doses and platforms provided internal cross-validation enabling to satisfactorily identify models including several variable parameters. The adopted procedure first led to establish a relatively simple but identifiable starting model, that was then progressively cautiously refined, up to the definition of a final model, which fitted all data reasonably well. The final best fit reproduced the proliferation process through cell cycle and subsequent generations in an objective way with the maximum detail allowed by the data. This kind of *in silico* dynamic rendering of the process is far more informative than the usual data inspection and may achieve a quantitative measure of the activity of the checkpoint responses underlying the data, revealing the relative impact of each phenomenon in play. In addition, it will enable to test *in silico*, in simulated experiments, the consequences of impairment of a given checkpoint.

Further theoretical studies could be undertaken for each step or subproblem of the optimization procedure. This could lead to a more formal approach. However the dependence on specific data sets of all steps of the procedure makes improbable, in our opinion, the definition of an “all-purpose” automatic model building routine. The method shown here suggests that the way to solve problems with complex models and complex data structures, might not be exclusively “mathematical” or “experimental”, but will emerge from a strict interaction of the two in each step of the procedure.

Acknowledgements The generous contribution of the Italian Association for Cancer Research (AIRC), Milan, Italy, is gratefully acknowledged.

References

1. Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprpto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., Rout, M.P., Sali, A.: Determining the architectures of macromolecular assemblies. *Nature*. **450**, 683–94. (2007)
2. Arino, O.: A survey of structured cell population dynamics. *Acta Biotheor.* **43**, 3–25 (1995)
3. Basse, B., Ubezio, P.: A generalised age and phase structured model of human tumour cell populations both unperturbed and exposed to a range of cancer therapies. *Bull. Math. Biol.* **69**, 1673–1690 (2007)

4. Chiorino, G., Metz, J.A., Tomasoni, D., Ubezio, P.: Desynchronization rate in cell populations: mathematical modeling and experimental data. *J. Theor. Biol.* **208**, 185–199 (2001)
5. Darzynkiewicz, Z., Traganos, F., Wlodkowic, D.: Impaired DNA damage response – An Achilles' heel sensitizing cancer to chemotherapy and radiotherapy *Eur. J. Pharmacol.* **625**, 143–150 (2009)
6. Forrester, H.B., Vidair, C.A., Albright, N., Ling, C.C., Dewey, W.C.: Using computerized video time lapse for quantifying cell death of X-irradiated rat embryo cells transfected with c-myc or c-Ha-ras. *Cancer. Res.* **59**, 931–939 (1999)
7. Gray, J.W., Dolbeare, F., Pallavicini, M.G., Beisker, W., Waldman, F.: Cell cycle analysis using flow cytometry. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.* **49**, 237–255 (1986)
8. Khan, I.A., Lupi, M., Campbell, L., Chappell, S.C., Brown, M.R., Wiltshire, M., Smith, P.J., Ubezio, P., Errington, R.J.: Interoperability of Time Series Cytometric Data: A Cross Platform Approach for Modeling Tumor Heterogeneity. *Cytometry A* **79A**, 214–226 (2011)
9. Lupi, M., Matera, G., Branduardi, D., D'Incalci, M., Ubezio, P.: Cytostatic and cytotoxic effects of topotecan decoded by a novel mathematical simulation approach. *Cancer Res.* **64**, 2825–2832 (2004)
10. Lupi, M., Matera, G., Natoli, C., Colombo, V., Ubezio, P.: The Contribution of p53 in the Dynamics of Cell Cycle Response to DNA Damage Interpreted by a Mathematical Model. *Cell Cycle* **6**, 943–950 (2007)
11. Mertens, H.D.T., Svergun, D.I.: Structural characterization of proteins and complexes using small-angle X-ray solution scattering. *J. Struct. Biol.* **172**, 128–141 (2010)
12. Montalenti, F., Sena, G., Cappella, P., Ubezio, P.: Simulating cancer-cell kinetics after drug treatment: Application to cisplatin on ovarian carcinoma. *Phys. Rev. E* **57**, 5877–5887 (1998)
13. Nurse, P.: A long twentieth century of the cell cycle and beyond. *Cell* **199**, 71–78 (2000)
14. Nurse, P., Masui, Y., Hartwell, L.: Understanding the cell cycle. *Nat. Med.* **4**, 1103–1106 (1998)
15. Spinelli, L., Torricelli, A., Ubezio, P., Basse, B.: Modelling the balance between quiescence and cell death in normal and tumour cell populations. *Math. Biosciences.* **202**, 349–370 (2006)
16. Steel, G.G.: *Growth Kinetics of Tumors: Cell Population Kinetics in Relation to the Growth and Treatment of Cancer*. Oxford University Press (1977)
17. Ubezio, P.: Unravelling the complexity of cell cycle effects of anticancer drugs in cell populations. *Discrete and Continuous Dynamical Systems-Series B* **4**, 323–335 (2004)
18. Ubezio, P., Lupi, M., Branduardi, D., Cappella, P., Cavallini, E., Colombo, V., Matera, G., Natoli, C., Tomasoni, D., D'Incalci, M.: Quantitative assessment of the complex dynamics of G_1 , S and G_2M checkpoint activities. *Cancer Res.* **69**, 5234–5240 (2009)
19. Ubezio, P., Rossotti, A.: Sensitivity of flow cytometric data to variations in cell cycle parameters. *Cell Tissue Kinet.* **20**, 507–517 (1987)
20. Webb, G.: *Theory of Nonlinear Age-Dependent Population Dynamics*. Marcel Dekker, New York (1985)