

Mathematical Models for Cell Migration and Proliferation Informed by Visualisation of the Cell Cycle

Sean T. Vittadello

B.Sc. (Hons) (Chemistry)

M.Sc.Stud. (Pure mathematics)

Ph.D. (Pure mathematics)

Principal supervisor: Professor Matthew Simpson
Queensland University of Technology

Associate supervisor: Professor Scott McCue
Queensland University of Technology

External supervisor: Professor Nikolas Haass
The University of Queensland

School of Mathematical Sciences
Faculty of Science and Engineering
Queensland University of Technology



2020

Submitted in fulfilment of the requirement for the degree of Doctor of Philosophy

Dedication

*This work is dedicated to Fiona, Lucius, and Felix,
for their unwavering support and encouragement.*

Migration and proliferation are fundamental functions of cells, and are essential for normal physiological processes including development, tissue renewal and repair, and immune surveillance. The misregulation of migration and proliferation contributes to many pathological processes, most notably cancer. Improving the understanding of cell migration and proliferation is therefore of great importance. In this thesis we develop and analyse new mathematical models of cell migration and proliferation, based on new experimental studies using cell cycle indicators to visualise progression through the cell cycle. Key outcomes include: new mathematical models that can describe both cell migration and cell cycle dynamics; the application of a multi-stage model to reconcile the normally-hidden cell synchronisation of subpopulations with the apparent exponential growth of the total cell population; a rigorous examination of the go-or-grow hypothesis using melanoma cells with cell-cycle-inhibiting drugs; and, a novel mathematical model for heterogeneous cell proliferation, where the total population consists of a subpopulation of slow-proliferating cells and a subpopulation of fast-proliferating cells, and switching between proliferative states occurs by the cellular processes of asymmetric cell division and influence from surrounding cells.

Statement of Original Authorship

The work contained in this thesis has not been previously submitted to meet requirements for an award at this or any other higher education institution. To the best of my knowledge and belief, the thesis contains no material previously published or written by another person except where due reference is made.

Signature: _____

Date: _____

Acknowledgements

I would like to express my sincere appreciation to Professor Mat Simpson, my principal supervisor, for providing me with the opportunity to pursue a Ph.D. under his supervision, for creating the opportunity to collaborate with experimental researchers Professor Nikolas Haass and Gency Gunasingh, and for sharing his formidable knowledge of mathematical biology in particular and applied mathematics more generally. At the very beginning of these studies I mentioned my interest in the integration of mathematics and nature. Little did I know the extent to which I would be able to pursue this interest by working as both a mathematician and an experimentalist. His ardour for mathematical biology is a significant influence on my motivation to continue to work at the interface of mathematics and biology.

Professor Scott McCue, my associate supervisor, always provided a unique and clear perspective on our work, which was essential for combining our mathematical and biological perspectives into a cohesive whole. I am very appreciative of his guidance and encouragement, and I have enjoyed our spirited discussions on mathematical matters.

Professor Nikolas Haass, my external supervisor, warmly welcomed me into his lab so that I felt like a regular member of his team. His encouragement, generosity with time, which yielded lengthy discussions on multiple topics on occasion, and his recognition of the effectiveness of mathematics in addressing questions in biology, are most appreciated.

Gency Gunasingh was tasked to teach me to independently perform some experiments in a cell culture laboratory. The fact that I obtained so much useful and interesting data is unnecessary proof of her impressive skills as an experimental scientist. No matter what issue arose in the lab, her ability to creatively solve problems ensured that my progress was never arrested for long. For all this, I have much appreciation.

Overall, I would like to express how thankful I am to Mat, Scott, Nikolas, and Gency for helping to make this collaboration work. Interdisciplinary research between experimental scientists and mathematicians poses many challenges, not least of which are the very different perspectives on the same subject, and the different hopes for the outcomes of the projects. Based on my experience, it seems that respect and appreciation for each others' fields is sufficient to result in effective and productive collaborations, evidenced by the research contained herein.

For financial support I would like to thank Professor Mat Simpson and Queensland University of Technology for funding my studies. I am also very grateful for the provision of travel grants to conferences, namely the Landahl Travel Grant to the Annual Meeting of the Society for Mathematical Biology, and the CSIRO-ANZIAM Student Support Scheme.

Contents

	Page
Dedication	i
Abstract	iii
Statement of Original Authorship	v
Acknowledgements	viii
1 Introduction	1
1.1 Overview	1
1.2 Research questions	5
1.3 Aims and outcomes of this thesis	8
1.4 Structure of this thesis	10
1.5 Statement of joint authorship	11
2 Mathematical models for cell migration with real-time cell cycle dynamics	17
Abstract	17
2.1 Introduction	18
2.2 Materials and methods	23
2.2.1 Experiments	23
2.2.1.1 Cell culture	23
2.2.1.2 Fluorescent ubiquitination-based cell cycle indicator (FUCCI)	23
2.2.1.3 Wound healing migration assay	23

2.2.2	Mathematical model	23
2.2.2.1	Numerical solutions	23
2.3	Results and Discussion	24
2.3.1	Experimental data	24
2.3.2	Model development	28
2.3.3	Model application	30
2.3.4	Analysis	31
2.3.4.1	Numerical solutions	31
2.3.4.2	Travelling wave analysis of the fundamental model	33
2.3.4.3	Dispersion relation	38
2.4	Conclusion	39
2.5	Supplementary Material	42
2.5.1	Numerical solutions of the mathematical model	42
2.5.2	Scratch assay of FUCCI-transduced WM983C melanoma cells	42
2.5.3	Numerical solutions demonstrating travelling wave behaviour	42

3 Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronisation in cell proliferation **45**

Abstract	45
3.1 Introduction	46
3.2 Results	48
3.2.1 Experimental data	48
3.2.2 Multi-stage mathematical model	49
3.3 Conclusion	50
3.4 Data accessibility	50
3.5 Supplementary Material	50
3.5.1 Experimental	50
3.5.1.1 Cell culture	50
3.5.1.2 Proliferation experiments	51
3.5.1.3 Fluorescent ubiquitination-based cell cycle indicator (FUCCI)	51
3.5.2 Image processing and analysis	52
3.5.2.1 Preprocessing	52
3.5.2.2 Segmentation	52
3.5.2.3 Analysis	52
3.5.3 Parameterisation of the exponential model	54

3.5.4	Parameterisation of the multi-stage mathematical model	55
3.5.4.1	Method for parameter estimation	55
3.5.4.2	Specific estimated parameters	57
3.5.4.2.1	C8161 cell line - Figure 3.1	57
3.5.4.2.2	C8161 cell line, different numbers of stages - Figure 3.2	58
3.5.4.2.3	C8161 cell line - Figure 3.3	60
3.5.4.2.4	C8161 cell line - Figure 3.4	62
3.5.4.2.5	WM983C cell line - Figure 3.5	64
3.5.4.2.6	WM983C cell line - Figure 3.6	66
3.5.4.2.7	WM983C cell line - Figure 3.7	68
3.5.4.2.8	1205Lu cell line - Figure 3.8	70
3.5.4.2.9	1205Lu cell line - Figure 3.9	72
3.5.4.2.10	1205Lu cell line - Figure 3.10	74
3.5.5	All experimental data	76
3.5.5.1	C8161 cell line	78
3.5.5.2	WM983C cell line	81
3.5.5.3	1205Lu cell line	84
4	Examining go-or-grow using fluorescent cell-cycle indicators and cell cycle-inhibiting drugs	87
	Abstract	87
4.1	Introduction	88
4.2	Results and Discussion	91
4.3	Supplementary Material	93
4.3.1	Materials and Methods	93
4.3.1.1	Experiments	93
4.3.1.1.1	Cell culture	93
4.3.1.1.2	Fluorescent ubiquitination-based cell cycle indicator	94
4.3.1.1.3	Proliferation assay: cycling cells	94
4.3.1.1.4	Proliferation assay: G1-arrested cells	94
4.3.1.1.5	Summary of experimental conditions	95
4.3.1.2	Image processing and analysis	95
4.3.1.2.1	Centroids of cells	96
4.3.1.2.2	Single-cell tracking	97
4.3.1.2.3	Trajectory selection	98

4.3.1.2.4	Trajectory authentication	101
4.3.2	Data analysis	102
4.3.2.1	Cell-cycle characteristics of each cell line	102
4.3.2.2	Cell trajectories	103
4.3.2.3	Cell migration: directionality	104
4.3.2.3.1	Drift velocity	104
4.3.2.3.2	Moment of inertia tensor	108
4.3.2.3.3	Temporal velocity autocorrelation function	109
4.3.2.4	Cell migration: diffusivities	112
5	A novel mathematical model of heterogeneous cell proliferation	119
	Abstract	119
5.1	Introduction	120
5.2	Model motivation	122
5.2.1	Biological considerations	122
5.2.2	Mathematical considerations	124
5.2.2.0.1	Distributed delays	124
5.2.2.0.2	Contact inhibition	125
5.2.2.0.3	Proliferation, switching between proliferative states, and apoptosis	125
5.3	Mathematical model	127
5.4	Main results	129
5.4.1	Non-negativity and boundedness	129
5.4.2	Existence and uniqueness	130
5.4.3	Local stability	132
5.5	Supporting numerical simulations	142
5.6	Discussion and outlook	145
5.7	Supplementary Material	148
5.7.1	Experimental	148
5.7.2	Experimental data and models for cell cycle durations	150
5.7.3	Complete proof of Theorem 5.2	152
5.7.4	Complete proof of Theorem 5.4	154
5.7.5	Alternative proof of Proposition 5.6	157
5.7.6	Graphical illustration of Cauchy's argument principle	158

6	Conclusions and future work	161
6.1	Summary	161
6.2	Main theme	165
6.3	Future work	166
6.3.1	A quantitative investigation of the effects of anti-cancer drugs on cell mi- gration and proliferation	166
6.3.2	Extension of our cell migration model to additional cell cycle phases . . .	166
6.3.3	Formal proof of existence of travelling wave solutions	166
6.3.4	Investigating the causes of inherent synchronisation	167
6.3.5	Investigating go-or-grow in 3-D	168
6.3.6	Mathematical model of heterogeneous cell proliferation with time-dependent induced switching	168
6.3.7	Mathematical model of heterogeneous cell proliferation with an explicit density-dependence of proliferative-state switching	168
6.3.8	Mathematical model of heterogeneous cell proliferation with spontaneous switching	168
6.3.9	Mathematical model of heterogeneous cell proliferation with spatial di- mensions	169
6.4	Additional co-authored publications and preprints motivated by this thesis . . .	169
6.5	Final remarks	169
	Bibliography	171

1.1 Overview

Migration and proliferation are fundamental functions of cells, essential for normal physiological processes including development, tissue renewal and repair, and immune surveillance [1–5]. Normally these functions are tightly regulated *in vivo* by a complex network of intracellular signals and the extracellular environment, so that migration is coordinated and directed [4, 6, 7] and proliferation initiates and ceases as required [8–11]. The misregulation of migration and proliferation contributes to many pathological processes, most notably cancer where cells can invade nearby tissue and proliferate in an uncontrolled manner [12–17]. Consequently, cell migration and proliferation are subjects of intensive investigation, both experimentally and theoretically.

Experimental studies of migration and proliferation are often performed in two-dimensional (2-D) cell culture, from which the acquired knowledge may be utilised to understand more complicated three-dimensional (3-D) *in vitro* and *in vivo* experiments, while necessarily accounting for the differences in the microenvironments [18] and the additional modes of migration in 3-D [19]. Indeed, there are several advantages in employing 2-D cell cultures: the observed cell migration is partly representative of cell migration *in vivo*; the assays are amenable to standard laboratory techniques, including live-cell microscopy; and, the relative ease of image analysis including cell counting and single-cell tracking [20–22]. Note that 2-D cell cultures require the cells to be *adherent*, or anchorage-dependent, meaning that the cells adhere to the surface of the cell culture flask or plate, which is typical for cell lines derived from animal tissue [23, Section 6.2]. There is no implication here that the cells are generally adhered to each other when in cell

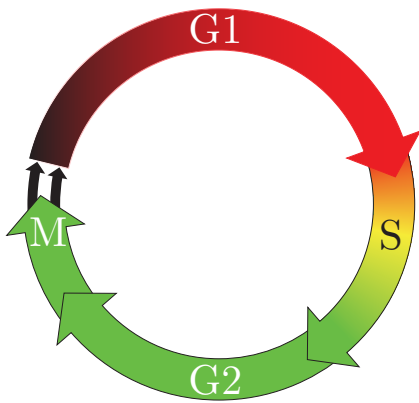


Figure 1.1: Schematic of the eukaryotic cell cycle, indicating the colour of FUCCI in each phase. During very early G1 phase there is no fluorescence as both FUCCI reporters are downregulated. As the cell progresses through G1 phase the red FUCCI reporter is upregulated and red fluorescence is observed. During the transition from G1 to S phase, called early S, both the red FUCCI reporter and the green FUCCI reporter are upregulated producing yellow. As the cell progresses through S/G2/M phase the red FUCCI reporter is downregulated and only the green FUCCI reporter is upregulated so that green fluorescence is observed.

culture. Adhesion is typically required for the cells to migrate, and also for regulation of the cell cycle and therefore proliferation [8].

For most eukaryotic cells, the cell cycle is a sequence of four discrete phases (Figure 1.1), namely gap 1 (G1), synthesis (S), gap 2 (G2) and mitosis (M). The phases G1, S, and G2 are collectively referred to as *interphase*, which refers to the part of the cell cycle between cell divisions. The primary function of the cell cycle is the replication of cellular DNA during S phase, followed by the division of the replicated chromosomes and cytoplasm into two daughter cells during M phase [24]. Progression through the cell cycle is tightly regulated in normal cells, which are subject to density-dependent contact inhibition producing reversible cell-cycle arrest [25, 26]. In cancer cells, however, the cell cycle is dysregulated [14] resulting in cell populations with proliferative heterogeneity, exemplified by the heterogeneous cell cycle speeds observed in tumours of solid cancers [27, 28]. In particular a small subpopulation of slow-proliferating cells is often present in tumours, and this subpopulation tends to survive anti-cancer drug treatment and can maintain the tumour by repopulating the fast-proliferating subpopulation [27, 29].

Although morphological changes associated with cell division, therefore the transition from M to G1 (Figure 1.2(a)), can be observed visually with a microscope, such distinct morphological changes are not observed during transitions between other cell cycle phases [30]. Therefore, alternative techniques are required to study these other cell cycle transitions. Fluorescent ubiquitination-based cell cycle indicator (FUCCI) technology [30] enables visualisation of the cell cycle progression in individual live cells (Figure 1.1). The FUCCI system consists of two fluorescent reporters in the cell nucleus: when the cell is in G1 phase the nucleus fluoresces red, and when the cell is in S/G2/M phase the nucleus fluoresces green. During the G1/S transition, called early S (eS), both reporters fluoresce and the nucleus appears yellow. Prior to the development of FUCCI it was difficult, if not impossible, to examine the cell cycle dynamics of individual cells beyond the M to G1 transition [30]. In contrast, FUCCI allows direct visualisation, in real time, of progression through the cell cycle. See Figure 1.2(b), which is the

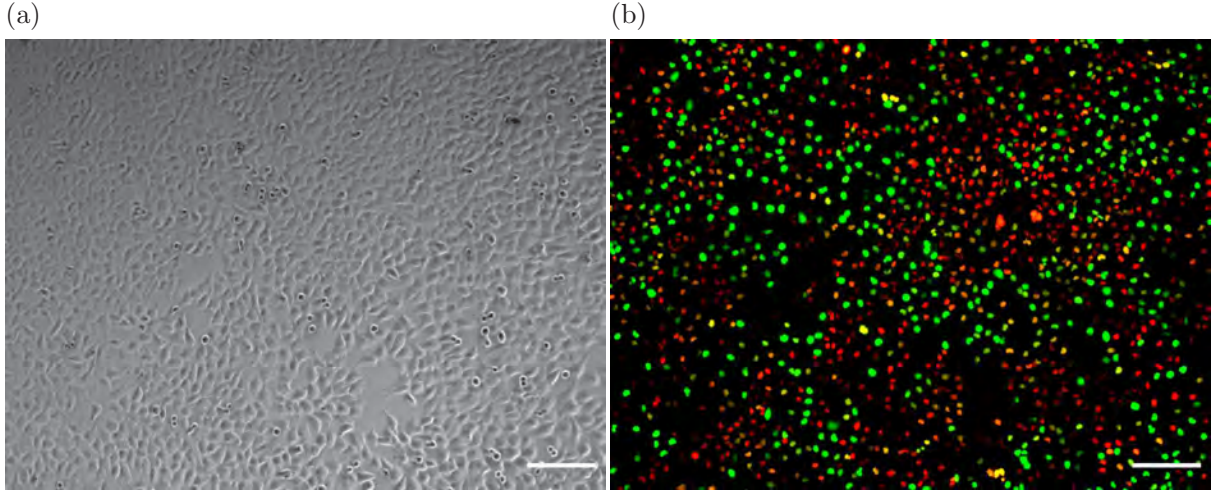


Figure 1.2: Experimental images of a proliferation assay with FUCCI-C8161 melanoma cells. (a) The phase contrast image reveals little information regarding the cell cycle phase in which each cell resides. (b) The corresponding FUCCI image provides detailed information on the cell cycle phase of each cell: red cells are in G1 phase, green cells are in S/G2/M phase, and yellow cells are in eS phase. Scale bar = 200 μm .

corresponding FUCCI image of Figure 1.2(a). This technology is particularly useful for research in cancer biology [13, 31–33], cell biology [34, 35], and stem cell biology [36, 37].

Mathematical models describing cell migration and proliferation in a population of cells often involve reaction-diffusion equations [38] based on the one-dimensional (1-D) Fisher–Kolmogorov–Petrovskii–Piskunov (FKPP) equation [38, 39],

$$\frac{\partial s}{\partial t} = D \frac{\partial^2 s}{\partial x^2} + \lambda s \left(1 - \frac{s}{K}\right), \quad (1.1)$$

where x is position, t is time, $s(x, t) \geq 0$ is the cell density, $D > 0$ is the diffusivity, $K > 0$ is the carrying-capacity density, and $\lambda > 0$ is the intrinsic growth rate. The density dependence of cell proliferation is described in the FKPP equation with a logistic source term. The FKPP equation and related adaptations, including stochastic analogues [40, 41], have been successfully used to model cell migration *in vitro* and *in vivo* [42–46]. A limitation of the FKPP equation is that it considers a single population of cells. For a more realistic situation, where the total population is composed of a number of distinct and interacting subpopulations, it is relevant to consider a model that involves a system of coupled equations that are often related to the FKPP equation [44, 47].

There are several key assumptions underlying the FKPP equation and related models regarding a cell population: cell migration, and therefore the diffusivity D , is independent of the cell cycle phase; monotonic growth of the total population implies monotonic growth for subpopulations; and the distribution of cell cycle durations is exponentially distributed. These

assumptions and more can all be tested with cell culture experiments using FUCCI-expressing cells. Indeed, FUCCI provides access to a large amount of detailed information regarding the cell cycle which was not previously available, so there exists the potential to combine this data with the development of novel mathematical models to provide new and interesting insights into cell migration and proliferation.

In this thesis we begin by developing mathematical models of cell migration with subpopulations based on the cell cycle phases highlighted by FUCCI. Our models are motivated by the FKPP equation (1.1), and consist of either two or three coupled partial differential equations, where each equation corresponds to a subpopulation of cells in a particular phase of the cell cycle that is highlighted by FUCCI. We demonstrate how our models can be used to simulate the spatial and temporal dynamics in a population of FUCCI cells by modelling our new experimental data obtained from scratch assays with FUCCI cells. The FKPP equation (1.1) has travelling wave solutions [38, 48], and we show that our model also has travelling wave solutions and derive analytical expressions for the minimum wave speed and a dispersion relation.

Our analysis of experimental data from cell culture assays with FUCCI cells reveals that the cell populations have oscillating subpopulations due to partial synchronisation. It is important to identify the presence of synchronisation, which may affect the experimental reproducibility of studies that aim to investigate cell cycle-dependent mechanisms such as changes in migration and drug response. Moreover, while the subpopulations in our assays are oscillatory the total cell population appears to grow exponentially, and we reconcile these observations using a multi-stage model of proliferation, which is based on the FKPP equation (1.1) without the diffusive term.

We then use FUCCI to examine the go-or-grow hypothesis, which states that adherent cells undergo reversible phenotype switching between migratory and proliferative states, with cells in the migratory state being more motile than cells in the proliferative state. If the go-or-grow hypothesis was confirmed it would have important implications for anti-cancer treatments that employ cell cycle-inhibiting drugs, which can induce cell-cycle arrest. Since an arrested cell is not proliferative, the cell's energy could be utilised for migration, potentially leading to an exacerbation of metastasis. We examine go-or-grow using 2-D *in vitro* assays with FUCCI melanoma cells and cell cycle-inhibiting drugs. Using FUCCI allows us to distinguish between cell migration in the phases G1 and S/G2/M. We compare migration by analysing the experimental data using single-cell tracking to calculate mean diffusivities for cells in different cell-cycle phases and in cell-cycle arrest.

Finally, we present a novel mathematical model of heterogeneous cell proliferation where the

total population consists of a subpopulation of slow-proliferating cells and a subpopulation of fast-proliferating cells. The model incorporates two cellular processes, asymmetric cell division and induced switching between proliferative states, which are important determinants for the heterogeneity of a cell population. As motivation for our model we provide experimental data with FUCCI cells which assists in illustrating the induced-switching process, and further FUCCI data on durations of cell cycle phases that guide our model choice. Our model is based on the logistic growth term of the FKPP equation (1.1), and consists of a system of two coupled delay differential equations with distributed time delays and the cell densities as functions of time.

1.2 Research questions

The work presented in this thesis is a combination of new experimental data using FUCCI-expressing cells, analysis of the data, and the development of novel mathematical models motivated by the data to examine cell migration and proliferation as a function of the cell cycle. In particular, we address the following four questions.

1. **How can we incorporate cell cycle dynamics into a continuum model of cell migration based on the information provided by FUCCI technology in relation to the cell cycle phase?**

2-D scratch assays are commonly used experimental models to study the invasive and proliferative behaviour of cancer cells. In combination with FUCCI, these experimental models can be used to examine the cell cycle dynamics of individual cells as a function of position within the scratch assay [13, 32, 33]. A major advantage of this method is that two fundamental phenomena associated with malignant invasion, namely cell proliferation and cell migration, can be characterised simultaneously. Indeed, FUCCI allows us to directly examine the spatial and temporal patterns of cell proliferation within a migrating population. To the best of our knowledge, there are no mathematical models in the literature that have been previously developed to describe cell migration with FUCCI technology.

We propose to develop a new mathematical model to predict new cell migration experiments with FUCCI-expressing cells. Our focus is on 2-D cell migration experiments as this is the most common experimental platform because of convenience, simplicity, and low cost [21]. 2-D cell migration assays are valuable because they are often used as high-throughput screening tools in conjunction with more sophisticated preclinical models [21].

By considering the cells in a particular phase of the cell cycle as indicated by FUCCI to comprise a distinct subpopulation, we can consider a mathematical model that consists of a system of coupled differential equations based on the FKPP equation (1.1).

Once we have developed our mathematical model, we will analyse the model for travelling wave solutions. It is well known that travelling wave solutions exist for the FKPP equation (1.1) [38, 48]. These solutions have the form of a constant shape, monotonically decreasing, wavefront, which propagates with a minimum speed, $c_{\min} = 2\sqrt{\lambda D}$ [38, 48]. Travelling wave solutions are interesting as experimental observations often exhibit moving fronts that can be thought of as a travelling wave [46]. The speed of a moving cell front is often the simplest quantitative measurement that can be obtained from an experiment [46, 49]. Therefore, understanding the relationship between the parameters in the mathematical model and the speed of the travelling wave solution is a useful way to help parameterise the mathematical model to match experimental observations.

2. Is a cell proliferation assay prepared using standard cell-culture procedures asynchronous?

Cell proliferation is essential for a range of normal and pathological processes. Many different mathematical models of proliferation have been proposed [50–56], and a key assumption often implicit in such models is that the cell population is asynchronous, meaning that the cells are distributed randomly among the cell cycle phases. In contrast, a population of cells is synchronous if the cells are in the same cell cycle phase and divide as a cohort in discrete stages.

We propose to test the assumption of cell asynchronicity with new experimental data from 2-D cell proliferation assays employing FUCCI-expressing cells, using multiple cell lines and replicates. By considering the subpopulations as indicated by FUCCI, synchronicity would be revealed as oscillations in these subpopulations. Generally, the presence of synchronisation may affect the reproducibility of experiments that investigate cell-cycle-dependent mechanisms, such as changes in migration and drug response.

3. What can FUCCI and cell-cycle-inhibiting drugs reveal about the go-or-grow hypothesis?

The go-or-grow hypothesis proposes that adherent cells reversibly switch between migratory and proliferative phenotypes [57], exhibiting higher motility in the migratory state as motile cells are not using energy for proliferation [57–61]. Previous experimental investigations of the go-or-grow hypothesis are conflicting, as some studies support the hypoth-

esis [57, 62, 63] while others refute it [13, 64, 65].

We propose to rigorously examine the go-or-grow hypothesis in new 2-D *in vitro* experiments using Fucci-expressing melanoma cells. Two key properties of cell lines subject to go-or-grow would be the ability to switch between invasive and proliferative phenotypes, and metastatic potential [57]. Studies of tumour heterogeneity in melanoma suggest that cells may reversibly switch between invasive and proliferative phenotypes [57]. Further, melanoma is highly metastatic, forms tumours that are very heterogeneous, and is well known to respond to drugs which induce G1 arrest [66, 67]. Melanoma cells are therefore a prime candidate for studying the go-or-grow hypothesis [57, 59, 68]. Fucci allows us to study cell motility in G1 separately from cell motility in S/G2/M [13, 69, 70]. Specifically, we can investigate cycling cells for differences in motility when the cells are in G1 compared with S/G2/M. Further, given the potential for an arrested cell to become more motile, we can use a cell cycle-inhibiting drug to effect G1 arrest, and compare the motility of the arrested cells with cycling cells. Single-cell tracking can then be used to obtain the trajectories of cells, from which cell diffusivities can be estimated.

4. How can data from proliferation experiments with Fucci-expressing cells motivate the development of a continuum model of heterogeneous cell proliferation?

Cell proliferation is the fundamental function of the cell cycle [10], which is a complex process regulated by both intracellular signals and the extracellular environment [9]. Despite all of the underlying complexity the cell cycle has two basic fates, either progression or arrest [10]. These two cellular fates form the basis of many mathematical models of cell proliferation in the literature, typically based on exponential growth [51, 54, 71, 72] or logistic growth [50, 52–54, 73, 74]. Exponential growth explicitly accounts for progression only, while logistic growth accounts for progression and density-dependent arrest, which can result from contact inhibition [75]. An important detail of the cell cycle not explicitly accounted for in exponential and logistic growth models is the distribution of cell cycle durations, which are always nonzero and exhibit considerable variation between different cell types and different extracellular environments [76, 77]. The main types of models which explicitly incorporate cell cycle duration are functional differential equation models [78–81] and multi-stage models [82–86].

We propose to develop a functional differential equation model for cell proliferation in which the cell population consists of a slow-proliferating subpopulation and a fast-proliferating

subpopulation, with cells able to switch between the subpopulations through two cellular processes, namely asymmetric cell division and induced switching of proliferative states by surrounding cells. New data from cell proliferation experiments with FUCCI-expressing cells can be used to inform an appropriate model for cell cycle durations and switching of proliferative states. While there are mathematical models in the literature that consider proliferative heterogeneity, they generally lack the flexibility of the proliferative-state dynamics that we consider: some models account for one proliferating subpopulation only, which undergoes asymmetric division, while the other subpopulations are quiescent or differentiated [87,88]; other models simply consider subpopulations with different proliferative states without any cells switching between the subpopulations [89].

1.3 Aims and outcomes of this thesis

The primary aim of this thesis is to investigate how new information revealed by FUCCI technology in 2-D *in vitro* cell culture experiments can inform the development of new mathematical models for cell migration and proliferation. To achieve this aim it is required to perform new cell culture experiments and analyse the resulting data, develop novel mathematical models motivated by the data, and analyse the solutions of the mathematical models.

This thesis addresses the following four objectives:

1. Incorporate cell cycle dynamics into a continuum model of cell migration based on the information provided by FUCCI technology in relation to the cell cycle phase.
2. Investigate whether cell proliferation assays prepared using standard cell-culture procedures are asynchronous, utilising FUCCI technology.
3. Rigorously examine the go-or-grow hypothesis using FUCCI technology and cell-cycle-inhibiting drugs.
4. Develop a continuum model of heterogeneous cell proliferation motivated by data from new proliferation experiments with FUCCI-expressing cells.

This thesis is by publication, consisting of three published articles and one manuscript under review. The Ph.D. candidate contributed significantly to all four manuscripts, recognised by first authorship of each manuscript. The work presented in this thesis fulfils the requirements for the award of a thesis by published papers at Queensland University of Technology.

This thesis is comprised of the following four articles:

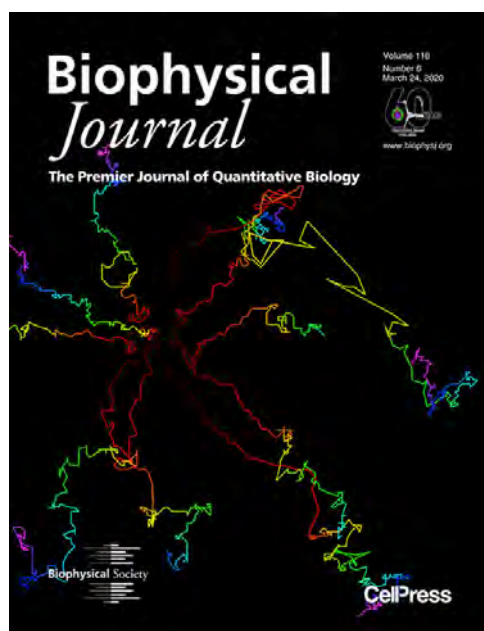
1. **Vittadello ST**, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*. 2018;114:1241–1253. doi:10.1016/j.bpj.2017.12.041.
2. **Vittadello ST**, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronization in cell proliferation. *Journal of the Royal Society Interface*. 2019;16:20190382. doi:10.1098/rsif.2019.0382.

This article is a Report, which is a short contribution with most of the details contained in the supplemental material.

3. **Vittadello ST**, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Examining go-or-grow using fluorescent cell-cycle indicators and cell-cycle-inhibiting drugs. *Biophysical Journal*. 2020;118:1243–1247. doi:10.1016/j.bpj.2020.01.036.

This article is a Biophysical Letter, which is a short contribution with most of the details contained in the supplemental material.

Our cover image was selected for the cover of the 24 March 2020 issue of the Biophysical Journal:



4. **Vittadello ST**, McCue SW, Gunasingh G, Haass NK, Simpson MJ. A novel mathematical model of heterogeneous cell proliferation. arXiv:2003.03024v2 preprint. 2020. <https://arxiv.org/abs/2003.03024v2>. Currently under review with the Journal of Mathematical Biology.

These four articles are reproduced in Chapters 2–5 of this thesis, except for some standardisation of the formatting and layout. Each of these four chapters is therefore self-contained and can be read independently of all other chapters. Consequently, there is a degree of overlap between Chapters 2–5 in regard to the background discussions of the experimental and modelling literature, and also our experimental methodology.

1.4 Structure of this thesis

The structure of this thesis is as follows.

Chapter 1 forms the introduction for this thesis, and provides an overview of cell migration and proliferation from both the experimental and modelling perspectives along with a review of the relevant literature. The research questions are then stated, as are the aims and outcomes of this thesis.

Chapter 2 contains Article 1, in which we develop a new mathematical model of cell migration that includes cell cycle dynamics as indicated by FUCCI. We present new experimental data from scratch assays using FUCCI-expressing melanoma cells, and show that the predictions of spatial and temporal patterns of cell density in the experiments can be described by the model. We show numerically that the model exhibits travelling wave solutions under appropriate conditions, and we derive and confirm an analytical expression for the minimum wave speed, as well as exploring how the wave speed depends on the spatial decay rate of the initial condition.

Chapter 3 contains Article 2, in which we present a suite of new experimental data showing that cell proliferation assays, prepared using standard methods thought to produce asynchronous cell populations, persistently exhibit inherent synchronization. Our experiments use FUCCI to reveal the normally-hidden cell synchronization, which would never be observed without these cell cycle indicators. On the other hand, our experimental data show that the total cell population appears to grow exponentially, as in an asynchronous population. We reconcile these seemingly inconsistent observations by employing a multi-stage mathematical model of cell proliferation that can replicate the oscillatory subpopulations.

Chapter 4 contains Article 3, in which we examine go-or-grow in new experiments consisting of two-dimensional *in vitro* assays using FUCCI-expressing melanoma cells and cell-cycle-inhibiting drugs. We analyze the experimental data using single-cell tracking to calculate mean diffusivities and compare motility between cells in different cell-cycle phases and in cell-cycle arrest.

Chapter 5 contains Article 4, in which we present a novel mathematical model for heterogeneous cell proliferation where the total population consists of a subpopulation of slow-proliferating cells and a subpopulation of fast-proliferating cells. The model incorporates two cellular processes, asymmetric cell division and induced switching between proliferative states, which are important determinants for the heterogeneity of a cell population. As motivation for our model we provide experimental data that assists in illustrating the induced-switching process, and further data on durations of cell cycle phases that guide our model choice. Our model consists of a system of two coupled nonlinear functional differential equations with the cell densities as functions of time. The distributed delays are bounded and allow for the choice of delay kernel. We analyse the model and prove the non-negativity and boundedness of solutions, the existence and uniqueness of solutions, and the local stability characteristics of the equilibrium points. Numerical simulations illustrate and support the theoretical findings.

Chapter 6 forms the conclusion of this thesis, in which we summarise our research findings from the perspective of our original objectives. We also discuss the possibilities for future work arising from this research, from both experimental and theoretical perspectives.

1.5 Statement of joint authorship

Here we outline the contributions of the Ph.D. candidate and the co-authors to each article. All co-authors consent to the presentation of this material in this thesis.

1. Chapter 2: Mathematical models for cell migration with real-time cell cycle dynamics

The corresponding article is:

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*. 2018;114:1241–1253. doi:10.1016/j.bpj.2017.12.041.

Abstract

The fluorescent ubiquitination-based cell cycle indicator, also known as FUCCI, allows the visualization of the G1 and S/G2/M cell cycle phases of individual cells. FUCCI consists of two fluorescent probes, so that cells in the G1 phase fluoresce red and cells in the S/G2/M phase fluoresce green. FUCCI reveals real-time information about cell cycle dynamics of individual cells, and can be used to explore how the cell cycle relates

to the location of individual cells, local cell density, and different cellular microenvironments. In particular, FUCCI is used in experimental studies examining cell migration, such as malignant invasion and wound healing. Here we present, to our knowledge, new mathematical models that can describe cell migration and cell cycle dynamics as indicated by FUCCI. The fundamental model describes the two cell cycle phases, G1 and S/G2/M, which FUCCI directly labels. The extended model includes a third phase, early S, which FUCCI indirectly labels. We present experimental data from scratch assays using FUCCI-transduced melanoma cells, and show that the predictions of spatial and temporal patterns of cell density in the experiments can be described by the fundamental model. We obtain numerical solutions of both the fundamental and extended models, which can take the form of traveling waves. These solutions are mathematically interesting because they are a combination of moving wavefronts and moving pulses. We derive and confirm a simple analytical expression for the minimum wave speed, as well as exploring how the wave speed depends on the spatial decay rate of the initial condition.

Statement of joint authorship:

- **Vittadello ST (Candidate):** performed the experiments, analysed the data, interpreted the data, analysed the mathematical model, implemented the numerical algorithm, composed all figures, wrote the manuscript, edited the manuscript
- **McCue SW:** interpreted the data, contributed analytic tools, edited the manuscript
- **Gunasingh G:** designed the experiments, oversaw the experimentation, interpreted the data, contributed analytic tools, edited the manuscript
- **Haass NK:** designed the project, directed the project, designed the experiments, interpreted the data, edited the manuscript
- **Simpson MJ:** devised the concept for the project, designed the project, directed the project, devised the mathematical model, devised the numerical algorithm, interpreted the data, edited the manuscript, prepared the cover letter, acted as corresponding author

2. Chapter 3: Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronization in cell proliferation

The corresponding article is:

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchroniza-

tion in cell proliferation. *Journal of the Royal Society Interface*. 2019;16:20190382. doi:10.1098/rsif.2019.0382.

Abstract

We present a suite of experimental data showing that cell proliferation assays, prepared using standard methods thought to produce asynchronous cell populations, persistently exhibit inherent synchronization. Our experiments use fluorescent cell cycle indicators to reveal the normally hidden cell synchronization, by highlighting oscillatory subpopulations within the total cell population. These oscillatory subpopulations would never be observed without these cell cycle indicators. On the other hand, our experimental data show that the total cell population appears to grow exponentially, as in an asynchronous population. We reconcile these seemingly inconsistent observations by employing a multi-stage mathematical model of cell proliferation that can replicate the oscillatory subpopulations. Our study has important implications for understanding and improving experimental reproducibility. In particular, inherent synchronization may affect the experimental reproducibility of studies aiming to investigate cell cycle-dependent mechanisms, including changes in migration and drug response.

Statement of joint authorship:

- **Vittadello ST (Candidate):** devised the concept for the project, designed the project, performed the experiments, analysed the data, interpreted the data, analysed the mathematical model, devised and implemented the numerical algorithms, composed all figures, wrote the manuscript, edited the manuscript
- **McCue SW:** interpreted the data, edited the manuscript
- **Gunasingh G:** designed the experiments, oversaw the experimentation, interpreted the data, edited the manuscript
- **Haass NK:** designed the project, directed the project, designed the experiments, interpreted the data, edited the manuscript
- **Simpson MJ:** devised the concept for the project, designed the project, directed the project, devised the mathematical model, interpreted the data, edited the manuscript, prepared the cover letter, acted as corresponding author

3. Chapter 4: Examining go-or-grow using fluorescent cell-cycle indicators and cell-cycle-inhibiting drugs

The corresponding article is:

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*. 2020;118:1243–1247. doi:10.1016/j.bpj.2020.01.036.

Abstract

The go-or-grow hypothesis states that adherent cells undergo reversible phenotype switching between migratory and proliferative states, with cells in the migratory state being more motile than cells in the proliferative state. Here, we examine go-or-grow in two-dimensional *in vitro* assays using melanoma cells with fluorescent cell-cycle indicators and cell-cycle-inhibiting drugs. We analyze the experimental data using single-cell tracking to calculate mean diffusivities and compare motility between cells in different cell-cycle phases and in cell-cycle arrest. Unequivocally, our analysis does not support the go-or-grow hypothesis. We present clear evidence that cell motility is independent of the cell-cycle phase and that nonproliferative arrested cells have the same motility as cycling cells.

Statement of joint authorship:

- **Vittadello ST (Candidate):** designed the project, performed the experiments, interpreted the data, devised the mathematical analysis of the data, analysed the data, composed all figures, wrote the manuscript, edited the manuscript, prepared the cover letter, acted as corresponding author
- **McCue SW:** interpreted the data, edited the manuscript
- **Gunasingh G:** designed the project, designed the experiments, oversaw the experimentation, interpreted the data, edited the manuscript
- **Haass NK:** devised the concept for the project, designed the project, directed the project, designed the experiments, interpreted the data, edited the manuscript
- **Simpson MJ:** devised the concept for the project, designed the project, directed the project, interpreted the data, edited the manuscript

4. Chapter 5: A novel mathematical model of heterogeneous cell proliferation

The corresponding article is:

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. A novel mathematical model of heterogeneous cell proliferation. arXiv:2003.03024v2 preprint. 2020. Currently under review with the Journal of Mathematical Biology.

Abstract

We present a novel mathematical model of heterogeneous cell proliferation where the total population consists of a subpopulation of slow-proliferating cells and a subpopulation of fastproliferating cells. The model incorporates two cellular processes, asymmetric cell division and induced switching between proliferative states, which are important determinants

for the heterogeneity of a cell population. As motivation for our model we provide experimental data that assist in illustrating the induced-switching process, and further data on durations of cell cycle phases that guide our model choice. Our model consists of a system of two coupled delay differential equations with distributed time delays and the cell densities as functions of time. The distributed delays are bounded and allow for the choice of delay kernel. We analyse the model and prove the non-negativity and boundedness of solutions, the existence and uniqueness of solutions, and the local stability characteristics of the equilibrium points. We find that the parameters for induced switching are bifurcation parameters and therefore determine the overall dynamics of the model. Numerical simulations illustrate and support the theoretical findings.

Statement of joint authorship:

- **Vittadello ST (Candidate):** devised the concept for the project, designed the project, directed the project, performed the experiments, analysed the data, interpreted the data, devised the mathematical model, analysed the mathematical model, devised and implemented the numerical algorithm, composed all figures, wrote the manuscript, edited the manuscript, prepared the cover letter, acted as corresponding author
- McCue SW: interpreted the data, edited the manuscript
- Gunasingh G: designed the experiments, oversaw the experimentation, interpreted the data, edited the manuscript
- Haass NK: designed the experiments, directed the project, interpreted the data, edited the manuscript
- Simpson MJ: interpreted the data, directed the project, edited the manuscript

Mathematical models for cell migration with real-time cell cycle dynamics

An article published in *Biophysical Journal*

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*. 2018; 114:1241–1253.

Abstract

The fluorescent ubiquitination-based cell cycle indicator, also known as FUCCI, allows the visualisation of the G1 and S/G2/M cell cycle phases of individual cells. FUCCI consists of two fluorescent probes, so that cells in the G1 phase fluoresce red and cells in the S/G2/M phase fluoresce green. FUCCI reveals real-time information about cell cycle dynamics of individual cells, and can be used to explore how the cell cycle relates to the location of individual cells, local cell density, and different cellular microenvironments. In particular, FUCCI is used in experimental studies examining cell migration, such as malignant invasion and wound healing. Here we present, to our knowledge, new mathematical models that can describe cell migration and cell cycle dynamics as indicated by FUCCI. The *fundamental* model describes the two cell cycle phases, G1 and S/G2/M, which FUCCI directly labels. The *extended* model includes a third phase, early S, which FUCCI indirectly labels. We present experimental data from scratch assays using FUCCI-transduced melanoma cells, and show that the predictions of spatial and temporal patterns of cell density in the experiments can be described by the fundamental model. We obtain numerical solutions of both the fundamental and extended models, which can take the form of traveling waves. These solutions are mathematically interesting because they are a

combination of moving wavefronts and moving pulses. We derive and confirm a simple analytical expression for the minimum wave speed, as well as exploring how the wave speed depends on the spatial decay rate of the initial condition.

2.1 Introduction

The cell cycle consists of a sequence of four distinct phases, namely: gap 1 (G1), synthesis (S), gap 2 (G2), and the mitotic (M) phase [90]. The phases G1, S and G2 are collectively referred to as “interphase”, and involve cell growth and preparation for division. After interphase, the cell enters the mitotic phase and divides into two daughter cells. Although morphological changes associated with cell division can be observed visually during the transition from M to G1, such distinct morphological changes are not possible during transitions between other cell cycle phases [30]. Therefore, different techniques are required to study these other cell cycle transitions.

Since 2008, fluorescent ubiquitination-based cell cycle indicator (FUCCI) technology [30] has enabled the visualisation of the cell cycle progression from G1 to S/G2/M in individual cells. The FUCCI system consists of two fluorescent probes in the cell nucleus, or cytoplasm, which emit red fluorescence when the cell is in the G1 phase, or green fluorescence when the cell is in the S/G2/M phase. Before the development of FUCCI it was difficult, if not impossible, to examine the cell cycle dynamics of individual cells beyond the M to G1 transition [30]. In contrast, FUCCI allows direct visualisation, in real time, of transitions in the cell cycle. This technology is particularly useful for research in cancer biology [13, 31–33], cell biology [34, 35], and stem cell biology [36, 37].

3-D spheroids and 2-D scratch assays are commonly used experimental models to study the invasive and proliferative behaviour of cancer cells. In combination with FUCCI, these experimental models can be used to examine the cell cycle dynamics of individual cells as a function of position within the spheroid or scratch assay [13, 32, 33]. A major advantage of this method is that two fundamental phenomena associated with malignant invasion, namely cell proliferation and cell migration, can be characterised simultaneously. Previous methods to examine the roles of cell migration and cell proliferation involve pretreating cells with antimitotic drugs, such as mitomycin-C [91]. A major limitation of these previous methods is that the application of the antimitotic drug is thought to suppress proliferation without interrupting migration. However, this assumption is questionable, and rarely examined [92]. The development of FUCCI technology obviates the need for such crude methods to isolate the roles of cell migration and cell proliferation. Instead, FUCCI allows us to directly examine the spatial and temporal patterns

of cell proliferation within a migrating population. To the best of our knowledge, there are no mathematical models in the literature that have been developed to describe cell migration with FUCCI technology. The focus of this work is on *cell migration*, by which we mean a moving front of a population of cells. These moving fronts are composed of a large number of individual cells that do not maintain cell-to-cell contacts. The formation of the moving front of cells is driven by a combination of cell motility and cell proliferation.

Cell migration involves diffusion, arising from random cell motility, and proliferation of cells [92]. Mathematical models describing these processes in a population of cells tend to involve reaction-diffusion equations [93], which are often based on the Fisher–Kolmogorov–Petrovskii–Piskunov (FKPP) equation [38, 39],

$$\frac{\partial s}{\partial t} = D \frac{\partial^2 s}{\partial x^2} + \lambda s(1 - s), \quad (2.1)$$

where $s(x, t) > 0$ is the cell density, $D > 0$ is the diffusion coefficient, and $\lambda > 0$ is the proliferation rate. Here, the dimensional cell density is scaled by the dimensional carrying capacity density, so that the maximum non-dimensional cell density is $s(x, t) = 1$. Carrying capacity limited proliferation of cells is described in Equation (2.1) with a logistic source term. Equation (2.1) has been successfully adapted to model many biological processes, such as *in vitro* cell migration [42, 45, 46]. A limitation of Equation (2.1) is that it considers a single population of cells. For a more realistic situation, where the total population is composed of a number of distinct, interacting subpopulations, it is relevant to consider a model that involves a system of coupled equations that are often related to Equation (2.1) [47].

The existence of travelling wave solutions for the FKPP equation is well known [38, 48]. Constant shape, monotonically decreasing wavefront travelling wave solutions propagate with a minimum speed, $c_{\min} = 2\sqrt{\lambda D}$ [38, 48]. Travelling wave solutions are of interest, because experimental observations tend to exhibit moving fronts that can be thought of as a travelling wave [46]. The speed of a moving cell front is often the simplest quantitative measurement that can be obtained from an experiment [46, 94]. Therefore, understanding the relationship between the parameters in the mathematical model and the speed of the travelling wave solution is a useful way to help parameterise the mathematical model to match experimental observations. Travelling wave solutions have also been observed in other mathematical models of cell migration [46, 47, 94, 95], as well as other reaction-diffusion models related to biological processes [96–99]. Most travelling wave solutions take the form of moving wavefronts, which have a monotone profile. Another type of travelling wave solution is a pulse, which is characterised by a nonmonotone profile [96, 100].

Here we present, to our knowledge, a new mathematical model of cell migration, which incorporates cell cycle dynamics based on the information provided by FUCCI technology in relation to the cell cycle phase. We consider the cells in a particular phase of the cell cycle to make up a distinct subpopulation, so our model consists of a system of coupled partial differential equations. To motivate our generalisation of Equation (2.1), we pay careful attention to the underlying biological features. This leads us to develop two different mathematical descriptions of cell migration with FUCCI technology:

- **Fundamental FUCCI model:** In the most fundamental format, FUCCI highlights a subpopulation of cells in the G1 phase as being red, and another subpopulation of cells in the S/G2/M phase as being green. Motivated by the ability to distinguish between these two phases of the cell cycle, we develop a mathematical model with two subpopulations: $v_r(x, t)$ and $v_g(x, t)$. Here, the $v_r(x, t)$ subpopulation corresponds to the red cells and the $v_g(x, t)$ subpopulation corresponds to the green cells. We refer to this model as the *fundamental* model.
- **Extended FUCCI model:** In some experimental descriptions, cell biologists identify an additional subpopulation that corresponds to the situation where both of the red and green probes are active simultaneously, giving rise to a third subpopulation that appears to be yellow. This overlap of the red and green fluorescence occurs during the early S, or eS, phase. Using experimental images, we find that the yellow subpopulation is more difficult to reliably identify than either the red or green subpopulations, as only a very small proportion of the cell population appear to be distinctly yellow. This is, in some sense, expected, because the yellow subpopulation results from the transient overlap of the G1 phase (red) and the S/G2/M phase (green). Despite this difficulty, we also develop another mathematical model which is capable of representing the three subpopulations: $u_r(x, t)$ is the red subpopulation, $u_y(x, t)$ is the yellow subpopulation, and $u_g(x, t)$ is the green subpopulation. We refer to this model as the *extended* model.

The fundamental and extended models are both related to the FKPP model with the dependent variable $s(x, t)$. To summarise the similarities and differences between the fundamental and extended models, the FKPP model, and the relationship between these mathematical models and the underlying biology, we report the dependent variables and their biological interpretation in Table 2.1.

In this work, we present information relating to both the fundamental and the extended

Table 2.1: Summary and comparison of the fundamental and extended models and the FKPP model.

Model	Equation reference	Dependent variables	Biological interpretation
FKPP	Equation (2.1)	$s(x, t)$	Total cell density is a function of position x and time t
Fundamental FUCCI	Equations (2.2) to (2.3)	$v_r(x, t)$	Cell density for cells in G1 (red) phase is a function of position x and time t
		$v_g(x, t)$	Cell density for cells in S/G2/M (green) phase is a function of position x and time t
		$s(x, t)$	Total cell density, $s(x, t) = v_r(x, t) + v_g(x, t)$, is a function of position x and time t
Extended FUCCI	Equations (2.4) to (2.6)	$u_r(x, t)$	Cell density for cells in G1 (red) phase is a function of position x and time t
		$u_y(x, t)$	Cell density for cells in eS (yellow) phase is a function of position x and time t
		$u_g(x, t)$	Cell density for cells in S/G2/M (green) phase is a function of position x and time t
		$s(x, t)$	Total cell density, $s(x, t) = u_r(x, t) + u_y(x, t) + u_g(x, t)$, is a function of position x and time t

models, but we focus on developing, to our knowledge, new mathematical analysis of the fundamental FUCCI model. Furthermore, we quantitatively apply the fundamental model, to our knowledge, to new experimental data. There are two types of parameters in our model, namely transition rates between phases of the cell cycle, and diffusion coefficients describing the rate of cell migration. The transition rates between phases of the cell cycle are estimated using experimental data from Haass et al. [13], who report data relating to the time spent in each phase of the cell cycle. Using new experimental data in the form of images of scratch assays of FUCCI-transduced melanoma cells, we extract quantitative information about cell density as a function of position and time, and compare these quantitative data with the predictions of our fundamental model. Because this is, to our knowledge, the first time that a mathematical model has been used to predict a cell migration experiment with FUCCI-transduced cells, we focus on 2-D cell migration experiments as this is the most common experimental platform because of convenience, simplicity, and low cost [21]. 2-D cell migration assays are valuable because they are often used as high-throughput screening tools in conjunction with more sophisticated preclinical models [21]. In addition to showing how these mathematical models can be used to predict the 2-D experiments, numerical solutions of the mathematical model show the formation of travelling wave solutions that are a combination of coupled wavefronts and pulses. We also derive an analytical expression for the minimum wave speed of the travelling waves, and we explore how these results are different from those of the standard FKPP equation, given in Equation (2.1).

This article is organised as follows. In the Materials and Methods, we detail the experimental protocol used to perform the scratch assays using FUCCI-transduced melanoma cells. We also outline the numerical analysis of our mathematical models. In the Results and Discussion, we present and discuss our mathematical models, and compare numerical solutions of the fundamental model with new data from scratch assays of melanoma cells. We then analyse our models numerically for travelling wave solutions, and for the fundamental model we derive an analytical expression for the minimum wave speed.

2.2 Materials and methods

2.2.1 Experiments

2.2.1.1 Cell culture

The human melanoma cell lines C8161 (kindly provided by Mary Hendrix, Chicago, IL, USA), 1205Lu and WM983C (both kindly provided by Meenhard Herlyn, Philadelphia, PA, USA) were genotypically characterized [101–104], grown as described [105] (using 4% fetal bovine serum instead of 2%), and authenticated by STR fingerprinting (QIMR Berghofer Medical Research Institute, Herston, Australia).

2.2.1.2 Fluorescent ubiquitination-based cell cycle indicator (FUCCI)

To generate stable melanoma cell lines expressing the FUCCI constructs, mKO2-hCdt1 (30-120) and mAG-hGem (1-110) [30] were subcloned into a replication-defective, self-inactivating lentiviral expression vector system as previously described [105]. The lentivirus was produced by co-transfection of human embryonic kidney 293T cells. High-titer viral solutions for mKO2-hCdt1 (30/120) and mAG-hGem (1/110) were prepared and used for co-transduction into the melanoma cell lines, and subclones were generated by single cell sorting [13, 32, 106].

2.2.1.3 Wound healing migration assay

Experiments were performed using the three melanoma cell lines C8161, 1205Lu and WM983C. For each cell line, three independent experiments were performed. FUCCI-transduced melanoma cells from each cell line were seeded in a 6-well plate to subconfluence. The seeding density was adjusted according to the doubling time for the cell line. The monolayer was scraped with a p200 pipette tip, and images were taken at regular time intervals.

2.2.2 Mathematical model

2.2.2.1 Numerical solutions

Numerical solutions of Equations (2.2) to (2.3) and Equations (2.4) to (2.6) are obtained on a domain, $0 \leq x \leq L$, with grid spacing Δx , and with uniform time steps of duration Δt . The details are in Supporting Material 1.

In this study, the initial condition takes one of three forms, depending on the purpose of the modelling exercise:

- **Modelling a scratch assay:** The first set of modelling results involves using the fundamental mathematical model to mimic a set of experimental data from a scratch assay. For this purpose, we take images from the experiment, manually count numbers of cells in each phase of the cell cycle, and use these numbers to specify $v_r(x, 0)$ and $v_g(x, 0)$. To simulate the experiment, we solve the governing equation numerically on a finite domain, $0 \leq x \leq L$, where L is chosen to match the physical dimension of the experimental image.
- **Exploring the minimum wave speed of travelling wave solutions:** The second set of modelling results involves studying long-time numerical solutions of the mathematical model to examine the possibility of travelling wave solutions. To ensure that we focus on the most biologically relevant travelling-wave solutions, we apply initial conditions with compact support. Further details are provided in the Results and Discussion.
- **Dispersion relation:** Having demonstrated the existence of travelling wave solutions, we then analyse how the long-time travelling wave speed depends on the decay rate of the initial condition for the fundamental model. Further details are provided in the Results and Discussion.

2.3 Results and Discussion

2.3.1 Experimental data

Our experimental data come from 2-D scratch assays performed with FUCCI-transduced melanoma cells. In particular, we use the C8161, 1205Lu and WM983C melanoma cell lines [13]. Here we provide analysis for the C8161 and 1205Lu cell lines, which have very different cell cycle dynamics. We use these two cell lines to demonstrate that our model can predict the cell density for cell lines with a wide range of transition rates. The analysis for the WM983C cell line is in Figure 2.5. The WM983C cell line has cell cycle dynamics intermediate between the C8161 and 1205Lu cell lines.

In these experiments, melanoma cells migrate into a gap created by scratching the cell monolayer, and cell proliferation acts to increase the density of the monolayer. Still images of the scratch assays are obtained at four time points after the scratch is made: 0, 6, 12 and, 18 h for the C8161 cell line, see Figure 2.1(A–D); 0, 16, 32, and 48 h for the 1205Lu cell line, see Figure 2.2(A–D); and 0, 16, 32, and 48 h for the WM983C cell line, see Figure 2.5(A–D). From these images, the nuclei of individual cells can be observed as red (G1 phase), yellow (eS phase) or green (S/G2/M phase). Over the time period of the experiments, the cells migrate into the gap, and cell proliferation is evident from the increasing density of cells behind the moving front.

A notable feature of each of the images is that very few cells appear to be distinctly yellow. Almost all of the yellow cells appear to be either partly green and yellow, or partly red and yellow. This ambiguity motivates us to work with the fundamental mathematical model, which treats the yellow eS phase as part of the red G1 and green S/G2/M phases.

Although the images in Figures 2.1(A–D), 2.2(A–D), and 2.5(A–D) provide some quantitative data about these particular scratch assays for these particular cell lines, the purpose of using a mathematical model is to provide significant generalisations beyond what is possible when working purely with experiments. For example, if we use this kind of data to parameterise a mathematical model, then we ought to be able to use the parameterised mathematical model to make predictions about varying different aspects of the experiment, such as changing the width of the scratch, the timescale of the experiment, or the initial density of the monolayer. To parameterise our mathematical model we require cell density data from the experiment. Although cells are free to move in two dimensions, the geometry of the experiment is such that the cell density is spatially uniform in the vertical, y , direction. We therefore quantify the cell density as a function of horizontal position, x , at various times, t [107]. Overall, this geometrical simplification allows us to describe the cell density data as a function of one spatial coordinate only, and we can therefore use a 1-D mathematical model to describe this kind of data [107].

To obtain cell density data from the images in Figures 2.1(A–D), 2.2(A–D), and 2.5(A–D), we divide each image into a series of vertical columns, each of width $50\text{ }\mu\text{m}$. We manually count the number of cells of each colour in each column. As previously discussed, there is some degree of ambiguity in classifying a cell as yellow, as almost all of the yellow cells appear to be a mixture of either red and yellow, or a mixture of green and yellow. This ambiguity is probably due to the fact that yellow arises from the transient overlap of red and green fluorescence. Consequently, we take the most straightforward approach and classify all of the cells as being either red or green. In this way we can work with just two subpopulations. Given the cell counts, we divide each cell count in each column by the area of that column to give the dimensional, column-averaged, cell density. These estimates of dimensional cell density are then converted into estimates of non-dimensional cell density by dividing through by the carrying capacity density, K . These data are provided in the Supporting Material. To estimate K [108], we assume that the cells are uniformly sized disks, and that the maximum monolayer density corresponds to hexagonal packing of cells. Hexagonal close packing corresponds to one cell at each vertex of the hexagon, and one cell at the centre of the hexagon, meaning that the hexagon contains the equivalent of three cells. The area of the hexagon is $3\sqrt{3}R^2/2$, where R is the circumradius. If the radius of the cells is a then, because $R = 2a$, the carrying capacity is given by $K = 1/(2\sqrt{3}a^2)$. The cell

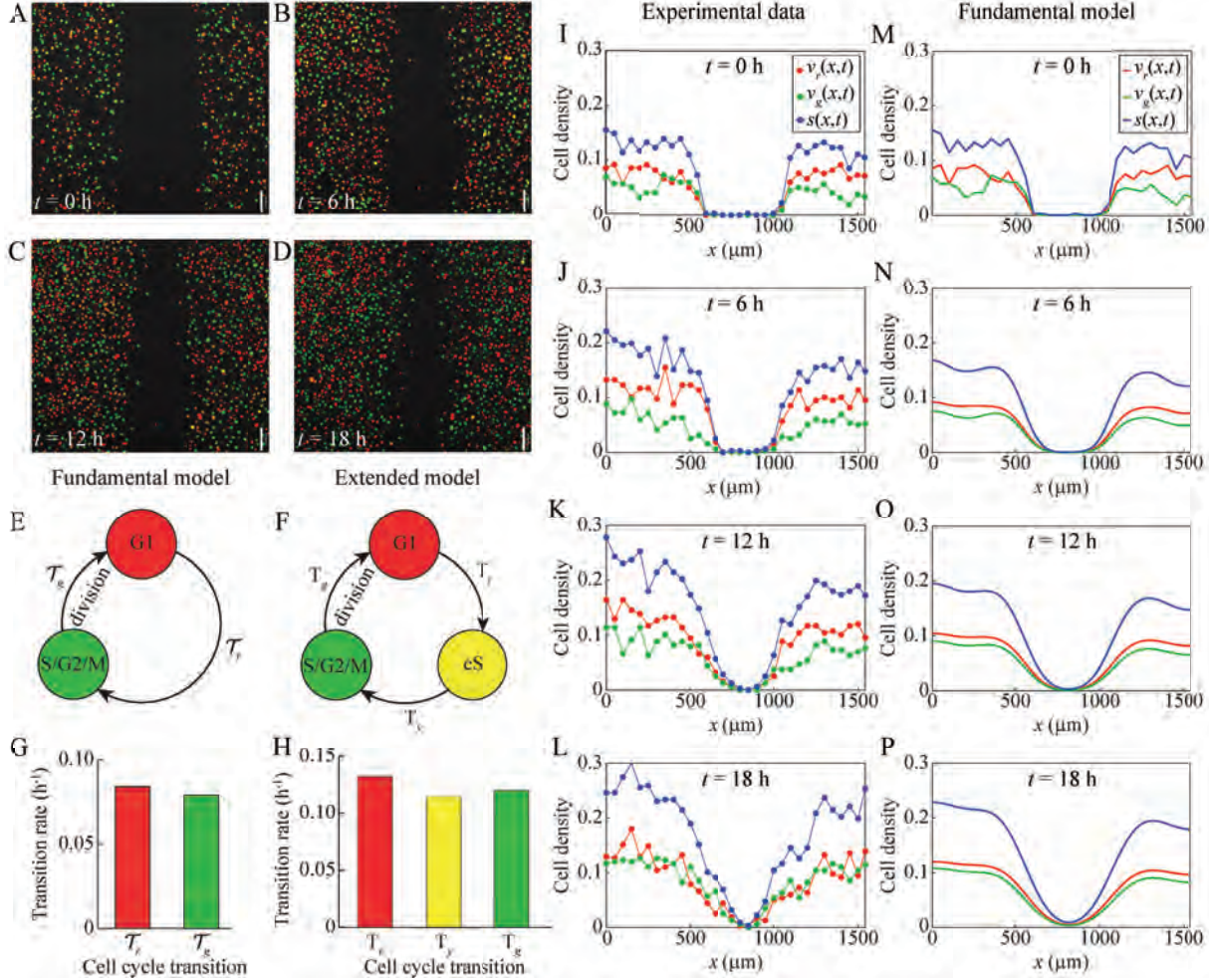


Figure 2.1: Comparison of experimental data and the fundamental model for a scratch assay of FUCCI-transduced C8161 melanoma cells. (A–D) Still images of a scratch assay with FUCCI-transduced C8161 melanoma cells at time 0 h, 6 h, 12 h and 18 h, respectively. Scale bar $150 \mu\text{m}$. (E) Schematic of the fundamental model with two subpopulations indicating the transitions between the cell cycle phases indicated by FUCCI. (F) Schematic of the extended model with three subpopulations indicating the transitions between the cell cycle phases indicated by FUCCI. (G) Estimated transition rates from one cell cycle phase to the next for the fundamental model with two subpopulations, based on data from the C8161 cell line from Figure 1C in [13]. (H) Estimated transition rates from one cell cycle phase to the next for the extended model with three subpopulations, based on data from the C8161 cell line from Figure 1C in [13]. (I–L) Experimental non-dimensional cell density data at 0 h, 6 h, 12 h and 18 h, respectively (based on images (A–D)). The cell density is treated as a function of x and t only, owing to the fact that the initial density does not depend on the vertical coordinate, y . (M–P) Numerical solutions of the fundamental model, Equations (2.2) to (2.3), at 0 h, 6 h, 12 h and 18 h. The numerical solutions are obtained with $\Delta x = 0.1 \mu\text{m}$, $\Delta t = 0.1$ h, $L = 1542 \mu\text{m}$, diffusion coefficients $\mathcal{D}_r = \mathcal{D}_g = 400 \mu\text{m}^2 \text{h}^{-1}$, transition rates $\kappa_r = 0.084 \text{h}^{-1}$ and $\kappa_g = 0.079 \text{h}^{-1}$, and initial conditions the same as for the experimental data.

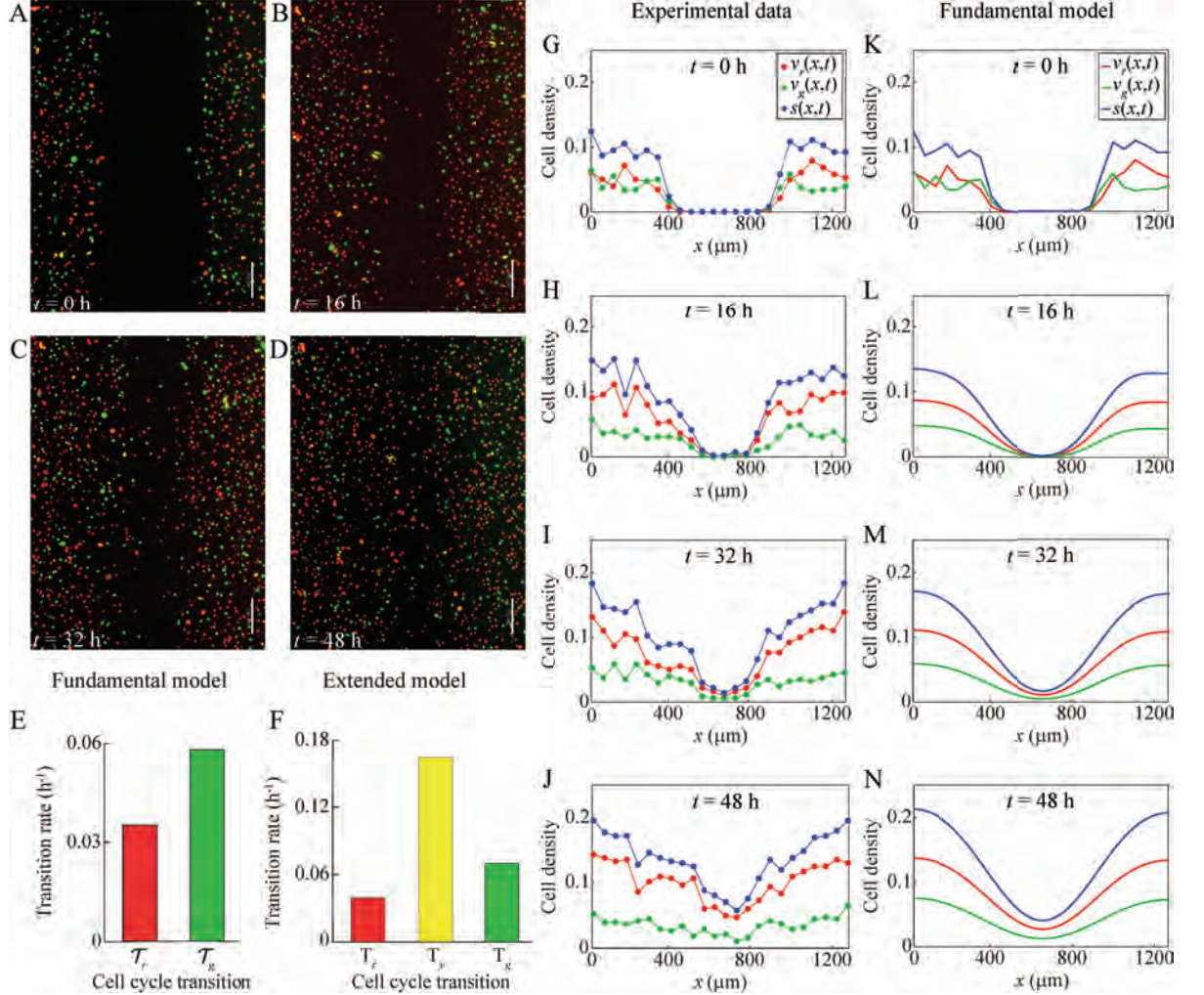


Figure 2.2: Comparison of experimental data and the fundamental model for a scratch assay of FUCCI-transduced 1205Lu melanoma cells. (A–D) Still images of a scratch assay with FUCCI-transduced 1205Lu melanoma cells at time 0 h, 16 h, 32 h and 48 h, respectively. Scale bar is $200 \mu\text{m}$. (E) Estimated transition rates from one cell cycle phase to the next for the fundamental model with two subpopulations, based on data from the 1205Lu cell line from Figure 1C in [13]. (F) Estimated transition rates from one cell cycle phase to the next for the extended model with three subpopulations, based on data from the 1205Lu cell line from Figure 1C in [13]. (G–J) Experimental non-dimensional cell density data at 0 h, 16 h, 32 h and 48 h, respectively (based on images (A–D)). The cell density is treated as a function of x and t only, owing to the fact that the initial density does not depend on the vertical coordinate, y . (K–N) Numerical solutions of the fundamental model, Equations (2.2) to (2.3), at 0 h, 16 h, 32 h and 48 h. The numerical solutions are obtained with $\Delta x = 0.1 \mu\text{m}$, $\Delta t = 0.1 \text{ h}$, $L = 1254 \mu\text{m}$, diffusion coefficients $\mathcal{D}_r = \mathcal{D}_g = 400 \mu\text{m}^2 \text{h}^{-1}$, transition rates $\kappa_r = 0.035 \text{ h}^{-1}$ and $\kappa_g = 0.058 \text{ h}^{-1}$, and initial conditions the same as for the experimental data.

diameter of C8161 melanoma cells is approximately $17\text{ }\mu\text{m}$ [109]. So, with $a = 8.5\text{ }\mu\text{m}$ we have $K = 0.004\text{ cells }\mu\text{m}^{-2}$, to one significant figure. We use the same value of K for the 1205Lu and WM983C melanoma cell lines, as with a microscope we observe that the cells from the three cell lines are of similar size.

Figure 2.1(I–L) shows the resulting experimental cell density profiles for C8161 cells at 0, 6, 12, and 18 h, respectively. Figure 2.2(G–J) shows the resulting experimental cell density profiles for 1205Lu cells at 0, 16, 32, and 48, respectively. Similar profiles for the WM983C cells are in Figure 2.5(G–J). These figures provide quantitative information about how the population of cells migrates into the gap, while simultaneously proliferating to increase the density of the spreading monolayer, as we observed qualitatively in Figures 2.1(A–D), 2.2(A–D), and 2.5(A–D). These cell density profiles, however, quantify the cell density of each subpopulation, and allow the changes in cell density over time to be determined quantitatively, which is not possible through visual interpretation of the images.

Movies of scratch assays associated with the three cell lines considered in this work are provided in Supporting Material.

2.3.2 Model development

We now describe the FUCCI scratch assays using the fundamental model. As previously explained, cells in the experiment move in two dimensions. However, the geometry of the experiment means that the cell density is spatially uniform in the vertical, y , direction. Therefore, we can model the experiment with a 1-D model, where the independent variables are time, t , and the horizontal coordinate, x [107].

As summarised in the Introduction, in the FUCCI system, red and green fluorescent proteins are fused to different regulators of the cell cycle so that a cell in G1 phase fluoresces red, and a cell in S/G2/M phase fluoresces green. During the cell cycle transition from G1 phase to S phase, referred to as eS phase, the red FUCCI signal decreases and the green FUCCI signal increases, producing varying shades of yellow fluorescence, ranging from darker yellow to lighter yellow. The indication of the eS (yellow) phase is therefore secondary, as it arises from the overlap of red and green fluorescence. In the experimental images, it is difficult to identify the cells in eS phase as very few cells appear distinctly yellow, rather appearing shades of either red or green. For these reasons, our fundamental mathematical model is

$$\frac{\partial v_r}{\partial t} = \mathcal{D}_r \frac{\partial^2 v_r}{\partial x^2} - \kappa_r v_r + 2\kappa_g v_g(1 - s), \quad (2.2)$$

$$\frac{\partial v_g}{\partial t} = \mathcal{D}_g \frac{\partial^2 v_g}{\partial x^2} - \kappa_g v_g(1 - s) + \kappa_r v_r, \quad (2.3)$$

where $v_r(x, t)$ and $v_g(x, t)$ are the non-dimensional cell densities of the coupled subpopulations corresponding to the G1 (red) and S/G2/M (green) phases of the cell cycle, respectively. The total density is $s(x, t) = v_r(x, t) + v_g(x, t)$. The rate at which cells in the G1 phase transition to the S/G2/M phase is κ_r , and the rate at which cells in the S/G2/M phase transition to the G1 phase is κ_g . The diffusion coefficients are \mathcal{D}_r for cells in phase G1, and \mathcal{D}_g for cells in phase S/G2/M. Although we are free to choose any realistic values for the diffusion coefficients, it is pertinent to set $\mathcal{D}_r = \mathcal{D}_g$, because we are considering subpopulations of cells of the same type which differ only with respect to their cell cycle phase. Indeed, cells from various melanoma cell lines, including C8161, 1205Lu and WM983C, appear to migrate independently of the cell cycle phases. For example, see the data in Figure 6 and Figure S3 in [13]. We will employ this biologically motivated simplifying assumption at various points in our analysis. Note that in Equation (2.2), the factor of two in the positive source term corresponds to a cell in phase S/G2/M undergoing division to produce two daughter cells in the phase G1, thereby doubling the local density.

Despite the challenges associated with the observation of the yellow eS phase, we also consider a mathematical model for three coupled subpopulations corresponding to G1, eS and S/G2/M. We refer to this model as our extended model:

$$\frac{\partial u_r}{\partial t} = D_r \frac{\partial^2 u_r}{\partial x^2} - k_r u_r + 2k_g u_g(1 - s), \quad (2.4)$$

$$\frac{\partial u_y}{\partial t} = D_y \frac{\partial^2 u_y}{\partial x^2} - k_y u_y + k_r u_r, \quad (2.5)$$

$$\frac{\partial u_g}{\partial t} = D_g \frac{\partial^2 u_g}{\partial x^2} - k_g u_g(1 - s) + k_y u_y. \quad (2.6)$$

Here, $u_r(x, t)$, $u_y(x, t)$ and $u_g(x, t)$ are the non-dimensional cell densities of the coupled subpopulations corresponding to the G1 (red), eS (yellow) and S/G2/M (green) phases of the cell cycle, respectively, with total cell density $s(x, t) = u_r(x, t) + u_y(x, t) + u_g(x, t)$. The G1 to eS transition rate is k_r , the eS to S/G2/M transition rate is k_y and the S/G2/M to G1 transition rate is k_g . The diffusion coefficients are D_r for cells in phase G1, D_y for cells in phase eS, and D_g for cells in phase S/G2/M. Once again, we have no restriction on the choice of the values for the diffusion coefficients. The subpopulations consist of cells of the same type, they are only in different phases of the cell cycle. So, as discussed above for the fundamental model, the most obvious choice is to set $D_r = D_y = D_g$. As in Equation (2.2), the factor of two in Equation (2.4) corresponds to a cell in phase S/G2/M undergoing division to produce two daughter cells in the phase G1.

2.3.3 Model application

We illustrate our mathematical model in Figures 2.1, 2.2, and 2.5, where we compare experimental data for scratch assays of FUCCI-transduced melanoma cells with the fundamental model Equations (2.2) to (2.3). Figure 2.1 corresponds to the C8161 melanoma cell line, Figure 2.2 to the 1205Lu melanoma cell line, and Figure 2.5 to the WM983C melanoma cell line.

Figure 2.1(E) describes, schematically, the fundamental model with two subpopulations, indicating the transitions between the cell cycle phases indicated by FUCCI. We denote the transitions between the cell cycle phases as

$$\mathcal{T}_r : G_1 \xrightarrow{\kappa_r} S/G_2/M \quad \text{and} \quad \mathcal{T}_g : S/G_2/M \xrightarrow{\kappa_g} G_1, \quad (2.7)$$

where κ_r is the G1 to S/G2/M transition rate, and κ_g is the S/G2/M to G1 transition rate. Figure 2.1(F) is a similar schematic for the extended model with three subpopulations, where we denote the transitions between the cell cycle phases as

$$T_r : G_1 \xrightarrow{k_r} eS, \quad T_y : eS \xrightarrow{k_y} S/G_2/M \quad \text{and} \quad T_g : S/G_2/M \xrightarrow{k_g} G_1, \quad (2.8)$$

where k_r , k_y and k_g are the G1 to eS, eS to S/G2/M and S/G2/M to G1 transition rates, respectively.

The estimated transition rates from one cell cycle phase to the next phase are based on FUCCI data for the C8161, 1205Lu and WM983C melanoma cell lines from Figure 1C in [13]. These data report the duration of time spent in each cell cycle phase for at least 20 individual cells. To estimate the transition rate from one cell cycle phase to the next, we first calculate the arithmetic mean of the data in [13], giving the mean times: t_r for the G1 phase, t_y for the eS phase, and t_g for the S/G2/M phase. We then estimate the transition rates for the extended model as $k_r = (\ln 2)/t_r$ for the G1 to eS transition, $k_y = (\ln 2)/t_y$ for the eS to S/G2/M transition, and $k_g = (\ln 2)/t_g$ for the S/G2/M to G1 transition. The factor of $\ln 2$ arises since this data corresponds to cells in a low density environment, and so the cells are likely to be proliferating exponentially. To obtain estimates of the transition rates for the fundamental model, we assume that half of the time spent in the eS phase contributes to the time spent in the G1 phase, and the other half of the time spent in the eS phase contributes to the time spent in the S/G2/M phase. This means that we have $\kappa_r = (\ln 2)/(t_r + t_y/2)$ for the G1 to S/G2/M transition, and $\kappa_g = (\ln 2)/(t_g + t_y/2)$ for the S/G2/M to G1 transition. The estimated transition rates for the fundamental and extended models are shown in Figure 2.1(G–

H), respectively, for the C8161 cell line, in Figure 2.2(E–F), respectively, for the 1205Lu cell line, and in Figure 2.5(E–F), respectively, for the WM983C cell line. We can express the transition rates for the fundamental model in terms of those for the extended model as $\kappa_r = 2k_r k_y / (k_r + 2k_y)$ and $\kappa_g = 2k_g k_y / (k_g + 2k_y)$.

Previous studies examining the migration of various melanoma cell lines suggest that the diffusion coefficients lie within the range $100\text{--}500 \mu\text{m}^2\text{h}^{-1}$ [110, 111]. Therefore, we will take an intermediate value and assume that the diffusivity of our melanoma cell lines is approximately $400 \mu\text{m}^2\text{h}^{-1}$.

For comparison with the experimental data for C8161 cells in Figure 2.1(I–L), we show the numerical solutions of the fundamental model, Equations (2.2) to (2.3), in Figure 2.1(M–P). Overall, the numerical solutions of Equations (2.2) to (2.3) in Figure 2.1(M–P) compare well with the corresponding experimental data in Figure 2.1(I–L). As time increases, the solution of the mathematical model shows the cell density profile spreading into the gap, and the cell density of each subpopulation increases throughout the domain due to proliferation. The relative size of each subpopulation also compares well between the numerical solutions and experimental data. Similarly, for comparison with the experimental data for 1205Lu cells in Figure 2.2(G–J), we show the numerical solutions of the fundamental model, Equations (2.2) to (2.3), in Figure 2.2(K–N). Once again, the numerical solutions of Equations (2.2) to (2.3) in Figure 2.2(K–N) compare well with the corresponding experimental data in Figure 2.2(G–J). A similar comparison for the WM983C cell line is in Figure 2.5.

We note that other studies use partial differential equations to model scratch assays [46, 47, 107]. None of these previous models, however, include the cell cycle phases of cells. Our results show potential for our model to successfully describe cell migration and proliferation along with cell cycle dynamics. With our model we can easily simulate experiments which would otherwise be expensive and time consuming. In particular, we can simulate experiments over longer periods of time, with different scratch widths, and with different parameters to accommodate different cell lines. Our model also provides quantitative data such as cell densities at any time and position within the domain.

2.3.4 Analysis

2.3.4.1 Numerical solutions

A key feature that can be observed in scratch assays initialised with a sufficiently wide scratch is the formation of a moving front of cells [46]. This is important, because similar observations are relevant to malignant invasion and wound healing [46]. The need to understand the key factors

that drive moving fronts of cells, motivates an analysis of our models for scenarios in which a single front propagates along a wide domain.

We now simulate solutions of our models on a much wider domain using a different initial condition to examine the existence of travelling wave solutions. There are many choices of initial condition; however, for the fundamental model, we set

$$v_r(x, 0) = v_g(x, 0) = \begin{cases} 0.5, & 0 \leq x < \xi, \\ 0, & \xi \leq x \leq L, \end{cases} \quad (2.9)$$

and for the extended model we set

$$u_r(x, 0) = \begin{cases} 0.4, & 0 \leq x < \xi, \\ 0, & \xi \leq x \leq L, \end{cases} \quad u_y(x, 0) = u_g(x, 0) = \begin{cases} 0.3, & 0 \leq x < \xi, \\ 0, & \xi \leq x \leq L, \end{cases} \quad (2.10)$$

where the choice of the constant ξ is not critical. Since the transition rates κ_r and κ_g have similar values for a given cell line, the subpopulations corresponding to G1 and S/G2/M will have similar densities in a population of freely cycling cells, so an initial condition for the fundamental model with equal densities for the two subpopulations is biologically reasonable. Correspondingly, the transition rates k_r , k_y , and k_g have similar values for a given cell line, so an initial condition for the extended model with similar densities for the three subpopulations is also biologically reasonable, noting that the yellow cells are due to the overlap of red and green fluorescence so are not all visually identifiable from the images. Because we use a numerical approach to explore travelling wave solutions, we set L to be sufficiently large so that the moving front does not interact with the boundary at $x = L$.

For this purpose, we obtain numerical solutions of the fundamental model, Equations (2.2) to (2.3), with typical solutions presented in Figure 2.3(A). These results suggest that the solutions develop into an interesting travelling wave profile, where the density of cells in the G1 phase forms a moving pulse, whereas the density of cells in the S/G2/M phase forms a moving wavefront. Furthermore, the total cell density profile also moves as a wavefront profile. Analogous solutions of the extended model, Equations (2.4) to (2.6), are shown in Figure 2.3(B), which have similar features except that there is an additional pulse arising from cells in the eS phase. We find that the existence of these travelling waves is robust, and does not depend on the values of the diffusion coefficients. Figure 2.6 shows solutions of the fundamental model for $\mathcal{D}_g \neq \mathcal{D}_r$.

The appearance of travelling wave solutions that take the form of a wavefront is not un-

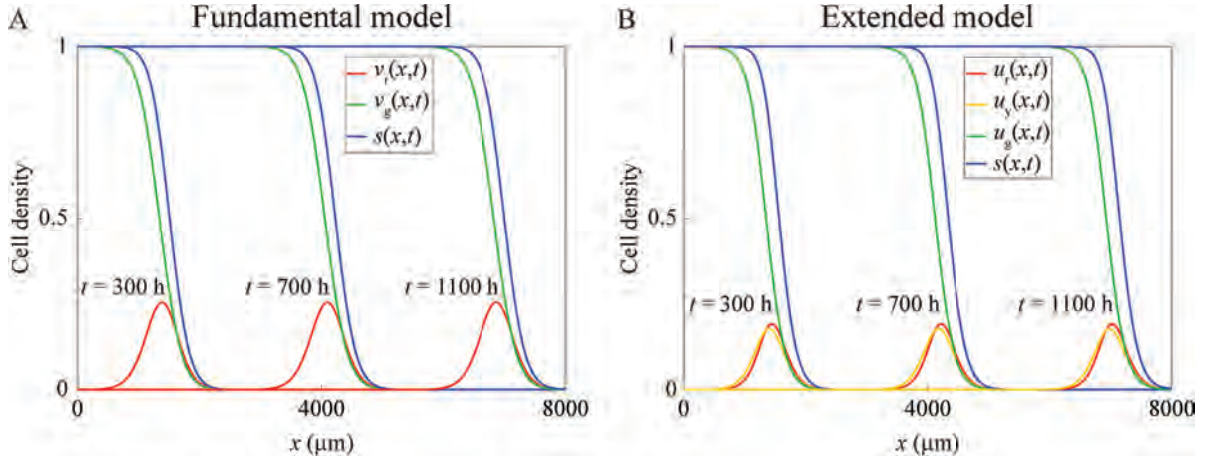


Figure 2.3: Numerical solutions demonstrating travelling wave behaviour for the fundamental and extended models. (A) Numerical solutions of the fundamental model, Equations (2.2) to (2.3), obtained with $\Delta x = 1 \mu\text{m}$, $\Delta t = 1 \text{ h}$, $L = 8000 \mu\text{m}$, $\mathcal{D}_r = \mathcal{D}_g = 400 \mu\text{m}^2 \text{ h}^{-1}$, $\kappa_r = \kappa_g = 0.08 \text{ h}^{-1}$, and the initial condition given by Equation (2.9) with $\xi = 10$. (B) Numerical solutions of the extended model, Equations (2.4) to (2.6), obtained with $\Delta x = 1 \mu\text{m}$, $\Delta t = 1 \text{ h}$, $L = 8000 \mu\text{m}$, $\mathcal{D}_r = \mathcal{D}_y = \mathcal{D}_g = 400 \mu\text{m}^2 \text{ h}^{-1}$, $k_r = k_y = k_g = 0.13 \text{ h}^{-1}$, and the initial condition given by Equation (2.10) with $\xi = 10$.

expected, as the partial differential equations in our model are related to the FKPP equation, Equation (2.1), which is well-known to exhibit travelling-wave solutions with a wavefront form [38, 48]. It is particularly interesting, however, that our model also exhibits travelling wave solutions with the form of a pulse, which are not observed for the FKPP equation. The pulses arise because only the cells near to the leading edge, where $s(x, t) < 1$, have the opportunity to cycle from S/G2/M to G1, which involves cell division and can be inhibited by crowding effects. Behind the wavefront where $s(x, t)$ approaches unity, cells do not have enough space to divide, and so these cells remain in the S/G2/M phase.

2.3.4.2 Travelling wave analysis of the fundamental model

We now analyse the fundamental model, Equations (2.2) to (2.3), with the aim of understanding how the parameters in the model relate to the speed of the travelling wave solutions. To simplify our analysis we non-dimensionalise the problem by defining the new variables $t^* = \kappa_g t$ and $x^* = x \sqrt{\kappa_g / \mathcal{D}_r}$, and the parameters

$$\mathcal{D} = \frac{\mathcal{D}_g}{\mathcal{D}_r} \quad \text{and} \quad \kappa = \frac{\kappa_r}{\kappa_g}, \quad (2.11)$$

to give the corresponding non-dimensional model

$$\frac{\partial v_r}{\partial t} = \frac{\partial^2 v_r}{\partial x^2} - \kappa v_r + 2v_g(1 - s), \quad (2.12)$$

$$\frac{\partial v_g}{\partial t} = \mathcal{D} \frac{\partial^2 v_g}{\partial x^2} - v_g(1 - s) + \kappa v_r, \quad (2.13)$$

in which the asterisks have been omitted for notational simplicity.

To examine the travelling wave solutions, we define the travelling wave coordinate, $z = x - ct$, where $c > 0$ is the wave speed associated with a travelling wave that propagates in the positive x -direction. We seek solutions of Equations (2.12) to (2.13) of the form $v_r(x, t) = U(z)$ and $v_g(x, t) = V(z)$. Such solutions, if they exist, correspond to travelling wave solutions. Substituting $U(z)$ and $V(z)$ into Equations (2.12) to (2.13) gives the system of ordinary differential equations

$$U'' + cU' - \kappa U + 2V(1 - U - V) = 0, \quad (2.14)$$

$$V'' + \frac{c}{\mathcal{D}}V' + \frac{\kappa}{\mathcal{D}}U - \frac{1}{\mathcal{D}}V(1 - U - V) = 0. \quad (2.15)$$

We want to find travelling-wave solutions which satisfy the conditions

$$U > 0, \quad \lim_{z \rightarrow -\infty} U(z) = 0 \quad \text{and} \quad \lim_{z \rightarrow \infty} U(z) = 0, \quad (2.16)$$

$$V > 0, \quad \lim_{z \rightarrow -\infty} V(z) = 1 \quad \text{and} \quad \lim_{z \rightarrow \infty} V(z) = 0. \quad (2.17)$$

Letting $W = U'$ and $X = V'$, we can write Equations (2.14) to (2.15) as a system of first-order equations:

$$U' = W, \quad (2.18)$$

$$V' = X, \quad (2.19)$$

$$W' = -cW + \kappa U - 2V(1 - U - V), \quad (2.20)$$

$$X' = -\frac{c}{\mathcal{D}}X - \frac{\kappa}{\mathcal{D}}U + \frac{1}{\mathcal{D}}V(1 - U - V). \quad (2.21)$$

The equilibrium points of Equations (2.18) to (2.21) are $(0, 0, 0, 0)$ and $(0, 1, 0, 0)$. Of all the solutions to Equations (2.18) to (2.21) in the four-dimensional phase space, (U, V, W, X) , we seek a heteroclinic orbit from $(0, 1, 0, 0)$ to $(0, 0, 0, 0)$ which has the physically-relevant property that $U > 0$ and $V > 0$.

The Jacobian of Equations (2.18) to (2.21) evaluated at $(0, 0, 0, 0)$ is

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \kappa & -2 & -c & 0 \\ -\kappa/\mathcal{D} & 1/\mathcal{D} & 0 & -c/\mathcal{D} \end{bmatrix}. \quad (2.22)$$

The eigenvalues of Equation (2.22) are the solutions of the corresponding characteristic equation

$$\mathcal{D}\lambda^4 + c(\mathcal{D} + 1)\lambda^3 + (c^2 - 1 - \mathcal{D}\kappa)\lambda^2 - c(1 + \kappa)\lambda - \kappa = 0. \quad (2.23)$$

To establish a condition for physically relevant travelling-wave solutions, where $U > 0$ and $V > 0$, we examine whether the solutions of Equation (2.23) are either complex or real. The analytical solutions of this quartic equation [112] are quite complicated when $\mathcal{D} \neq 1$. Because we are considering subpopulations of the same cell type which differ only with respect to cell cycle phase, the biologically relevant case is when the two diffusion coefficients are equal, leading to $\mathcal{D} = 1$. In this case, by defining

$$\alpha^\pm(\kappa, c) = 2\kappa + c^2 + 2 \pm 2\sqrt{\kappa^2 + 6\kappa + 1}, \quad (2.24)$$

we can express the solutions of the quartic simply as

$$\lambda_1^\pm = -\frac{1}{2}c \pm \frac{1}{2}\sqrt{\alpha^-(\kappa, c)} \quad \text{and} \quad \lambda_2^\pm = -\frac{1}{2}c \pm \frac{1}{2}\sqrt{\alpha^+(\kappa, c)}. \quad (2.25)$$

The eigenvalues λ_2^+ and λ_2^- are always real, and λ_1^+ and λ_1^- are real when $\alpha^-(\kappa, c) \geq 0$. If $\alpha^-(\kappa, c) < 0$, however, then λ_1^+ and λ_1^- are complex eigenvalues which yield solutions with oscillatory behaviour, which necessarily involves negative cell densities. Because we are interested in travelling wave solutions for which U and V remain positive, we shall therefore require that $\alpha^-(\kappa, c) \geq 0$. It is then reasonable to suspect that, for a given $\kappa \geq 0$, the value of c such that $\alpha^-(\kappa, c) = 0$ is the minimum wave speed for the travelling waves, which we denote as

$$c_{\min}(\kappa) = \sqrt{-2\kappa - 2 + 2\sqrt{\kappa^2 + 6\kappa + 1}}. \quad (2.26)$$

The minimum wave speed is bounded above, and in fact $c_{\min}(\kappa) \rightarrow 2^-$ as $\kappa \rightarrow \infty$.

Equation (2.26) shows that the minimum speed of the travelling wave solution depends on κ , which is the ratio of the time the cells spend in phase S/G2/M to the time the cells spend in

phase G1. In other words, the minimum wave speed depends on the cell cycle dynamics of the particular cells under consideration. We observe here that $\alpha^-(\kappa, c) \geq 0$ is a necessary condition for the existence of travelling waves. We have not demonstrated, however, that this condition is sufficient for the existence of travelling waves. This would require a formal proof of existence for the travelling waves, which is beyond the scope of this work.

Observe that if $\kappa > 0$ and $c \geq c_{\min}$ then λ_1^+ , λ_1^- and λ_2^- are real and negative, and λ_2^+ is real and positive. Therefore, the equilibrium point $(0, 0, 0, 0)$ is hyperbolic and has a 3-D stable manifold and a 1-D unstable manifold. The presence of a stable manifold at the point $(0, 0, 0, 0)$ is necessary for the existence of a heteroclinic orbit and the real eigenvalues allow for this orbit to correspond to physically relevant travelling wave solutions with $U > 0, V > 0$. This analysis does not constitute a formal proof of existence for the travelling waves, however it does show that our observations are consistent with their existence.

The Jacobian of the system Equations (2.18) to (2.21) evaluated at $(0, 1, 0, 0)$ is

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \kappa + 2 & 2 & -c & 0 \\ -(\kappa + 1)/\mathcal{D} & -1/\mathcal{D} & 0 & -c/\mathcal{D} \end{bmatrix}. \quad (2.27)$$

The eigenvalues of Equation (2.27) are the solutions of the corresponding characteristic equation

$$\mathcal{D}\lambda^4 + c(\mathcal{D} + 1)\lambda^3 + (c^2 + 1 - \mathcal{D}\kappa - 2\mathcal{D})\lambda^2 - c(1 + \kappa)\lambda + \kappa = 0. \quad (2.28)$$

Once again, the analytical solutions of this quartic equation [112] are quite complicated when $\mathcal{D} \neq 1$. For the biologically-relevant case $\mathcal{D} = 1$, however, the solutions are

$$\lambda_3^\pm = \frac{1}{2}(-c \pm \sqrt{c^2 + 4}) \quad \text{and} \quad \lambda_4^\pm = \frac{1}{2}(-c \pm \sqrt{c^2 + 4\kappa}). \quad (2.29)$$

If $\kappa > 0$ then λ_3^- and λ_4^- are real and negative, and λ_3^+ and λ_4^+ are real and positive. Therefore, the equilibrium point $(0, 1, 0, 0)$ is hyperbolic and has a 2-D stable manifold and a 2-D unstable manifold. The existence of an unstable manifold at the point is necessary for the presence of a travelling wave solution.

Figure 2.4(A) compares $c_{\min}(\kappa)$ in Equation (2.26) with the wave speed estimates obtained from the numerical solutions of the partial differential equations [94]. The numerical solutions are obtained using initial conditions with compact support, so we would expect the resulting travelling waves to have the minimum wave speed [38]. We observe that the numerically esti-

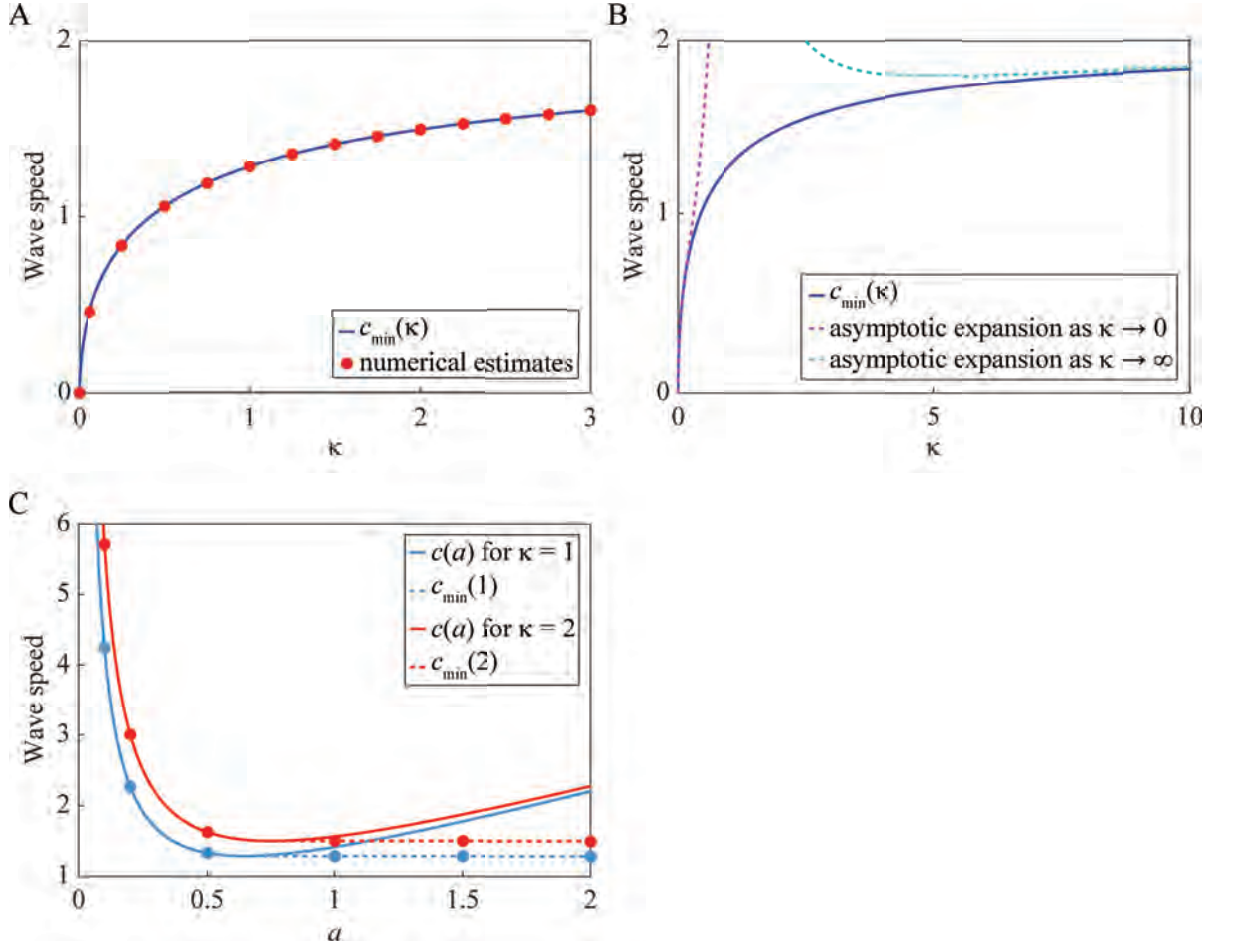


Figure 2.4: Minimum wave speed and the dispersion relation. (A) Comparison of $c_{\min}(\kappa)$ from Equation (2.26) with the numerically-estimated wave speed for $s(x, t) = v_r(x, t) + v_g(x, t)$. Numerical solutions of Equations (2.12) to (2.13) are obtained with $\Delta x = 0.1$, $\Delta t = 0.001$ and $\mathcal{D} = 1$. Further, the initial condition is Equation (2.9) with $\xi = 10$. For $\kappa = 0$ there is no travelling wave, so we set $c = 0$. The numerical solutions are considered beginning with $\kappa = 0.06$, and then with increasing values of κ from 0.25 to 3, with increments of 0.25. From these numerical solutions we estimate the wave speed for $s(x, t) = v_r(x, t) + v_g(x, t)$ by using linear interpolation to find the position corresponding to $s(x, t) = 0.5$ on the wave for various times [94]. (B) Asymptotic expansions for $c_{\min}(\kappa)$ as $\kappa \rightarrow 0$ and $\kappa \rightarrow \infty$. (C) Comparison of c from Equation (2.35) with the wavespeed estimated using numerical solutions and with c_{\min} from Equation (2.26). Solutions are given for $\kappa = 1$ (blue) and $\kappa = 2$ (red). The continuous curves show c from Equation (2.35). The dots represent the wave speed from numerical solutions obtained with $\Delta x = 0.1$, $\Delta t = 0.001$, $\mathcal{D} = 1$, and initial conditions of the form Equation (2.36) with $a = 0.1, 0.2, 0.5, 1, 1.5$ and 2 . The dotted horizontal lines show c_{\min} from Equation (2.26).

mated wave speed is consistent with Equation (2.26) over the range of κ we consider. Therefore, we have numerical evidence to strongly support the claim that the minimum speed is given by Equation (2.26).

In Figure 2.4(B) we show the asymptotic expansions of Equation (2.26) as $\kappa \rightarrow 0$ and $\kappa \rightarrow \infty$:

$$c_{\min}(\kappa) = 2\kappa^{\frac{1}{2}} - 2\kappa^{\frac{3}{2}} + 5\kappa^{\frac{5}{2}} + O\left(\kappa^{\frac{7}{2}}\right) \quad \text{as } \kappa \rightarrow 0, \quad (2.30)$$

$$c_{\min}(\kappa) = 2 - \frac{2}{\kappa} + \frac{5}{\kappa^2} + O\left(\frac{1}{\kappa^3}\right) \quad \text{as } \kappa \rightarrow \infty. \quad (2.31)$$

Thus, $c_{\min}(\kappa)$ behaves like $2\kappa^{\frac{1}{2}}$ when κ is small and like $2 - 2/\kappa$ when κ is large. We discuss the connection with the FKPP equation in the Conclusion.

2.3.4.3 Dispersion relation

Here we investigate the relationship between the initial conditions for Equations (2.12) to (2.13) with $\mathcal{D} = 1$, and the speed of the resulting travelling wave solution. Our approach is to examine the leading edge of the travelling wave, assuming that the initial conditions at infinity have an exponential form [38].

At the leading edge of the evolving waves, $s(x, t) = v_r(x, t) + v_g(x, t) \ll 1$, so we can linearise Equations (2.12) to (2.13) to give

$$\frac{\partial \tilde{v}_r}{\partial t} = \frac{\partial^2 \tilde{v}_r}{\partial x^2} - \kappa \tilde{v}_r + 2\tilde{v}_g, \quad (2.32)$$

$$\frac{\partial \tilde{v}_g}{\partial t} = \frac{\partial^2 \tilde{v}_g}{\partial x^2} - \tilde{v}_g + \kappa \tilde{v}_r. \quad (2.33)$$

Assuming initial conditions of the form $\tilde{v}_r(x, 0) \sim Ae^{-ax}$ and $\tilde{v}_g(x, 0) \sim Be^{-bx}$ as $x \rightarrow \infty$ for arbitrary positive constants a , b , A and B , we seek travelling wave solutions satisfying Equations (2.32) to (2.33) of the form

$$\tilde{v}_r(x, t) = Ae^{-a(x-ct)} \quad \text{and} \quad \tilde{v}_g(x, t) = Be^{-b(x-ct)}, \quad (2.34)$$

corresponding to the leading edges of the pulse and wavefront solutions, respectively. Substituting Equation (2.34) into Equations (2.32) to (2.33) and solving for c gives

$$c = a - \frac{1}{a} + \frac{1}{2a}(1 - \kappa + \sqrt{\kappa^2 + 6\kappa + 1}). \quad (2.35)$$

The dispersion relation, Equation (2.35), depends only on a and κ , so we can obtain the travelling wave solutions with the form Equation (2.34) from initial conditions of the form

$$v_r(x, 0) = 0 \quad \text{and} \quad v_g(x, 0) = \begin{cases} 1, & 0 \leq x < \xi, \\ e^{-a(x-\xi)}, & \xi \leq x \leq L, \end{cases} \quad (2.36)$$

for constants ξ and $a > 0$. Note that for large a , this initial condition for v_g is approximately a Heaviside function. For a given κ , we observe that the minimum wave speed according to the dispersion relation, Equation (2.35), is equal to the minimum wave speed c_{\min} , given by Equation (2.26).

In Figure 2.4(C) we compare c from Equation (2.35) with estimates of the wave speed from numerical solutions of the governing partial differential equations, and with c_{\min} from Equation (2.26). The value of c given by the dispersion relation tends to infinity as both $a \rightarrow 0$ and $a \rightarrow \infty$, and has a unique minimum value for a given κ . The dispersion relation for the FKPP equation has similar properties [38]. For a given κ , the numerical estimates of the wave speed agree with Equation (2.35) for increasing values of $a > 0$, until the minimum wave speed is obtained. As we further increase a , the numerical estimates of the wave speed remain constant, at the minimum value of the wave speed. Once again, the dispersion relation for the FKPP equation has similar properties [38].

2.4 Conclusion

Here we present, to our knowledge, the first mathematical model of cell migration that can be used to quantitatively describe experiments using FUCCI technology, which highlights the spatial and temporal distribution of individual cells in different parts of the cell cycle. The fundamental model consists of two coupled partial differential equations, each of which governs the subpopulations of cells corresponding to the two phases of the cell cycle that are directly labelled by FUCCI. Our study suggests that the model can describe cell migration and cell proliferation in a way that highlights the spatial and temporal distribution of two subpopulations. In particular, we show that the model can describe the dynamics of scratch assays performed with cells highlighted by FUCCI. This is a useful outcome, as scratch assays are routinely employed to study cell migration, for example in the context of malignant invasion [107] and wound healing [46]. Although a typical scratch assay may require several days to grow cells, and to perform and record the experiment, our model can simulate such an experiment on a single desktop computer in a few seconds. In addition, we can easily vary the parameters in our model to simulate experiments over any period of time, any scratch width, any geometry, and any cell line, provided that we have some information about the amount of time that is spent in each phase of the cell cycle. Therefore, this kind of computational modelling tool can provide valuable information to assist in the design and interpretation of these kinds of experiments conducted with FUCCI.

In this work, we use numerical results to demonstrate the existence of travelling wave solutions. Furthermore, our analysis shows that the minimum wave speed depends on the ratio of the time spent in each of the G1 and S/G2/M phases. Another outcome of this study is that we derive an analytical expression for the minimum wave speed as a function of this ratio which,

in dimensional variables, can be written as

$$c_{\min} = \sqrt{2\mathcal{D}_r \left(-\kappa_r - \kappa_g + \sqrt{\kappa_r^2 + 6\kappa_r\kappa_g + \kappa_g^2} \right)}. \quad (2.37)$$

This relationship is based on the biologically reasonable assumption that $\mathcal{D}_r = \mathcal{D}_g$, where \mathcal{D}_r is the diffusion coefficient for cells in phase G1 and \mathcal{D}_g is the diffusion coefficient for cells in phase S/G2/M. Further, κ_r is the transition rate from G1 to S/G2/M and κ_g is the transition rate from S/G2/M to G1. It then follows from Equations (2.30) to (2.31) that $c_{\min} \sim 2\sqrt{\kappa_r}\mathcal{D}_r(1 - \kappa_r/\kappa_g)$ as $\kappa_r/\kappa_g \rightarrow 0$, and $c_{\min} \sim 2\sqrt{\kappa_g}\mathcal{D}_g(1 - \kappa_g/\kappa_r)$ as $\kappa_g/\kappa_r \rightarrow 0$. Therefore, when $\kappa_r/\kappa_g \ll 1$, so that cells spend much more time in phase G1 compared with phase S/G2/M, the minimum wave speed obtained from our fundamental FUCCI model, Equations (2.2) to (2.3), approaches the minimum wave speed obtained from the FKPP equation, Equation (2.1). A similar observation holds for the case $\kappa_g/\kappa_r \ll 1$, corresponding to the situation where the cells spend much more time in phase S/G2/M compared with phase G1.

Travelling wave solutions are of great practical interest as cell migration tends to exhibit travelling wave characteristics [46]. The analytical expression we derive for the minimum wave speed is of practical interest, as a moving front of cells can be thought of as acting like a travelling wave solution, so our expression can provide a prediction for the speed of the moving front in experimental studies. The travelling wave solutions of the fundamental model are mathematically interesting because they are a combination of moving wavefronts and moving pulses. Monotonically decreasing wavefront solutions are well known for the FKPP equation, and because our models are related to the FKPP equation, it is not surprising that wavefront solutions are also observed in our study. It is interesting, however, that travelling wave solutions of our models involve pulses, which are not features of the travelling wave solutions of the FKPP equation.

There are many possibilities for future work arising from this study. An area of particular interest would involve using our models to quantitatively study how the migration of melanoma cells and the cell cycle for melanoma cells are affected by the application of particular melanoma drugs. Indeed, there is still much to learn regarding the effects of introduced drugs on melanoma cell activity [13]. These kinds of drugs often act to arrest the cell cycle, thereby preventing melanoma proliferation. Another feature that could be examined is the role of contact inhibition and cell cycle arrest, because it is accepted that cells in relatively high-density environments can undergo cell cycle arrest. Indeed, our model does not account for this phenomenon, because we are interested in low to moderate cell densities. Another way that this study could be extended is to consider additional phases of the cell cycle. This is of interest because a very recent extension

of FUCCI technology, referred to as FUCCI4 [113], highlights all four cell cycle phases G1, S, G2 and M. If extended to four coupled partial differential equations, our modelling framework could be used to quantitatively describe cell migration where individual cells are highlighted using FUCCI4.

2.5 Supplementary Material

2.5.1 Numerical solutions of the mathematical model

The numerical solutions of Equations (2.2) to (2.3) and Equations (2.4) to (2.6) are obtained using an implicit finite difference approximation. In particular, we use a central difference approximation for the spatial derivative term, and a backward Euler approximation for the temporal derivative [114]. The spatial domain, $0 \leq x \leq L$, is uniformly discretised with grid spacing Δx . No-flux boundary conditions are implemented at both $x = 0$ and $x = L$. The temporal domain is uniformly discretised with time steps of duration Δt . The resulting non-linear system of equations is solved using the Thomas algorithm [114], applying Picard iteration at each time step until the maximum absolute change in the given dependent variable across the spatial grid is less than a specified tolerance, ϵ .

2.5.2 Comparison of experimental data and the fundamental model for a scratch assay of FUCCI-transduced WM983C melanoma cells.

In Fig 2.5 we present experimental data for the WM983C melanoma cell line, together with corresponding numerical solutions of the fundamental model. The numerical solutions compare well with the experimental data.

2.5.3 Numerical solutions demonstrating travelling wave behaviour for the fundamental model when $\mathcal{D} \neq 1$.

In Fig 2.6 we present numerical solutions of the fundamental model, Equations (2.2) to (2.3), with $\mathcal{D} \neq 1$. In Fig 2.6A, $\mathcal{D} = 0.5$, and in Fig 2.6B, $\mathcal{D} = 2$. These solutions are qualitatively the same as for $\mathcal{D} = 1$, see Fig 2.3, demonstrating that the existence of these travelling waves is robust, and does not depend on the value of $\mathcal{D}_g/\mathcal{D}_r$.

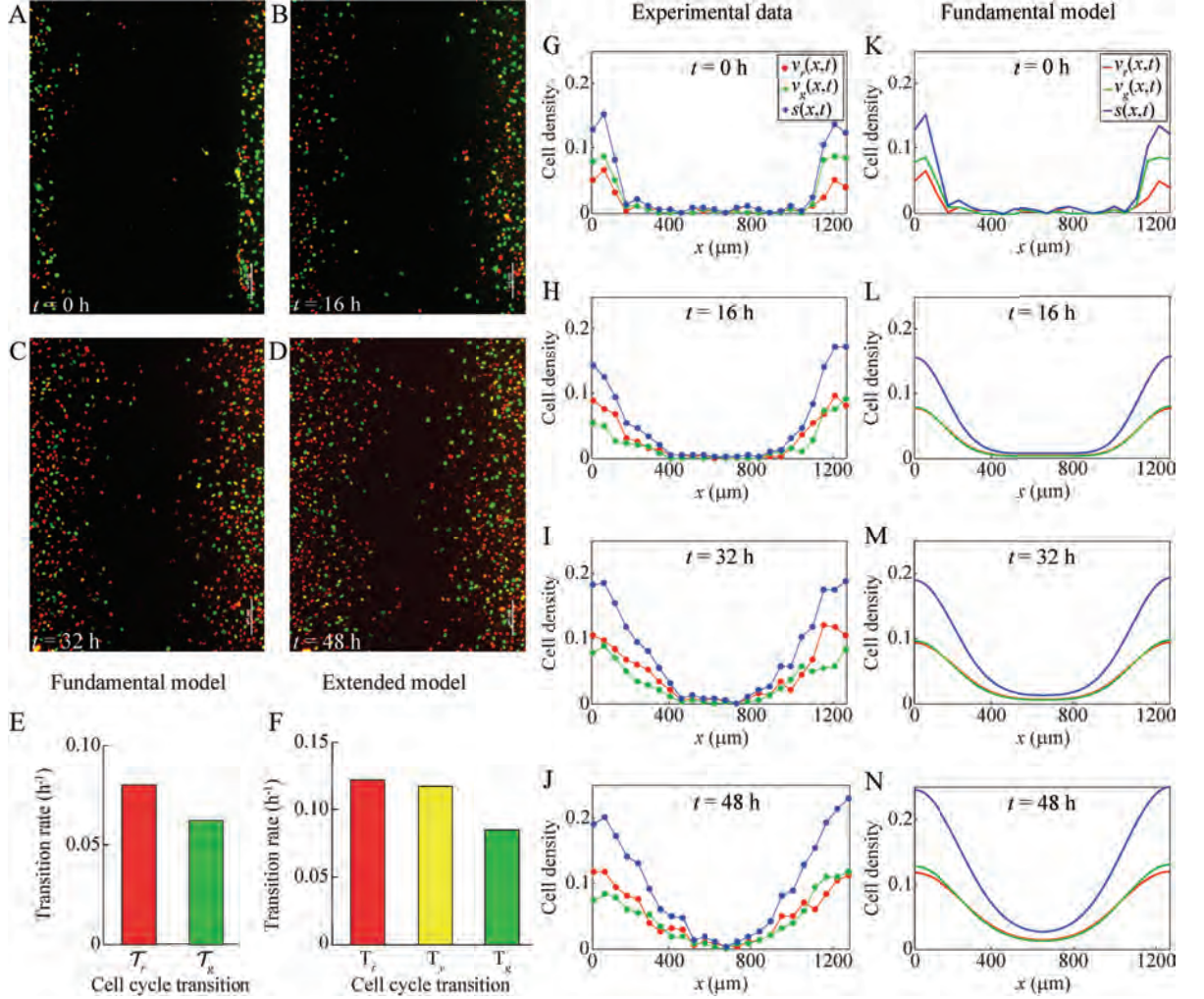


Figure 2.5: Comparison of experimental data and the fundamental model for a scratch assay of FUCCI-transduced WM983C melanoma cells. (A)–(D) Still images of a scratch assay with FUCCI-transduced WM983C melanoma cells at time 0 h, 16 h, 32 h and 48 h, respectively. Scale bar is $200 \mu\text{m}$. (E) Estimated transition rates from one cell cycle phase to the next for the fundamental model with two subpopulations, based on data from the WM983C cell line from Fig 1C in [13]. (F) Estimated transition rates from one cell cycle phase to the next for the extended model with three subpopulations, based on data from the WM983C cell line from Fig 1C in [13]. (G)–(J) Experimental three-dimensional cell density data at 0 h, 16 h, 32 h and 48 h, respectively (based on images (A)–(D)). The cell density is treated as a function of x and t only, owing to the fact that the initial density does not depend on the vertical coordinate, y . (K)–(N) Numerical solutions of the fundamental model, Equations (2.2) to (2.3), at 0 h, 16 h, 32 h and 48 h. The numerical solutions are obtained with $\Delta x = 0.1 \mu\text{m}$, $\Delta t = 0.1 \text{ h}$, $L = 1254 \mu\text{m}$, diffusion coefficients $\mathcal{D}_r = \mathcal{D}_g = 400 \mu\text{m}^2 \text{h}^{-1}$, transition rates $\kappa_r = 0.080 \text{ h}^{-1}$ and $\kappa_g = 0.062 \text{ h}^{-1}$, and initial conditions the same as for the experimental data.

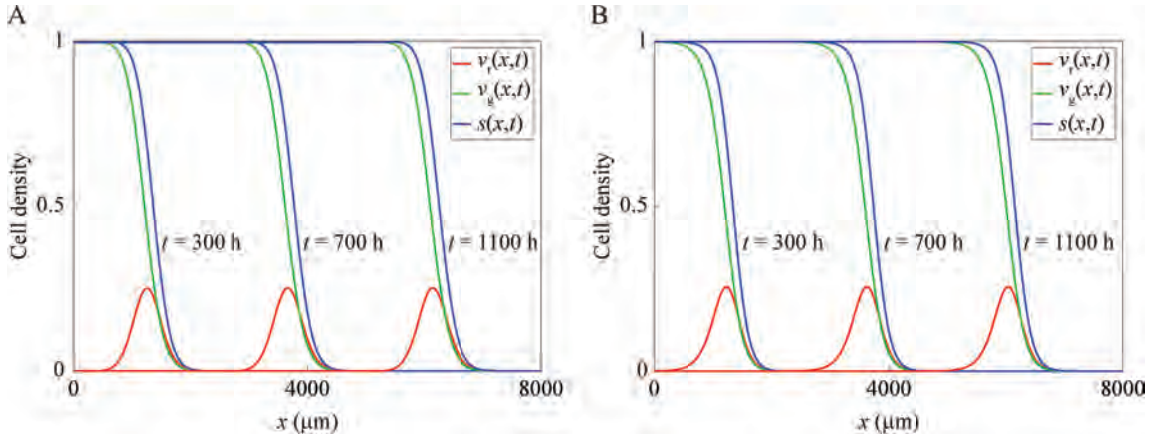


Figure 2.6: Numerical solutions of the fundamental model, Equations (2.2) to (2.3), demonstrating travelling wave behaviour with $\mathcal{D} \neq 1$. (A) Solutions obtained with $\Delta x = 0.1 \mu\text{m}$, $\Delta t = 0.1$ h, $L = 8000 \mu\text{m}$, $\mathcal{D}_r = 400 \mu\text{m}^2 \text{h}^{-1}$, $\mathcal{D}_g = 200 \mu\text{m}^2 \text{h}^{-1}$, $\kappa_r = \kappa_g = 0.08 \text{h}^{-1}$, and the initial condition given by Equation (2.9) with $\xi = 10$. (B) Solutions obtained with $\Delta x = 0.1 \mu\text{m}$, $\Delta t = 0.1$ h, $L = 8000 \mu\text{m}$, $\mathcal{D}_r = 200 \mu\text{m}^2 \text{h}^{-1}$, $\mathcal{D}_g = 400 \mu\text{m}^2 \text{h}^{-1}$, $\kappa_r = \kappa_g = 0.08 \text{h}^{-1}$, and the initial condition given by Equation (2.9) with $\xi = 10$.

Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronisation in cell proliferation

An article published as a *Report* in the *Journal of the Royal Society Interface*

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronisation in cell proliferation. *Journal of the Royal Society Interface*. 2019; 16:20190382.

Abstract

We present a suite of experimental data showing that cell proliferation assays, prepared using standard methods thought to produce asynchronous cell populations, persistently exhibit inherent synchronisation. Our experiments use fluorescent cell cycle indicators to reveal the normally-hidden cell synchronisation by highlighting oscillatory subpopulations within the total cell population. These oscillatory subpopulations would never be observed without these cell cycle indicators. On the other hand, our experimental data show that the total cell population appears to grow exponentially, as in an asynchronous population. We reconcile these seemingly inconsistent observations by employing a multi-stage mathematical model of cell proliferation that can replicate the oscillatory subpopulations. Our study has important implications for understanding and improving experimental reproducibility. In particular, inherent synchronisation may affect the experimental reproducibility of studies aiming to investigate cell cycle-dependent mechanisms, including changes in migration and drug response.

3.1 Introduction

Cell proliferation is essential for a range of normal and pathological processes. Many different mathematical models of proliferation have been proposed [50–56]. It is often assumed that cells proliferate exponentially,

$$\frac{dM(t)}{dt} = \lambda M(t), \quad M(t) = M(0)e^{\lambda t}, \quad (3.1)$$

where $M(t)$ is the number of cells at time t and $\lambda > 0$ is the proliferation rate.

The eukaryotic cell cycle consists of four phases in sequence, namely gap 1 (G1), synthesis (S), gap 2 (G2) and mitosis (M) (Figure 3.1(a)). A key assumption implicit in Equation (3.1) is that the cell population is asynchronous, meaning that the cells are distributed randomly among the cell cycle phases (Figure 3.1(b)), yielding a constant per capita growth rate, $(1/M(t)) dM(t)/dt = \lambda$. In contrast, a population of cells is synchronous if the cells are in the same cell cycle phase (Figure 3.1(c)), or partially synchronous if only a subpopulation of cells is synchronous (Figure 3.1(d)). In this case, the synchronous cells divide as a cohort in discrete stages producing a variable per capita growth rate. In addition to the implicit assumption of asynchronicity, classical exponential growth models, and generalisations thereof [115], do not account for subpopulations, and predict monotonic population growth.

Here we provide new experimental data from two-dimensional cell proliferation assays in which the cell growth appears exponential as in Equation (3.1). Unexpectedly, however, we observe oscillatory subpopulations arising from a phenomenon we refer to as *inherent* synchronisation. We reveal the normally-hidden inherent synchronisation by identifying subpopulations based on cell cycle phase, employing fluorescent ubiquitination-based cell cycle indicator (FUCCI) [30]. FUCCI enables visualisation of the cell cycle of individual live cells via two sensors: when the cell is in G1 the nucleus fluoresces red, and when the cell is in S/G2/M the nucleus fluoresces green. During the G1/S transition, called early S (eS), both sensors fluoresce and the nucleus appears yellow (Figure 3.1(a)). We explain these seemingly inconsistent observations by applying a multi-stage mathematical model for cell proliferation.

Previous studies of cell synchronisation utilising FUCCI induce the synchronisation using methods including serum starvation, cell cycle-inhibiting drugs, environmental pH, or contact inhibition [13, 33, 116–119]. Our assays are prepared using a standard method [33] normally thought to produce asynchronous populations, and we take utmost care to ensure that there is no induced synchronisation in our cell cultures due to serum starvation, low pH, or contact inhibition (Supporting Information 1). Over three cell lines and four independent experiments,

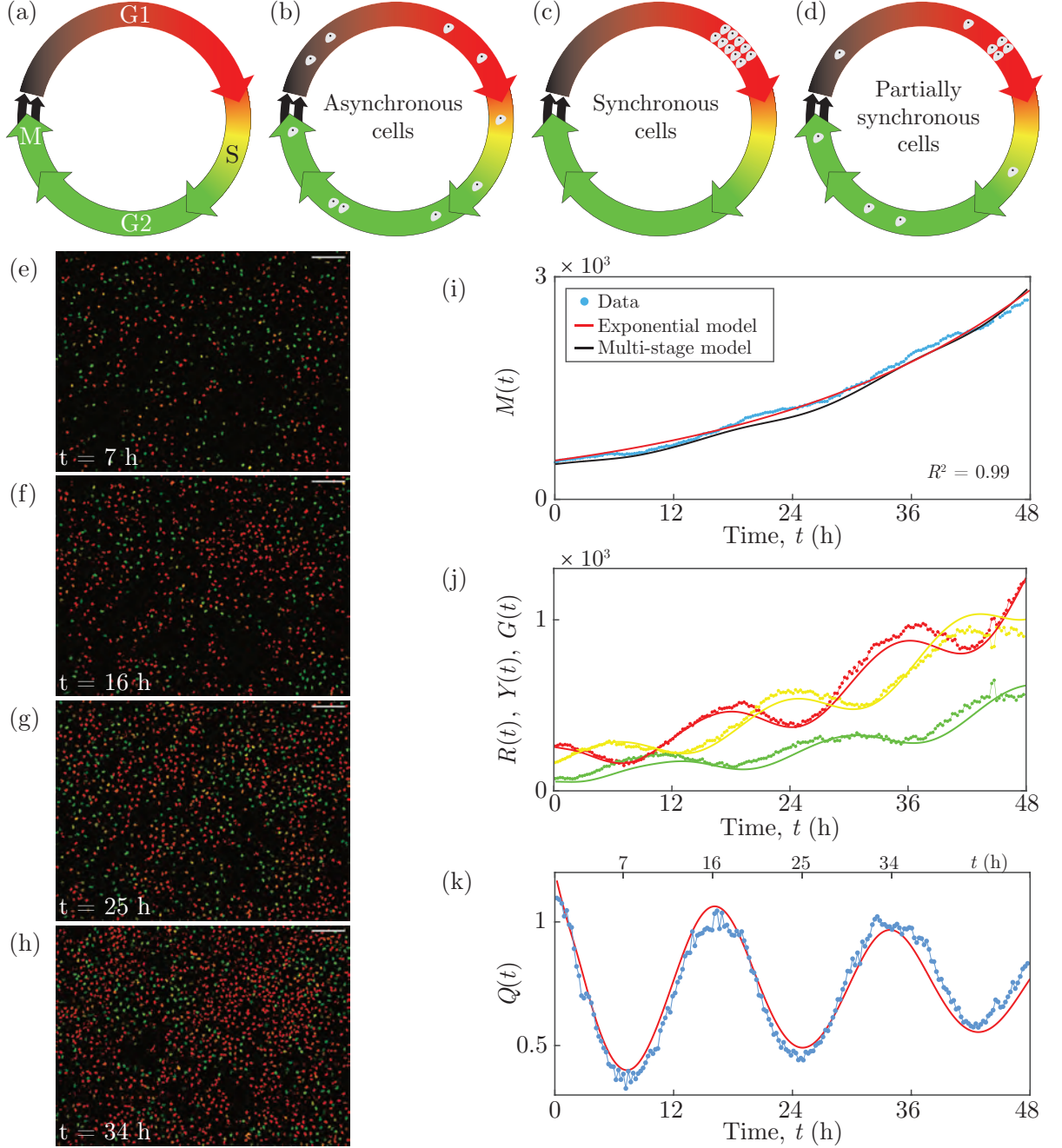


Figure 3.1: C8161 experimental data and multi-stage model solution. (a) The cell cycle, indicating the colour of FUCCI in each phase. (b)–(d) Asynchronous, synchronous and partially synchronous cells. (e)–(h) Images of a proliferation assay with FUCCI-C8161 cells. Scale bar $200 \mu\text{m}$. (i) $M(t)$. Linear regression of $\ln M(t)$ versus t gives $R^2 = 0.99$. (j) $R(t)$, $Y(t)$ and $G(t)$. (k) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

however, we consistently observe inherent synchronisation.

Neglecting synchronous subpopulations can have important implications for experiment reproducibility. For example, the accurate experimental evaluation of cell cycle-inhibiting drugs is highly dependent on the cell cycle distribution of the cell population [33, 120]. In a partially synchronous population, the drug may have a delayed or advanced effect compared with an asynchronous population, depending on the cell cycle position of the synchronous cells. Generally, the presence of synchronisation may affect the reproducibility of experiments that investigate cell cycle-dependent mechanisms, such as changes in migration and drug response. Revealing any synchronisation with quantitative techniques like FUCCI will lead to a better understanding of these mechanisms.

3.2 Results

3.2.1 Experimental data

Our experimental data are time-series images from two-dimensional proliferation assays using three melanoma cell lines C8161, WM983C and 1205Lu [13, 84, 85], which have mean cell-cycle durations of approximately 18, 27 and 36 h, respectively [13]. Four independent experiments are performed for each cell line. Live-cell images are acquired at 15 minute intervals over 48 h.

Images from one position in a single well of a FUCCI-C8161 proliferation assay at 7, 16, 25 and 34 h show red, yellow or green nuclei corresponding to the phases G1, eS or S/G2/M (Figure 3.1(e)–(h)). We quantify the population growth by counting the total number of cells in each image (Supporting Information 1), to give $M(t)$ at time t (Figure 3.1(i)). The total number of cells appears to grow exponentially over 48 h, supported by the best fit of Equation (3.1) (Supporting Information 1) since we have $R^2 = 0.99$ from the linear regression of $\ln M(t)$ versus t . The temporal variations in the numbers of cells in the subpopulations $R(t)$, $Y(t)$ and $G(t)$ with red, yellow or green nuclei (Figure 3.1(j)), where $M(t) = R(t) + Y(t) + G(t)$, are oscillatory. In an asynchronous population, the subpopulations would exhibit monotone growth. The oscillations we observe, however, reveal that the cells are partially synchronous.

To explore the inherent synchronisation further, we group cells in eS and S/G2/M together, since eS is part of S, and consider the ratio $Q(t) = R(t)/(Y(t) + G(t))$ (Figure 3.1(k)). Synchronisation is clearly evident in the oscillatory nature of $Q(t)$. Note the troughs at 7 and 25 h and the peaks at 16 and 34 h are separated by 18 h, which is the approximate cell cycle time for C8161. We can visualise the oscillations in these two subpopulations (Figure 3.1(e)–(h)), where the ratio of the number of red cells to the number of yellow and green cells is lower at 7 and 25

h and higher at 16 and 34 h. Equation (3.1) and related generalisations [115] cannot account for the oscillations in these subpopulations. Similar observations are made for further examples of this cell line, and the two additional cell lines (Supporting Information 1). We quantitatively confirm the presence of oscillations in $Q(t)$, arising from inherent synchronisation, for all 90 data sets by calculating the discrete Fourier transform of the $Q(t)$ signal, and identifying the distinct dominant frequencies (Supporting Information 1). These results confirm that all 90 experimental replicates display oscillatory subpopulations that are inconsistent with traditional exponential and logistic growth models.

3.2.2 Multi-stage mathematical model

We employ a multi-stage model of cell proliferation [86] which can describe synchronous populations. The model assumes that the cell cycle durations follow a hypoexponential distribution, which consists of a series of independent exponential distributions with different rates. To apply this model we partition the cell cycle into k stages, P_i for $i = 1, \dots, k$, where the duration of each P_i is exponentially distributed with mean μ_i . If \mathcal{T} is the mean cell cycle time then $\sum_{i=1}^k \mu_i = \mathcal{T}$. The stages P_i do not necessarily correspond to phases of the cell cycle, but instead are a mathematical device which allows control over the variance of cell-cycle phase durations in the multi-stage model, whereby more stages correspond to less variance in the phase durations for a cell population. If we let the transition rates be $\lambda_i = 1/\mu_i$ and consider the partitioned cell cycle $P_1 \xrightarrow{\lambda_1} P_2 \xrightarrow{\lambda_2} \dots \xrightarrow{\lambda_{k-1}} P_k \xrightarrow{\lambda_k} 2P_1$, we arrive at a system of differential equations describing the mean population $M_i(t)$ in each stage [86]:

$$\frac{dM_i(t)}{dt} = \begin{cases} 2\lambda_k M_k(t) - \lambda_1 M_1(t), & \text{for } i = 1, \\ \lambda_{i-1} M_{i-1}(t) - \lambda_i M_i(t), & \text{for } i = 2, \dots, k. \end{cases} \quad (3.2)$$

Note that $M(t) = \sum_{i=1}^k M_i(t)$. If $k = 1$, Equation (3.2) simplifies to Equation (3.1). Within the 48 h duration of our experiments, none of the cell lines exhibits contact inhibition of proliferation, consistent with the typical loss of contact inhibition in cancer cells [14]. Consequently, a carrying capacity is not incorporated into the model.

We solve Equation (3.2) numerically with the forward Euler method, and estimate the parameters by fitting the solution to our experimental data (Supporting Information 1). Using 18 stages for each of the three cell cycle phases described by FUCCI, giving $k = 54$, we obtain $M(t)$ (Figure 3.1(i)), $R(t)$, $Y(t)$, $G(t)$ (Figure 3.1(j)) and $Q(t)$ (Figure 3.1(k)) which all correspond well with the experimental data. In particular, the multi-stage model replicates the oscillations

in $R(t)$, $Y(t)$, $G(t)$ and $Q(t)$, a feature that is not possible with traditional exponential models. While the multi-stage model can replicate the oscillatory subpopulations, the model is unable to predict all features of the inherent synchronisation in a cell proliferation experiment due to variable initial conditions, as the inherent synchronisation is a stochastic phenomenon which likely arises from cell division and intercellular interactions. The model can, however, be used to predict general features of the inherent synchronisation of each cell line.

3.3 Conclusion

Our new experimental data demonstrate that cell populations may appear to grow exponentially despite subpopulations exhibiting oscillatory growth arising from normally-hidden inherent synchronisation. We use standard experimental methods thought to produce asynchronous populations; however, all of our proliferation assays exhibit inherent synchronisation. We use FUCCI to track cell-cycle progression, which is necessary to confirm cell synchronisation. As the standard exponential growth model cannot account for subpopulations with oscillating growth, we use a multi-stage mathematical model of cell proliferation to replicate oscillations in population growth. Our results are important because revealing any synchronisation will help to better understand cell cycle-dependent mechanisms, such as changes in migration and drug response. Without quantitative techniques like FUCCI to probe the cell cycle, synchronisation and its effects on experimental outcomes and reproducibility may remain hidden.

3.4 Data accessibility

All experimental data are available in the electronic supplementary material documents.

All algorithms required to replicate this work are available on GitHub at <https://github.com/ProfMJSimpson/Vittadello2019>.

3.5 Supplementary Material

3.5.1 Experimental

Here we provide details of the experimental methodology that we use for our cell proliferation experiments involving three melanoma cell lines.

3.5.1.1 Cell culture

The human melanoma cell lines C8161 (kindly provided by Mary Hendrix, Chicago, IL, USA), WM983C and 1205Lu (both kindly provided by Meenhard Herlyn, Philadelphia, PA, USA) were

genotypically characterised [101–104], grown as described [105] (using 4% fetal bovine serum instead of 2%), and authenticated by STR fingerprinting (QIMR Berghofer Medical Research Institute, Herston, Australia).

We maintain the cell cultures to prevent any induced synchronisation from cell cycle arrest in G1 phase. In general, such induced synchronisation can occur through various experimental conditions, namely contact inhibition of proliferation at relatively high population densities [119], decreased pH of the growth medium due to the concentration of acidic cell-metabolites such as lactic acid [118], and reduced availability of nutrients such as serum [33]. We prevent induced synchronisation by passaging the cells every three days, and on the day prior to setting up an experiment, to maintain a subconfluent cell density and a fresh growth medium, so that the cell culture conditions are never such that they cause G1 arrest.

3.5.1.2 Proliferation experiments

Cells are seeded from subconfluent culture flasks onto a 24-well plate at a density of 10^4 cells cm^{-2} , with 2.5 ml of medium per well, which is 2.5 times the volume of the standard protocol. After incubating the plate for 24 h at 37°C with 5% CO_2 , live-cell images are acquired at 15 minute intervals over 48 h at six different positions of the well. Four independent experiments are performed for each cell line.

Our preliminary experiments used a standard 1 ml of medium in each well, however the cells started to arrest in G1 around 48 hours after seeding, which is likely due to decreased pH of the medium from the lactic acid concentration. We therefore performed a large number of preliminary tests in an attempt to prevent the cells from arresting in G1 during the experiment, and we found that this is possible by increasing the volume of medium in each well to the reasonable maximum of 2.5 ml, given the volume of each well is 3 ml. The result of the increased volume of medium is that the cells do not begin to arrest in G1 until close to 72 hours following seeding, which provides us with almost 48 hours of imaging using cells that have minimal G1 arrest.

3.5.1.3 Fluorescent ubiquitination-based cell cycle indicator (FUCCI)

To generate stable melanoma cell lines expressing the FUCCI constructs, mKO2-hCdt1 (30-120) and mAG-hGem (1-110) [30] were subcloned into a replication-defective, self-inactivating lentiviral expression vector system as previously described [105]. The lentivirus was produced by co-transfection of human embryonic kidney 293T cells. High-titer viral solutions for mKO2-hCdt1 (30/120) and mAG-hGem (1/110) were prepared and used for co-transduction into the

melanoma cell lines, and subclones were generated by single cell sorting [13, 32, 106].

3.5.2 Image processing and analysis

The microscopy data consist of multi-channel time-series stacks which are processed and analysed automatically with Fiji/ImageJ and MATLAB as described below.

3.5.2.1 Preprocessing

To maximise the accuracy in identifying particles, which in our case are cell nuclei, we enhance the quality of the microscopy images using ImageJ as follows.

1. Import the time-series stack with the Bio-Formats Importer plugin, splitting the red and green channels.
2. Apply five iterations of Subtract Background with rolling-ball radius of 5 pixels.
3. Apply Enhance Contrast with the Equalize Histogram option selected.
4. Apply the Gaussian Blur filter with $\sigma = 1$.

3.5.2.2 Segmentation

We now identify the particles in the processed images using ImageJ.

1. Apply Auto-thresholding using the Yen method, selecting the option to ‘calculate the threshold for each image’.
2. The resulting binary images are then refined by applying:
 - (a) Watershed;
 - (b) Fill Holes;
 - (c) Open, with iterations = 10 and count = 5;
 - (d) Watershed.

3.5.2.3 Analysis

For every image in the segmented binary time-series stacks we count the number of particles in each of the red and green channels using ImageJ. We then use MATLAB to determine which particles are yellow.

1. For each of the red and green channels, apply Analyze Particles in ImageJ with sizes in the range $5-\infty$ pixels² and the option ‘limit to threshold’ selected. Output the stack position and the centroid of every particle in each channel.
2. We now need to determine which particles are red, yellow or green. A particle is red if it appears in the red channel, and there is no corresponding particle in the green channel. Similarly, a particle is green if it appears in the green channel, and there is no corresponding particle in the red channel. A particle is then yellow if it appears in both the red and green channels. Identifying whether a particle appears in both the red and green channels is complicated by the possible alteration of the shape of the particle during image processing. While we process every image in exactly the same way, the original microscopy images may have different signal-to-noise ratios between the red and green channels. Consequently, there may be a difference in the shape of a particle depending on the channel in which it is viewed, and thereby a difference in the centroid of the particle in each channel. We therefore use MATLAB to determine which particles are red, yellow or green, using the stack position and centroid of each particle, as follows.
 - (a) We first find the yellow particles using the stack position and centroid of each particle, so choose a particle, in turn, from the red channel.
 - (b) Search the green channel for a corresponding particle such that the Euclidean distance between the centroids of the two particles is not greater than 3 pixels, noting that the pixel size in our images is $1.8150 \mu\text{m}$. This distance allows for a location error of the centroids of the red and green particles, whereby the centroids may be translated up to one pixel from the original centroid of the yellow particle in the unprocessed images. Placing the original yellow centroid at the centre of a 3×3 grid of pixels, the red and green centroids from the processed images may be located at any of the nine pixels in the grid.
 - (c) Once all of the yellow particles are found, the red particles are all of the particles in the red channel which are not yellow. Similarly, the green particles are all of the particles in the green channel which are not yellow.

3.5.3 Parameterisation of the exponential model

To estimate the parameters of the exponential model Equation (3.1) when fitting the model solution to the experimental data for the total number of cells, we use the `fit` function and `exp1` model [121] in MATLAB. The parameter estimates, with 95% confidence intervals, are:

- **C8161 cell line - Figure 3.1(i)**

$$M(0) = 524.3 \text{ (515.1, 533.4) and } \lambda = 0.03504 \text{ h}^{-1} \text{ (0.03456, 0.03551).}$$

- **C8161 cell line - Figure 3.3(e)**

$$M(0) = 386.4 \text{ (382.2, 390.5) and } \lambda = 0.0316 \text{ h}^{-1} \text{ (0.0313, 0.0319).}$$

- **C8161 cell line - Figure 3.4(e)**

$$M(0) = 401 \text{ (393.8, 408.2) and } \lambda = 0.03573 \text{ h}^{-1} \text{ (0.03525, 0.03622).}$$

- **WM983C cell line - Figure 3.5(e)**

$$M(0) = 247.7 \text{ (244.3, 251.2) and } \lambda = 0.02541 \text{ h}^{-1} \text{ (0.02501, 0.02581).}$$

- **WM983C cell line - Figure 3.6(e)**

$$M(0) = 366.4 \text{ (362.8, 370) and } \lambda = 0.01917 \text{ h}^{-1} \text{ (0.01888, 0.01946).}$$

- **WM983C cell line - Figure 3.7(e)**

$$M(0) = 158 \text{ (155.4, 160.7) and } \lambda = 0.0175 \text{ h}^{-1} \text{ (0.01699, 0.01801).}$$

- **1205Lu cell line - Figure 3.8(e)**

$$M(0) = 215.9 \text{ (214.1, 217.7) and } \lambda = 0.01932 \text{ h}^{-1} \text{ (0.01907, 0.01958).}$$

- **1205Lu cell line - Figure 3.9(e)**

$$M(0) = 249.1 \text{ (246.5, 251.7) and } \lambda = 0.01934 \text{ h}^{-1} \text{ (0.01903, 0.01965).}$$

- **1205Lu cell line - Figure 3.10(e)**

$$M(0) = 266.6 \text{ (263.6, 269.6) and } \lambda = 0.01926 \text{ h}^{-1} \text{ (0.01893, 0.0196).}$$

3.5.4 Parameterisation of the multi-stage mathematical model

Here we describe our methodology for estimating the parameters of the multi-stage mathematical model Equation (3.2) and fitting the model solution to the experimental data.

3.5.4.1 Method for parameter estimation

The multi-stage model requires specification of the number of stages, the transition rates from each stage to the successive stage, and the initial population in each stage. In this work, we aim to achieve the best fit of the model to our data while keeping the number of model parameters with distinct values to a minimum.

We partition the phases G1, eS and S/G2/M into the same number of stages, N . The mean durations of the phases G1, eS and S/G2/M are denoted by L_r , L_y and L_g , respectively. The transition rates between successive stages are set equal within each phase to N/L_r in G1, N/L_y in eS, and N/L_g in S/G2/M. For each $i = 1, \dots, N$ we denote the mean number of cells at time t in stage i of G1 as $R_i(t)$, of eS as $Y_i(t)$, and of S/G2/M as $G_i(t)$. Therefore, $R(t) = \sum_{i=1}^N R_i(t)$, $Y(t) = \sum_{i=1}^N Y_i(t)$, and $G(t) = \sum_{i=1}^N G_i(t)$. The parameters that we need to estimate are the components of the vector

$$\mathbf{x} = [R_1(0) \dots R_N(0) \quad Y_1(0) \dots Y_N(0) \quad G_1(0) \dots G_N(0) \quad L_r \quad L_y \quad L_g]. \quad (\text{S1})$$

The parameters in Equation (S1) are either numbers of cells or phase durations, which are all non-negative, so we require our optimisation algorithm to accept bound constraints. To find estimates for these parameters we use the MATLAB nonlinear least-squares solver `lsqnonlin` [122] with the trust-region-reflective algorithm [123], which allows for bound constraints of the parameters. In the following, a dependent variable has the subscript ‘model’ or ‘data’ to distinguish between model and data values of the variable. With non-negative weights w_2, \dots, w_7 , we define the vector objective function

$$\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \quad w_2 \mathbf{f}_2(\mathbf{x}) \quad w_3 \mathbf{f}_3(\mathbf{x}) \quad w_4 \mathbf{f}_4(\mathbf{x}) \quad w_5 \mathbf{f}_5(\mathbf{x}) \quad w_6 \mathbf{f}_6(\mathbf{x}) \quad w_7 \mathbf{f}_7(\mathbf{x})] \quad (\text{S2})$$

as the concatenation of the weight-scaled vectors

$$\mathbf{f}_1(\mathbf{x}) = [(Q_{\text{model}}(\mathbf{x}) - Q_{\text{data}})(t_1) \quad \dots \quad (Q_{\text{model}}(\mathbf{x}) - Q_{\text{data}})(t_n)], \quad (\text{S3})$$

$$\mathbf{f}_2(\mathbf{x}) = [(R_{\text{model}}(\mathbf{x}) - R_{\text{data}})(t_1) \quad \dots \quad (R_{\text{model}}(\mathbf{x}) - R_{\text{data}})(t_n)], \quad (\text{S4})$$

$$\mathbf{f}_3(\mathbf{x}) = [(Y_{\text{model}}(\mathbf{x}) - Y_{\text{data}})(t_1) \quad \dots \quad (Y_{\text{model}}(\mathbf{x}) - Y_{\text{data}})(t_n)], \quad (\text{S5})$$

$$\mathbf{f}_4(\mathbf{x}) = \left[(G_{\text{model}}(\mathbf{x}) - G_{\text{data}})(t_1) \quad \dots \quad (G_{\text{model}}(\mathbf{x}) - G_{\text{data}})(t_n) \right], \quad (\text{S6})$$

$$\mathbf{f}_5(\mathbf{x}) = \left[(G_{\text{model}}(\mathbf{x}) - G_{\text{data}})(t_n) \right], \quad (\text{S7})$$

$$\mathbf{f}_6(\mathbf{x}) = \left[\mathcal{T} - L_r - L_y - L_g \right], \quad (\text{S8})$$

$$\mathbf{f}_7(\mathbf{x}) = \left[(M_{\text{model}}(\mathbf{x}) - M_{\text{data}})(t_1) \quad \dots \quad (M_{\text{model}}(\mathbf{x}) - M_{\text{data}})(t_n) \right], \quad (\text{S9})$$

where:

1. $Q_{\text{model}}(\mathbf{x})$ is the ratio of the number of cells in G1 to the number of cells in eS/S/G2/M, from the model solution;
2. Q_{data} is the ratio of the number of cells in G1 to the number of cells in eS/S/G2/M, from the data;
3. $R_{\text{model}}(\mathbf{x})$, $Y_{\text{model}}(\mathbf{x})$ and $G_{\text{model}}(\mathbf{x})$ are the subpopulations of cells in G1, eS and S/G2/M, respectively, from the model solution;
4. R_{data} , Y_{data} and G_{data} are the subpopulations of cells in G1, eS and S/G2/M, respectively, from the data;
5. $M_{\text{model}}(\mathbf{x})$ is the total cell population, from the model solution;
6. M_{data} is the total cell population, from the data;
7. $t_1 < \dots < t_n$ are the data time points over 48 hours;
8. \mathcal{T} is the cell cycle time.

The vector \mathbf{f}_1 is used to fit the model to the ratio data, and the vectors $\mathbf{f}_2, \dots, \mathbf{f}_4$ are used to fit the model to the three subpopulations corresponding to G1, eS and S/G2/M. The vector \mathbf{f}_5 fits the model to the S/G2/M subpopulation data at the final time point, and is only required if the cells are starting to arrest in G1 near the end of the experiment due to the decreased pH of the growth medium. The vector \mathbf{f}_6 constrains the estimated phase durations to sum to the expected cell cycle time, and is generally required only when there are an insufficient number of oscillations in $Q_{\text{data}}(t)$ to bound the estimated cell-phase durations to physically realistic values. The vector \mathbf{f}_7 is used to fit the model to the total population data, and is often not required as a good fit usually follows from fitting to the subpopulations.

Note that the weights in the objective function Equation (S2) are specified prior to optimising the estimates of the parameters in Equation (S1). The weights differ between data sets in order to obtain the best fit of the multi-stage model to each data set.

3.5.4.2 Specific estimated parameters

Here we provide a summary of the estimated parameters of the multi-stage model Equation (3.2) corresponding to Figure 1, along with additional data sets from the C8161 (Figures 3.3–3.4), WM983C (Figures 3.5–3.7) and 1205Lu (Figure 3.8–3.10) cell lines. The correspondence between these figures and the data sets in Figures 3.5.5.1–3.5.5.2 is:

- Figure 3.1 - First plot in Well 1 of Experiment 2, Figures 3.5.5.1 and 3.5.5.1;
- Figure 3.3 - Second plot in Well 1 of Experiment 1, Figures 3.5.5.1 and 3.5.5.1;
- Figure 3.4 - First plot in Well 1 of Experiment 3, Figures 3.5.5.1 and 3.5.5.1;
- Figure 3.5 - Second plot in Well 1 of Experiment 4, Figures 3.5.5.1 and 3.5.5.2;
- Figure 3.6 - Fourth plot in Well 1 of Experiment 1, Figures 3.5.5.1 and 3.5.5.2;
- Figure 3.7 - Third plot in Well 1 of Experiment 2, Figures 3.5.5.1 and 3.5.5.2;
- Figure 3.8 - First plot in Well 1 of Experiment 4, Figures 3.5.5.2 and 3.5.5.2;
- Figure 3.9 - Fourth plot in Well 1 Experiment 4, Figures 3.5.5.2 and 3.5.5.2;
- Figure 3.10 - Fifth plot in Well 2 of Experiment 4, Figures 3.5.5.2 and 3.5.5.2.

3.5.4.2.1 C8161 cell line - Figure 3.1 The experimentally-determined mean cell cycle time for C8161 is approximately $\mathcal{T} = 18$ h [13]. We partition each cell cycle phase into $N = 18$ stages, giving a total of $k = 54$ stages for the complete cell cycle. In each phase we set the first half of the stages, totalling 9 stages, to have equal numbers of cells, and the second half of the stages to have equal numbers of cells. We therefore only require a total of six distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-4}\mathbf{f}_2(\mathbf{x}) \ 10^{-4}\mathbf{f}_3(\mathbf{x}) \ 10^{-4}\mathbf{f}_4(\mathbf{x}) \ 10^{-3}\mathbf{f}_5(\mathbf{x})]$. Starting with the parameters $R_i(0) = Y_i(0) = G_i(0) = 0.5$ for $i = 1, \dots, N$, and $L_r = L_y = L_g = 6$, we obtain the parameterisation

$$\begin{aligned}
 R_i(0) &= \begin{cases} 15.48 & \text{for } i = 1, \dots, 9, \\ 12.93 & \text{for } i = 10, \dots, 18, \end{cases} & Y_i(0) &= \begin{cases} 12.94 & \text{for } i = 1, \dots, 9, \\ 5.41 & \text{for } i = 10, \dots, 18, \end{cases} \\
 & & L_r &= 6.14 \text{ h}, \\
 G_i(0) &= \begin{cases} 2.51 & \text{for } i = 1, \dots, 9, \\ 3.48 & \text{for } i = 10, \dots, 18, \end{cases} & L_y &= 7.43 \text{ h}, \\
 & & L_g &= 4.42 \text{ h}.
 \end{aligned} \tag{S10}$$

All parameter estimates given in this document are presented to two decimal places. Note that $L_r + L_y + L_g = 17.99$ h, in good agreement with the observed cell cycle time of 18 h.

3.5.4.2.2 C8161 cell line, different numbers of stages - Figure 3.2 In Figure 3.2 we compare solutions of the multi-stage model for $N = 2, 6, 10$ and 14 stages per phase, with the ratio Q_{data} . In fitting the model solution we use the same parameters as for Figure 3.1, except

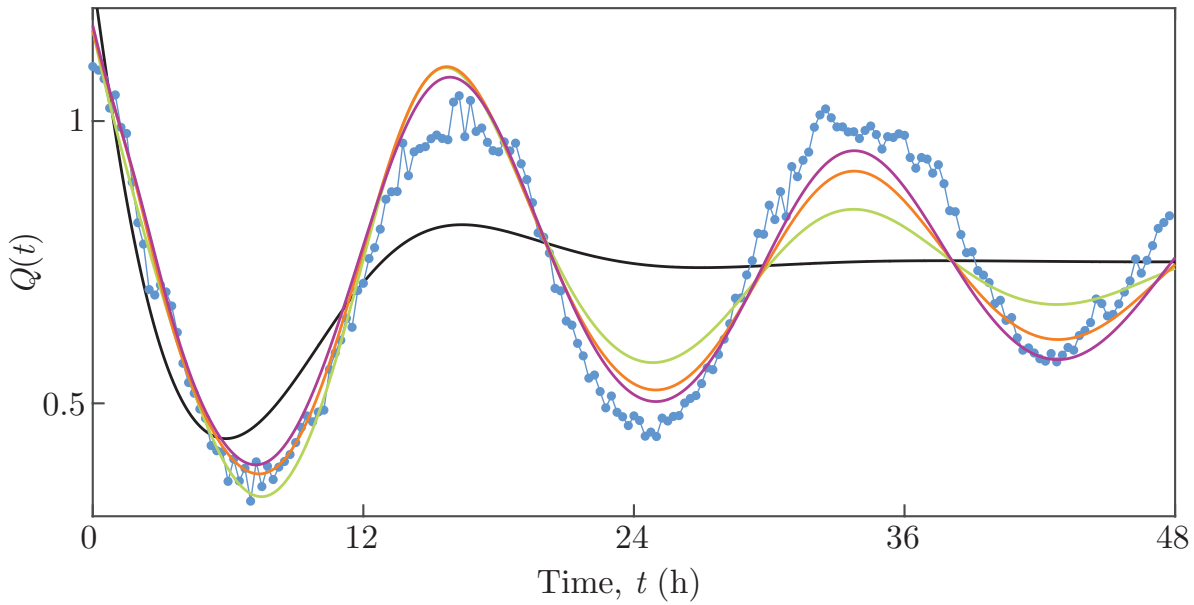


Figure 3.2: $Q(t)$ for C8161 experimental data and multi-stage model solutions with different numbers of stages. Experimental data are shown as discs and the model solutions as curves. The model solutions with 2, 6, 10 and 14 stages are the black, green, orange and purple curves, respectively.

the number of stages differ. For each number of stages, in each phase we set the first half of the stages to have equal numbers of cells and the second half of the stages to have equal numbers of cells, so that we therefore only require a total of 6 distinct population parameters. Starting with the parameters $R_i(0) = Y_i(0) = G_i(0) = 0.5$ for $i = 1, \dots, N$, and $L_r = L_y = L_g = 6$, the parameterisation for $N = 14$ is

$$\begin{aligned}
 R_i(0) &= \begin{cases} 17.89 & \text{for } i = 1, \dots, 9, \\ 18.62 & \text{for } i = 10, \dots, 18, \end{cases} & Y_i(0) &= \begin{cases} 16.28 & \text{for } i = 1, \dots, 9, \\ 6.77 & \text{for } i = 10, \dots, 18, \end{cases} \\
 G_i(0) &= \begin{cases} 2.39 & \text{for } i = 1, \dots, 9, \\ 5.78 & \text{for } i = 10, \dots, 18, \end{cases} & L_r &= 6.17 \text{ h}, \\
 & & L_y &= 7.27 \text{ h}, \\
 & & L_g &= 4.64 \text{ h},
 \end{aligned} \tag{S11}$$

the parameterisation for $N = 10$ is

$$\begin{aligned}
 R_i(0) &= \begin{cases} 16.60 & \text{for } i = 1, \dots, 9, \\ 33.58 & \text{for } i = 10, \dots, 18, \end{cases} & Y_i(0) &= \begin{cases} 20.81 & \text{for } i = 1, \dots, 9, \\ 9.61 & \text{for } i = 10, \dots, 18, \end{cases} \\
 & & L_r &= 6.22 \text{ h}, \\
 G_i(0) &= \begin{cases} 0 & \text{for } i = 1, \dots, 9, \\ 12.81 & \text{for } i = 10, \dots, 18, \end{cases} & L_y &= 7.06 \text{ h}, \\
 & & L_g &= 4.92 \text{ h},
 \end{aligned} \tag{S12}$$

the parameterisation for $N = 6$ is

$$\begin{aligned}
 R_i(0) &= \begin{cases} 16.42 & \text{for } i = 1, \dots, 9, \\ 66.47 & \text{for } i = 10, \dots, 18, \end{cases} & Y_i(0) &= \begin{cases} 44.97 & \text{for } i = 1, \dots, 9, \\ 0 & \text{for } i = 10, \dots, 18, \end{cases} \\
 & & L_r &= 6.28 \text{ h}, \\
 G_i(0) &= \begin{cases} 0 & \text{for } i = 1, \dots, 9, \\ 26.34 & \text{for } i = 10, \dots, 18, \end{cases} & L_y &= 6.81 \text{ h}, \\
 & & L_g &= 5.21 \text{ h},
 \end{aligned} \tag{S13}$$

and the parameterisation for $N = 2$ is

$$\begin{aligned}
 R_i(0) &= \begin{cases} 195.22 & \text{for } i = 1, \dots, 9, \\ 107.41 & \text{for } i = 10, \dots, 18, \end{cases} & Y_i(0) &= \begin{cases} 241.05 & \text{for } i = 1, \dots, 9, \\ 0 & \text{for } i = 10, \dots, 18, \end{cases} \\
 & & L_r &= 6.91 \text{ h}, \\
 G_i(0) &= \begin{cases} 0 & \text{for } i = 1, \dots, 9, \\ 0 & \text{for } i = 10, \dots, 18, \end{cases} & L_y &= 7.11 \text{ h}, \\
 & & L_g &= 5.79 \text{ h}.
 \end{aligned} \tag{S14}$$

The corresponding solutions of the multi-stage model are shown in Figure 3.2.

Considering parameterisations of the model whereby in each phase we set the first half of the stages to have equal numbers of cells and the second half of the stages to have equal numbers of cells, the oscillations decay at a faster rate for a smaller number of stages per phase. A higher number of stages produces a hypoexponential distribution with lower variance, resulting in oscillations which are sustained for longer. Consequently, fewer than 18 stages per phase results in a model solution with a poorer fit.

3.5.4.2.3 C8161 cell line - Figure 3.3 The experimentally-determined mean cell cycle time for C8161 is approximately $\mathcal{T} = 18$ h [13]. We partition each cell cycle phase into $N = 40$ stages, giving a total of $k = 120$ stages for the complete cell cycle. In each phase we set the first half of the stages, totalling 20 stages, to have equal numbers of cells, and the second half of the stages to have equal numbers of cells. We therefore only require a total of six distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-3}\mathbf{f}_2(\mathbf{x}) \ 10^{-3}\mathbf{f}_3(\mathbf{x}) \ 10^{-3}\mathbf{f}_4(\mathbf{x})]$. Trialling different parameters chosen randomly and uniformly from $(0.1, 1)$ for $R_i(0)$, $Y_i(0)$ and $G_i(0)$, and from $(4, 8)$ for L_r , L_y and L_g , we obtain the parameterisation

$$\begin{aligned}
 R_i(0) &= \begin{cases} 2.56 & \text{for } i = 1, \dots, 20, \\ 1.99 & \text{for } i = 21, \dots, 40, \end{cases} & Y_i(0) &= \begin{cases} 5.37 & \text{for } i = 1, \dots, 20, \\ 4.13 & \text{for } i = 21, \dots, 40, \end{cases} \\
 & & L_r &= 5.47 \text{ h}, \\
 G_i(0) &= \begin{cases} 0.90 & \text{for } i = 1, \dots, 20, \\ 0.98 & \text{for } i = 21, \dots, 40, \end{cases} & L_y &= 8.67 \text{ h}, \\
 & & L_g &= 4.57 \text{ h}.
 \end{aligned} \tag{S15}$$

Note that $L_r + L_y + L_g = 18.71$ h, in good agreement with the observed cell cycle time of 18 h.

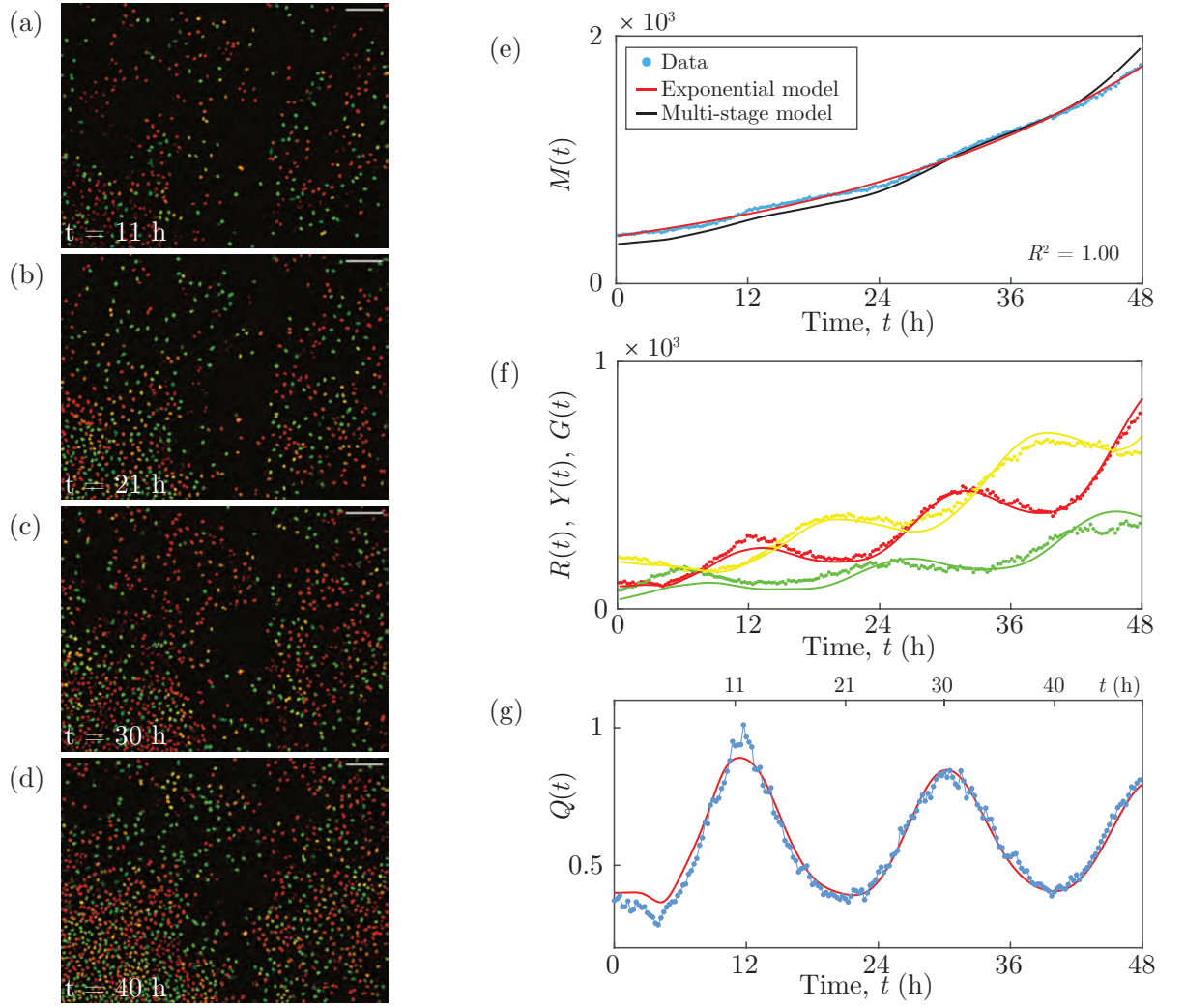


Figure 3.3: C8161 experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with FUCCI-C8161 cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.4 C8161 cell line - Figure 3.4 The experimentally-determined mean cell cycle time for C8161 is approximately $\mathcal{T} = 18$ h [13]. We partition each cell cycle phase into $N = 40$ stages, giving a total of $k = 120$ stages for the complete cell cycle. In each phase we set the first half of the stages, totalling 20 stages, to have equal numbers of cells, and the second half of the stages to have equal numbers of cells. We therefore only require a total of six distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-4}\mathbf{f}_2(\mathbf{x}) \ 10^{-4}\mathbf{f}_3(\mathbf{x}) \ 10^{-4}\mathbf{f}_4(\mathbf{x}) \ 10^{-3}\mathbf{f}_5(\mathbf{x})]$. Trialling different parameters chosen randomly and uniformly from $(0.1, 1)$ for $R_i(0)$, $Y_i(0)$ and $G_i(0)$, and from $(4, 8)$ for L_r , L_y and L_g , we obtain the parameterisation

$$\begin{aligned} R_i(0) &= \begin{cases} 3.21 & \text{for } i = 1, \dots, 20, \\ 4.44 & \text{for } i = 21, \dots, 40, \end{cases} & Y_i(0) &= \begin{cases} 5.03 & \text{for } i = 1, \dots, 20, \\ 2.59 & \text{for } i = 21, \dots, 40, \end{cases} \\ & & L_r &= 6.29 \text{ h}, \\ G_i(0) &= \begin{cases} 1.31 & \text{for } i = 1, \dots, 20, \\ 1.34 & \text{for } i = 21, \dots, 40, \end{cases} & L_y &= 7.30 \text{ h}, \\ & & L_g &= 4.47 \text{ h}. \end{aligned} \tag{S16}$$

Note that $L_r + L_y + L_g = 18.06$ h, in good agreement with the observed cell cycle time of 18 h.

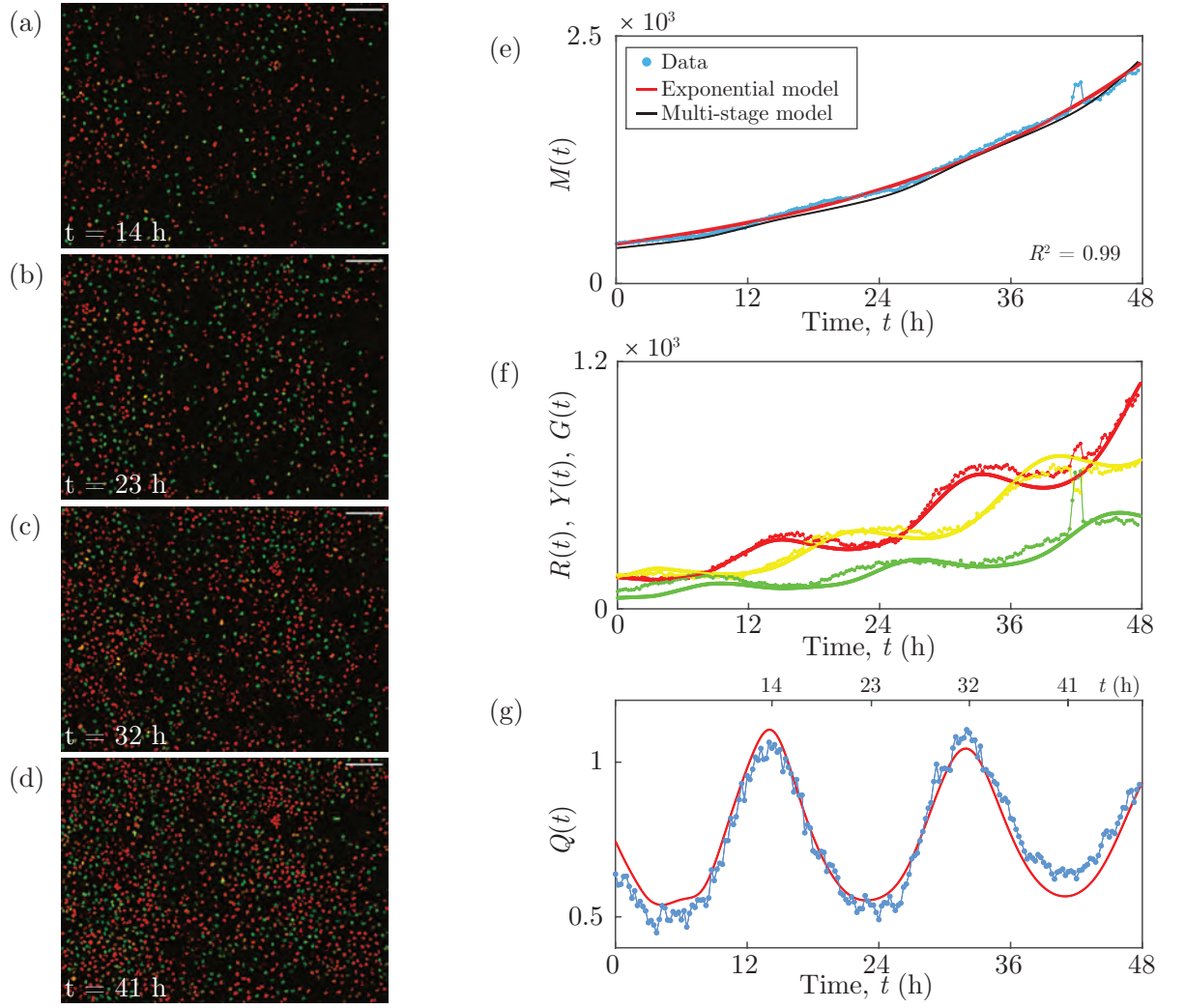


Figure 3.4: C8161 experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with FUCCI-C8161 cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.5 WM983C cell line - Figure 3.5 The experimentally-determined mean cell cycle time for WM983C is $\mathcal{T} = 27$ h [13]. We partition each cell cycle phase into $N = 10$ stages, giving a total of $k = 30$ stages for the complete cell cycle. From the start of each phase we set every 2 successive stages to have equal numbers of cells. We therefore only require a total of 15 distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-3}\mathbf{f}_2(\mathbf{x}) \ 10^{-3}\mathbf{f}_3(\mathbf{x}) \ 10^{-3}\mathbf{f}_4(\mathbf{x}) \ 10\mathbf{f}_6(\mathbf{x})]$. Starting with the parameters

$$\begin{aligned}
 R_i(0) &= \begin{cases} 0 & \text{for } i = 1, 2, \\ 0 & \text{for } i = 3, 4, \\ 29.23 & \text{for } i = 5, 6, \\ 19.66 & \text{for } i = 7, 8, \\ 0 & \text{for } i = 9, 10, \end{cases} & Y_i(0) &= \begin{cases} 33.02 & \text{for } i = 1, 2, \\ 0 & \text{for } i = 3, 4, \\ 0 & \text{for } i = 5, 6, \\ 0 & \text{for } i = 7, 8, \\ 0 & \text{for } i = 9, 10, \end{cases} \\
 G_i(0) &= \begin{cases} 11.04 & \text{for } i = 1, 2, \\ 0 & \text{for } i = 3, 4, \\ 4.29 & \text{for } i = 5, 6, \\ 24.41 & \text{for } i = 7, 8, \\ 0 & \text{for } i = 9, 10, \end{cases} & L_r &= 10.28 \text{ h}, \\
 & & L_y &= 3.87 \text{ h}, \\
 & & L_g &= 12.85 \text{ h},
 \end{aligned} \tag{S17}$$

we obtain the same parameterisation Equation (S17). Note that $L_r + L_y + L_g = 27.00$ h, in good agreement with the observed cell cycle time of 27 h.

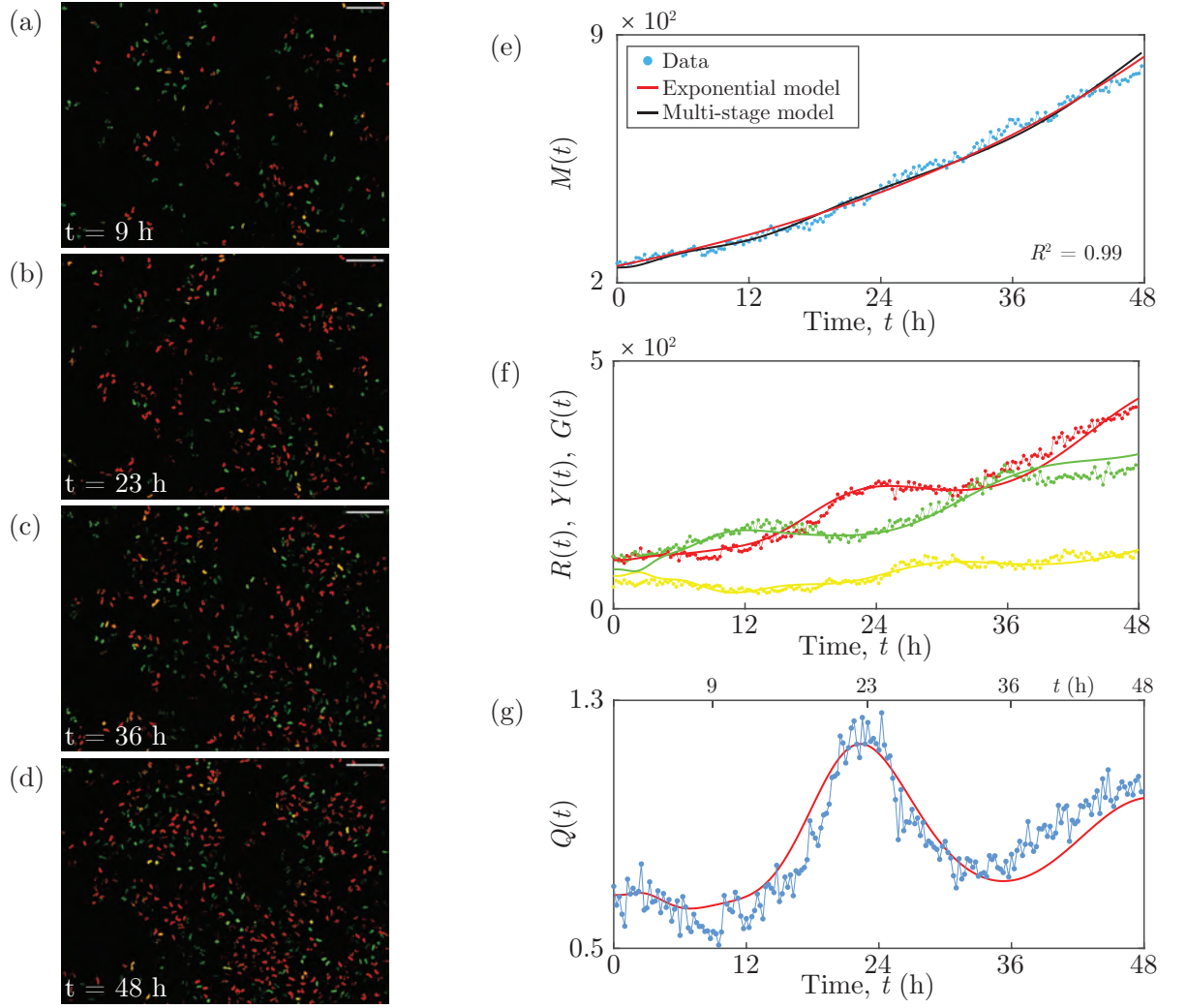


Figure 3.5: WM983C experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with Fucci-WM983C cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.6 WM983C cell line - Figure 3.6 The experimentally-determined mean cell cycle time for WM983C is $\mathcal{T} = 27$ h [13]. We partition each cell cycle phase into $N = 10$ stages, giving a total of $k = 30$ stages for the complete cell cycle. From the start of each phase we set every 5 successive stages to have equal numbers of cells. We therefore only require a total of 6 distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-3}\mathbf{f}_2(\mathbf{x}) \ 10^{-3}\mathbf{f}_3(\mathbf{x}) \ 10^{-3}\mathbf{f}_4(\mathbf{x}) \ 10^{-3}\mathbf{f}_7(\mathbf{x})]$. Trialling different parameters chosen randomly and uniformly from $(0.1, 1)$ for $R_i(0)$, $Y_i(0)$ and $G_i(0)$, and from $(4, 20)$ for L_r , L_y and L_g , we obtain the parameterisation

$$\begin{aligned}
 R_i(0) &= \begin{cases} 12.82 & \text{for } i = 1, \dots, 5, \\ 13.13 & \text{for } i = 6, \dots, 10, \end{cases} & Y_i(0) &= \begin{cases} 7.86 & \text{for } i = 1, \dots, 5, \\ 8.47 & \text{for } i = 6, \dots, 10, \end{cases} \\
 & & L_r &= 9.06 \text{ h}, \\
 G_i(0) &= \begin{cases} 18.10 & \text{for } i = 1, \dots, 5, \\ 7.51 & \text{for } i = 6, \dots, 10, \end{cases} & L_y &= 6.50 \text{ h}, \\
 & & L_g &= 15.79 \text{ h}.
 \end{aligned} \tag{S18}$$

Note that $L_r + L_y + L_g = 31.35$ h, in good agreement with the observed cell cycle time of 27 h.

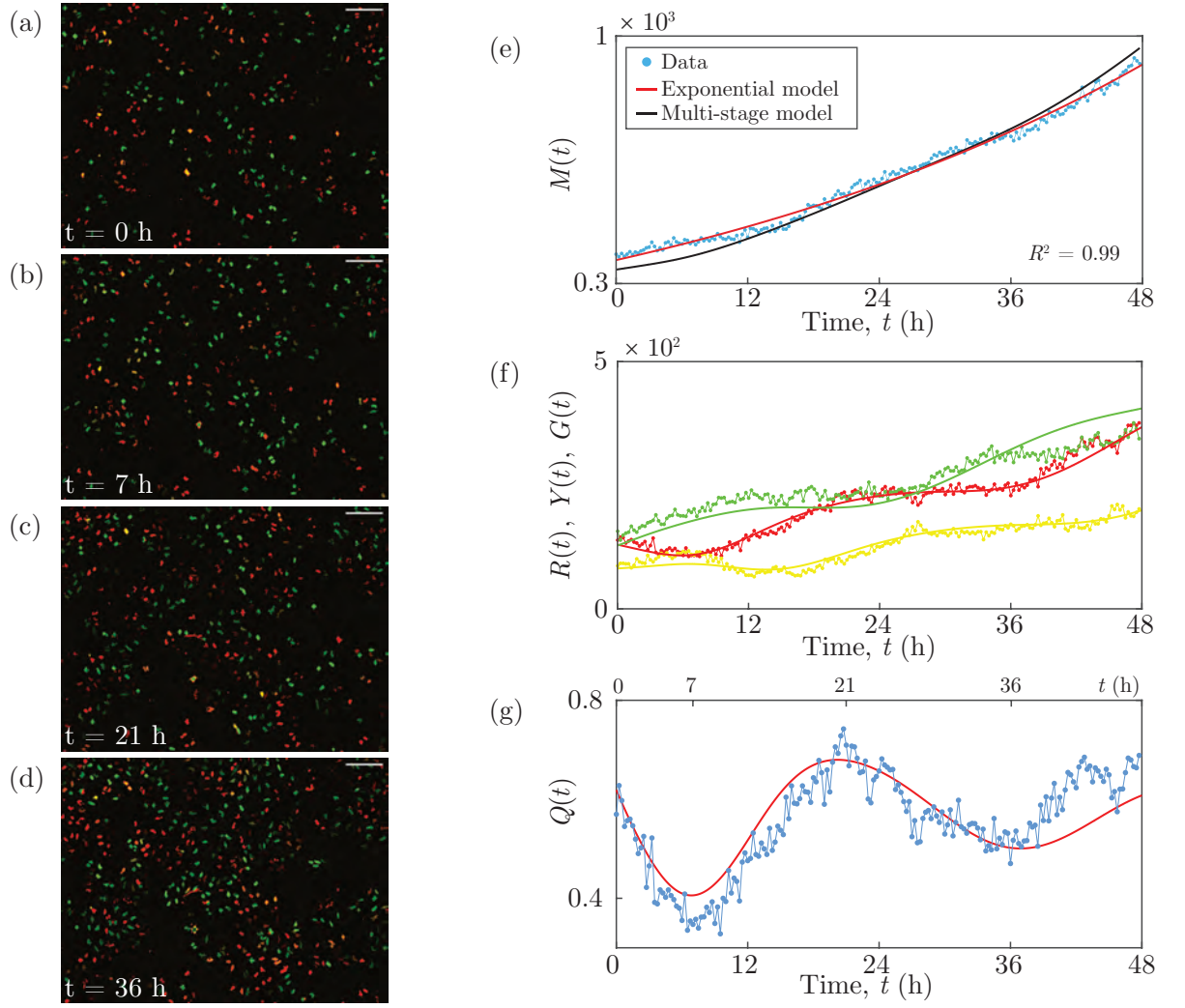


Figure 3.6: WM983C experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with Fucci-WM983C cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.7 WM983C cell line - Figure 3.7 The experimentally-determined mean cell cycle time for WM983C is $\mathcal{T} = 27$ h [13]. We partition each cell cycle phase into $N = 20$ stages, giving a total of $k = 60$ stages for the complete cell cycle. From the start of each phase we set every 10 successive stages to have equal numbers of cells. We therefore only require a total of 6 distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-3}\mathbf{f}_2(\mathbf{x}) \ 10^{-3}\mathbf{f}_3(\mathbf{x}) \ 10^{-3}\mathbf{f}_4(\mathbf{x}) \ 10^{-2}\mathbf{f}_5(\mathbf{x})]$. Trialling different parameters chosen randomly and uniformly from $(0.1, 1)$ for $R_i(0)$, $Y_i(0)$ and $G_i(0)$, and from $(4, 20)$ for L_r , L_y and L_g , we obtain the parameterisation

$$\begin{aligned} R_i(0) &= \begin{cases} 3.47 & \text{for } i = 1, \dots, 10, \\ 2.35 & \text{for } i = 11, \dots, 20, \end{cases} & Y_i(0) &= \begin{cases} 1.10 & \text{for } i = 1, \dots, 10, \\ 2.09 & \text{for } i = 11, \dots, 20, \end{cases} \\ L_r &= 9.22 \text{ h}, & & \\ G_i(0) &= \begin{cases} 2.68 & \text{for } i = 1, \dots, 10, \\ 1.64 & \text{for } i = 11, \dots, 20, \end{cases} & L_y &= 5.46 \text{ h}, \\ & & L_g &= 14.44 \text{ h}. \end{aligned} \tag{S19}$$

Note that $L_r + L_y + L_g = 29.12$ h, in good agreement with the observed cell cycle time of 27 h.

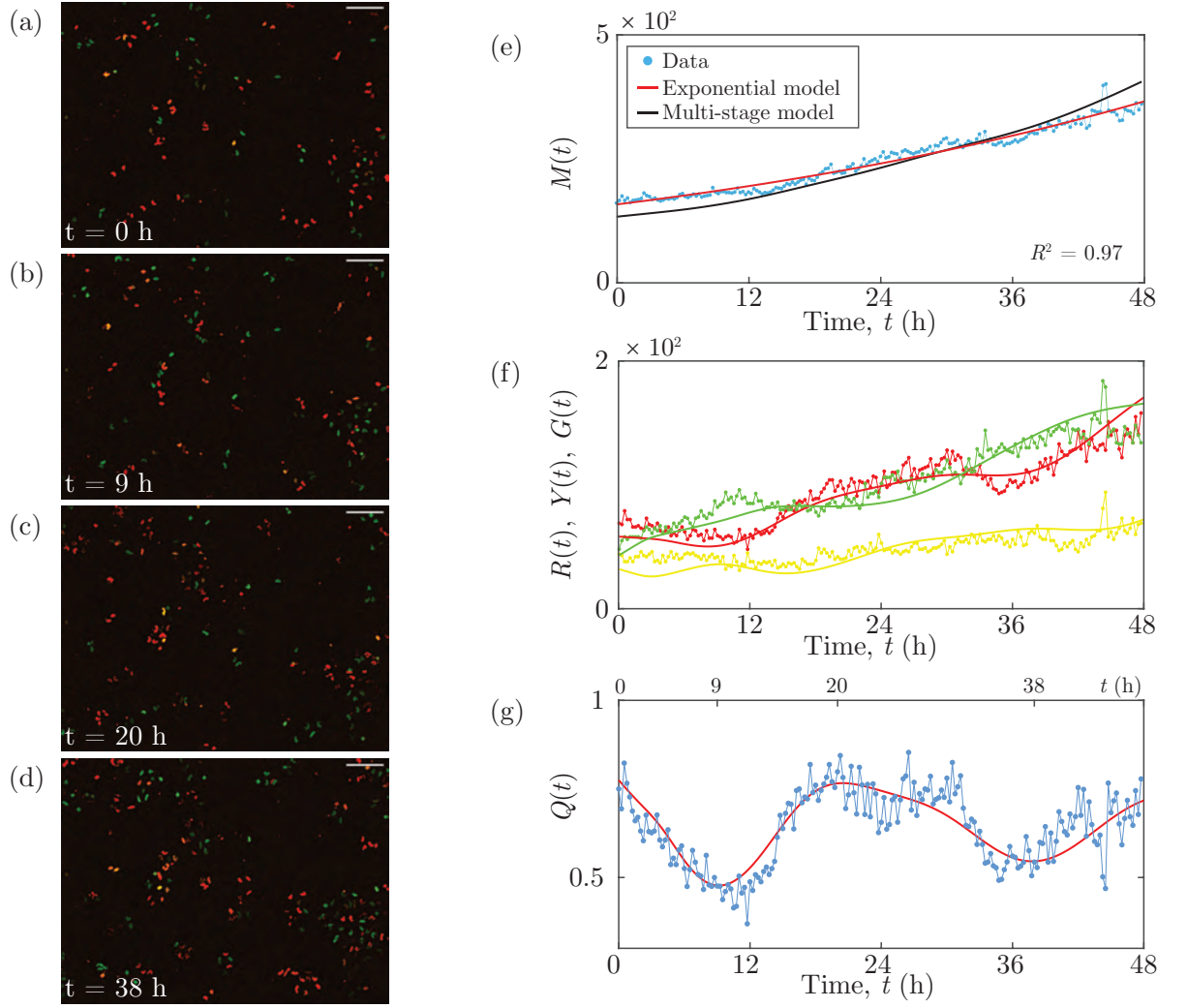


Figure 3.7: WM983C experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with Fucci-WM983C cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.8 1205Lu cell line - Figure 3.8 The experimentally-determined mean cell cycle time for 1205Lu is $\mathcal{T} = 36$ h [13]. We partition each cell cycle phase into $N = 20$ stages, giving a total of $k = 60$ stages for the complete cell cycle. From the start of each phase we set every 5 successive stages to have equal numbers of cells. We therefore only require a total of 12 distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-2}\mathbf{f}_2(\mathbf{x}) \ 10^{-2}\mathbf{f}_3(\mathbf{x}) \ 10^{-2}\mathbf{f}_4(\mathbf{x}) \ 0.5 \mathbf{f}_6(\mathbf{x})]$. Starting with the parameters

$$\begin{aligned}
 R_i(0) &= \begin{cases} 0.27 & \text{for } i = 1, \dots, 5, \\ 0 & \text{for } i = 6, \dots, 10, \\ 21.78 & \text{for } i = 11, \dots, 15, \\ 0 & \text{for } i = 16, \dots, 20, \end{cases} & Y_i(0) &= \begin{cases} 4.73 & \text{for } i = 1, \dots, 5, \\ 5.39 & \text{for } i = 6, \dots, 10, \\ 0 & \text{for } i = 11, \dots, 15, \\ 2.49 & \text{for } i = 16, \dots, 20, \end{cases} \\
 G_i(0) &= \begin{cases} 3.15 & \text{for } i = 1, \dots, 5, \\ 5.17 & \text{for } i = 6, \dots, 10, \\ 2.08 & \text{for } i = 11, \dots, 15, \\ 0.45 & \text{for } i = 16, \dots, 20, \end{cases} & L_r &= 20.97 \text{ h}, \\
 & & L_y &= 10.07 \text{ h}, \\
 & & L_g &= 10.52 \text{ h},
 \end{aligned} \tag{S20}$$

we obtain the parameterisation

$$\begin{aligned}
 R_i(0) &= \begin{cases} 0 & \text{for } i = 1, \dots, 5, \\ 13.79 & \text{for } i = 6, \dots, 10, \\ 5.73 & \text{for } i = 11, \dots, 15, \\ 4.19 & \text{for } i = 16, \dots, 20, \end{cases} & Y_i(0) &= \begin{cases} 9.39 & \text{for } i = 1, \dots, 5, \\ 0 & \text{for } i = 6, \dots, 10, \\ 0 & \text{for } i = 11, \dots, 15, \\ 5.04 & \text{for } i = 16, \dots, 20, \end{cases} \\
 G_i(0) &= \begin{cases} 2.31 & \text{for } i = 1, \dots, 5, \\ 0 & \text{for } i = 6, \dots, 10, \\ 3.52 & \text{for } i = 11, \dots, 15, \\ 0.63 & \text{for } i = 16, \dots, 20, \end{cases} & L_r &= 19.61 \text{ h}, \\
 & & L_y &= 7.99 \text{ h}, \\
 & & L_g &= 10.76 \text{ h},
 \end{aligned} \tag{S21}$$

Note that $L_r + L_y + L_g = 38.36$ h, in good agreement with the observed cell cycle time of 36 h.

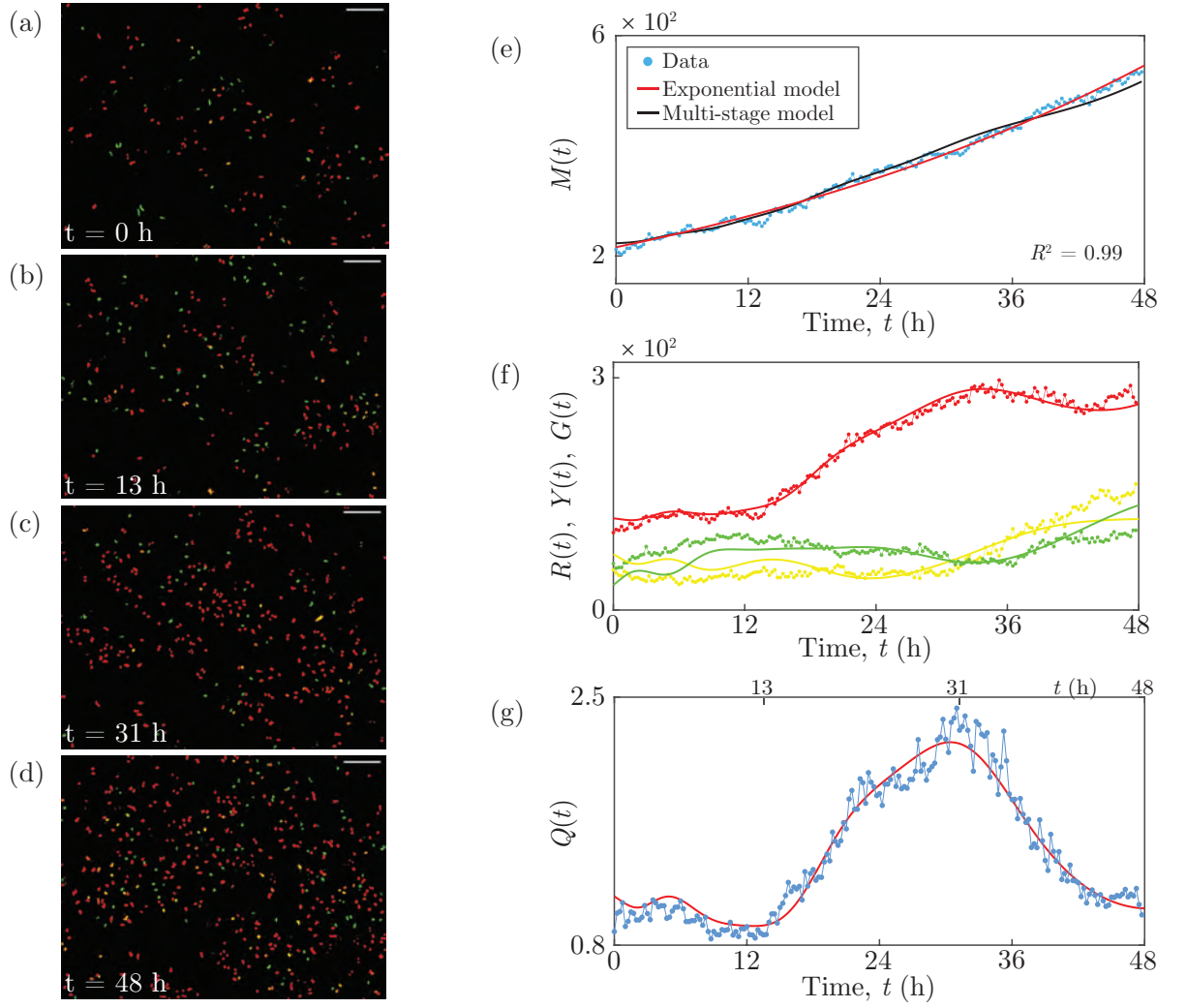


Figure 3.8: 1205Lu experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with Fucci-1205Lu cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.9 1205Lu cell line - Figure 3.9 The experimentally-determined mean cell cycle time for 1205Lu is $\mathcal{T} = 36$ h [13]. We partition each cell cycle phase into $N = 20$ stages, giving a total of $k = 60$ stages for the complete cell cycle. From the start of each phase we set every 4 successive stages to have equal numbers of cells. We therefore only require a total of 15 distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 10^{-2}\mathbf{f}_2(\mathbf{x}) \ 10^{-2}\mathbf{f}_3(\mathbf{x}) \ 10^{-2}\mathbf{f}_4(\mathbf{x}) \ 0.6\mathbf{f}_6(\mathbf{x})]$. Trialling different parameters chosen randomly and uniformly from $(0.1, 1)$ for $R_i(0)$, $Y_i(0)$ and $G_i(0)$, and from $(4, 20)$ for L_r , L_y and L_g , we obtain the parameterisation

$$\begin{aligned}
 R_i(0) &= \begin{cases} 0 & \text{for } i = 1, \dots, 4, \\ 23.75 & \text{for } i = 5, \dots, 8, \\ 0 & \text{for } i = 9, \dots, 12, \\ 13.54 & \text{for } i = 13, \dots, 16, \\ 0.92 & \text{for } i = 17, \dots, 20, \end{cases} & Y_i(0) &= \begin{cases} 12.55 & \text{for } i = 1, \dots, 4, \\ 0 & \text{for } i = 5, \dots, 8, \\ 0 & \text{for } i = 9, \dots, 12, \\ 0 & \text{for } i = 13, \dots, 16, \\ 0.67 & \text{for } i = 17, \dots, 20, \end{cases} \\
 G_i(0) &= \begin{cases} 9.94 & \text{for } i = 1, \dots, 4, \\ 0 & \text{for } i = 5, \dots, 8, \\ 3.24 & \text{for } i = 9, \dots, 12, \\ 1.46 & \text{for } i = 13, \dots, 16, \\ 0 & \text{for } i = 17, \dots, 20, \end{cases} & L_r &= 19.20 \text{ h}, \\
 & & L_y &= 7.98 \text{ h}, \\
 & & L_g &= 10.89 \text{ h},
 \end{aligned} \tag{S21}$$

Note that $L_r + L_y + L_g = 38.07$ h, in good agreement with the observed cell cycle time of 36 h.

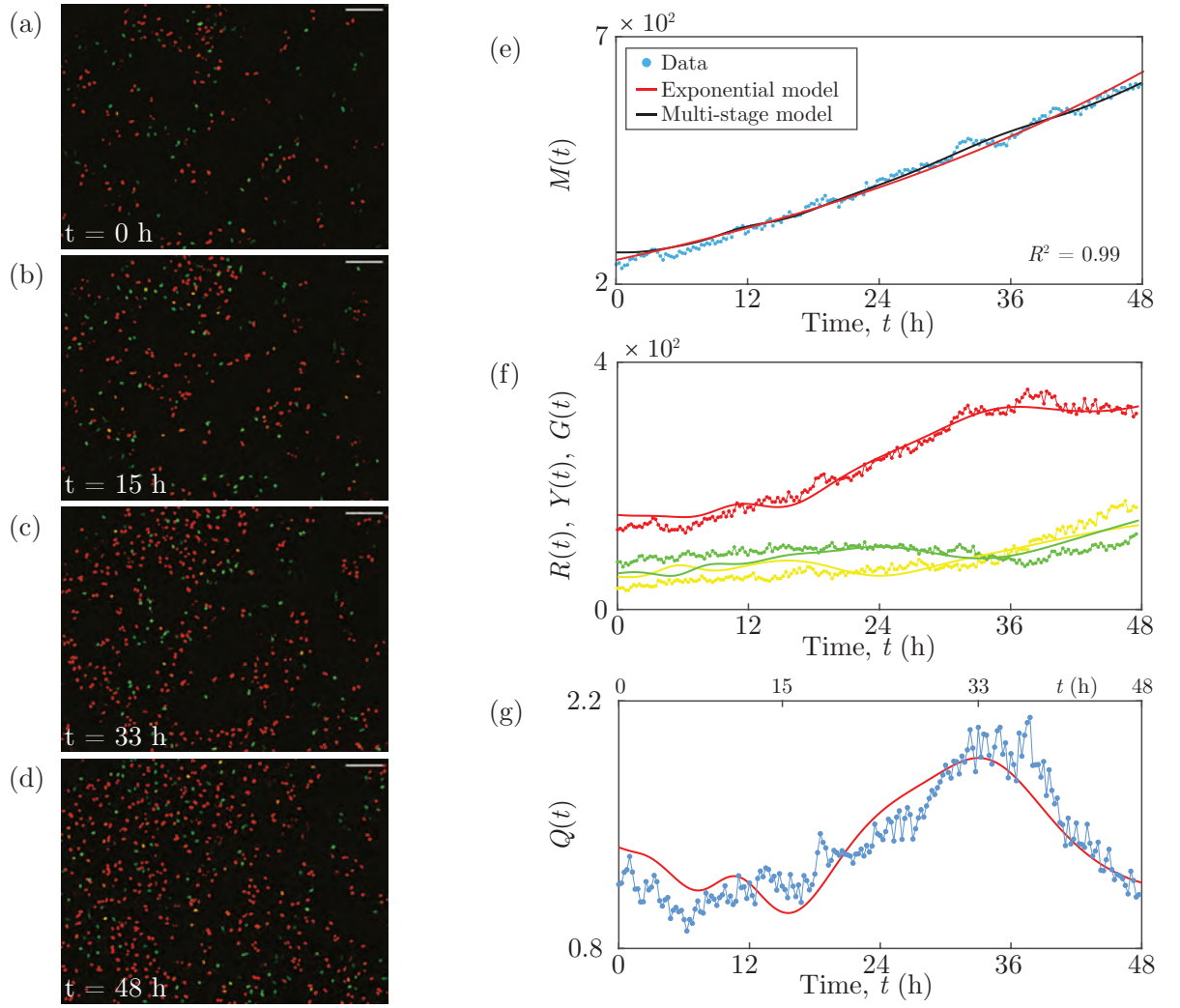


Figure 3.9: 1205Lu experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with FUCCI-1205Lu cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.4.2.10 1205Lu cell line - Figure 3.10 The experimentally-determined mean cell cycle time for 1205Lu is $\mathcal{T} = 36$ h [13]. We partition each cell cycle phase into $N = 12$ stages, giving a total of $k = 36$ stages for the complete cell cycle. From the start of each phase we set every 3 successive stages to have equal numbers of cells. We therefore only require a total of 12 distinct population parameters.

The vector objective function is $\mathbf{F}(\mathbf{x}) = [\mathbf{f}_1(\mathbf{x}) \ 3 \times 10^{-3} \mathbf{f}_2(\mathbf{x}) \ 3 \times 10^{-3} \mathbf{f}_3(\mathbf{x}) \ 3 \times 10^{-3} \mathbf{f}_4(\mathbf{x}) \ 0.2 \mathbf{f}_6(\mathbf{x})]$. Trialling different parameters chosen randomly and uniformly from $(0.1, 1)$ for $R_i(0)$, $Y_i(0)$ and $G_i(0)$, and from $(4, 20)$ for L_r , L_y and L_g , we obtain the parameterisation

$$\begin{aligned}
 R_i(0) &= \begin{cases} 0 & \text{for } i = 1, \dots, 3, \\ 24.89 & \text{for } i = 4, \dots, 6, \\ 25.23 & \text{for } i = 7, \dots, 9, \\ 0 & \text{for } i = 10, \dots, 12, \end{cases} & Y_i(0) &= \begin{cases} 20.14 & \text{for } i = 1, \dots, 3, \\ 0 & \text{for } i = 4, \dots, 6, \\ 0 & \text{for } i = 7, \dots, 9, \\ 0 & \text{for } i = 10, \dots, 12, \end{cases} \\
 G_i(0) &= \begin{cases} 2.68 & \text{for } i = 1, \dots, 3, \\ 20.35 & \text{for } i = 4, \dots, 6, \\ 0 & \text{for } i = 7, \dots, 9, \\ 0 & \text{for } i = 10, \dots, 12, \end{cases} & L_r &= 19.18 \text{ h}, \\
 & & L_y &= 9.08 \text{ h}, \\
 & & L_g &= 10.99 \text{ h},
 \end{aligned} \tag{S22}$$

Note that $L_r + L_y + L_g = 39.25$ h, in good agreement with the observed cell cycle time of 36 h.

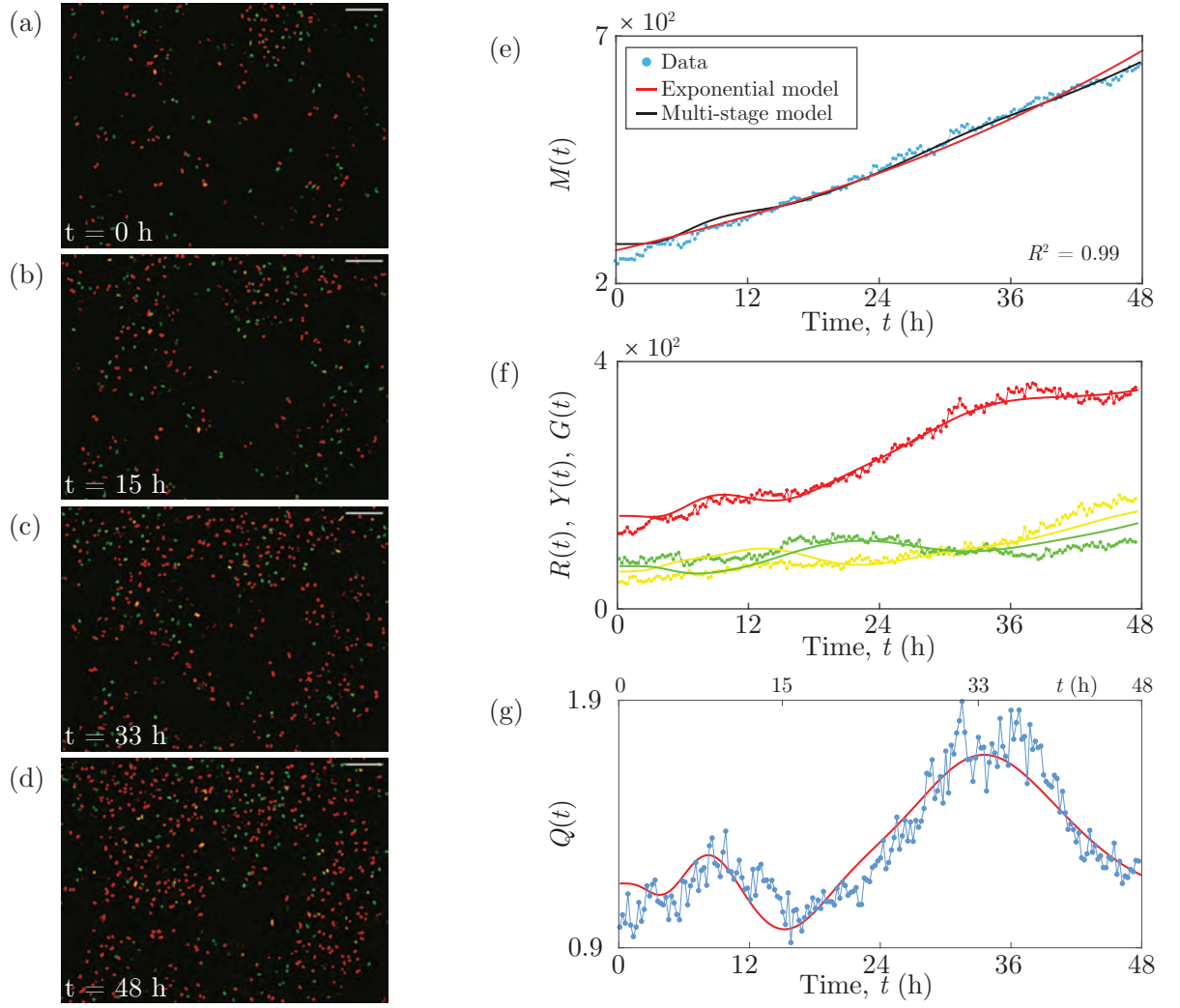


Figure 3.10: 1205Lu experimental data and multi-stage model solution. (a)–(d) Images of a proliferation assay with FUCCI-1205Lu cells. Scale bar $200 \mu\text{m}$. (e) $M(t)$. Linear regression of $\ln M(t)$ versus t provides R^2 . (f) $R(t)$, $Y(t)$ and $G(t)$. (g) $Q(t)$. Experimental data are shown as discs and the model solutions as curves.

3.5.5 All experimental data

Here we provide all of our new experimental data in the following forms:

- The total number of cells $M(t)$;
- The ratio $Q(t)$ of the number of cells in G1 to the combined number of cells in eS and S/G2/M;
- The discrete Fourier transform of $Q(t)$.

These data are obtained from the three cell lines C8161, WM983C and 1205Lu, and four independent experiments. In Experiments 1–3 we use one well of a 24-well plate, and in Experiment 4 we use two wells of a 24-well plate. From each well we obtain time-series stacks at six different positions. These data are available in Supporting Information 2–4, in the form of the number of cells in each phase, G1, eS and S/G2/M, at each time point.

In every experiment, the population growth $M(t)$ appears to be exponential, and the ratio $Q(t)$ reveals the presence of inherent synchronisation. In a given well, the six different positions can exhibit different degrees of inherent synchronisation. Further, the synchronisation can be out of phase between the different positions in a given well, and between the different wells.

Note that for some of the data there are a couple of consecutive time points which show a much higher total number of cells than expected, and a corresponding lower ratio in the ratio data. This is due to a large decrease in the signal-to-noise ratio in the green channel at these time points. The specific cause of this is unknown, however fluorescence microscopy is subject to such variations in the signal-to-noise ratio at times. As there is such a large reduction in the signal-to-noise ratio, it is not possible to reduce the unwanted noise without compromising the signal quality.

We provide the discrete Fourier transforms of $Q(t)$ for every data set to quantitatively demonstrate the existence of oscillating subpopulations in our experimental data. The transforms are obtained using the fast Fourier transform `fft` function [124] in MATLAB, without spectral interpolation. For clarity, we would like the amplitude of the Fourier transform to be zero at zero frequency, so we apply the transform to the time series $Q(t) - \overline{Q(t)}$, where $\overline{Q(t)}$ is the mean value of the time series. The transformed data are presented as single-sided spectra showing the magnitude of the Fourier transform, $A(f)$, as a function of frequency, f , where $0 \text{ h}^{-1} \leq f \leq 2 \text{ h}^{-1}$. Note that the Nyquist frequency is 2 h^{-1} .

The Fourier transforms all show dominant frequencies corresponding to periods of either 16, 24 or 48 h, which clearly indicate the presence of oscillations in $Q(t)$ for each of our data sets,

in accordance with the existence of inherent synchronisation. These periods are a reasonable approximation of the experimental cell-cycle durations of the cell lines. To increase the resolution of the frequencies, and thereby obtain better estimates of the periods of the oscillations, a time interval between the time-series images which is less than 15 minutes is required.

3.5.5.1 C8161 cell line

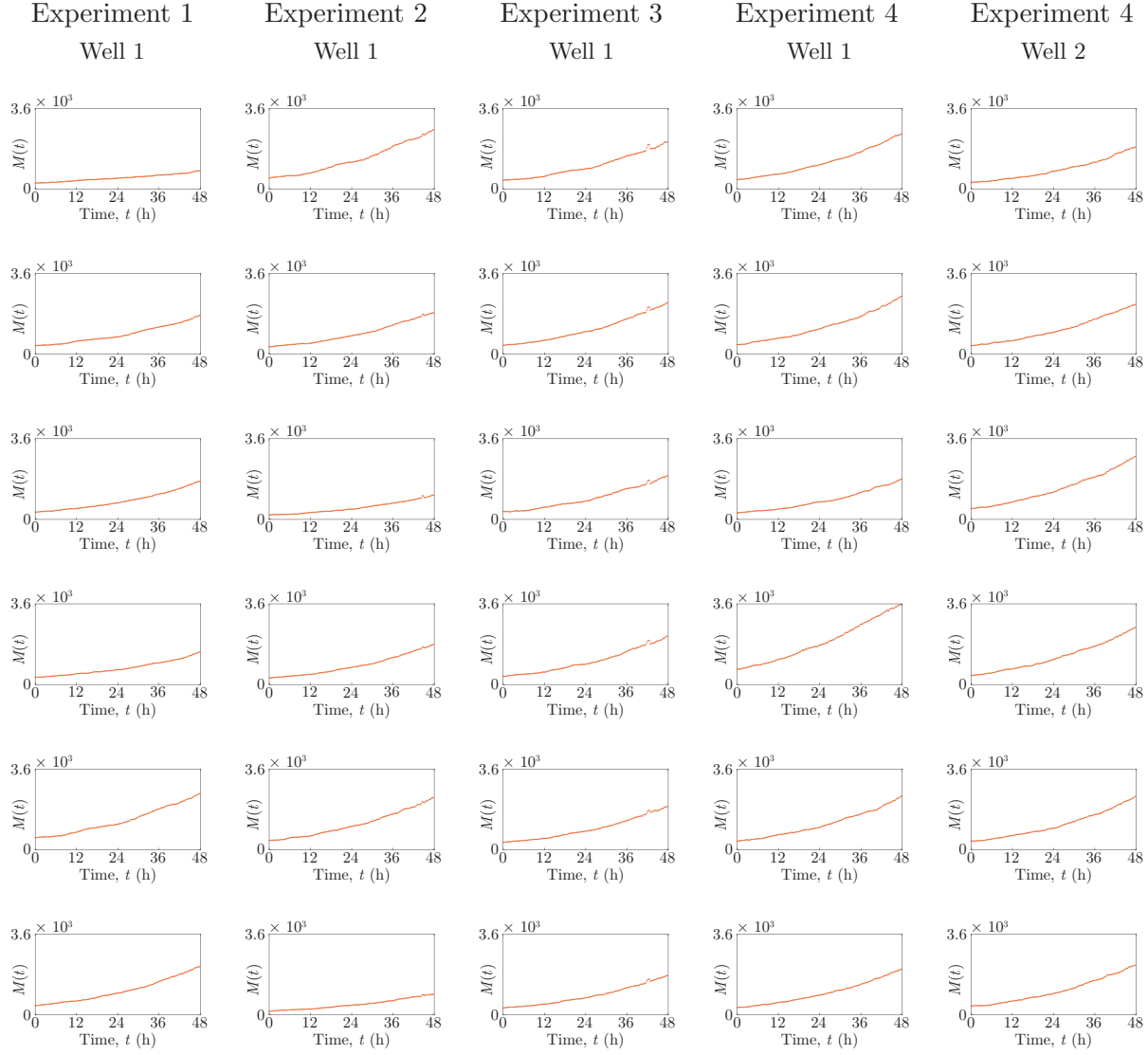


Figure 3.11: C8161 experimental data. Total number of cells $M(t)$. Each of the six rows corresponds to a different position in the specified well.

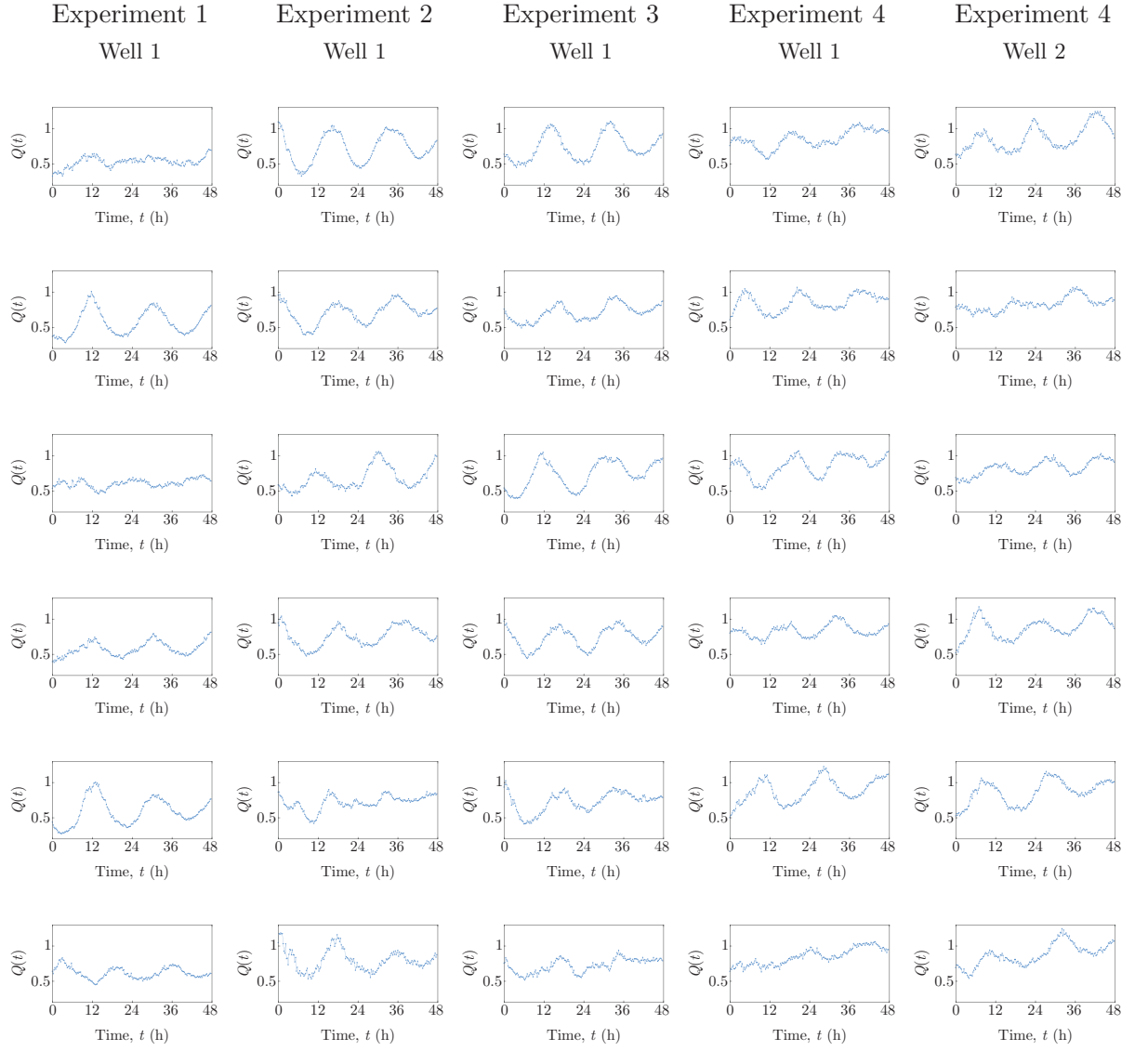


Figure 3.12: C8161 experimental data. Ratio $Q(t)$ of the number of cells in G1 to the number of cells in eS and S/G2/M. Each of the six rows corresponds to a different position in the specified well.

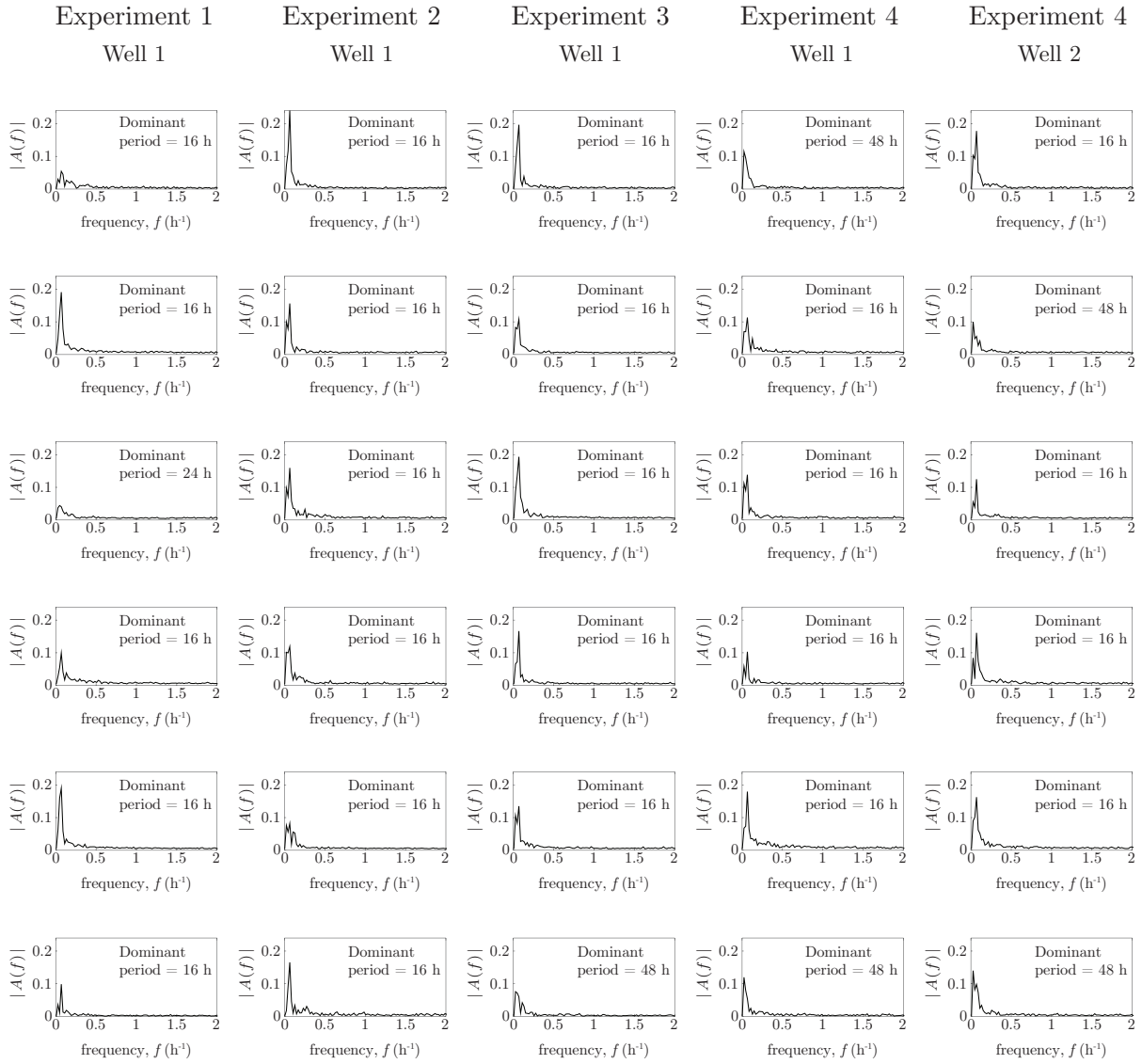


Figure 3.13: C8161 experimental data. Magnitude of the Fourier transform, $A(f)$, of the ratio $Q(t) - \bar{Q}(t)$, as a function of frequency, f . Each of the six rows corresponds to a different position in the specified well.

3.5.5.2 WM983C cell line

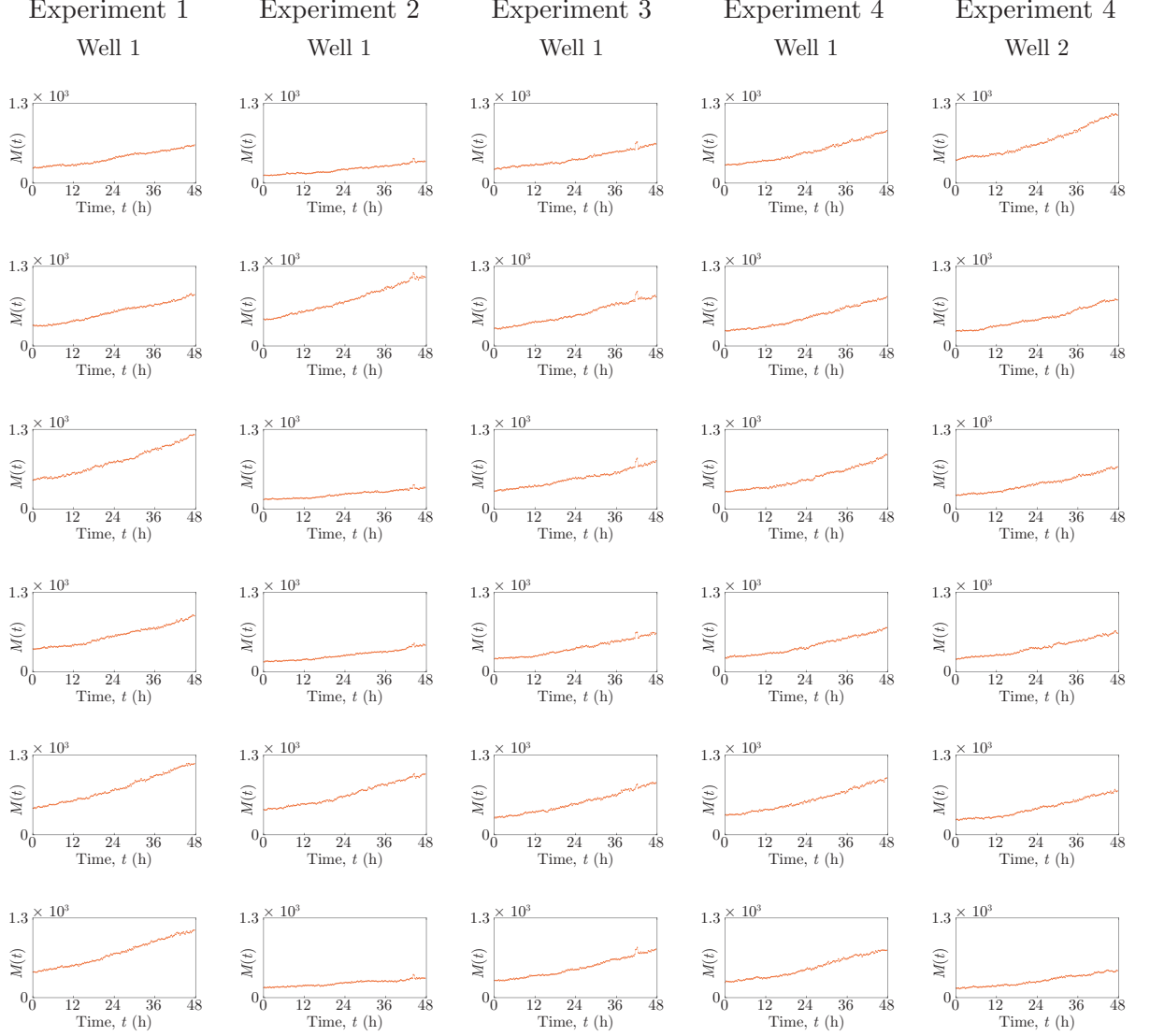


Figure 3.14: WM983C experimental data. Total number of cells $M(t)$. Each of the six rows corresponds to a different position in the specified well.

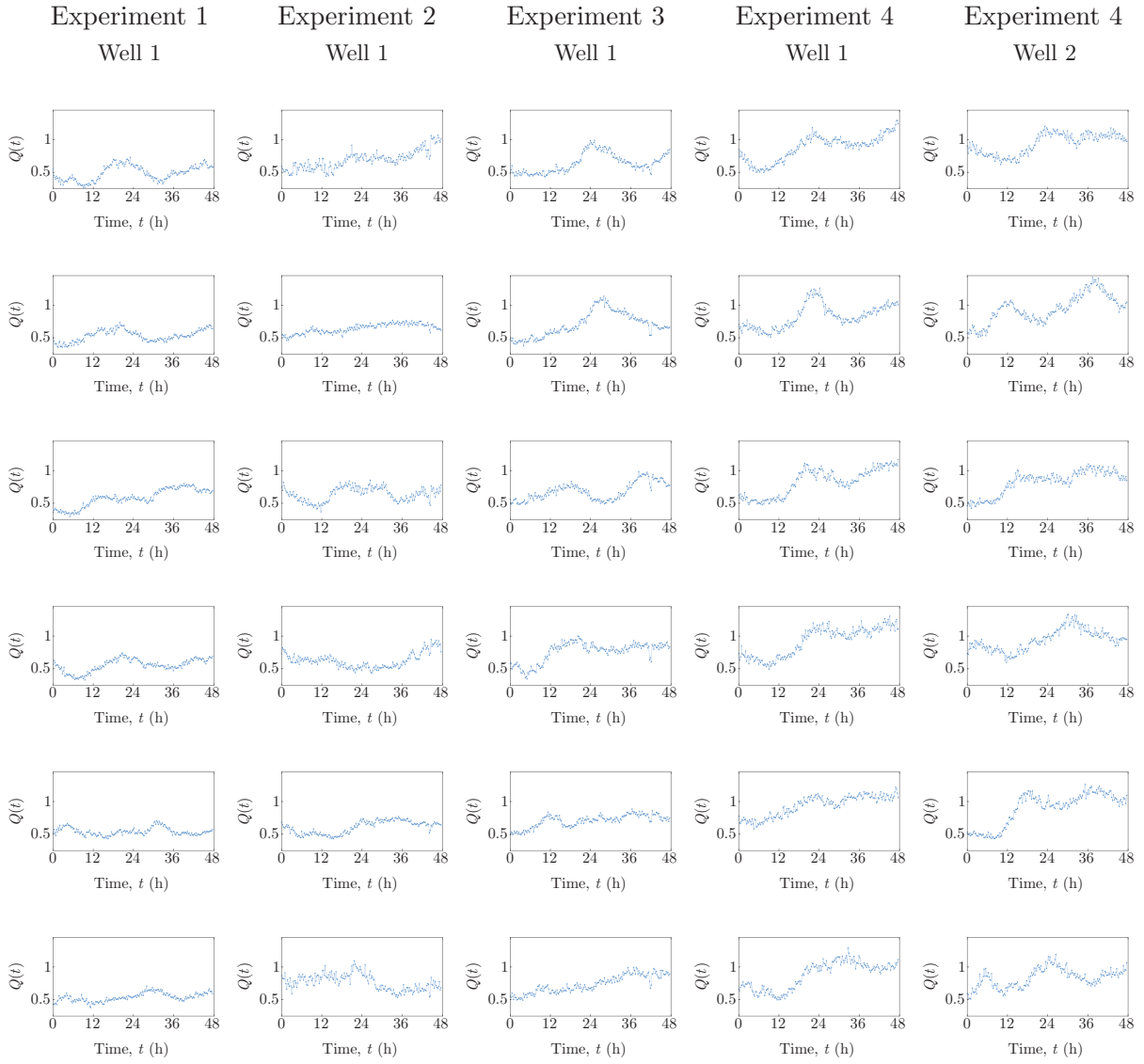


Figure 3.15: WM983C experimental data. Ratio $Q(t)$ of the number of cells in G1 to the number of cells in eS and S/G2/M. Each of the six rows corresponds to a different position in the specified well.

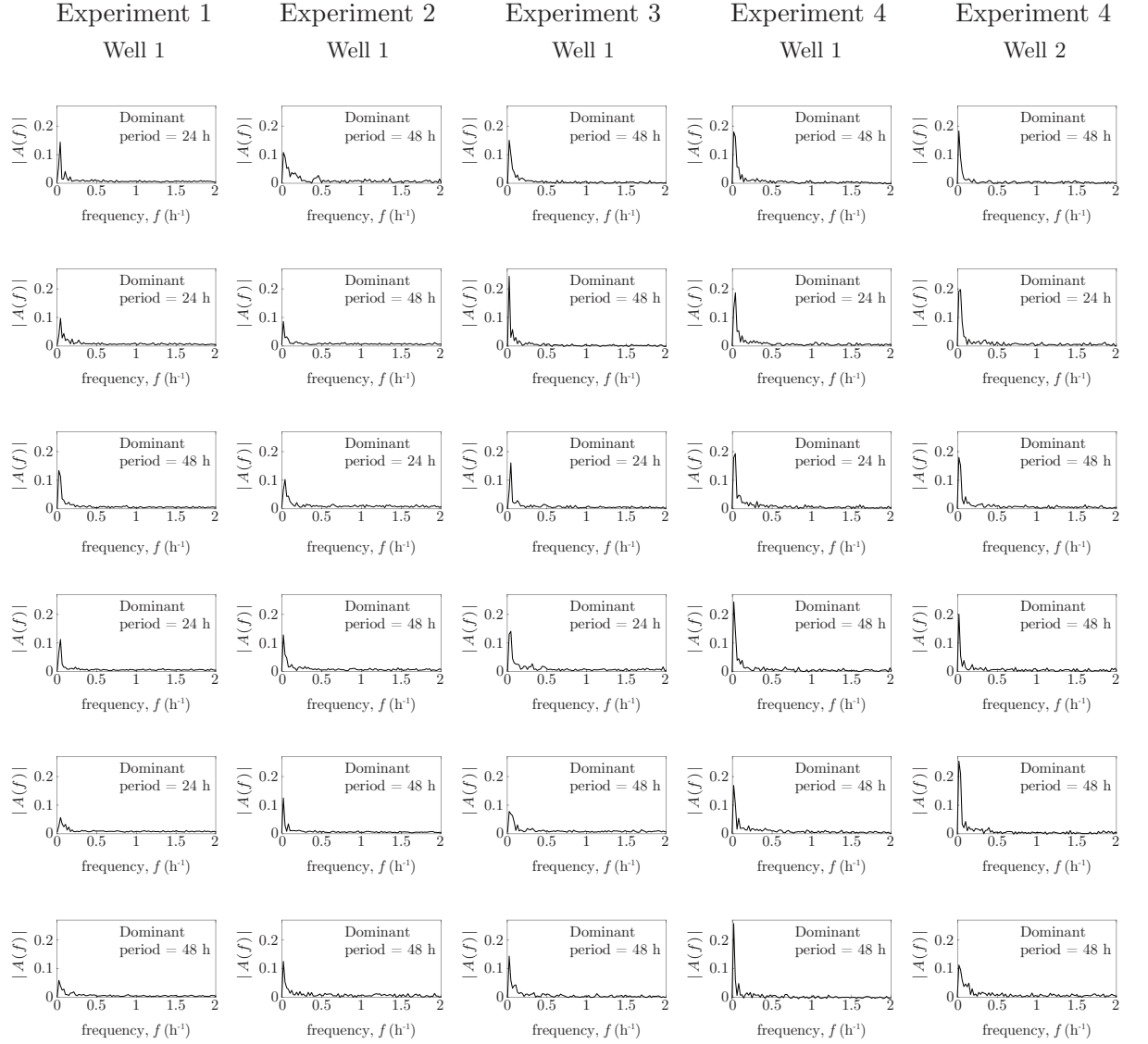


Figure 3.16: WM983C experimental data. Magnitude of the Fourier transform, $A(f)$, of the ratio $Q(t) - \overline{Q(t)}$, as a function of frequency, f . Each of the six rows corresponds to a different position in the specified well.

3.5.5.3 1205Lu cell line

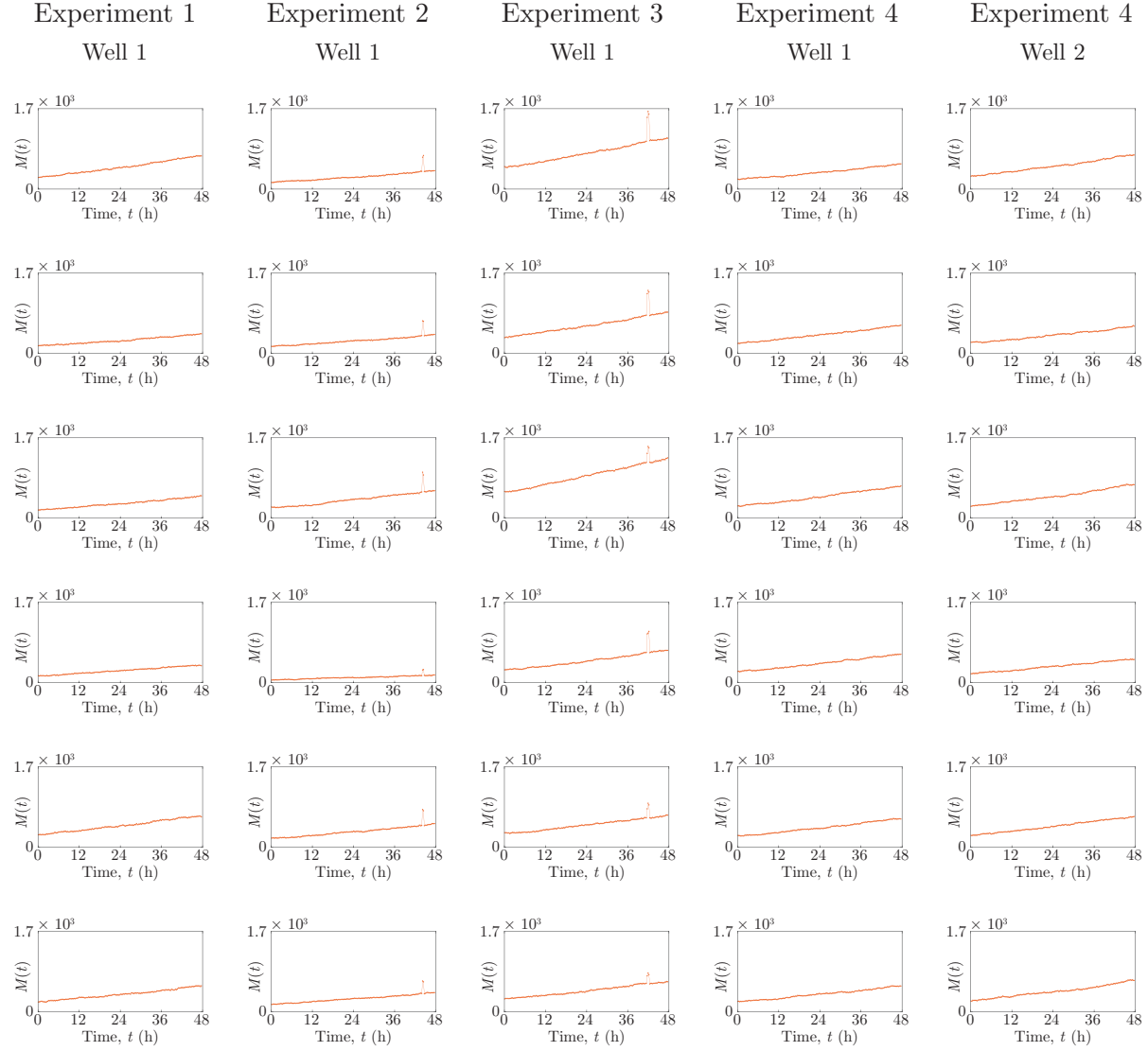


Figure 3.17: 1205Lu experimental data. Total number of cells $M(t)$. Each of the six rows corresponds to a different position in the specified well.

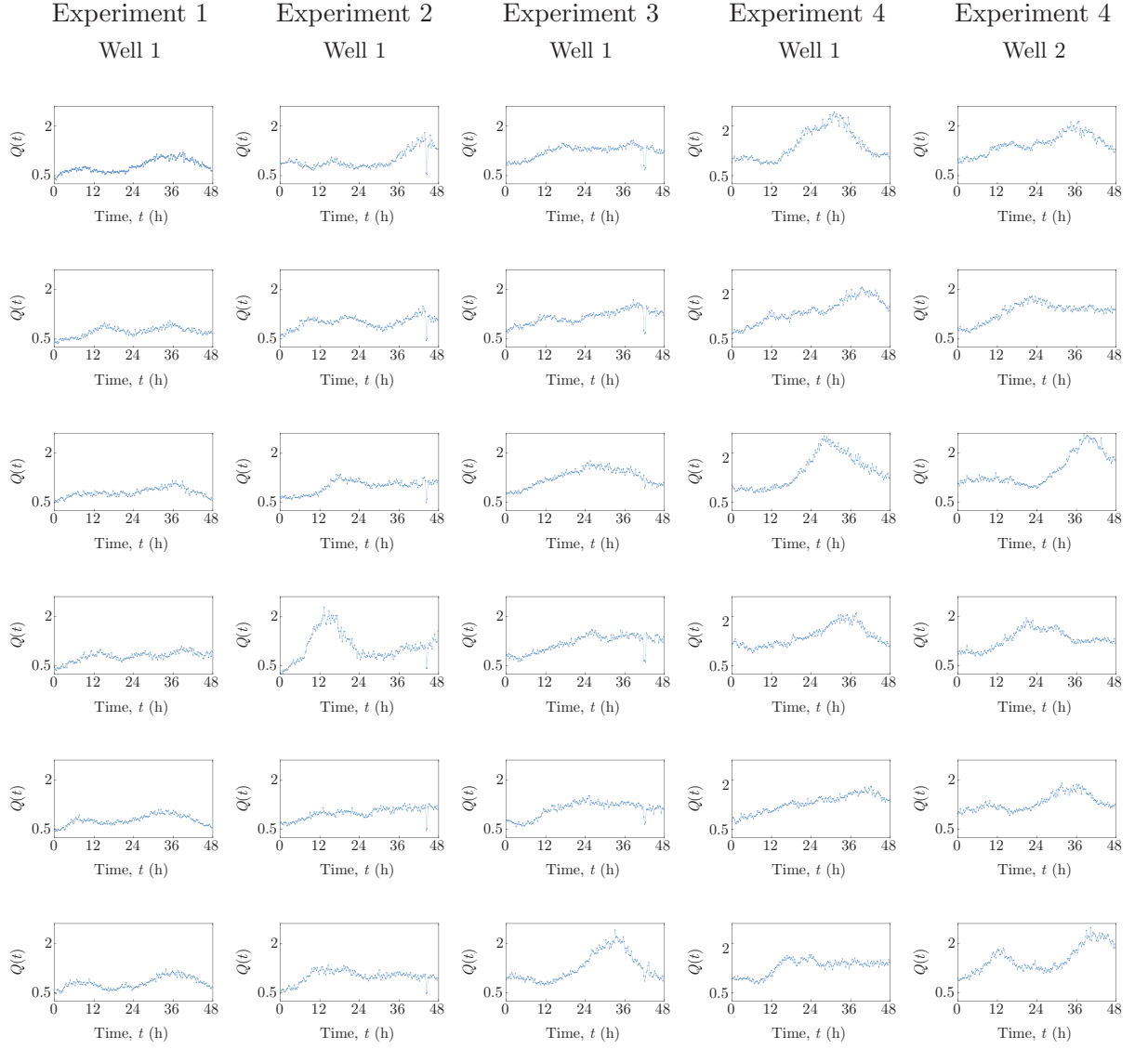


Figure 3.18: 1205Lu experimental data. Ratio $Q(t)$ of the number of cells in G1 to the number of cells in eS and S/G2/M. Each of the six rows corresponds to a different position in the specified well.

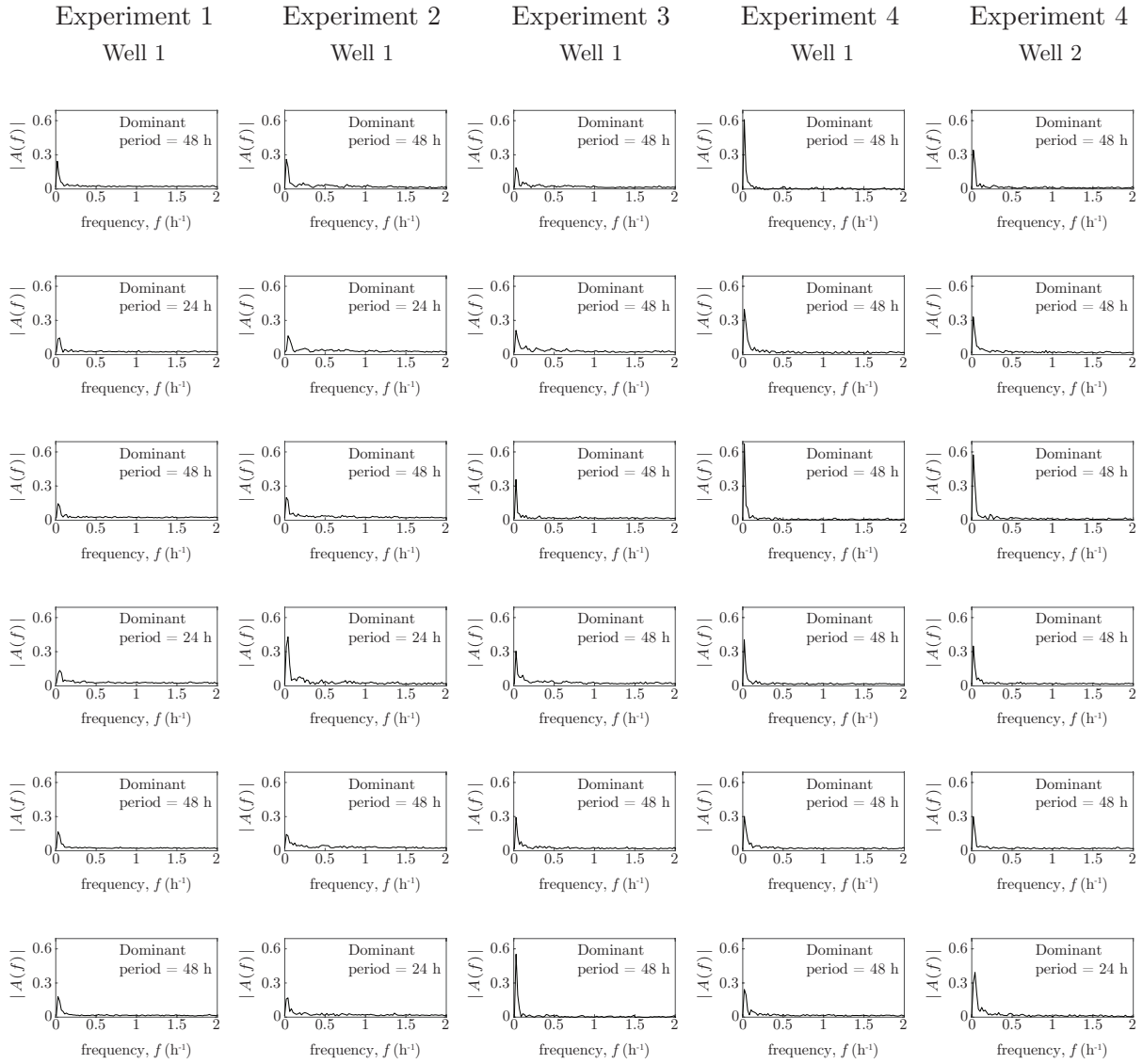


Figure 3.19: 1205Lu experimental data. Magnitude of the Fourier transform, $A(f)$, of the ratio $Q(t) - \overline{Q(t)}$, as a function of frequency, f . Each of the six rows corresponds to a different position in the specified well.

Examining go-or-grow using fluorescent cell-cycle indicators and cell cycle-inhibiting drugs

An article published as a *Letter* in *Biophysical Journal*

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Examining go-or-grow using fluorescent cell-cycle indicators and cell cycle-inhibiting drugs. *Biophysical Journal*. 2020;118:1243–1247.

Our cover image was selected for the cover of the 24 March 2020 issue of the Biophysical Journal.

Abstract

The go-or-grow hypothesis states that adherent cells undergo reversible phenotype switching between migratory and proliferative states, with cells in the migratory state being more motile than cells in the proliferative state. Here we examine go-or-grow in 2-D *in vitro* assays using melanoma cells with fluorescent cell-cycle indicators and cell cycle-inhibiting drugs. We analyse the experimental data using single-cell tracking to calculate mean diffusivities, and compare motility between cells in different cell-cycle phases and in cell-cycle arrest. Unequivocally, our analysis does not support the go-or-grow hypothesis. We present clear evidence that cell motility is independent of the cell-cycle phase, and non-proliferative arrested cells have the same motility as cycling cells.

4.1 Introduction

The *go-or-grow* hypothesis, also referred to as the *phenotype switching model* or the *migration/proliferation dichotomy*, proposes that adherent cells reversibly switch between migratory and proliferative phenotypes [57], exhibiting higher motility in the migratory state as motile cells are not using energy for proliferation [57–61]. Previous experimental investigations of the *go-or-grow* hypothesis are conflicting, as some studies support the hypothesis [57, 62, 63] while others refute it [13, 64, 65].

Go-or-grow was initially proposed as an explanation for the apparent mutual exclusivity of migration and proliferation for astrocytoma cells, first in 2-D *in vitro* experiments [63], and later for *in vivo* investigations [62]. In these early studies, claims for evidence of *go-or-grow* are based on the comparison of the subpopulation of cells at the perimeter of the cell population, where cells are considered to be invasive, with the subpopulation of cells in the central region, where cells are considered non-invasive. Data suggest that the proliferation rate is lower at the perimeter and higher in the centre, leading to the assertion that the more migratory cells are less proliferative. The experimental data, however, only indicate that the subpopulation at the perimeter is less proliferative as a whole compared with the centre, and therefore we cannot conclude definitively that the more migratory cells are less proliferative.

To test for evidence of *go-or-grow* it is necessary to look at the single-cell level, as in subsequent studies [13, 64, 65], where single-cell tracking is used with single-cell migration measured in terms of the net displacement of the cell trajectory. These three studies, none of which support *go-or-grow*, involve 2-D and 3-D *in vitro* experiments with medulloblastoma cells [65], 2-D *in vitro* experiments with mesothelioma, melanoma, and lung cancer cells [64], and 2-D and 3-D *in vitro* experiments with melanoma cells [13]. Studies of tumour heterogeneity in melanoma suggest that cells may reversibly switch between invasive and proliferative phenotypes [57]. As melanoma is highly metastatic, forms tumours that are very heterogeneous, and is well known to respond to MAPK inhibitors which induce G1 arrest [66, 67], melanoma cells are a prime candidate for studying the *go-or-grow* hypothesis.

Confirmation of *go-or-grow* would have important implications for anti-cancer treatments employing cell cycle-inhibiting drugs. For most eukaryotic cells, the cell cycle is a sequence of four discrete phases (Fig. 4.1a), namely gap 1 (G1), synthesis (S), gap 2 (G2) and mitosis (M). Cell-cycle arrest (Fig. 4.1d), which occurs when progression through the cell cycle halts [24], can be induced by cell cycle-inhibiting drugs [13, 33, 120]. An arrested cell is not proliferative, so the cell's energy could be utilised for migration, potentially leading to an exacerbation of

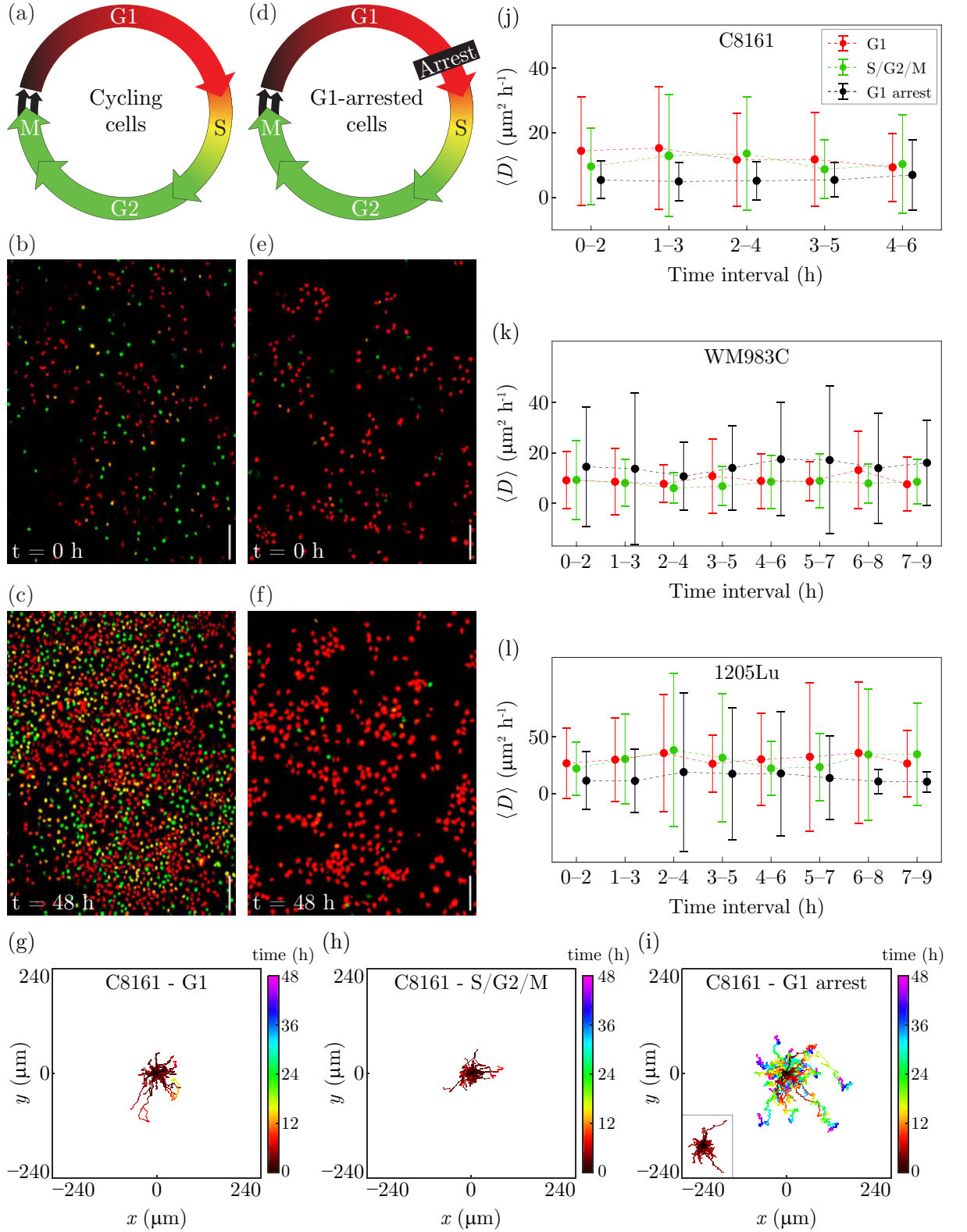


Figure 4.1: Experimental data and mean diffusivities. (a) The cell cycle, indicating the colour of FUCCI in each phase. (b)–(c) Experimental images of cycling C8161 cells; cell counts at 0 and 48 h are 331 and 1878, respectively. (d) The cell cycle, indicating the colour of FUCCI in each phase together with arrest in G1. (e)–(f) Experimental images of C8161 cells in G1 arrest (30 nM trametinib); cell counts at 0 and 48 h are 261 and 469, respectively. (g)–(i) 50 cell trajectories of G1 cycling, S/G2/M cycling and G1-arrested (30 nM trametinib) C8161 cells, respectively. (j)–(l) No difference in mean diffusivity, $\langle D \rangle$, for C8161, WM983C and 1205Lu cells, respectively. For each 2-h time interval, $\langle D \rangle$ is the mean of all individual diffusivities D corresponding to cells with trajectories within the time interval. In each case we show $\langle D \rangle$, and report the variability using $\langle D \rangle$ plus or minus the sample standard deviation. Data for each experimental condition are offset with respect to the time-interval axis for clarity. Scale bar 200 μm .

metastasis [59].

The go-or-grow hypothesis also has important implications for mathematical models of collective cell invasion in a population of migratory and proliferative cells. Such models of cell invasion are often based on the Fisher–Kolmogorov–Petrovskii–Piskunov (FKPP) equation [38,39,51,74],

$$\frac{\partial u}{\partial t} = D \frac{\partial^2 u}{\partial x^2} + \lambda u \left(1 - \frac{u}{K}\right), \quad (4.1)$$

where x is position, t is time, $u(x, t) > 0$ is cell density, $D > 0$ is the diffusivity, $\lambda > 0$ is the proliferation rate, and $K > 0$ is the carrying-capacity density. Eq. (4.1) and related adaptations, including stochastic analogues [40,41], have been successfully used to model cell migration *in vitro* and *in vivo* [43–46,85]. A key assumption underlying these models is that D is independent of the cell-cycle phase, which may not hold if cells are subject to go-or-grow as then a cycling, therefore non-arrested, cell may become less motile as it progresses through the cell cycle and nears cell division [13].

In this work we rigorously examine the go-or-grow hypothesis for adherent melanoma cells, for which phenotype switching between migratory and proliferative states is proposed to occur [57]. We use melanoma cell lines in this study as melanoma is the *prototype* for the phenotype switching model and is highly responsive to G1 arrest-inducing MEK inhibitors, such as trametinib. Melanoma cells are therefore an ideal candidate for studying go-or-grow [57,59,68]. Our experimental data are obtained from single-cell tracking in 2-D *in vitro* assays. We conduct our experiments in 2-D as it is the natural situation in which to commence a new experimental study, before utilising the knowledge gained in more complicated 3-D or *in vivo* experiments. Indeed, experimental studies of cell migration are often conducted in 2-D *in vitro* assays for several reasons: the observed cell migration is partly representative of cell migration *in vivo*; the assays are amenable to standard laboratory techniques, such as live-cell microscopy; and the relative ease of image analysis, such as cell counting and single-cell tracking [20–22]. Further, cell migration in 3-D may be affected by the properties of a 3-D matrix, which is not present in 2-D assays. For example, cell migration in 3-D through constricting pores can damage the nucleus and thereby cause a delay in cell division as the nucleus undergoes repair, which could be interpreted incorrectly as evidence for go-or-grow [125].

We employ fluorescent ubiquitination-based cell cycle indicator (FUCCI) [30], which consists of two reporters enabling visualisation of the cell cycle of individual live cells: when the cell is in G1 the nucleus fluoresces red, and when the cell is in S/G2/M the nucleus fluoresces green (Fig. 4.1(a)). During early S, called eS, both of the red and green reporters are active producing yellow. FUCCI allows us to study cell motility in G1 separately from cell motility in

S/G2/M [13, 69, 70, 85]. Specifically, we investigate cycling cells for differences in motility when the cells are in G1 compared with S/G2/M. Further, given the potential for an arrested cell to become more motile, we use a cell cycle-inhibiting drug to effect G1 arrest, and compare the motility of the arrested cells with cycling cells. Note that FUCCI does not provide delineation of S, G2, and M, so our motility measurements for these phases are combined into S/G2/M.

Our methodology for examining go-or-grow is novel in a number of ways. We induce G1 arrest in cells to determine whether non-proliferative cells have higher motility than cycling cells. We use experimental data to show that our three cell lines have distinctly different cell-cycle durations, ratios of duration in G1 to S/G2/M, and migration characteristics, all of which may affect motility under the go-or-grow hypothesis. Importantly, the data set we generate and analyse is large: for each cell line and experimental condition we randomly sample 50 single-cell trajectories for analysis out of more than 10^3 trajectories. In total, we analyse 450 carefully-collected trajectories for evidence of go-or-grow. Using these trajectories we carefully estimate diffusivities by first accounting for anisotropy in the cell migration, so that our estimates are based on time frames for which the cells are undergoing free diffusion.

4.2 Results and Discussion

Our data consist of time-series images, acquired every 15 min for 48 h, from 2-D proliferation assays using the melanoma cell lines C8161, WM983C and 1205Lu [13, 82, 84, 85], which have respective mean cell-cycle durations of approximately 21, 23 and 37 h [13]. The cell lines have very different ratios of durations in G1 to S/G2/M (Supplementary Material). Fig. 4.1(b)–(c) shows images of an assay with cycling C8161 cells at 0 and 48 h, illustrating the red, yellow and green nuclei corresponding to cells in G1, eS and S/G2/M, respectively. For comparison, Fig. 4.1(e)–(f) shows images of an assay with G1-arrested C8161 cells treated with the cell cycle-inhibiting drug trametinib (30 nM), illustrating that most cells are arrested in G1, appearing red. We use the lowest possible concentration of trametinib to induce G1 arrest for the experiment duration to minimise other effects. Consequently, each cell eventually returns to cycling, illustrated by the small proportion of green cells (Fig. 4.1(e)–(f)). These few green cells will eventually divide with both daughter cells arresting in G1. We quantitatively confirm the G1 arrest by comparing the cell counts between the experiments with cycling cells and arrested cells. For the cycling cells there is a 5.7-fold increase in the number of cells over 48 h (Fig. 4.1(b)–(c)), whereas there is only a 1.8-fold increase in the number of arrested cells over 48 h (Fig. 4.1(e)–(f)). The 1.8-fold increase in the population of G1-arrested cells is expected as we use the lowest possible concentration of trametinib. Consequently, a small subpopulation of cells may not be arrested at the

start of the experiment, and cells may recommence cycling during the experiment, producing a small increase in the population.

For each cell line, we employ single-cell tracking to obtain 50 trajectories of cells for each experimental condition: (i) G1 cycling; (ii) S/G2/M cycling; and (iii) G1 arrest. Each trajectory is selected randomly without replacement from the set of all trajectories for a given cell line and experimental condition. For the cycling cells, trajectories are recorded for the complete duration of the G1 or S/G2/M phase. For the G1-arrested cells, the duration of the trajectory corresponds to the maximum duration that the cell is arrested within the 48-h duration of the experiment (Supplementary Material).

In Fig. 4.1(g)–(i) we visualise the trajectories for cycling C8161 in G1 and S/G2/M, and C8161 in G1-arrest. The trajectories are translated so that their initial positions are at the origin. The trajectories of the G1-arrested cells are generally much longer than those for the cycling cells, as the arrested cells reside in G1 for a much longer duration than cycling cells reside in G1 or S/G2/M. Specifically, the approximate mean duration of cycling C8161 cells in G1 is 5 h, in S/G2/M is 6 h [13], and for cells in G1 arrest during the 48 h of the experiment is 34 h (Supplementary Material). Therefore, to easily compare the trajectories of G1-arrested cells with cycling cells in G1, we show within the inset the truncated trajectories of the G1 arrested cells. The trajectories are truncated to a duration equal to the mean duration of the corresponding trajectories for cycling cells in G1. Based on these data, the migration is isotropic, without any drift, and independent of the cell cycle phase. We now quantify these observations.

For each cell line and experimental condition, we find that the cell migration is isotropic and directional persistence dissipates within a relatively short lag time of 1 h (Supplementary Material). From each individual cell trajectory we estimate D , using the mean square displacement as a function of lag time, within 2-h time intervals. The intervals begin at the initial point of the trajectory, $t = 0$ h, with successive intervals offset by 1 h. We always use lag times from 1–2 h to guarantee the absence of persistence (Supplementary Material). We then calculate the mean diffusivity $\langle D \rangle$ for each time interval by averaging our estimates of D for those trajectories that extend to the end of that interval.

Fig. 4.1(j)–(l) shows, for each cell line, $\langle D \rangle$ for successive time intervals. From these data we arrive at clear conclusions (Supplementary Material), none of which are consistent with the go-or-grow hypothesis:

- For each cell line and experimental condition, there is little variation in $\langle D \rangle$ over time, indicating that there is no appreciable change in motility during each cell-cycle phase and during G1 arrest (Supplementary Material).

- For each cell line, there is little variation in $\langle D \rangle$ between cycling cells in G1, cycling cells in S/G2/M, and G1-arrested cells. The lack of variability in $\langle D \rangle$ is remarkable, and clearly demonstrates that cells in G1 are not more motile than cells in S/G2/M, and that G1-arrested cells at no time become more migratory than the cycling cells.
- Even though our three cell lines have very different proliferation and migration characteristics (Supplementary Material), our estimate of $\langle D \rangle$ is remarkably consistent across the three very different cell lines.

In summary, our analysis of cell migration in 2-D assays using three melanoma cell lines does not support the go-or-grow hypothesis. We find that cell motility is independent of the cell-cycle phase, so that the implication from go-or-grow that cells are more motile in G1 than in S/G2/M when they are nearing cell division is not supported by our data. Notably, there is no change in cell motility when we effect drug-induced G1 arrest in the cells, again displaying a lack of support for the go-or-grow hypothesis.

4.3 Supplementary Material

4.3.1 Materials and Methods

4.3.1.1 Experiments

Our experimental data consist of microscopy time-series images of two-dimensional cell proliferation assays using the three melanoma cell lines C8161, WM983C and 1205Lu [13, 82, 84, 85], which have cell cycle durations of approximately 21, 23 and 37 h, respectively [13]. Here we discuss in detail our experimental set-up and analysis of the time-series images.

4.3.1.1.1 Cell culture The human melanoma cell lines C8161 (kindly provided by Mary Hendrix, Chicago, IL, USA), WM983C and 1205Lu (both kindly provided by Meenhard Herlyn, Philadelphia, PA, USA) were genotypically characterised [101–104], grown as described in [106], and authenticated by STR fingerprinting (QIMR Berghofer Medical Research Institute, Herston, Australia).

We maintain the cell cultures to prevent any induced synchronisation from cell cycle arrest in G1 phase. In general, such induced synchronisation can occur through various culture conditions, namely contact inhibition of proliferation at relatively high population densities [119], decreased pH of the growth medium due to the concentration of acidic cell-metabolites such as lactic acid [118], and reduced availability of nutrients such as serum [33]. We prevent induced

synchronisation by passaging the cells every three days, and on the day prior to setting up an experiment, to maintain a subconfluent cell density and a fresh growth medium.

4.3.1.1.2 Fluorescent ubiquitination-based cell cycle indicator To generate stable melanoma cell lines expressing the FUCCI constructs, mKO2-hCdt1 (30-120) and mAG-hGem (1-110) [30] were subcloned into a replication-defective, self-inactivating lentiviral expression vector system as previously described [105]. The lentivirus was produced by co-transfection of human embryonic kidney 293T cells. High-titer viral solutions for mKO2-hCdt1 (30/120) and mAG-hGem (1/110) were prepared and used for co-transduction into the melanoma cell lines, and subclones were generated by single-cell sorting [13, 32, 106].

4.3.1.1.3 Proliferation assay: cycling cells Cells are seeded from subconfluent culture flasks onto a 24-well plate at a density of 10^4 cells cm^{-2} , with 2.5 ml of medium per well, which is 2.5 times the volume of the standard protocol. After incubating the plate for 24 h at 37 °C with 5% CO_2 , live-cell images are acquired at 15 min intervals over 48 h at six different positions within the well. For each cell line we performed four independent experiments, with one technical replicate in three experiments and two technical replicates in the fourth experiment. Therefore, for each cell line we have 30 sets of time-series images.

Our preliminary experiments used a standard 1 ml of medium in each well, however the cells started to arrest in G1 around 48 h after seeding, which is likely due to decreased pH of the medium from the lactic acid concentration. G1 arrest is visually obvious when viewing a time series of images of FUCCI cells as the cycling cells are either green or become green according to the duration of G1, while the G1-arrested cells remain red. We therefore performed preliminary tests in an attempt to prevent the cells from arresting in G1 during the experiment, and found that this is possible by increasing the volume of medium in each well to the workable maximum of 2.5 ml, given the volume of each well is 3 ml. The result of the increased volume of medium is that the cells do not begin to arrest in G1 until around 72 h following seeding, which provides us with almost 48 h of imaging using cells that have minimal G1 arrest. Our definition of minimal G1 arrest is that there is no visually detectable arrest due to culture conditions, such as low extracellular pH, throughout the time series images except perhaps within the last hour of the experiment.

4.3.1.1.4 Proliferation assay: G1-arrested cells The experimental set up is identical to that for the cycling cells, except that the MEK inhibitor trametinib, which causes G1 cell-cycle arrest, is added at a concentration of 30 nM. Trametinib is added 3.5 h after seeding, to allow

the cells to attach to the plate before becoming affected by the drug. In order to minimise other effects of trametinib we employ the minimum concentration of the drug that causes G1 arrest. Preliminary testing found that the optimal concentration for our cell lines is 30 nM. For each cell line, we performed two independent experiments, with two technical replicates in each experiment. Therefore, for each cell line we have 24 sets of time-series images.

4.3.1.1.5 Summary of experimental conditions Our experiments provide nine experimental conditions with which to examine the go-or-grow hypothesis:

- | | | |
|------------------------------|-------------------------------|-------------------------------|
| (1) C8161 in G1, cycling | (4) WM983C in G1, cycling | (7) 1205Lu in G1, cycling |
| (2) C8161 in S/G2/M, cycling | (5) WM983C in S/G2/M, cycling | (8) 1205Lu in S/G2/M, cycling |
| (3) C8161 in G1, arrested | (6) WM983C in G1, arrested | (9) 1205Lu in G1, arrested |

4.3.1.2 Image processing and analysis

Our microscopy data consist of multi-channel time-series stacks of images obtained from 2-D proliferation assays. FUCCI cells in G1 phase appear in the red channel and in S/G2/M phase appear in the green channel. We consider nine experimental conditions: namely, for each of the three cell lines C8161, WM983C and 1205Lu we have cycling cells in G1, cycling cells in S/G2/M, and arrested cells in G1.

It is standard terminology in the microscopy field to refer to an object of interest, such as the image of a cell or cell nucleus, within a microscopy image as a *spot* [126, 127]. Note that the spots in the red and green channels of our microscopy images correspond to the cell nuclei and not to the whole cell. Nevertheless, throughout this document we refer to the red and green spots interchangeably as cell nuclei or cells.

The microscopy data are processed and analysed with Fiji/ImageJ and MATLAB with as much automation as possible. Two main procedures are required. The first procedure is to determine, in each image, the centroids of the spots corresponding to cells in the G1, eS and S/G2/M phases. This requires images which have been processed and segmented to remove as much of the background noise as possible. The second procedure is to obtain 50 cell trajectories for each of the nine experimental conditions using single-cell tracking, and to then use the spot centroids from the first procedure to authenticate the spots identified from tracking. The authentication is used to remove spots in the trajectories which correspond to background noise rather than a cell, and to remove spots from the G1 (red) and S/G2/M (green) trajectories

which are actually yellow and therefore appear in both of the red and green channels.

4.3.1.2.1 Centroids of cells The first procedure, which is completely automated within ImageJ and MATLAB, is to identify the spots in the images which correspond to cells and not background noise. This is achieved by preprocessing the images to maximise the signal-to-noise ratio, then thresholding the images to remove the remaining background noise, and finally determining the centroids of all spots in the images. The following describes this in detail.

Preprocessing: To maximise the accuracy in identifying spots, which in our case are images of cell nuclei, we enhance the quality of the microscopy images using ImageJ as follows.

1. Import the time-series stack with the Bio-Formats Importer plug-in, splitting the red and green channels.
2. Apply five iterations of Subtract Background with rolling-ball radius of 5 pixels.
3. Apply Enhance Contrast with the Equalize Histogram option selected.
4. Apply the Gaussian Blur filter with $\sigma = 1$.

Segmentation: We now identify the spots in the processed images using ImageJ.

1. Apply Auto-thresholding using the Yen method, selecting the option to “calculate the threshold for each image”.
2. The resulting binary images are then refined by applying the following commands in the prescribed order.
 - (a) Watershed.
 - (b) Fill Holes.
 - (c) Open, with iterations = 10 and count = 5.
 - (d) Watershed.

Analysis: For every image in the segmented binary time-series stacks we count the number of spots in each of the red and green channels using ImageJ. We then use MATLAB to determine which spots are yellow.

1. For each of the red and green channels, apply Analyze Particles in ImageJ with size range $5\text{--}\infty$ pixels² and the option “limit to threshold” selected. Output the stack position and the centroid of every spot in each channel.
2. We now need to determine which spots are red, yellow or green. A spot is red if it appears in the red channel and there is no corresponding spot in the green

channel. Similarly, a spot is green if it appears in the green channel and there is no corresponding spot in the red channel. A spot is then yellow if it appears in both of the red and green channels.

Identifying whether a spot is yellow, and therefore appearing in both of the red and green channels, is complicated by the possible alteration of the shape of the spot during image processing. While we process every image in exactly the same way, the original microscopy images may have different signal-to-noise ratios between the red and green channels. Consequently, there may be a subtle difference in the shape of a spot depending on the channel in which it is viewed, and thereby a difference in the centroid of the spot in each channel.

We therefore use MATLAB to determine which spots are red, yellow or green, using the stack position and centroid of each spot. The first step is to locate the yellow spots by choosing each red spot in turn and searching the green channel for a corresponding spot in the same stack position as the red spot. Once the yellow spots are identified, the remaining spots are either red or green. Details of this procedure follow below.

- (a) Choose a spot from the red channel and then search the green channel, in the same stack position as the red spot, for a corresponding spot such that the Euclidean distance between the centroids of the two spots is not greater than 3 pixels, noting that the pixel size in our images is $1.8150 \mu\text{m}$. This distance allows for a location error of the centroids of the red and green spots, whereby the centroids may be translated up to one pixel from the original centroid of the yellow spot in the unprocessed images. Placing the original yellow centroid at the centre of a 3×3 grid of pixels, the red and green centroids from the processed images may be located at any of the nine pixels in the grid. If the green channel has a spot within the specified distance of the red spot then the spot is yellow.
- (b) Once all of the yellow spots are found, the red spots are all of the spots in the red channel which are not yellow. Similarly, the green spots are all of the spots in the green channel which are not yellow.

4.3.1.2.2 Single-cell tracking For each of the nine experimental conditions we require 50 cell trajectories, which we obtain from one randomly-chosen time series, of which there are 30 time-series for cycling cells and 24 time series for G1-arrested cells. We employ single-cell tracking to obtain the trajectories from the red and green channels of our time-series images. The tracking is performed with images which are not thresholded to remove background noise,

as thresholding produces binary images in which the spots have lost most of the distinctive features which are required for tracking. Consequently, it is necessary to determine which of the spots identified by the tracking software correspond to cells rather than background noise, and we achieve this by comparing the positions of the tracked spots with the cell centroids that are obtained from thresholded images.

Cell tracking is performed with the TrackMate plug-in for ImageJ [128], using the default settings unless otherwise specified. The tracking is mostly automated within ImageJ, so that the trajectories are identified automatically with TrackMate, however when we select a trajectory for further analysis we manually correct any segmentation or tracking errors. The cell-tracking process is described in detail below.

Preprocessing: Within ImageJ, we first prepare the microscopy images for tracking, which includes reducing background noise in the images.

1. Import the time-series stack with the Bio-Formats Importer plug-in, splitting the red and green channels.
2. Apply five iterations of Subtract Background with rolling-ball radius of 5 pixels.
3. For the green channel, apply Enhance Contrast with no options selected, other than to process all slices.
4. For the red channel, modify the Brightness/Contrast settings by applying Reset to restore them to the original values, then apply Set with the “Minimum displayed value” equal to 0 and the “Maximum displayed value” equal to 90. Then select Apply to apply the LUT (lookup table) to all stack slices.

Tracking: We use the TrackMate plug-in to find all trajectories in the time-series images, with the following settings.

1. Segmentation is by the Laplacian of Gaussian detection (LoG) algorithm, with the settings detailed in Table 4.1, and we use neither the median filter nor sub-pixel localisation.
2. We do not filter any spots with initial thresholding of spot quality.
3. For the spot tracker, we use the Linear Assignment Problem (LAP) tracker with the settings detailed in Table 4.1.

4.3.1.2.3 Trajectory selection Once all of the trajectories in the time series of an experimental condition are identified, we randomly choose trajectories without replacement until 50

trajectories are selected which satisfy prescribed requirements. The requirements for deciding whether to keep or discard a chosen trajectory depend on the particular experimental condition.

Cycling cells in G1 or S/G2/M: If the chosen trajectory consists of the complete phase under consideration, therefore either G1 or S/G2/M, then we keep the trajectory. If the trajectory does not consist of the complete phase, such as the trajectory leaves the image area, begins prior to the first frame, or ends after the last frame, then we discard it. If the chosen trajectory has a segmentation or tracking error, and can be manually edited to produce the full trajectory, then we keep it, otherwise we discard it. As our cell culture conditions ensure minimal G1 arrest, and we randomly choose 50 trajectories from more than 10^3 trajectories throughout the complete time series, there is little probability of choosing a cell which is in G1 arrest rather than a cycling cell in G1.

G1-arrested cells: We need to be able to distinguish between a cell that is arrested in G1 and a cycling cell that is in G1. For each cell line, the time-series images for the G1-arrested cells indicate that almost all of the cells are in G1 phase at a given time, which is very distinct from the cycling cells for which there are cells in all phases of the cell cycle at any given time. It is reasonable to conclude that a randomly-chosen trajectory of a cell from a time series of G1-arrested cells is associated with a G1-arrested cell. Therefore, we keep a chosen trajectory unless it leaves the image area within a duration less than 2 h, as a trajectory of duration less than this is not suitable for our analysis. Most trajectories of G1-arrested cells are much longer than 2 h. If the chosen trajectory has a segmentation or tracking error then we manually edit it.

Table 4.1: Settings for the ImageJ plugin TrackMate, for each experimental condition. The LoG detector settings are for spot segmentation, and the frame-to-frame linking and track segment gap-closing settings are for spot tracking.

	C8161 G1	C8161 S/G2/M	C8161 G1 arrest	WM983C G1	WM983C S/G2/M	WM983C G1 arrest	1205Lu G1	1205Lu S/G2/M	1205Lu G1 arrest
LoG detector settings:									
Estimated blob diameter (μm)	14	19	22	28	30	26	16	20	26
Threshold	900	300	900	400	200	900	1200	400	900
Frame-to-frame linking:									
Maximum distance (μm)	50	50	50	50	50	50	50	50	50
Feature penalties - estimated diameter	2	0	2	2	1	2	0	0	2
Feature penalties - total intensity	2	0	2	2	1	2	0	0	2
Feature penalties - visibility	1	0	1	1	1	1	0	0	1
Feature penalties - quality	2	0	2	2	1	2	0	0	2
Track segment gap-closing:									
Maximum distance (μm)	50	50	50	50	50	50	50	50	50
Maximum frame gap	1	1	1	1	1	1	1	1	1

4.3.1.2.4 Trajectory authentication Each selected trajectory consists of a sequence of spots from either the red or the green channel. Since cells in eS phase appear in both of the red and green channels, and we only want trajectories corresponding to cells in the G1 or S/G2/M phases, we need to remove any spots from the trajectories which are yellow. This is accomplished in MATLAB by comparing the spot positions from tracking with the known centroids of the cells in G1, eS and S/G2/M obtained previously from the thresholded images. Since we are comparing the coordinates of each spot with the centroids of the cells in the same stack position of the time series, it is important to note that while ImageJ usually numbers the frames of a time series starting at one, TrackMate uses frame numbering beginning at zero, so it may be necessary to renumber the frames for each spot of a trajectory to ensure comparison within the same frame.

We set a tolerance of 10 μm when comparing the spot coordinates with the centroids to allow for the variability associated with estimating positions of spots within images which have been processed differently. Given that the cells have diameters much larger than 10 μm , for example C8161 cells have a typical diameter of more than 16 μm [109], the tolerance is relatively small but large enough to find the cell corresponding to a spot.

G1 trajectories: Here we detail the authentication of G1 trajectories.

1. Given a trajectory in the red channel, choose each spot in the trajectory in turn.
2. Calculate the Euclidean distance between the coordinates of the spot and the centroids of the cells in the red channel that are in the same stack position as the spot. If there is a cell in eS phase that is within 10 μm of the spot then we identify the spot as yellow, otherwise the spot corresponds to a cell in G1.
3. Note that any spots identified as yellow will occur towards the end of the trajectory in forward time, corresponding to the cell transitioning from G1 to eS. Therefore, we need to truncate the trajectory so that it corresponds only to the cell in G1 phase. Given that spots corresponding to cells in eS may not always appear yellow due to fluctuations in fluorescence intensity in either channel, we additionally identify a red spot as yellow if it is immediately between two previously identified yellow spots. So, to remove all yellow spots from the trajectory we begin at the final time point of the trajectory and progress backwards in time, removing spots which are yellow including any spots which are immediately between two previously identified yellow spots. The process terminates when a sequence of two red spots is reached, neither of which are identified as yellow.

S/G2/M trajectories: Here we detail the authentication of S/G2/M trajectories.

1. Given a trajectory in the green channel, choose each spot in the trajectory in turn.
2. Calculate the Euclidean distance between the coordinates of the spot and the centroids of the cells in the green channel that are in the same stack position as the spot. If there is a cell in eS phase that is within 10 μm of the spot then we identify the spot as yellow, otherwise the spot corresponds to a cell in S/G2/M.
3. Note that any spots identified as yellow will occur towards the beginning of the trajectory in forward time, corresponding to the cell transitioning from eS to S/G2/M. Therefore, we need to truncate the trajectory so that it corresponds only to the cell in S/G2/M phase. Given that spots corresponding to cells in eS may not always appear yellow due to fluctuations in fluorescence intensity in either channel, we additionally identify a green spot as yellow if it is immediately between two previously identified yellow spots. So, to remove all yellow spots from the trajectory we begin at the initial time point of the trajectory and progress forwards in time, removing spots which are yellow including any spots which are immediately between two previously identified yellow spots. The process terminates when a sequence of two green spots is reached, neither of which are identified as yellow.

4.3.2 Data analysis

Here we provide the detailed analysis of the trajectory data for the C8161, WM983C and 1205Lu cell lines.

4.3.2.1 Cell-cycle characteristics of each cell line

In Table 4.2 we provide for each cell line the mean durations of G1 phase, S/G2/M phase, and the total cell cycle for cycling cells [13], and the mean duration of induced G1 arrest (30nM trametinib).

The G1-arrest duration corresponds to the mean duration that the cells are arrested in G1 during the 48 h of the experiment. Note that the cells are likely to be arrested for longer than we can observe during 48 h, so the mean duration that we report corresponds to our observations during the experiment and not necessarily to the total duration for which the cells are arrested. Comparing the long durations of G1 arrest with the relatively short G1 durations of the cycling cells provides quantitative confirmation of the cell-cycle arrest. Indeed, in Figure 4.2 we show histograms of the durations of G1 arrest for the 50 cells of each cell line.

	C8161	WM983C	1205Lu
G1-arrest duration (h)	34 ± 15	34 ± 15	42 ± 11
G1 duration (h)	5 ± 3	6 ± 2	18 ± 9
S/G2/M duration (h)	6 ± 2	8 ± 3	10 ± 4
Total cell-cycle duration (h)	21 ± 8	23 ± 10	37 ± 16
Ratio of G1 to S/G2/M durations	0.9	0.7	1.8
Ratio of G1 to total durations	0.3	0.2	0.5
Ratio of S/G2/M to total durations	0.3	0.4	0.3

Table 4.2: Cell-cycle characteristics for the C8161, WM983C and 1205Lu cell lines. Data for the cycling cells are from [13], and correspond to 20 cells for each cell line. Data for the G1-arrested cells (30 nM trametinib) are from our current study, and correspond to 50 cells for each cell line. The duration of each cell cycle phase and cell cycle arrest is the mean for the corresponding cell population, and the error is one standard deviation from the mean. The data for the G1-arrest durations correspond to the durations that the cells are arrested during the 48 h of the experiments.

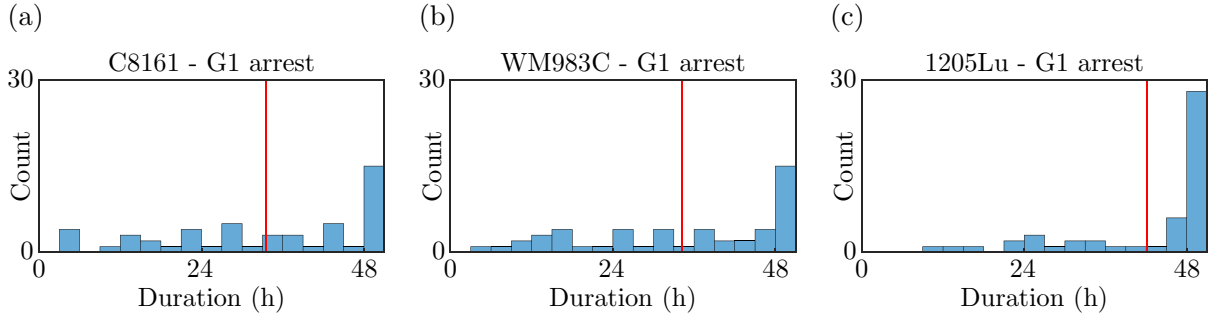


Figure 4.2: Histograms of the durations of G1 arrest (30 nM trametinib) for each cell line. Each histogram corresponds to 50 cells. The vertical red bars indicate the mean durations.

The data demonstrate that the three cell lines have very different cell-cycle characteristics. In particular, the cell-cycle durations increase in the order C8161, WM983C and 1205Lu. Note that 1205Lu has a relatively long cell cycle, which corresponds to the relatively slow proliferation rate that is observed for this cell line. Further, the ratio of G1 to S/G2/M durations illustrates that C8161 cells spend around the same amount of time in G1 and S/G2/M, WM983C cells spend more time in S/G2/M than G1, and 1205Lu cells spend more time in G1 than S/G2/M, highlighting that these cell lines have very different cell cycles.

4.3.2.2 Cell trajectories

Figure 4.3 shows trajectories for the WM983C and 1205Lu cell lines, similar to the trajectories for C8161 in Figure 1(g)–(i) in the main document. The insets for Figures 4.3(c) and (f) show the truncated trajectories for comparison with the cycling cells in G1 given in Figures 4.3(a) and (d), respectively. The duration of the truncated trajectories is equal to the mean duration of the corresponding trajectories of the cycling cells in G1.

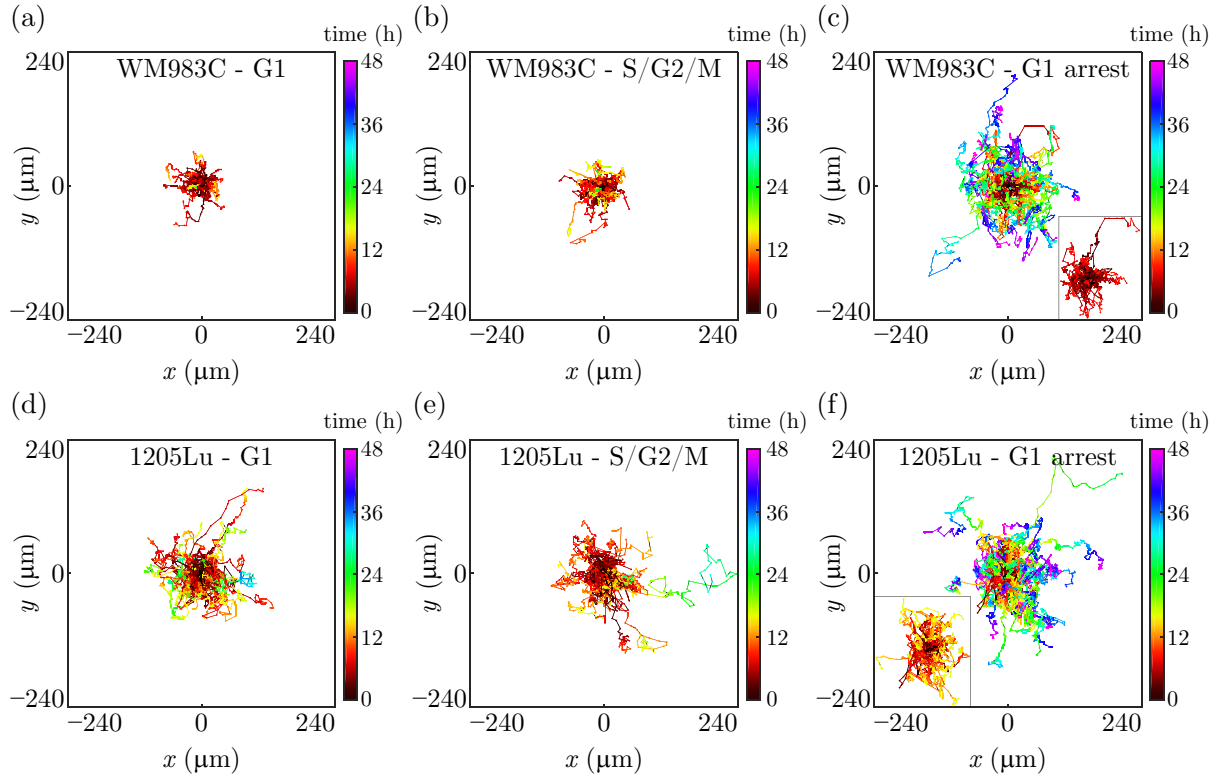


Figure 4.3: Cell trajectories of the WM983C and 1205Lu cell lines. (a)–(c) 50 cell trajectories of G1 cycling, S/G2/M cycling and G1-arrested (30 nM trametinib) WM983C cells, respectively. (d)–(f) 50 cell trajectories of G1 cycling, S/G2/M cycling and G1-arrested (30 nM trametinib) 1205Lu cells, respectively.

4.3.2.3 Cell migration: directionality

We investigate whether the cell migration has directionality by identifying any global anisotropy or directional autocorrelation of the cell trajectories. To reliably identify the presence or absence of anisotropy in cell migration, it is necessary to test for directionality with more than one measure [129,130]. We employ the mean drift velocity and the moment of inertia tensor, which demonstrate that cell migration is globally isotropic for all experimental conditions. We then consider the directional persistence of individual cells over short time intervals, which may not be revealed by measures of global anisotropy, using the temporal velocity autocorrelation function.

4.3.2.3.1 Drift velocity The drift velocity is the mean velocity of the cell from the initial to the final position, and is a measure of the directionality of the trajectory. Specifically, the drift velocity is given by

$$\mathbf{v} = (v_x, v_y), \quad v_x = \frac{x(t_f) - x(t_i)}{t_f - t_i}, \quad v_y = \frac{y(t_f) - y(t_i)}{t_f - t_i}, \quad (\text{S1})$$

where $(x(t), y(t))$ is the cell position at time t , t_i is the time at the initial cell position, and t_f is the time at the final cell position. By determining the drift velocities in the x - and y -directions for all trajectories, and then finding the corresponding mean values, we can quantify whether the migrating cells have global directionality. If the mean drift velocity of a cell population is non-zero, then to estimate the diffusivities it is first necessary to subtract the drift from the original cell displacement [74].

Figure 4.4 shows histograms of the drift velocities, and in Table 4.3 we statistically characterise the drift velocity data using the mean, median, standard deviation (SD), mean absolute deviation (MAD), the p -value of the one-sample t -test, the p -value of the one-sample sign test, kurtosis, skewness, and the p -value of the Anderson-Darling test for normality.

We use the one-sample t -test to determine whether the mean of each drift-velocity component differs statistically from a population with mean of $0 \mu\text{m h}^{-1}$. With a cut-off of 0.05 for statistical significance, the p -values indicate that, except in the case of v_x for WM983C in S/G2/M, the results are not statistically significant. Regarding v_x for WM983C in S/G2/M, we find that the results are not statistically significant with a cut-off of 0.05 when we hypothesise a population mean of $-0.3 \mu\text{m h}^{-1}$. Over the complete 48 h of the experiments, a drift of magnitude $0.3 \mu\text{m h}^{-1}$ corresponds to a displacement of magnitude $14.4 \mu\text{m}$, which is less than the diameter of a typical melanoma cell, such as $16.44 \mu\text{m}$ for C8161 [109]. Therefore, we conclude that there is no drift in any of our migration experiments. We arrive at the same conclusion with a similar analysis using the one-sample sign test and the median of each drift-velocity component.

Kurtosis is a measure of how light-tailed or heavy-tailed a distribution is relative to the normal distribution. The value of the kurtosis for a normal distribution is 3, which indicates that our drift-velocity data are generally heavy-tailed.

Skewness is a measure of the asymmetry of a distribution, so that a symmetric distribution such as the normal distribution has a skewness of zero. Our drift-velocity data generally have skewness values that differ from zero, indicating that the data have a degree of asymmetry.

The Anderson-Darling test provides a measure for whether each set of drift-velocity data is a sample from a normal distribution. Assuming that the sample data are normally distributed, the calculated p -value is the conditional probability of observing data at least as extreme as the sample data under consideration. Therefore, the p -value is a measure of the strength of evidence against the assumption that the sample data are from a normal distribution. Applying the test to our data, with a cut-off of 0.05 for statistical significance, we conclude that around half of the data sets are unlikely to be sampled from a normal distribution. We therefore cannot assume in general that the drift-velocity data are distributed normally.

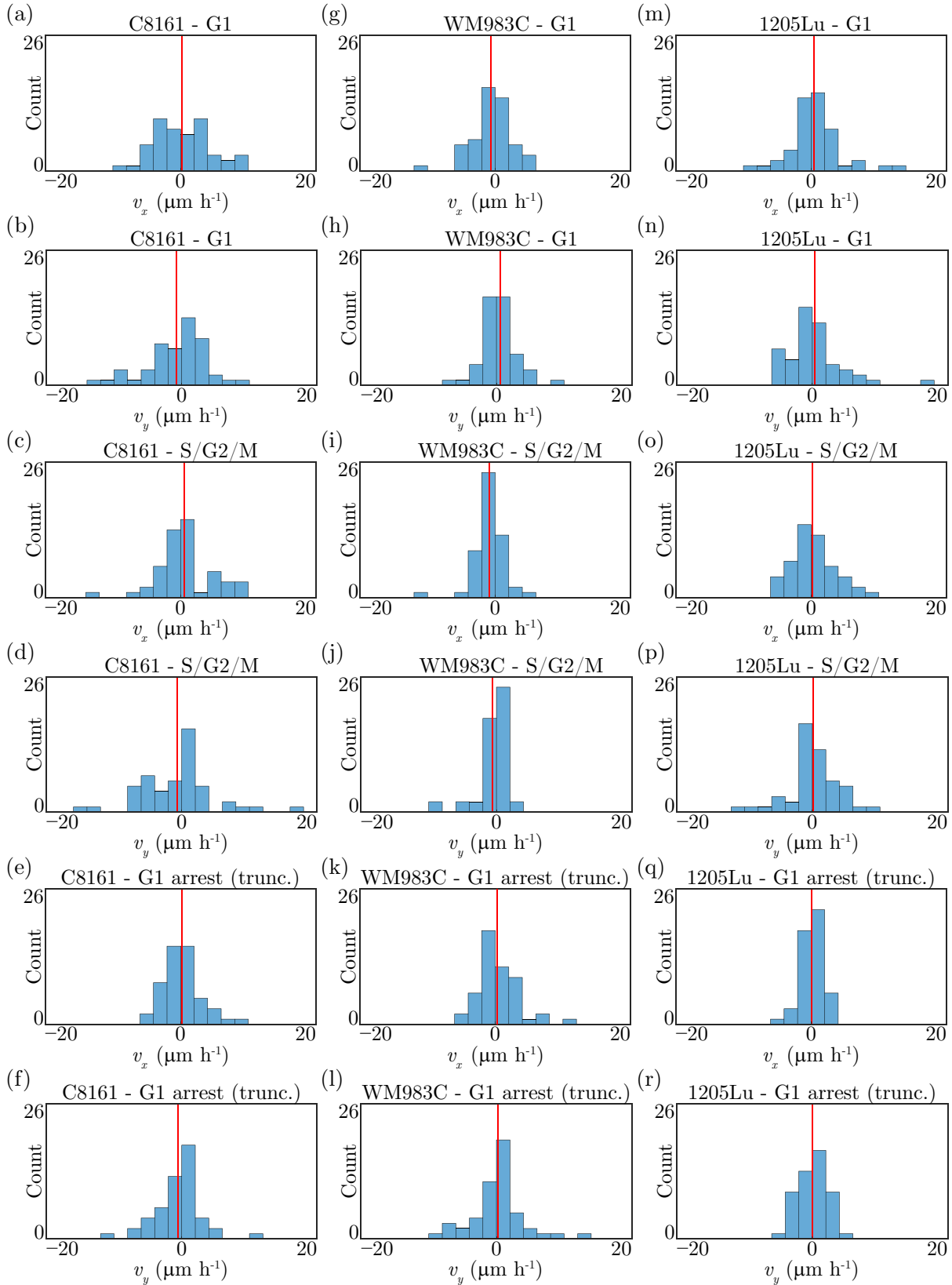


Figure 4.4: Drift velocity. (a) v_x for C8161 - G1. (b) v_y for C8161 - G1. (c) v_x for C8161 - S/G2/M. (d) v_y for C8161 - S/G2/M. (e) v_x for C8161 - G1 arrest (30nM trametinib). (f) v_y for C8161 - G1 arrest (30nM trametinib). (g) v_x for WM983C - G1. (h) v_y for WM983C - G1. (i) v_x for WM983C - S/G2/M. (j) v_y for WM983C - S/G2/M. (k) v_x for WM983C - G1 arrest (30nM trametinib). (l) v_y for WM983C - G1 arrest (30nM trametinib). (m) v_x for 1205Lu - G1. (n) v_y for 1205Lu - G1. (o) v_x for 1205Lu - S/G2/M. (p) v_y for 1205Lu - S/G2/M. (q) v_x for 1205Lu - G1 arrest (30nM trametinib). (r) v_y for 1205Lu - G1 arrest (30nM trametinib). The vertical red bars indicate the mean drift velocity.

Table 4.3: Statistical characterisation of the drift-velocity data.

		Mean ($\mu\text{m h}^{-1}$)	Median ($\mu\text{m h}^{-1}$)	Standard deviation ($\mu\text{m h}^{-1}$)	Mean absolute deviation ($\mu\text{m h}^{-1}$)	One- sample <i>t</i> -test <i>p</i> -value	One- sample sign test <i>p</i> -value	Kurtosis	Skewness	Anderson- Darling test <i>p</i> -value
C8161 - G1	v_x	0.20	-0.15	4.1	3.3	1.0	0.74	3.0	0.20	0.95
	v_y	-0.79	0.33	4.3	3.3	0.20	0.89	3.6	-0.73	0.0012
C8161 - S/G2/M	v_x	0.59	0.16	4.2	3.0	0.33	0.89	4.4	-0.083	0.019
	v_y	-0.71	0.00	5.3	3.7	0.34	0.65	5.5	0.40	0.0054
C8161 - G1 arrest (truncated)	v_x	0.24	-0.16	2.8	2.1	0.55	1.0	4.7	0.81	0.30
	v_y	-0.50	-0.059	3.5	2.6	0.32	0.77	4.8	0.064	0.15
WM983C - G1	v_x	-0.66	-0.40	2.9	2.1	0.11	0.39	4.9	-0.76	0.20
	v_y	0.52	0.54	2.5	1.8	0.15	0.67	6.0	0.23	0.058
WM983C - S/G2/M	v_x	-0.90	-0.57	2.2	1.5	0.0060	0.0021	10	-1.4	0.0014
	v_y	-0.56	0.051	2.3	1.6	0.088	1.0	7.0	-1.7	0.00050
WM983C - G1 arrest (truncated)	v_x	0.33	-0.21	2.9	2.2	0.43	0.67	5.2	1.1	0.023
	v_y	0.21	0.32	3.9	2.6	0.71	0.47	5.3	0.51	0.0027
1205Lu - G1	v_x	0.43	0.21	3.8	2.7	0.43	0.57	5.5	0.81	0.0091
	v_y	0.33	-0.068	4.2	2.9	0.58	0.67	7.0	1.5	0.011
1205Lu - S/G2/M	v_x	0.24	-0.068	3.3	2.7	0.61	1.0	2.9	0.35	0.49
	v_y	0.084	0.057	3.7	2.6	0.87	1.0	4.9	-0.60	0.024
1205Lu - G1 arrest (truncated)	v_x	0.13	0.27	1.7	1.4	0.61	0.48	3.9	-0.62	0.79
	v_y	-0.017	0.055	2.3	1.8	0.96	0.78	2.7	-0.15	0.61

4.3.2.3.2 Moment of inertia tensor The moment of inertia tensor \mathbf{I} can be used to reliably identify anisotropy in cell migration [129–131]. Here we consider the mean moment of inertia tensor $\bar{\mathbf{I}}$, which quantifies the anisotropy of the mean of all cell trajectories.

For the 50 trajectories of each experimental condition we consider all of the displacements that occur during the 15 min time interval between images in the time series. We denote these displacements by (x_i, y_i) , for $i = 1, \dots, N$, where N is the total number of displacements. With the start point at the origin, we consider a unit mass at the end point of each displacement. The two-dimensional mean moment of inertia tensor is then given by

$$\bar{\mathbf{I}} = \begin{pmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{pmatrix}, \quad \text{where} \quad I_{xx} = \sum_{i=1}^N y_i^2, \quad I_{xy} = I_{yx} = -\sum_{i=1}^N x_i y_i \quad \text{and} \quad I_{yy} = \sum_{i=1}^N x_i^2. \quad (\text{S2})$$

The eigenvalues $\bar{\lambda}_1$ and $\bar{\lambda}_2$ of $\bar{\mathbf{I}}$, which are the principal moments of inertia, indicate directionality in cell migration. By convention $\bar{\lambda}_1 \geq \bar{\lambda}_2$, so that $\bar{\lambda}_1 = \bar{\lambda}_2$ corresponds to isotropy, and $\bar{\lambda}_1 > \bar{\lambda}_2$ corresponds to anisotropy. The ratio $\bar{\lambda}_1/\bar{\lambda}_2 \geq 1$ therefore quantifies the degree of anisotropy, or directionality, in cell migration. Table 4.4 provides the ratio $\bar{\lambda}_1/\bar{\lambda}_2$ for cycling cells in G1 and S/G2/M, and for cells arrested in G1 with trametinib (30 nM), for each of the three cell lines. The results indicate that $\bar{\lambda}_1/\bar{\lambda}_2$ is almost equal to one for all experimental conditions, hence we

	$\bar{\lambda}_1/\bar{\lambda}_2$
C8161 - G1	1.18 ± 0.09
C8161 - S/G2/M	1.39 ± 0.12
C8161 - G1 arrest	1.10 ± 0.05
C8161 - G1 arrest (truncated)	1.03 ± 0.07
WM983C - G1	1.15 ± 0.07
WM983C - S/G2/M	1.06 ± 0.08
WM983C - G1 arrest	1.22 ± 0.06
WM983C - G1 arrest (truncated)	1.29 ± 0.10
1205Lu - G1	1.04 ± 0.04
1205Lu - S/G2/M	1.32 ± 0.07
1205Lu - G1 arrest	1.03 ± 0.03
1205Lu - G1 arrest (truncated)	1.06 ± 0.06

Table 4.4: Ratio of the eigenvalues for the mean moment of inertia tensor. The error estimates are obtained by bootstrapping the sets of x - and y -displacements, each with 10^4 iterations, and calculating $\bar{\lambda}_1/\bar{\lambda}_2$ for each re-sampled data set, whereby the error is then taken as the standard deviation of the 10^4 values of $\bar{\lambda}_1/\bar{\lambda}_2$.

conclude that the cell migration is essentially globally isotropic for all conditions. This result is consistent with the finding of no drift in all experiments. We also include the results for the

G1-arrested cells over the full 48 h of the experiments, which show that there is no anisotropy over the complete experiment duration.

The error estimates in Table 4.4 are obtained by bootstrapping the set of all x -displacements and the set of all y -displacements, each with 10^4 iterations. Each re-sampling of the displacements is used to calculate $\bar{\mathbf{I}}$ and $\bar{\lambda}_1/\bar{\lambda}_2$. The standard deviation of the 10^4 values of $\bar{\lambda}_1/\bar{\lambda}_2$ is then taken to be the error.

4.3.2.3.3 Temporal velocity autocorrelation function The temporal velocity autocorrelation function (TVAF) provides a measure for directional persistence of cell migration [132–134]. We say that the velocity $\mathbf{v}(t)$ of a cell at time t is equal to the mean velocity of the cell between the time points t and $t + \Delta t$, where t is a time point of the experimental time-series images and Δt is the time interval between images, here 15 min. If $\mathbf{r}(t)$ is the position of the cell at time t , then we have $\mathbf{v}(t) = (\mathbf{r}(t + \Delta t) - \mathbf{r}(t))/\Delta t$. The TVAF is given by

$$C_v(\tau) = \langle \mathbf{v}(t + \tau) \cdot \mathbf{v}(t) \rangle, \quad (\text{S3})$$

where, for a given lag time τ , the averaging is over all possible time points t of a cell trajectory and over all cell trajectories. The normalised, and dimensionless, TVAF is

$$\hat{C}_v(\tau) = \frac{C_v(\tau)}{C_v(0)} = \frac{\langle \mathbf{v}(t + \tau) \cdot \mathbf{v}(t) \rangle}{\langle \mathbf{v}(t) \cdot \mathbf{v}(t) \rangle}. \quad (\text{S4})$$

For a population of cells in which the migration can be described as a random walk, \hat{C}_v will be 1 at $\tau = 0$ and 0 for all $\tau > 0$, revealing no correlation in cell migration. A persistent random walk is revealed by a continuously decaying profile for \hat{C}_v , since the correlation reduces over time. To estimate the decorrelation time for the persistence, we fit to \hat{C}_v either a single exponential function, f_1 , or a sum of two exponential functions, f_2 , which we write generally as

$$f_n(\tau) = \sum_{i=1}^n A_i e^{-\tau/\mathcal{T}_i}, \text{ for } n = 1 \text{ or } 2, \quad (\text{S5})$$

and constants A_i and \mathcal{T}_i , where the latter are the decorrelation times. When $n = 2$ we adopt the convention that $\mathcal{T}_1 < \mathcal{T}_2$. We estimate the decorrelation times \mathcal{T}_i by fitting Equation (S5) to \hat{C}_v in Equation (S4) evaluated with experimental data, using the `fit` function and either the `exp1` or the `exp2` models [121] in MATLAB. A best fit is obtained with $n = 2$ for the cell lines C8161 and 1205Lu, and $n = 1$ for the WM983C cell line, shown in Figure 4.5 and Table 4.5. Therefore, for the cell lines C8161 and 1205Lu, a sum of two exponential functions is found to

	\mathcal{T}_1 (h)	\mathcal{T}_2 (h)
C8161 - G1	0.10 ± 0.03	4 ± 1
C8161 - S/G2/M	0.18 ± 0.02	3 ± 3
C8161 - G1 arrest (truncated)	0.010 ± 0.003	4.0 ± 0.6
WM983C - G1	0.11 ± 0.01	-
WM983C - S/G2/M	0.010 ± 0.001	-
WM983C - G1 arrest (truncated)	0.09 ± 0.01	-
1205Lu - G1	0.13 ± 0.01	2.2 ± 0.7
1205Lu - S/G2/M	0.003 ± 0.003	0.15 ± 0.08
1205Lu - G1 arrest (truncated)	0.006 ± 0.003	0.15 ± 0.02

Table 4.5: Estimated decorrelation times of directional persistence. The temporal velocity autocorrelation function data is fit with a sum of two exponential functions for C8161 and 1205Lu, and is fit with a single exponential function for WM983C. The error estimates are obtained by a bootstrapping process which produces 10^4 estimates of \hat{C}_v , from which decorrelation times are estimated by fitting Equation (S5). The error is then taken to be the standard deviation of the 10^4 estimates for the decorrelation times.

fit the TVAF data best, whereas for the WM983C cell line, a best fit is obtained with a single exponential function.

Both the exponential model and the sum of two exponentials model have previously been observed as the functional forms for the TVAF in adherent cell migration on two-dimensional surfaces [134,135]. Our results suggest that the WM983C cells exhibit a single slow decorrelation time, while the C8161 and 1205Lu cells exhibit both a slow and a fast decorrelation time. For the 1205Lu cell line, the G1-arrested cells appear to decorrelate faster than the cycling cells in G1. These observations illustrate that similar cell types, in this case melanoma cells, can have distinctly different migratory behaviour both with and without drug-induced cell cycle arrest.

The estimated error in the decorrelation times, indicated in Table 4.5, is obtained using a bootstrapping process. For each lag time τ , we re-sample the data set $\{\mathbf{v}(t + \tau) \cdot \mathbf{v}(t) \mid \text{for all possible } t\}$ with replacement, to give a re-sampled data set of the same size. Using the mean values of each of these re-sampled data sets, we calculate \hat{C}_v and then fit Equation (S5), thereby obtaining new estimates for the decorrelation times. The re-sampling process is performed 10^4 times to obtain 10^4 estimates for the decorrelation times, and the standard deviation of these estimates is taken to be the error in the original estimates of the decorrelation times.

Fitting a function which is a sum of two exponentials to data is well known to be ill-conditioned [136]. Consequently, the estimated errors in the decorrelation times may be proportionally larger when using a model which is the sum of two exponentials compared with a single exponential model. This has minimal consequence for our purposes as we only need an estimate

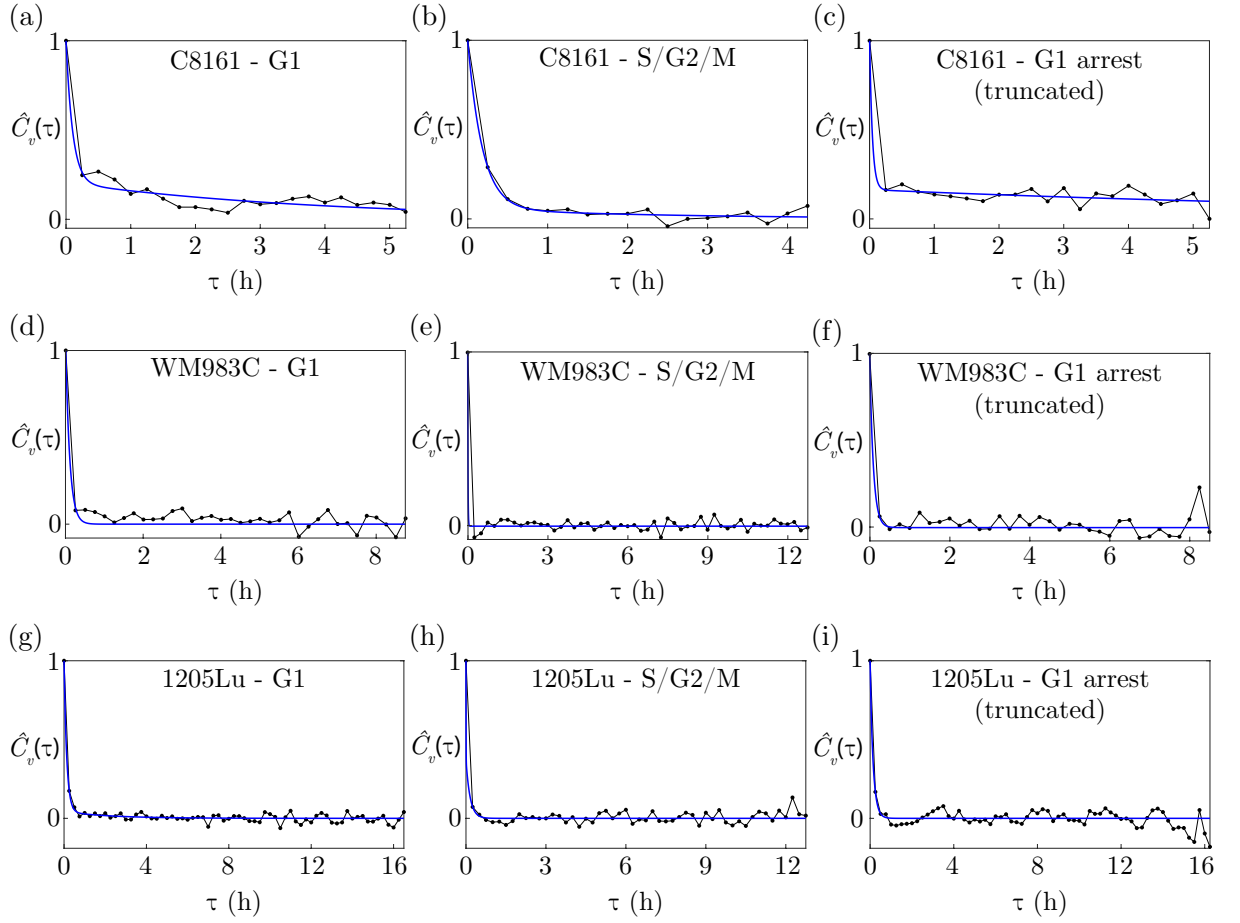


Figure 4.5: Best fit of the exponential function, f_1 , or the sum of two exponential functions, f_2 , in Equation (S5) to the normalised temporal velocity autocorrelation function, $\hat{C}_v(\tau)$, in Equation (S4) evaluated with experimental data. (a)–(c) G1 cycling, S/G2/M cycling, and truncated G1-arrested (30nM trametinib) C8161 cells, respectively, for which f_2 fits best. (d)–(f) G1 cycling, S/G2/M cycling, and truncated G1-arrested (30nM trametinib) WM983C cells, respectively, for which f_1 fits best. (g)–(i) G1 cycling, S/G2/M cycling, and truncated G1-arrested (30nM trametinib) 1205Lu cells, respectively, for which f_2 fits best. The black discs correspond to the data and the blue curves to the best fit of f_1 or f_2 .

of the decorrelation times as a guide for the smallest lag time after which the cell migration is essentially random, allowing us to then estimate diffusivities.

Based on the decorrelation times in Table 4.5, we note that most of the correlation is lost within 1 h, for all cell lines. While 1205Lu and particularly C8161 each have an associated slow decorrelation time which results in residual correlation beyond 1 h, it is not practical to use longer lag times for estimating the diffusivities due to the limitations imposed by the durations of the G1 and S/G2/M phases. Any residual correlation beyond a lag time of 1 h, however, is relatively small and would have minimal consequence for our estimation of diffusivities.

4.3.2.4 Cell migration: diffusivities

The migration of an adherent cell tends to exhibit directional persistence, as we demonstrate with the temporal velocity autocorrelation function (TVAF) applied to our three melanoma cell lines. Cell migration is therefore often modelled as a persistent random walk (PRW) [74,137,138]. There are, however, numerous experimental studies that do not support the PRW as a model for cell migration, as the characteristics of the observed migration differ from predictions of the PRW models [61,134,135,139]. In particular, PRW models predict that the TVAF has a simple exponential form [134,135]. For our cell lines, however, only one cell line has a TVAF with an exponential form, while the other two cell lines have a TVAF with the form of a sum of two exponentials. Therefore, we do not employ a PRW model to estimate diffusivities. Rather, our diffusivity estimates are obtained over time periods for which the directional persistence is lost and the cells are undergoing free diffusion.

We estimate the diffusivities using the mean square displacement (MSD), which has various definitions [140–142]. Here we discuss our methodology in detail. Consider a cell trajectory in 2-D obtained from a time series of N images, where Δt is the time interval between successive positions. In our case, $\Delta t = 0.25$ h. We denote the sequence of cell positions as $\{\mathbf{r}_i\}_{i=1}^N$. Let $1 \leq M \leq N - 1$ be the integer such that $M\Delta t$ is the specified minimum lag time. For the lag time $\tau_n = n\Delta t$, where $n = M, \dots, N - 1$, there are $N - n$ non-zero forward displacements along the cell trajectory. The MSD, denoted $\rho(\tau_n)$, is given by

$$\rho(\tau_n) = \frac{1}{N - n} \sum_{i=1}^{N-n} \|\mathbf{r}_{i+n} - \mathbf{r}_i\|^2. \quad (\text{S6})$$

We have established that our cell migration experiments are globally isotropic, and that the directional persistence of the cells dissipates within 1 h. Therefore, we may regard the cell migration as free diffusion for lag times of at least 1 h, which corresponds to $M = 4$. We may then estimate the diffusivity D of the cell by fitting the equation [141]

$$\rho(\tau) = 4D\tau, \text{ for lag time } \tau, \quad (\text{S7})$$

to the set of data points $\{(\tau_n, \rho(\tau_n)) \mid n = M, \dots, N - 1\}$. We fit Equation (S7) to our data using the `fitlm` linear regression function in MATLAB [143] with no constant term, therefore passing through the origin.

For each cell line and experimental condition we have 50 cell trajectories. For each trajectory we calculate D , using the minimum lag time of 1 h, within the 2-h time intervals 0–2 h, 1–3

h, 2–4 h and so on, successively offset by 1 h up until the end of the trajectory. To illustrate the variability in D for the 50 cell trajectories for each cell line and experimental condition we show in Figure 4.6 the histograms of D within the time interval 0–2 h. For each 2-h time

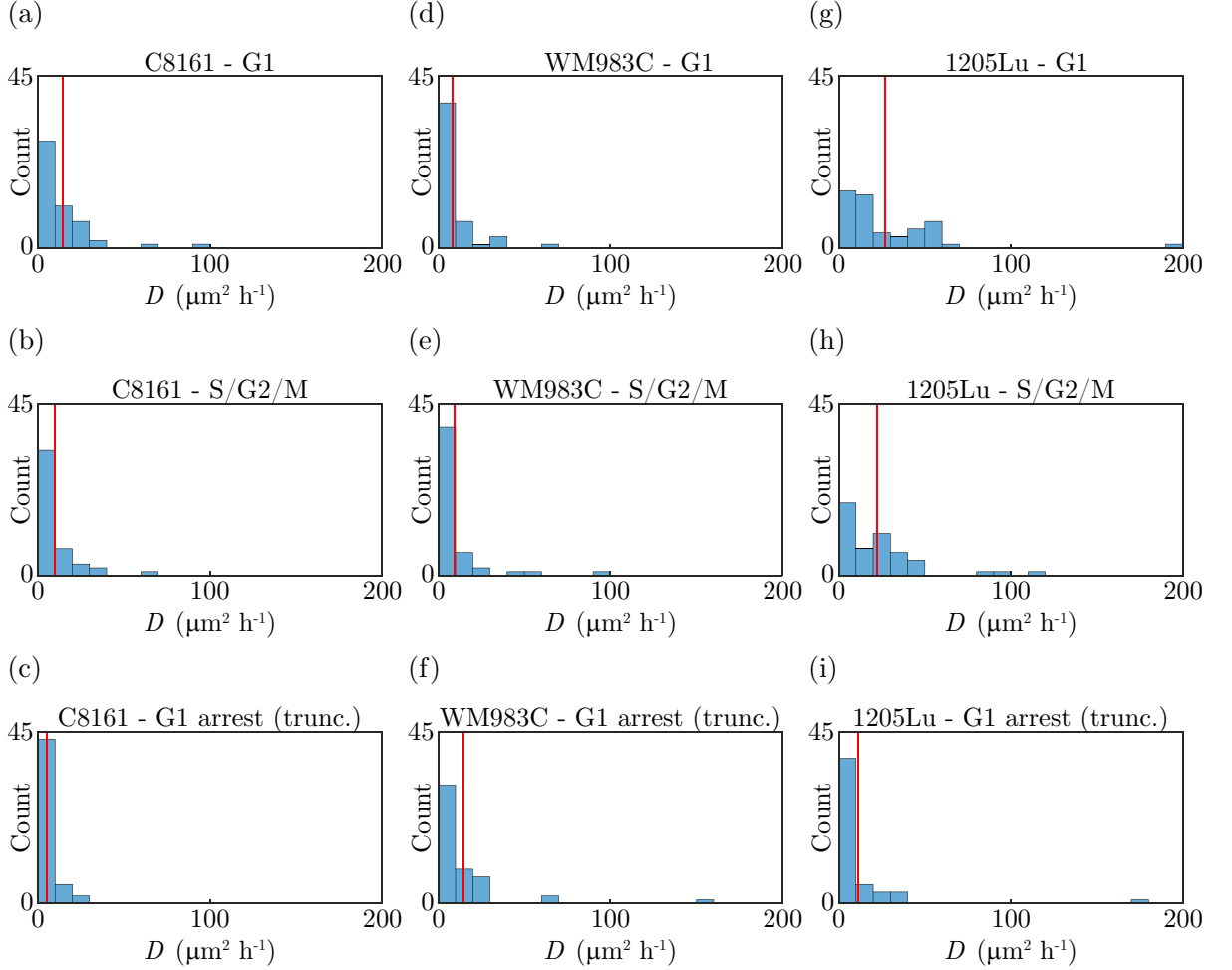


Figure 4.6: Histograms of the diffusivities D within the time interval 0–2 h for each cell line and experimental condition. Each histogram corresponds to 50 cell trajectories. The vertical red bars indicate $\langle D \rangle$.

interval we then calculate the mean diffusivity $\langle D \rangle$ from the individual D corresponding to all trajectories that extend to the end of the interval. Note that the mean may be taken over fewer than 50 trajectories due to the differing durations of the trajectories. Figure 1(j)–(l) in the main document shows a comparison of $\langle D \rangle$ as a function of time interval between cycling cells in G1, cycling cells in S/G2/M, and G1-arrested cells (30 nM trametinib), for each of the cell lines C8161, WM983C and 1205Lu.

For each cell line we have three independent samples of diffusivities corresponding to three experimental conditions: cycling cells in G1, cycling cells in S/G2/M, and G1-arrested cells (30 nM trametinib). To compare the mean diffusivities between each pair of diffusivity samples we employ the permutation test, which is a non-parametric significance test which makes no

assumptions about the population distributions such as normality and equal variance. For each cell line we are interested in three null hypotheses, H_0 , where we denote the mean diffusivities of cycling cells in G1 as $\langle D \rangle_{G1}$, of cycling cells in S/G2/M as $\langle D \rangle_{S/G2/M}$, and of G1-arrested cells as $\langle D \rangle_{G1\text{-arrest}}$:

1. $H_0 : \langle D \rangle_{G1} = \langle D \rangle_{S/G2/M}$, that is, the mean diffusivities are the same for cells in G1 and cells in S/G2/M. Since we are interested in whether the mean diffusivities are equal, we test if $\langle D \rangle_{G1}$ is significantly greater than $\langle D \rangle_{S/G2/M}$ and if $\langle D \rangle_{S/G2/M}$ is significantly greater than $\langle D \rangle_{G1}$, which is therefore a two-sided test.
2. $H_0 : \langle D \rangle_{G1} \geq \langle D \rangle_{G1\text{-arrest}}$, that is, the mean diffusivity for G1-arrested cells is not larger than the mean diffusivity for cycling cells in G1. Since the G1-arrested cells are not progressing through the cell cycle, the go-or-grow hypothesis suggests that they may be more motile, with correspondingly larger mean diffusivities, than the cycling cells. So we test whether $\langle D \rangle_{G1\text{-arrest}}$ is significantly greater than $\langle D \rangle_{G1}$, which is therefore a one-sided test.
3. $H_0 : \langle D \rangle_{S/G2/M} \geq \langle D \rangle_{G1\text{-arrest}}$, that is, the mean diffusivity for G1-arrested cells is not larger than the mean diffusivity for cycling cells in S/G2/M. Since the G1-arrested cells are not progressing through the cell cycle, the go-or-grow hypothesis suggests that they may be more motile, with correspondingly larger mean diffusivities, than the cycling cells. So we test whether $\langle D \rangle_{G1\text{-arrest}}$ is significantly greater than $\langle D \rangle_{S/G2/M}$, which is therefore a one-sided test.

In Table 4.6 we provide the p -values, with the corresponding effect sizes in parentheses, obtained from the permutation test for each time interval. The effect size is defined as the magnitude of the difference of the two sample means divided by the mean of the two sample standard deviations. If we use a threshold of 0.01 for statistical significance, we note that in almost all cases the p -values indicate a lack of significance, so we cannot reject the corresponding null hypotheses. These tests don't prove that the null hypotheses are true, however we are unable to reject them based on our current data. For WM983C the hypothesis tests indicate that $\langle D \rangle_{G1\text{-arrest}}$ is significantly greater than $\langle D \rangle_{G1}$ for two of the eight time intervals, and $\langle D \rangle_{G1\text{-arrest}}$ is significantly greater than $\langle D \rangle_{S/G2/M}$ for three of the eight time intervals. The particular threshold of 0.01 for statistical significance is arbitrary, and there is increasing support to employ a threshold of 0.005 for new scientific discoveries in order to reduce false positives and improve the reproducibility of research [144]. If we employ a threshold of 0.005 for statistical significance then the hypothesis tests indicate that, for WM983C, $\langle D \rangle_{G1\text{-arrest}}$ is not significantly greater than $\langle D \rangle_{G1}$ for any time

interval, and $\langle D \rangle_{\text{G1-arrest}}$ is significantly greater than $\langle D \rangle_{\text{S/G2/M}}$ for only two of the eight time intervals. While statistical significance can be a helpful guide, the ultimate interest in biology must be the biological importance of an effect [145], on which significance testing provides no information. In our case, the significance testing indicates an effect in only a relatively few time intervals for WM983C. The data in Figure 1(j)–(l) and Figure 4.6 demonstrate that the mean diffusivities for each cell line are remarkably similar between the different experimental conditions, where typically one mean diffusivity is within one standard deviation of another mean diffusivity. Therefore, there appears to be no biological difference between the mean diffusivities of the three experimental conditions.

We conclude that:

- For each cell line and experimental condition, $\langle D \rangle$ has little variation over time, hence motility is essentially constant during each cell-cycle phase and during G1 arrest.
- For each cell line there is remarkably little variation in $\langle D \rangle$ between cycling cells in G1, cycling cells in S/G2/M, and G1-arrested cells. Therefore, cell motility is the same whether the cells are in G1, S/G2/M or G1-arrested.
- The three cell lines have very different proliferation and migration characteristics, however $\langle D \rangle$ is remarkably consistent across the cell lines.

In Figure 4.7 we show $\langle D \rangle$ as a function of time interval over the full 48 h of the experiments for all three cell lines in G1-arrest (30 nM trametinib). The data in Figure 4.7 demonstrate the remarkably low variability of $\langle D \rangle$ over the experimental duration of 48 h. For each cell line, the cells are arrested in G1 and so are not in a proliferative state, however we observe essentially no change in motility over two days.

Time interval (h)	0–2	1–3	2–4	3–5	4–6	5–7	6–8	7–9
C8161								
$H_0 : \langle D \rangle_{G1} = \langle D \rangle_{S/G2/M}$	0.11 (0.34)	0.57 (0.12)	0.63 (0.12)	0.44 (0.26)	0.89 (0.077)	-	-	-
$H_0 : \langle D \rangle_{G1} \geq \langle D \rangle_{G1\text{-arrest}}$	1.0 (0.80)	1.0 (0.83)	1.0 (0.64)	1.0 (0.64)	0.76 (0.22)	-	-	-
$H_0 : \langle D \rangle_{S/G2/M} \geq \langle D \rangle_{G1\text{-arrest}}$	0.99 (0.47)	1.0 (0.64)	1.0 (0.72)	0.95 (0.46)	0.77 (0.26)	-	-	-
WM983C								
$H_0 : \langle D \rangle_{G1} = \langle D \rangle_{S/G2/M}$	0.77 (0.063)	0.92 (0.021)	0.49 (0.14)	0.18 (0.29)	0.86 (0.039)	0.68 (0.099)	0.094 (0.39)	0.50 (0.18)
$H_0 : \langle D \rangle_{G1} \geq \langle D \rangle_{G1\text{-arrest}}$	0.046 (0.35)	0.11 (0.27)	0.055 (0.35)	0.12 (0.25)	0.0069 (0.56)	0.021 (0.50)	0.40 (0.080)	0.0065 (0.67)
$H_0 : \langle D \rangle_{S/G2/M} \geq \langle D \rangle_{G1\text{-arrest}}$	0.11 (0.27)	0.099 (0.29)	0.012 (0.48)	0.0035 (0.59)	0.0054 (0.54)	0.025 (0.41)	0.038 (0.41)	0.0046 (0.58)
1205Lu								
$H_0 : \langle D \rangle_{G1} = \langle D \rangle_{S/G2/M}$	0.42 (0.17)	0.94 (0.016)	0.83 (0.043)	0.59 (0.13)	0.27 (0.25)	0.50 (0.19)	0.92 (0.023)	0.34 (0.22)
$H_0 : \langle D \rangle_{G1} \geq \langle D \rangle_{G1\text{-arrest}}$	1.0 (0.55)	1.0 (0.58)	0.90 (0.28)	0.80 (0.21)	0.88 (0.26)	0.97 (0.37)	1.0 (0.69)	1.0 (0.85)
$H_0 : \langle D \rangle_{S/G2/M} \geq \langle D \rangle_{G1\text{-arrest}}$	0.99 (0.44)	1.0 (0.57)	0.91 (0.28)	0.86 (0.25)	0.62 (0.11)	0.90 (0.29)	1.0 (0.70)	1.0 (0.9)

Table 4.6: Permutation tests on mean diffusivities $\langle D \rangle$ for the C8161, WM983C and 1205Lu cell lines. For each experimental condition, cycling G1, cycling S/G2/M and G1 arrest, and relevant time interval, we provide the p -value and corresponding effect size (in parentheses) obtained from the permutation test.

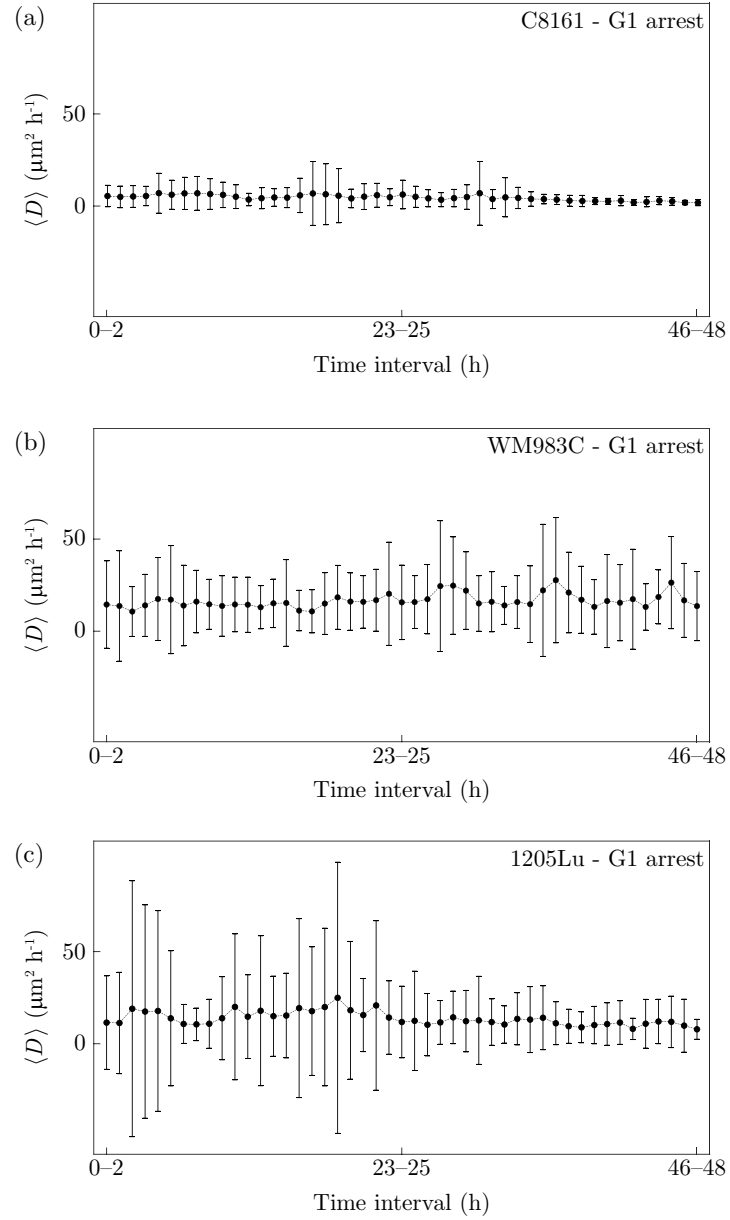


Figure 4.7: Mean diffusivities for C8161, WM983C and 1205Lu in G1 arrest (30nM trametinib) over 48 h. (a)–(c) $\langle D \rangle$ for C8161, WM983C and 1205Lu, respectively, for the range of time intervals over the full experiment duration of 48 h. In each case we show $\langle D \rangle$, and report the variability using $\langle D \rangle$ plus or minus the sample standard deviation.

A novel mathematical model of heterogeneous cell proliferation

A preprint submitted to arXiv

Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. A novel mathematical model of heterogeneous cell proliferation. arXiv:2003.03024v2 preprint. 2020. Currently under review with the Journal of Mathematical Biology.

Abstract

We present a novel mathematical model of heterogeneous cell proliferation where the total population consists of a subpopulation of slow-proliferating cells and a subpopulation of fast-proliferating cells. The model incorporates two cellular processes, asymmetric cell division and induced switching between proliferative states, which are important determinants for the heterogeneity of a cell population. As motivation for our model we provide experimental data that illustrate the induced-switching process. Our model consists of a system of two coupled delay differential equations with distributed time delays and the cell densities as functions of time. The distributed delays are bounded and allow for the choice of delay kernel. We analyse the model and prove the non-negativity and boundedness of solutions, the existence and uniqueness of solutions, and the local stability characteristics of the equilibrium points. We find that the parameters for induced switching are bifurcation parameters and therefore determine the long-term behaviour of the model. Numerical simulations illustrate and support the theoretical findings, and demonstrate the primary importance of transient dynamics for understanding the evolution of many experimental cell populations.

5.1 Introduction

Cell proliferation is the fundamental function of the cell cycle [10], which is a complex process regulated by both intracellular signals and the extracellular environment [9]. Such complexity necessitates that mathematical models of cell proliferation are often restricted to details that are most pertinent to the experimental situation under consideration. The main requirement is that the model must account for progression through the cell cycle in a manner relevant to the cell population and the surrounding environment. Despite all of the underlying complexity the cell cycle has two basic fates, either progression or arrest [10]. These two cellular fates form the basis of many mathematical models of cell proliferation in the literature, typically based on exponential growth [51, 54, 71, 72] or logistic growth [50, 52–54, 73, 74]. Exponential growth explicitly accounts for progression only, while logistic growth accounts for progression and density-dependent arrest, which can result from contact inhibition [75].

An important detail of the cell cycle not explicitly accounted for in exponential and logistic growth models is the duration of the cell cycle, which is always nonzero and exhibits considerable variation between different cell types and different extracellular environments [76, 77, 82, 146]. From a modelling perspective the cell cycle duration is a positive time delay between two sequential cell proliferation events. There are two main types of models which incorporate time delays: one involves functional differential equations [78–81, 147–150], of which delay differential equations are a specific type; and multi-stage models [82–86]. Models incorporating time delays are consistent with the kinetics of cell proliferation, and can result in a better qualitative and quantitative fit of the model to experimental data [79]. The inclusion of a time delay must be based on whether the improved model fit outweighs the increase in model complexity arising from additional parameters and, for functional differential equations, an infinite-dimensional state space. Models with time delays are particularly relevant when the transient dynamics of a cell population are of interest, especially when modelling slow-proliferating cells. We briefly note that age-structured models [151–155], which are related to delay differential equations, provide another approach to incorporating realistic cell cycle durations into models of cell population growth.

In this article we introduce a delay differential equation model for cell proliferation in which the cell population consists of a slow-proliferating subpopulation and a fast-proliferating subpopulation. The cells can switch between the *proliferative states* of slow and fast proliferation through two cellular processes: asymmetric cell division and induced switching of proliferative states by surrounding cells. Our model is motivated by the *proliferative heterogeneity*, with

respect to cell cycle duration, of tumours, which are often composed of a large proportion of fast-proliferating cells and a small proportion of slow-proliferating cells which can repopulate the fast-proliferating subpopulation [27, 29]. The slow-proliferating subpopulation is sometimes considered to be quiescent, or arrested, however it is possible that this subpopulation is actually in a very-slow-proliferating state [156, 157]. Experimental studies have found slow-proliferating cells with cell cycle durations greater than four weeks, whereas the predominant fast-proliferating cells have cell cycle durations of around 48 hours [28].

In the literature there are various mathematical models that consider proliferative heterogeneity. Some models account for one proliferating subpopulation [150], which may undergo asymmetric division [87, 88], while the other subpopulations are quiescent or differentiated. Other models consider subpopulations with different proliferative states without any cells switching between the subpopulations [89]. Our model, and our mathematical analysis of the model, are novel in several ways: (1) for each subpopulation we model a distribution of cell cycle durations using a distributed delay with an arbitrary delay kernel on a bounded interval, which allows us to freely choose an appropriate proliferative state for each subpopulation; (2) cells can switch between the slow- and fast-proliferating subpopulations through two important processes, either during cell division or induced by surrounding cells; (3) we provide formal proofs of existence, uniqueness, non-negativity, and boundedness of the solutions for our model under appropriate initial conditions; (4) the local stability of all equilibrium points is characterised and bifurcation parameters identified, involving the analysis of an interesting transcendental characteristic equation; (5) numerical simulations are provided which illustrate and support the theoretical results, and demonstrate the importance of considering the transient dynamics of experimental cell populations.

The remainder of this article is organised as follows. In Section 5.2 we discuss the biological and mathematical motivations for our model, which we then present in Section 5.3. Our main analytical results are in Section 5.4 in the form of three theorems: Theorem 5.2 for non-negativity and boundedness of solutions, Theorem 5.4 for the existence and uniqueness of solutions, and Theorem 5.5 for the local stability of the equilibrium points. Some examples of numerical simulations of our model are provided in Section 5.5, illustrating the long-term dynamics described in Theorem 5.5, and demonstrating the importance of the transient dynamics. Finally, in Section 5.6 we summarise our results, discuss the utility of our model to describe experimental scenarios, and note some possibilities for future work.

5.2 Model motivation

In this section we discuss the biological and mathematical considerations that motivate the development of our model.

5.2.1 Biological considerations

The eukaryotic cell cycle (Figure 5.1) is a sequence of four phases, namely gap 1 (G1), synthesis (S), gap 2 (G2) and mitosis (M). The primary function of the cell cycle is the replication

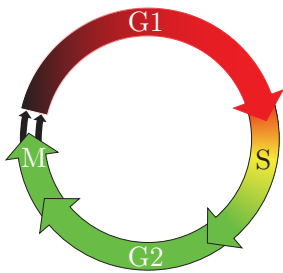


Figure 5.1: Schematic of the eukaryotic cell cycle, indicating the colour of fluorescent ubiquitination-based cell cycle indicator (FUCCI), see [30], in each phase. During very early G1 phase there is no fluorescence as both FUCCI reporters are downregulated. In G1 phase the red FUCCI reporter is upregulated and red fluorescence is observed. During the transition from G1 to S phase, called early S, both the red and green FUCCI reporters are upregulated producing yellow. Through S/G2/M phase the red FUCCI reporter is downregulated and only the green FUCCI reporter is upregulated so that green fluorescence is observed

of cellular DNA during S phase, followed by the division of the replicated chromosomes and cytoplasm into two daughter cells during M phase [24]. Progression through the cell cycle is tightly regulated in normal cells, which are subject to density-dependent contact inhibition producing reversible cell-cycle arrest [25, 26]. In cancer cells, however, cell cycle regulation is generally lost [14] resulting in cell populations with proliferative heterogeneity, as exemplified by tumours of solid cancers [27, 28]. In particular a small subpopulation of slow-proliferating cells is often present in tumours, and this subpopulation tends to survive anticancer drug treatment and can maintain the tumour by repopulating the fast-proliferating subpopulation [27, 29].

Experimental studies have revealed the highly dynamic nature of intratumoural heterogeneity, which can cause adverse outcomes from cancer therapy, notably drug resistance [13, 33, 66, 158]. Therefore the nonequilibrium, or transient, state of a tumour tends to be of greater relevance than the equilibrium states. The main purpose of our model is to provide insight into the transient dynamics of intratumoural heterogeneity, specifically with regard to cells switching their proliferative states through cellular mechanisms.

The range of mechanisms leading to proliferative heterogeneity in cancer cell populations are not completely understood, although asymmetric cell division is known to be partly responsible [159, 160]. Asymmetric cell division is a normal process of stem cell proliferation, required for development and the maintenance of tissue homeostasis, whereby a stem cell divides to produce one daughter stem cell, called *self renewal*, and a second daughter cell that will undergo differentiation. In contrast, symmetric division of a stem cell produces either two daughter stem

cells or two daughter cells that will both undergo differentiation [159]. It is known that cancer cells can utilise the pathway of asymmetric cell division to produce heterogeneous populations of cancer cells that support survival of the cancer [160–162].

Another important mechanism contributing to proliferative heterogeneity in cancer cell populations is cell-induced switching between the slow- and fast-proliferating states, which occurs through cell–cell signalling and direct contact between cells [163, 164]. The possibility that cells can switch their proliferative state, either through asymmetric cell division or induced switching by surrounding cells, means that the growth rate of a cell population is highly dependent on the influence of each of these two processes.

To illustrate the concept of induced switching we show in Figure 5.2 a series of our experimental images from a two-dimensional proliferation assay using Fucci-C8161 melanoma cells [13, 106]. See Electronic Supplementary Material for further details. While this experiment

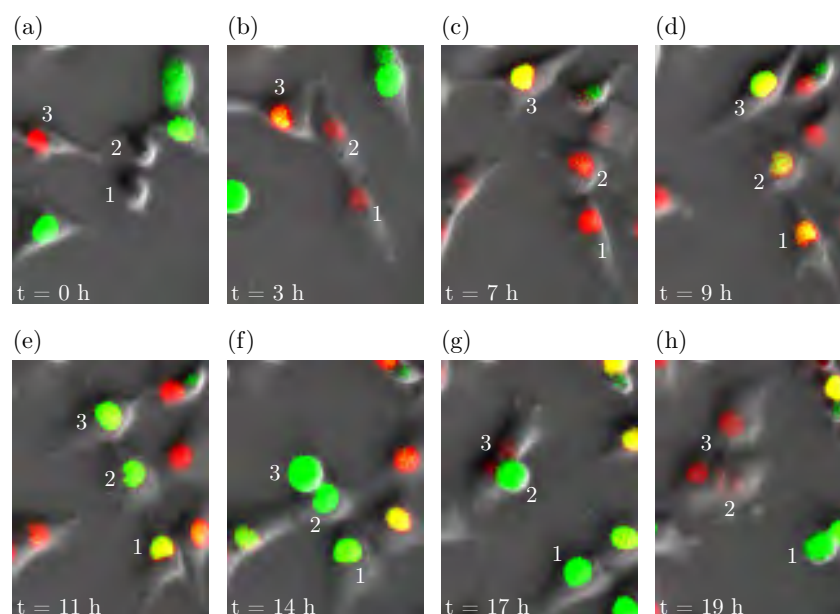


Figure 5.2: Experimental images from a proliferation assay using Fucci-C8161 melanoma cells, illustrating the possibility of induced switching between proliferative states, as discussed in the text.

is not explicitly designed to study slow- and fast-proliferating subpopulations, it is possible that switching of cell cycle speeds occurs and can be observed by careful inspection of the time-series images. For example, consider the cells labelled 1, 2 and 3 throughout the images. Cells 1 and 2 are daughter cells from the same parent cell and are in an early stage of G1, while cell 3 is a daughter cell from a different parent cell and is in a later stage of G1 at time 0 hours (Figure 5.2(a)). At time 3 hours we observe cell 3 interact closely with cell 2 and not cell 1 (Figure 5.2(b)). At times 7 and 9 hours the three cells continue to progress through the cell cycle with no close interaction between cell 3 and cells 1 and 2 (Figure 5.2(c)–(d)). At 11 hours,

cell 3 interacts closely with cell 2 again; cell 2 is in S/G2/M phase, which is further through the cell cycle than cell 1 which is in eS phase (Figure 5.2(e)). At time 14 hours cell 3, which is still close to cell 2, is in M phase and is undergoing mitotic rounding in preparation for cell division (Figure 5.2(f)). At 17 hours, cell 3 undergoes division to produce two daughter cells, cell 2 is undergoing mitotic rounding in preparation for division, whereas cell 1 is in an earlier stage of the cell cycle (Figure 5.2(g)). At time 19 hours cell 2 has divided to produce two daughter cells, whereas cell 1 has only just undergone mitotic rounding in preparation for division (Figure 5.2(h)). These experimental observations illustrate the possibility that cells can switch between states of slow and fast proliferation, induced by surrounding cells. In summary, it seems plausible that cell 2 progresses through the cell cycle faster than cell 1 because of the interactions with cell 3. Indeed, for a given cell line, G1 phase tends to have the most variable duration of the cell cycle phases [76], and cells 1 and 2 appear to progress through G1 at a similar rate (Figure 5.2(a)–(c)), so it is possible that asymmetric division does not account for the overall variation in cell cycle duration between cells 1 and 2.

5.2.2 Mathematical considerations

Delay differential equations are often used when the evolution of the process to be modelled depends on the history of the process, represented as a time delay which may be discrete [165–167], distributed [150, 168–171], or, more generally, state-dependent [148]. We employ a system of two coupled nonlinear delay differential equations to model the transient dynamics of cell proliferation in a population consisting of slow- and fast-proliferating cells. The time delays are distributed so that only cells of a certain age can proliferate according to an appropriate probability distribution, or *delay kernel*, and our system is therefore of integro-differential type. An alternative to modelling cell proliferation with delay differential equations is to use a multi-stage model, however this is not suitable for modelling the cell proliferation scenario that we consider here. Indeed, cell cycle durations in the multi-stage model are hypoexponentially distributed, and we want to allow for more general distributions of cell cycle duration. Further, the multi-stage model can be difficult to parameterise due to the large number of stages and hence parameters required to represent the cell cycle as stages.

5.2.2.0.1 Distributed delays Standard deterministic mathematical models of cell proliferation, such as exponential and logistic growth models, are based on cell cycle durations with an exponential distribution, which allows for a relatively large probability of arbitrarily small cell cycle durations. Experimental investigations, however, suggest that the duration of the cell

cycle, and in particular each cell cycle phase, is not exponentially distributed, rather the hypoexponential distribution is often found to be a reasonable approximation [76, 77, 82, 83, 86]. Our experimental data support these observations (Figure S1, Electronic Supplementary Material). Ordinary differential equations therefore tend to overestimate the cell population growth rate, and may not qualitatively and quantitatively represent the transient growth dynamics of cell populations, particularly for slow-proliferating cells.

The distributions of cell cycle durations for all cell lines are naturally bounded, so unbounded distributions such as the hypoexponential distribution are unrealistic since they theoretically have a nonzero probability of arbitrarily large values. For this reason we consider only bounded distributions for their greater biological realism, and we otherwise allow complete generality for the distributions. Unbounded distributions can be left- or right-truncated to form bounded distributions, and while our model assumes a left bound of zero, alternative left bounds are easily incorporated. Our distributed delays have the following form. Let X denote either the slow-proliferating cells S or fast-proliferating cells F , then the distributed delay $\bar{X}(t)$ is defined by

$$\bar{X}(t) = \int_0^{U_X} X(t-z) g_X(z) dz, \quad (5.1)$$

where the upper limit of integration $U_X \in (0, \infty)$ corresponds to the maximum possible duration of the cell cycle, and g_X is the delay kernel which is normalised so that $\int_0^{U_X} g_X(z) dz = 1$. The delay kernel is the probability density for the distribution of the cell cycle durations. $\bar{X}(t)$ is a weighted average over the past population densities $X(t-z)$, and corresponds to the subpopulation of cells at time t that are ready to divide.

5.2.2.0.2 Contact inhibition Normal cells are subject to contact inhibition [75], so we assume that the growth rate of each subpopulation at time t has a logistic density dependence given by $(1 - P(t)/K)$, where P is the total population density and K is the carrying capacity density. For cancer cells, however, contact inhibition may be lost [14] resulting in density-independent growth, so in this case we could set the carrying capacity to infinity in the logistic-growth terms. We can keep the carrying capacity finite in the induced-switching terms, as increasing the cell density beyond a finite carrying capacity would, realistically, increase the probability of cell-induced switching between proliferative states due to surrounding cells.

5.2.2.0.3 Proliferation, switching between proliferative states, and apoptosis Given the cells that are able to undergo division based on the time delays and density constraints, the intrinsic growth rates $0 < r_S \leq r_F$ for slow- and fast-proliferating cells, respectively, determine

the cells that are parent cells and divide at time t . Parent cells can divide symmetrically, where the daughter cells have the same proliferative state as the parent cell, or asymmetrically, where a daughter cell can have a proliferative state different from the parent cell. When a subpopulation of slow-proliferating cells divides to produce twice as many daughter cells, the parameter α_S determines the proportion of these daughter cells that are also slow-proliferating cells, so the proportion $1 - \alpha_S$ of the daughter cells are fast-proliferating cells. Similarly, the parameter α_F determines the proportion of daughter cells from fast-proliferating cells that are also fast-proliferating cells, so the proportion $1 - \alpha_F$ of the daughter cells are slow-proliferating cells. We only consider asymmetric cell division that is self-renewing, that is $\alpha_S, \alpha_F \in [\frac{1}{2}, 1]$, so that the division of a parent cell produces at least one daughter cell with the same proliferative state as the parent, which is the relevant process for cancer cells [159–162]. While the analysis of our model is valid for $\alpha_S, \alpha_F \in [0, 1]$, excluding Theorem 5.2 for non-negativity, our biological motivation necessitates that $\alpha_S, \alpha_F \in [\frac{1}{2}, 1]$.

Induced switching allows for a cell to switch proliferative states at any position of the cell cycle, induced by surrounding cells with a different proliferative state. We make the reasonable assumption that a cell is increasingly likely to switch proliferative states as the density of cells with a different proliferative state increases. This modelling approach would be most realistic when the cell population has a uniform spatial distribution, such as in the proliferation assay in Figure 5.2. The parameter β_S corresponds to the per capita interaction strength of fast-proliferating cells to induce slow-proliferating cells to switch to fast proliferation. Similarly, the parameter β_F corresponds to the per capita interaction strength of slow-proliferating cells to induce fast-proliferating cells to switch to slow proliferation.

Because we focus on growing tumours, for which proliferation outcompetes cell death, and mechanisms of switching between proliferative states, we simplify our model by not incorporating apoptosis.

5.3 Mathematical model

The total cell population consists of the two subpopulations of slow-proliferating cells and fast-proliferating cells, where $P(t) \geq 0$, $S(t) \geq 0$, and $F(t) \geq 0$ are the respective cell densities, so that $P(t) = S(t) + F(t)$. The model is

$$\begin{aligned} \frac{dS(t)}{dt} = & \underbrace{(2\alpha_S - 1)r_S \int_0^{U_S} S(t-z) g_S(z) dz}_{\text{slow-proliferating cells from asymmetric division of slow-proliferating cells}} \underbrace{\left(1 - \frac{(S(t) + F(t))}{K}\right)}_{\text{contact inhibition of proliferation}} \\ & + \underbrace{2(1 - \alpha_F)r_F \int_0^{U_F} F(t-z) g_F(z) dz}_{\text{slow-proliferating cells from asymmetric division of fast-proliferating cells}} \underbrace{\left(1 - \frac{(S(t) + F(t))}{K}\right)}_{\text{contact inhibition of proliferation}} \\ & - \underbrace{\beta_S S(t) \frac{F(t)}{K}}_{\text{induced switching of slow-proliferating cells to fast-proliferating cells}} + \underbrace{\beta_F F(t) \frac{S(t)}{K}}_{\text{induced switching of fast-proliferating cells to slow-proliferating cells}}, \end{aligned} \quad (5.2)$$

$$\begin{aligned} \frac{dF(t)}{dt} = & \underbrace{2(1 - \alpha_S)r_S \int_0^{U_S} S(t-z) g_S(z) dz}_{\text{fast-proliferating cells from asymmetric division of slow-proliferating cells}} \underbrace{\left(1 - \frac{(S(t) + F(t))}{K}\right)}_{\text{contact inhibition of proliferation}} \\ & + \underbrace{(2\alpha_F - 1)r_F \int_0^{U_F} F(t-z) g_F(z) dz}_{\text{fast-proliferating cells from asymmetric division of fast-proliferating cells}} \underbrace{\left(1 - \frac{(S(t) + F(t))}{K}\right)}_{\text{contact inhibition of proliferation}} \\ & + \underbrace{\beta_S S(t) \frac{F(t)}{K}}_{\text{induced switching of slow-proliferating cells to fast-proliferating cells}} - \underbrace{\beta_F F(t) \frac{S(t)}{K}}_{\text{induced switching of fast-proliferating cells to slow-proliferating cells}}, \end{aligned} \quad (5.3)$$

where the parameters satisfy

$$\text{Intrinsic growth rates:} \quad r_S, r_F \in (0, \infty) \text{ with } r_S \leq r_F, \quad (5.4)$$

$$\text{Proportion of symmetric divisions:} \quad \alpha_S, \alpha_F \in [\tfrac{1}{2}, 1], \quad (5.5)$$

$$\text{Maximum cell cycle durations:} \quad U_S, U_F \in (0, \infty), \quad (5.6)$$

$$\text{Induced switching rates:} \quad \beta_S, \beta_F \in [0, \infty), \quad (5.7)$$

and the mean values of the delay kernels g_S and g_F satisfy

$$\int_0^{U_S} z g_S(z) dz \geq \int_0^{U_F} z g_F(z) dz, \quad (5.8)$$

so that the mean cell cycle duration for the slow-proliferating cells is not smaller than the mean cell cycle duration for the fast-proliferating cells.

As functional differential equations depend on the solution and perhaps derivatives of the solution at past times it is necessary to specify a function for the initial condition, called the *history function*. Defining $\widehat{U} = \max\{U_S, U_F\}$, the history function $\phi = (\phi_S, \phi_F)$ for our model satisfies

$$\phi \in C([- \widehat{U}, 0], \mathbb{R}_{>0}^2), \quad (5.9)$$

$$\phi_S + \phi_F \in C([- \widehat{U}, 0], (0, K)), \quad (5.10)$$

where $C([- \widehat{U}, 0], \mathbb{R}_{>0}^2)$ is the space of continuous functions on $[- \widehat{U}, 0]$ into $\mathbb{R}_{>0}^2$ and $C([- \widehat{U}, 0], (0, K))$ is the space of continuous functions on $[- \widehat{U}, 0]$ into $(0, K)$. Note that state space is therefore an infinite-dimensional function space. For bounded delays the state space is typically the Banach space $C([-h, 0], \mathbb{R}^n)$, for some $h \in (0, \infty)$, of continuous functions $\chi: [-h, 0] \rightarrow \mathbb{R}^n$ on the closed interval $[-h, 0]$ under the supremum norm $\|\cdot\|$ defined by $\|\chi\| = \sup\{\|\chi(t)\|_2 \mid t \in [-h, 0]\}$, where $\|\cdot\|_2$ is the Euclidean norm on \mathbb{R}^n . In our case, state space C is defined by

$$C = C([- \widehat{U}, 0], \mathbb{R}^2). \quad (5.11)$$

Finally, we note that adding Equations (5.2) and (5.3) gives

$$\frac{dP(t)}{dt} = \left(r_S \int_0^{U_S} S(t-z) g_S(z) dz + r_F \int_0^{U_F} F(t-z) g_F(z) dz \right) \left(1 - \frac{P(t)}{K} \right), \quad (5.12)$$

so, from the perspective of the whole population, asymmetric division and induced switching have no net effects. Moreover, if we consider the total population as composed of cells in the same proliferative state with $r_S = r_F$, and $g_S(z) = g_F(z) = \delta(z)$ is the Dirac kernel for zero delay, then Equation (5.12) reduces to the logistic growth model (Electronic Supplementary Material).

5.4 Main results

We now discuss our analysis of Equations (5.2) and (5.3), namely non-negativity and boundedness of solutions, existence and uniqueness of solutions, and local stability analysis of the equilibrium points. A solution for Equations (5.2) and (5.3) means the following [172, 173]:

Definition 5.1 (Solution). Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions that satisfy (5.4)–(5.10). A *solution* for the system is a function $(S, F) \in C([- \widehat{U}, u), \mathbb{R}_{\geq 0}^2)$ for some $u \in [0, \infty]$ such that:

- S and F are differentiable on $(0, u)$ and right-differentiable at 0;
- (S, F) satisfies Equations (5.2) and (5.3) for $t \in [0, u)$.

Additionally, (S, F) is a solution with *initial condition* $\phi = (\phi_S, \phi_F) \in C([- \widehat{U}, 0], \mathbb{R}_{> 0}^2)$ if

- $(S, F)|_{[- \widehat{U}, 0]} = \phi$.

5.4.1 Non-negativity and boundedness

Since the dependent variables $S(t)$ and $F(t)$ in Equations (5.2) and (5.3) represent cell densities they must assume non-negative values at all times. Further, the densities of normal cells are bounded above by the carrying capacity density K arising from contact inhibition. Cancer cells typically have unregulated growth due to the loss of contact inhibition, however continued growth depends on environmental conditions such as nutrient availability, so it is reasonable to assume that the density of cancer cells is also bounded by a carrying capacity.

Theorem 5.2 (Non-negativity and boundedness of solutions).

Let (S, F) be a solution for the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions that satisfy (5.4)–(5.10). Then $S(t), F(t) \in [0, K]$ for all $t > 0$, therefore the solutions S and F are non-negative and bounded.

We give an elementary proof of this theorem as it facilitates understanding of the non-negativity and boundedness of the solutions. We first require a lemma.

Lemma 5.3. *Let (S, F) be a solution for the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions that satisfy (5.4)–(5.10). If there exists $T > 0$ such that the distributed delays satisfy the relations $\overline{S}(t) \geq 0$ and $\overline{F}(t) \geq 0$ for all $t \in (0, T]$ then $P(t) \in [0, K]$ for all $t \in (0, T]$.*

Proof. If $P(t_1) > K$ for some $t_1 \in (0, T]$ then, since $P(0) < K$ by Equation (5.10), we may assume without loss of generality that $dP(t)/dt|_{t=t_1} > 0$. Since Equation (5.12) gives $dP(t)/dt|_{t=t_1} \leq 0$, we have a contradiction. It follows that $P(t) \leq K$ for all $t \in (0, T]$. Similarly, if $P(t_2) < 0$ for some $t_2 \in (0, T]$ then, since $P(0) > 0$ by Equation (5.10), we may assume without loss of generality that $dP(t)/dt|_{t=t_2} < 0$. Since Equation (5.12) gives $dP(t)/dt|_{t=t_2} \geq 0$, we have a contradiction. It follows that $P(t) \geq 0$ for all $t \in (0, T]$. \square

Theorem 5.2. It suffices to prove that $S(t) \geq 0$ and $F(t) \geq 0$ for all $t > 0$, for then it follows from Lemma 5.3 that $S(t) \leq K$ and $F(t) \leq K$ for all $t > 0$. Define t_1 and t_2 by $t_1 = \inf\{t > 0 \mid S(t) < 0\}$ and $t_2 = \inf\{t > 0 \mid F(t) < 0\}$. We consider the infima in the extended real numbers, so that $t_1, t_2 \in (0, +\infty]$, where either infimum is equal to $+\infty$ if the corresponding set is empty. The proof consists of four separate cases which are proved similarly. We demonstrate one case here, and provide the complete proof in the Electronic Supplementary Material.

Case 1: Let $\beta_S - \beta_F \geq 0$ and suppose $S(t) < 0$ for some $t > 0$.

Note that $t_1 \in \mathbb{R}$. If $t_1 < t_2$ then choose $t_3 \in (t_1, t_2)$ such that $S(t_3) < 0$, $dS(t)/dt|_{t=t_3} < 0$, and the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3 and since $F(t_3) \geq 0$, Equation (5.2) gives $dS(t)/dt|_{t=t_3} \geq 0$, a contradiction.

If $t_1 \geq t_2$ then, since $F(0) > 0$ and $F(t_2) = 0$, there exists $t_3 \in (0, t_2)$ such that $dF(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.3) gives $dF(t)/dt|_{t=t_3} \geq 0$, a contradiction. We conclude that $S(t) \geq 0$ for all $t > 0$. \square

5.4.2 Existence and uniqueness

We begin by introducing some simplifying notation. For ρ in the state space C we denote the component functions by ρ_S and ρ_F so that $\rho = (\rho_S, \rho_F)$, and then define $\bar{\rho}_S$ and $\bar{\rho}_F$ by

$$\bar{\rho}_S = \int_0^{U_S} \rho_S(-z) g_S(z) dz \quad \text{and} \quad \bar{\rho}_F = \int_0^{U_F} \rho_F(-z) g_F(z) dz. \quad (5.13)$$

Now we define $f: C \rightarrow \mathbb{R}^2$ by

$$f(\rho) = \begin{bmatrix} \left((2\alpha_S - 1)r_S \bar{\rho}_S + 2(1 - \alpha_F)r_F \bar{\rho}_F \right) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \\ - \frac{(\beta_S - \beta_F)}{K} \rho_S(0) \rho_F(0) \\ \left(2(1 - \alpha_S)r_S \bar{\rho}_S + (2\alpha_F - 1)r_F \bar{\rho}_F \right) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \\ + \frac{(\beta_S - \beta_F)}{K} \rho_S(0) \rho_F(0) \end{bmatrix}. \quad (5.14)$$

Note that f is continuous. If (S, F) is a solution for Equations (5.2) and (5.3), $t \geq 0$, and we define $(S_t, F_t) \in C([- \widehat{U}, 0], \mathbb{R}^2)$ by $(S_t(r), F_t(r)) = (S(t+r), F(t+r))$ for $r \in [- \widehat{U}, 0]$, then $f((S_t, F_t)) = [dS(t)/dt, dF(t)/dt]^T$ from (5.2) and (5.3).

Theorem 5.4 (Existence and uniqueness of solutions).

Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions which satisfy (5.4)–(5.10). Then there exists a unique solution $(S, F) \in C([- \widehat{U}, \infty), \mathbb{R}_{\geq 0}^2)$ of (5.2) and (5.3).

Proof. Here we give an outline of the proof. The complete proof is provided in the Electronic Supplementary Material. We first show that f defined in Equation (5.14) satisfies the following Lipschitz condition on every bounded subset of C : for all $M > 0$ there exists $L > 0$ such that for every $\rho, \psi \in C([- \widehat{U}, 0], \mathbb{R}^2)$ with $\|\rho\|, \|\psi\| \leq M$ we have $\|f(\rho) - f(\psi)\|_2 \leq L\|\rho - \psi\|$.

To further simplify the notation in Equation (5.14) we define $\kappa_1 = (2\alpha_S - 1)r_S$, $\kappa_2 = 2(1 - \alpha_F)r_F$, $\kappa_3 = 2(1 - \alpha_S)r_S$, $\kappa_4 = (2\alpha_F - 1)r_F$, and $\kappa_5 = (\beta_S - \beta_F)/K$. Now,

$$f(\rho) - f(\psi) = \begin{bmatrix} (\kappa_1(\overline{\rho_S} - \overline{\psi_S}) + \kappa_2(\overline{\rho_F} - \overline{\psi_F})) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K}\right) \\ + (\kappa_1\overline{\psi_S} + \kappa_2\overline{\psi_F}) \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K}\right) \\ + \kappa_5(\psi_S - \rho_S)(0)\psi_F(0) + \kappa_5(\psi_F - \rho_F)(0)\rho_S(0) \\ (\kappa_3(\overline{\rho_S} - \overline{\psi_S}) + \kappa_4(\overline{\rho_F} - \overline{\psi_F})) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K}\right) \\ + (\kappa_3\overline{\psi_S} + \kappa_4\overline{\psi_F}) \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K}\right) \\ + \kappa_5(\rho_S - \psi_S)(0)\psi_F(0) + \kappa_5(\rho_F - \psi_F)(0)\rho_S(0) \end{bmatrix}$$

so, using the triangle inequality, we obtain

$$\|f(\rho) - f(\psi)\|_2 \leq \left(\left(\sqrt{\kappa_1^2 + \kappa_3^2} + \sqrt{\kappa_2^2 + \kappa_4^2} \right) \left(1 + \frac{4M}{K}\right) + 2\sqrt{2}|\kappa_5|M \right) \|\rho - \psi\|,$$

so we can set L to be

$$L = \left(\sqrt{\kappa_1^2 + \kappa_3^2} + \sqrt{\kappa_2^2 + \kappa_4^2} \right) \left(1 + \frac{4M}{K}\right) + 2\sqrt{2}|\kappa_5|M$$

and then f satisfies the Lipschitz condition. Then [173, Page 32, Theorem 3.7] provides local existence and uniqueness of solutions for the system (5.2) and (5.3). Since our solutions of interest are bounded by Theorem 5.2, it follows from [173, Page 37, Proposition 3.10] that the solutions are continuable to all positive time. \square

5.4.3 Local stability

Here we consider the local stability analysis of the equilibrium points for the system in (5.2) and (5.3), and show that β_S and β_F are bifurcation parameters with bifurcation point when $\beta_S = \beta_F$. We will prove the following theorem.

Theorem 5.5 (Local stability).

Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions that satisfy (5.4)–(5.10).

- *When $\beta_S \neq \beta_F$ the system has the three equilibrium points $(0, 0)$, $(0, K)$, and $(K, 0)$ with the following properties:*
 - *$(0, 0)$ is locally unstable.*
 - *If $\beta_S > \beta_F$ then $(K, 0)$ is locally unstable and $(0, K)$ is locally stable.*
 - *If $\beta_S < \beta_F$ then $(K, 0)$ is locally stable and $(0, K)$ is locally unstable.*
- *When $\beta_S = \beta_F$ the system has infinitely many equilibrium points corresponding to the line segment joining $(K, 0)$ and $(0, K)$, all of which are locally stable.*

The parameters β_S and β_F are therefore bifurcation parameters.

Note that, since equilibrium points for delay differential equations are functions in a Banach space, we have $(0, 0), (0, K), (K, 0) \in C([- \hat{U}, \infty), \mathbb{R}_{\geq 0}^2)$.

The proof of Theorem 5.5 follows immediately from Propositions 5.6, 5.8, 5.9 and 5.10. We begin by non-dimensionalising Equations (5.2) and (5.3), and then linearising the non-dimensional system about the equilibrium points. Denoting the dimensionless variables with a caret, we define $\hat{t} = r_F t$, $\hat{S}(\hat{t}) = S(t)/K$ and $\hat{F}(\hat{t}) = F(t)/K$. We also define the dimensionless parameters $r = r_S/r_F$ and $\beta = (\beta_S - \beta_F)/r_F$. Equations (5.2) and (5.3) then become, dropping the caret notation for simplicity,

$$\begin{aligned} \frac{dS(t)}{dt} = & \left((2\alpha_S - 1)r \int_0^{U_S} S(t - r_F z) g_S(z) dz + 2(1 - \alpha_F) \int_0^{U_F} F(t - r_F z) g_F(z) dz \right) \\ & \times (1 - S(t) - F(t)) - \beta S(t)F(t), \end{aligned} \quad (5.15)$$

$$\begin{aligned} \frac{dF(t)}{dt} = & \left(2(1 - \alpha_S)r \int_0^{U_S} S(t - r_F z) g_S(z) dz + (2\alpha_F - 1) \int_0^{U_F} F(t - r_F z) g_F(z) dz \right) \\ & \times (1 - S(t) - F(t)) + \beta S(t)F(t). \end{aligned} \quad (5.16)$$

Since S and F are cell densities, hence non-negative, we only consider equilibrium points $(S^*, F^*) \in C([- \widehat{U}, \infty), \mathbb{R}_{\geq 0}^2)$. To find the equilibrium points we substitute $S = S^*$ and $F = F^*$ into (5.15) and (5.16) to give

$$0 = ((2\alpha_S - 1)rS^* + 2(1 - \alpha_F)F^*)(1 - S^* - F^*) - \beta S^* F^*, \quad (5.17)$$

$$0 = (2(1 - \alpha_S)rS^* + (2\alpha_F - 1)F^*)(1 - S^* - F^*) + \beta S^* F^*, \quad (5.18)$$

hence $(S^*, F^*) = (0, 0)$, $(1, 0)$ or $(0, 1)$ when $\beta \neq 0$. When $\beta = 0$ the equilibrium points consist of the two lines $(S^*, F^*) = (u, 1 - u)$ for all $u \in \mathbb{R}$ and $(S^*, F^*) = (u, -ru)$ for all $u \in \mathbb{R}$, for which the non-negative points are $(S^*, F^*) = (u, 1 - u)$ for all $u \in [0, 1]$ and $(S^*, F^*) = (0, 0)$.

To examine the local stability of the equilibrium points (S^*, F^*) we linearise the system in Equations (5.15) and (5.16) about each point. Defining $x(t) = S(t) - S^*$ and $y(t) = F(t) - F^*$ we obtain the linearised system:

$$\begin{aligned} \frac{dx(t)}{dt} &= \left((2\alpha_S - 1)r \left(\int_0^{U_S} x(t - r_F z) g_S(z) dz + S^* \right) \right. \\ &\quad \left. + 2(1 - \alpha_F) \left(\int_0^{U_F} y(t - r_F z) g_F(z) dz + F^* \right) \right) \\ &\quad \times (1 - x(t) - y(t) - S^* - F^*) - \beta(x(t) + S^*)(y(t) + F^*) \\ &\sim \left((2\alpha_S - 1)r \int_0^{U_S} x(t - r_F z) g_S(z) dz + 2(1 - \alpha_F) \int_0^{U_F} y(t - r_F z) g_F(z) dz \right) \\ &\quad \times (1 - S^* - F^*) \\ &\quad + \left((2\alpha_S - 1)rS^* + 2(1 - \alpha_F)F^* \right) (1 - x(t) - y(t) - S^* - F^*) \\ &\quad - \beta(x(t)F^* + y(t)S^* + S^*F^*), \quad \text{as } x(t), y(t) \rightarrow 0, \end{aligned} \quad (5.19)$$

$$\begin{aligned} \frac{dy(t)}{dt} &= \left(2(1 - \alpha_S)r \left(\int_0^{U_S} x(t - r_F z) g_S(z) dz + S^* \right) \right. \\ &\quad \left. + (2\alpha_F - 1) \left(\int_0^{U_F} y(t - r_F z) g_F(z) dz + F^* \right) \right) \\ &\quad \times (1 - x(t) - y(t) - S^* - F^*) + \beta(x(t) + S^*)(y(t) + F^*) \\ &\sim \left(2(1 - \alpha_S)r \int_0^{U_S} x(t - r_F z) g_S(z) dz + (2\alpha_F - 1) \int_0^{U_F} y(t - r_F z) g_F(z) dz \right) \\ &\quad \times (1 - S^* - F^*) \\ &\quad + \left(2(1 - \alpha_S)rS^* + (2\alpha_F - 1)F^* \right) (1 - x(t) - y(t) - S^* - F^*) \\ &\quad + \beta(x(t)F^* + y(t)S^* + S^*F^*), \quad \text{as } x(t), y(t) \rightarrow 0. \end{aligned} \quad (5.20)$$

By the Principle of Linearised Stability [174, Page 240, Theorem 6.8] it suffices to consider the

stability of the equilibrium points for the linearisation in Equations (5.19) and (5.20).

Proposition 5.6 (Equilibrium point $(0, 0)$).

Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions which satisfy (5.4)–(5.10). Then $(0, 0)$ is locally unstable.

Proposition 5.6 follows immediately from Proposition 5.7, in which we analyse the transcendental characteristic equation associated with $(0, 0)$ of the linearised system (5.19) and (5.20) to show that the characteristic equation has at least one zero in \mathbb{C} with positive real part. While there are alternative methods for proving Proposition 5.6 (Electronic Supplementary Material), we consider the direct approach of analysing the associated transcendental characteristic equation to have mathematical relevance for other studies involving delay differential equations. Indeed, transcendental characteristic equations are generally difficult to analyse [174, Chapter XI], so new analysis of such equations is of mathematical interest.

For $(S^*, F^*) = (0, 0)$, Equations (5.19) and (5.20) become

$$\frac{dx(t)}{dt} = (2\alpha_S - 1)r \int_0^{U_S} x(t - r_F z) g_S(z) dz + 2(1 - \alpha_F) \int_0^{U_F} y(t - r_F z) g_F(z) dz, \quad (5.21)$$

$$\frac{dy(t)}{dt} = 2(1 - \alpha_S)r \int_0^{U_S} x(t - r_F z) g_S(z) dz + (2\alpha_F - 1) \int_0^{U_F} y(t - r_F z) g_F(z) dz. \quad (5.22)$$

Equations (5.21) and (5.22) have a solution of the form

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\lambda t}, \quad \text{where } \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \in \mathbb{C}^2 \text{ is nonzero and } \lambda \in \mathbb{C}, \quad (5.23)$$

so substitution gives

$$e^{\lambda t} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} (2\alpha_S - 1)r \int_0^{U_S} e^{-\lambda r_F z} g_S(z) dz & 2(1 - \alpha_F) \int_0^{U_F} e^{-\lambda r_F z} g_F(z) dz \\ 2(1 - \alpha_S)r \int_0^{U_S} e^{-\lambda r_F z} g_S(z) dz & (2\alpha_F - 1) \int_0^{U_F} e^{-\lambda r_F z} g_F(z) dz \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\lambda t}. \quad (5.24)$$

To ensure that $(c_1, c_2)^\top \neq 0$ we must have the characteristic equation

$$\begin{aligned}
 G(\lambda) &= \begin{vmatrix} (2\alpha_S - 1)r \int_0^{U_S} e^{-\lambda r_F z} g_S(z) dz - \lambda & 2(1 - \alpha_F) \int_0^{U_F} e^{-\lambda r_F z} g_F(z) dz \\ 2(1 - \alpha_S)r \int_0^{U_S} e^{-\lambda r_F z} g_S(z) dz & (2\alpha_F - 1) \int_0^{U_F} e^{-\lambda r_F z} g_F(z) dz - \lambda \end{vmatrix} \\
 &= \lambda^2 - \lambda \left((2\alpha_S - 1)r \int_0^{U_S} e^{-\lambda r_F z} g_S(z) dz + (2\alpha_F - 1) \int_0^{U_F} e^{-\lambda r_F z} g_F(z) dz \right) \\
 &\quad + (2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} e^{-\lambda r_F(z+v)} g_F(v) g_S(z) dv dz
 \end{aligned} \tag{5.25}$$

equal to zero. The zeros of the transcendental equation $G(\lambda)$ are the eigenvalues. Proposition 5.7 shows that $G(\lambda)$ has at least one zero with positive real part, so the equilibrium point $(0, 0)$ is locally unstable. In the proof of Proposition 5.7 we consider three cases for $G(\lambda)$ in Equation (5.25), depending on whether $2\alpha_S + 2\alpha_F - 3$ is negative, zero, or positive. To understand the physical interpretation of $2\alpha_S + 2\alpha_F - 3$, first note that $2\alpha_S + 2\alpha_F - 3 = (2\alpha_S - 1) - 2(1 - \alpha_F)$. Referring to Equation (5.15), $(2\alpha_S - 1) - 2(1 - \alpha_F)$ is the difference between the proportion of slow-proliferating parent cells that produce slow-proliferating daughter cells beyond self renewal, and the proportion of fast-proliferating parent cells that produce slow-proliferating daughter cells, at a given time. A similar interpretation follows by referring to Equation (5.16) and noting that $2\alpha_S + 2\alpha_F - 3 = (2\alpha_F - 1) - 2(1 - \alpha_S)$. When $2\alpha_S + 2\alpha_F - 3$ is negative or zero we use the intermediate value theorem to prove that $G(\lambda)$ has a zero in \mathbb{R} which is positive. When $2\alpha_S + 2\alpha_F - 3$ is positive, however, we require a different approach involving Cauchy's argument principle, which we now outline.

Let Ω be a non-empty connected open set, let Γ be a closed curve in Ω with positive, or counter-clockwise, orientation which is homologous to zero with respect to Ω , and let h be a meromorphic function on Ω with no zeros or poles on Γ . Then Cauchy's argument principle is [175, Page 152, Theorem 18]

$$\frac{1}{2\pi i} \oint_{\Gamma} \frac{h'(\lambda)}{h(\lambda)} d\lambda = \mathcal{Z} - \mathcal{P}, \tag{5.26}$$

where \mathcal{Z} is the number of zeros of h inside Γ and \mathcal{P} is the number of poles of h inside Γ , including multiplicities.

Now, let Γ be a piecewise differentiable closed curve in \mathbb{C} with positive orientation that does not pass through the point z_0 . Then the *index* of z_0 with respect to Γ , denoted $\text{Ind}_{\Gamma}(z_0)$, is defined by [175, Page 115]

$$\text{Ind}_{\Gamma}(z_0) = \frac{1}{2\pi i} \oint_{\Gamma} \frac{dz}{z - z_0}. \tag{5.27}$$

Note that $\text{Ind}_\Gamma(z_0)$ is also referred to as the *winding number* of Γ with respect to z_0 . By substituting $z = h(\lambda)$ into Equation (5.26) and using Equation (5.27) we arrive at the standard observation

$$\mathcal{Z} - \mathcal{P} = \frac{1}{2\pi i} \oint_{h(\Gamma)} \frac{dz}{z} = \text{Ind}_{h(\Gamma)}(0), \quad (5.28)$$

where the term on the right of the last equality is the winding number of the closed curve $h(\Gamma)$ with respect to the origin. Therefore, the number of zeros minus the number of poles of h inside Γ , including multiplicities, can be determined by calculating the winding number of the image $h(\Gamma)$ with respect to the origin.

Note that $G(\lambda)$ in Equation (5.25) is a holomorphic function, hence meromorphic, on \mathbb{C} , and has no poles. Therefore, the number of zeros of $G(\lambda)$ inside a contour Γ which satisfies the conditions for Equations (5.26) and (5.27) is equal to the winding number of $G(\Gamma)$ with respect to the origin. We will be considering rectangular contours with positive orientation in the right half-plane, and we need to know that $G(\lambda)$ is not identically zero in the region bounded by the closed contour. Our contour will bound an interval of the positive real axis arbitrarily close to, but excluding, the origin, and with arbitrary upper bound. For a contour which bounds sufficiently large positive real numbers, and by considering $\text{Re}(G(\lambda))$ on the positive real axis, we can see that $G(\lambda)$ is not identically zero in the region bounded by the contour. Furthermore, by [176, Page 208, Theorem 10.18] it follows that the zeros of $G(\lambda)$ are isolated and countable, so we can always choose an appropriate rectangular contour which does not pass through a zero of $G(\lambda)$. In the Electronic Supplementary Material we graphically illustrate our application of Cauchy's argument principle. We now state and prove the required proposition.

Proposition 5.7. *The transcendental characteristic equation $G(\lambda)$ in Equation (5.25) has at least one zero in \mathbb{C} with positive real part.*

Proof. We consider three cases according to whether $2\alpha_S + 2\alpha_F - 3$ is negative, zero or positive.

Case 1: If $2\alpha_S + 2\alpha_F - 3 < 0$ then, since $G(0) = (2\alpha_S + 2\alpha_F - 3)r < 0$ and $\lim_{\lambda \rightarrow \infty} G|_{\mathbb{R}}(\lambda) = \lim_{\lambda \rightarrow \infty} \lambda^2 = \infty$, it follows that G has a positive real zero by the intermediate value theorem.

Case 2: If $2\alpha_S + 2\alpha_F - 3 = 0$ then G factors as $G(\lambda) = \lambda H(\lambda)$, where

$$H(\lambda) = \lambda - \left((2\alpha_S - 1)r \int_0^{U_S} e^{-\lambda r_F z} g_S(z) dz + (2\alpha_F - 1) \int_0^{U_F} e^{-\lambda r_F z} g_F(z) dz \right). \quad (5.29)$$

Since $2\alpha_S + 2\alpha_F - 3 = 0$ it follows that $2\alpha_S - 1 < 0$ implies $\alpha_F > 1$, so $2\alpha_S - 1 \geq 0$. Similarly, $2\alpha_F - 1 \geq 0$. Further, if both $2\alpha_S - 1 = 0$ and $2\alpha_F - 1 = 0$ then $2\alpha_S + 2\alpha_F - 2 = 0$, contradicting

$2\alpha_S + 2\alpha_F - 3 = 0$. It therefore follows that $H(0) = -((2\alpha_S - 1)r + (2\alpha_F - 1)) < 0$. Since $\lim_{\lambda \rightarrow \infty} H|_{\mathbb{R}}(\lambda) = \lim_{\lambda \rightarrow \infty} \lambda = \infty$, it follows that H has a positive real zero by the intermediate value theorem.

Case 3: Suppose now that $2\alpha_S + 2\alpha_F - 3 > 0$. We employ Cauchy's argument principle to show that the holomorphic function $G(\lambda)$ has a zero with positive real part. Let Γ be the simple closed contour with positive orientation in the right half-plane of the complex plane defined piecewise as follows:

$$\Gamma_1 : (m(1-t) + (N/2)t) - iN, \quad 0 \leq t \leq 1, \quad (5.30)$$

$$\Gamma_2 : ((N/2)(1-t) + (3N/2)t) - iN, \quad 0 \leq t \leq 1, \quad (5.31)$$

$$\Gamma_3 : (3N/2) + i((-N)(1-t) + Nt), \quad 0 \leq t \leq 1, \quad (5.32)$$

$$\Gamma_4 : ((3N/2)(1-t) + (N/2)t) + iN, \quad 0 \leq t \leq 1, \quad (5.33)$$

$$\Gamma_5 : ((N/2)(1-t) + mt) + iN, \quad 0 \leq t \leq 1, \quad (5.34)$$

$$\Gamma_6 : m + i(N(1-t) + (-N)t), \quad 0 \leq t \leq 1, \quad (5.35)$$

where we fix $m > 0$ to be arbitrarily small and $N > 0$ to be arbitrarily large. Note that $\Gamma = \bigcup_{j=1}^6 \Gamma_j$ is rectangular, with vertices at $m - iN$, $3N/2 - iN$, $3N/2 + iN$, and $m + iN$. Since the zeros of a holomorphic function that is not identically zero are isolated and countable, we can choose m arbitrarily small and N arbitrarily large while ensuring $G(\lambda)$ is nonzero on Γ . Therefore, since $G(\lambda)$ has no poles, the number of zeros of $G(\lambda)$ inside Γ is equal to the index of the image contour $G(\Gamma)$ with respect to the origin, $\text{Ind}_{G(\Gamma)}(0)$. We now begin our application of the argument principle. Since we only need to show the existence of one zero with a positive real part, it suffices to prove that $\text{Ind}_{G(\Gamma)}(0) \geq 1$. Specifically, we show that the image contour $G(\Gamma)$ crosses the positive real axis at least once in a counter-clockwise direction while encircling the origin, and doesn't cross the positive real axis in a clockwise direction.

We will traverse Γ for one cycle in a counter-clockwise direction beginning with Γ_1 , and determine when $G(\Gamma)$ crosses the positive real axis. For this it is helpful to consider the real and imaginary parts of $G(\lambda)$, so evaluating $G(\lambda)$ at the arbitrary complex number $\lambda = a + ib$

we have

$$\begin{aligned}
 \operatorname{Re}(G(\lambda)) &= a^2 - b^2 - a(2\alpha_S - 1)r \int_0^{U_S} e^{-ar_F z} \cos(br_F z) g_S(z) dz \\
 &\quad - a(2\alpha_F - 1) \int_0^{U_F} e^{-ar_F z} \cos(br_F z) g_F(z) dz \\
 &\quad - b(2\alpha_S - 1)r \int_0^{U_S} e^{-ar_F z} \sin(br_F z) g_S(z) dz \\
 &\quad - b(2\alpha_F - 1) \int_0^{U_F} e^{-ar_F z} \sin(br_F z) g_F(z) dz \\
 &\quad + (2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} e^{-ar_F(z+v)} \cos(br_F(z+v)) g_F(v) g_S(z) dv dz, \tag{5.36}
 \end{aligned}$$

and

$$\begin{aligned}
 \operatorname{Im}(G(\lambda)) &= 2ab + a(2\alpha_S - 1)r \int_0^{U_S} e^{-ar_F z} \sin(br_F z) g_S(z) dz \\
 &\quad + a(2\alpha_F - 1) \int_0^{U_F} e^{-ar_F z} \sin(br_F z) g_F(z) dz \\
 &\quad - b(2\alpha_S - 1)r \int_0^{U_S} e^{-ar_F z} \cos(br_F z) g_S(z) dz \\
 &\quad - b(2\alpha_F - 1) \int_0^{U_F} e^{-ar_F z} \cos(br_F z) g_F(z) dz \\
 &\quad - (2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} e^{-ar_F(z+v)} \sin(br_F(z+v)) g_F(v) g_S(z) dv dz. \tag{5.37}
 \end{aligned}$$

In the following argument we shall generally use that m is arbitrarily small and N is arbitrarily large without further comment.

Consider $G(\lambda)$ along Γ_1 , where $b = -N$ and a increases from m to $N/2$. For sufficiently large N and for all $a \in [m, N/2]$, $\operatorname{Re}(G(\lambda)) < 0$ and $\operatorname{Re}(G(\lambda))$ is dominated by N^2 . At the end of Γ_1 and for sufficiently large N , $\operatorname{Im}(G(\lambda)) < 0$ and $\operatorname{Im}(G(\lambda))$ is dominated by N^2 . So $G(\Gamma_1)$ starts in the left half-plane and ends in the third quadrant.

Consider $G(\lambda)$ along Γ_2 , where $b = -N$ and a increases from $N/2$ to $3N/2$. At the end of Γ_2 and for sufficiently large N , $\operatorname{Re}(G(\lambda)) > 0$ and $\operatorname{Re}(G(\lambda))$ is dominated by N^2 . For sufficiently large N and for all $a \in [N/2, 3N/2]$, $\operatorname{Im}(G(\lambda)) < 0$ and $\operatorname{Re}(G(\lambda))$ is dominated by N^2 . So $G(\Gamma_2)$ starts in the third quadrant and ends in the fourth quadrant.

Consider $G(\lambda)$ along Γ_3 , where $a = 3N/2$ and b increases from $-N$ to N . For sufficiently large N , $\operatorname{Re}(G(\lambda)) > 0$ and $\operatorname{Re}(G(\lambda))$ is dominated by N^2 . At the end of Γ_3 and for sufficiently large N , $\operatorname{Im}(G(\lambda)) > 0$ and $\operatorname{Im}(G(\lambda))$ is dominated by N^2 . So $G(\Gamma_3)$ starts in the fourth quadrant and ends in the first quadrant. The image contour $G(\Gamma)$ has now crossed the positive real axis in a counter-clockwise direction.

Consider $G(\lambda)$ along Γ_4 , where $b = N$ and a decreases from $3N/2$ to $N/2$. At the end of Γ_4 and for sufficiently large N , $\operatorname{Re}(G(\lambda)) < 0$ and $\operatorname{Re}(G(\lambda))$ is dominated by N^2 . For sufficiently large N and for all $a \in [3N/2, N/2]$, $\operatorname{Im}(G(\lambda)) > 0$ and $\operatorname{Im}(G(\lambda))$ is dominated by N^2 . So $G(\Gamma_4)$ starts in the first quadrant and ends in the second quadrant.

Consider $G(\lambda)$ along Γ_5 , where $b = N$ and a decreases from $N/2$ to m . At the end of Γ_5 and for sufficiently large N , $\operatorname{Re}(G(\lambda)) < 0$ and $\operatorname{Re}(G(\lambda))$ is dominated by N^2 . $\operatorname{Im}(G(\lambda))$ could be positive or negative. So $G(\Gamma_5)$ starts in the second quadrant and ends in the left half-plane.

Consider $G(\lambda)$ along Γ_6 , where $a = m$ and b decreases from N to $-N$, which completes one circuit around Γ in a counter-clockwise direction. If we fix N to be as large as required then we can choose m sufficiently small so that along Γ_6 the Equations (5.36) and (5.37) are approximated arbitrarily closely by the equations

$$\begin{aligned} \operatorname{Re}(G(b)) = & -b^2 - b(2\alpha_S - 1)r \int_0^{U_S} \sin(br_F z) g_S(z) dz \\ & - b(2\alpha_F - 1) \int_0^{U_F} \sin(br_F z) g_F(z) dz \\ & + (2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} \cos(br_F(z+v)) g_F(v) g_S(z) dv dz, \end{aligned} \quad (5.38)$$

$$\begin{aligned} \operatorname{Im}(G(b)) = & -b(2\alpha_S - 1)r \int_0^{U_S} \cos(br_F z) g_S(z) dz \\ & - b(2\alpha_F - 1) \int_0^{U_F} \cos(br_F z) g_F(z) dz \\ & - (2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} \sin(br_F(z+v)) g_F(v) g_S(z) dv dz. \end{aligned} \quad (5.39)$$

For notational simplicity, define the functions f_1 , f_2 , f_3 , g_1 , g_2 , and g_3 by

$$f_1(b) = -b(2\alpha_S - 1)r \int_0^{U_S} \sin(br_F z) g_S(z) dz, \quad (5.40)$$

$$f_2(b) = -b(2\alpha_F - 1) \int_0^{U_F} \sin(br_F z) g_F(z) dz, \quad (5.41)$$

$$f_3(b) = (2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} \cos(br_F(z+v)) g_F(v) g_S(z) dv dz, \quad (5.42)$$

$$g_1(b) = -b(2\alpha_S - 1)r \int_0^{U_S} \cos(br_F z) g_S(z) dz, \quad (5.43)$$

$$g_2(b) = -b(2\alpha_F - 1) \int_0^{U_F} \cos(br_F z) g_F(z) dz, \quad (5.44)$$

$$g_3(b) = -(2\alpha_S + 2\alpha_F - 3)r \int_0^{U_S} \int_0^{U_F} \sin(br_F(z+v)) g_F(v) g_S(z) dv dz, \quad (5.45)$$

so that Equations (5.38) and (5.39) become

$$\operatorname{Re}(G(b)) = -b^2 + f_1(b) + f_2(b) + f_3(b), \quad (5.46)$$

$$\operatorname{Im}(G(b)) = g_1(b) + g_2(b) + g_3(b). \quad (5.47)$$

Now, consider decreasing b from N to $-N$. The curves $f_1(b) + ig_1(b)$ and $f_2(b) + ig_2(b)$ have the same orientation as the spiral $-b \sin(b) - ib \cos(b)$, which is traversed counter-clockwise as b decreases. Similarly, the curve $f_3(b) + ig_3(b)$ has the same orientation as the circle $\cos(b) - i \sin(b)$, which is also counter-clockwise. Note that for discrete delays the curves $f_1(b) + ig_1(b)$ and $f_2(b) + ig_2(b)$ are spirals and the curve $f_3(b) + ig_3(b)$ is a circle. The sum of these three curves, $\sum_{k=1}^3 f_k(b) + ig_k(b)$, has counter-clockwise orientation, and the $-b^2$ term in $\operatorname{Re}(G(b))$ translates these curves along the negative real axis. It follows that if $G(b)$ encircles the origin as b decreases from N to $-N$ then it does so in a counter-clockwise direction. In particular, $G(b)$ does not encircle the origin in a clockwise direction. So, for sufficiently small m , $G(\Gamma_6)$ can only encircle the origin in a counter-clockwise direction.

To ensure that the image contour $G(\Gamma)$ encircles the origin at least once, note that $G(\Gamma_6)$ crosses the positive real axis in a counter-clockwise direction at approximately the point $(2\alpha_S + 2\alpha_F - 3)r > 0$ for sufficiently small m , so it follows that $G(\Gamma_6)$ must cross the negative real axis at a point closer to the start of Γ_6 . Therefore, the symmetry of $\operatorname{Im}(G(\lambda))$ with respect to the real axis implies that $G(\Gamma)$ completes a cycle around the origin before the end of $G(\Gamma_6)$. We conclude that $\operatorname{Ind}_{G(\Gamma)}(0) \geq 1$, and our proof is complete. \square

Proposition 5.8 (Equilibrium point $(1, 0)$ when $\beta_S - \beta_F \neq 0$).

Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions which satisfy (5.4)–(5.10). For all $\beta_S, \beta_F \in \mathbb{R}$, $(1, 0)$ is locally stable when $\beta_S - \beta_F < 0$ and locally unstable when $\beta_S - \beta_F > 0$.

Proof. For $(S^*, F^*) = (1, 0)$, Equations (5.19) and (5.20) become

$$\frac{dx(t)}{dt} = (2\alpha_S - 1)r(-x(t) - y(t)) - \beta y(t), \quad (5.48)$$

$$\frac{dy(t)}{dt} = 2(1 - \alpha_S)r(-x(t) - y(t)) + \beta y(t). \quad (5.49)$$

Equations (5.48) and (5.49) have a solution of the form in Equation (5.23), so substitution gives

$$e^{\lambda t} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -(2\alpha_S - 1)r & -(2\alpha_S - 1)r - \beta \\ -2(1 - \alpha_S)r & -2(1 - \alpha_S)r + \beta \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\lambda t}. \quad (5.50)$$

To ensure that $(c_1, c_2)^\top \neq 0$ we must have the characteristic equation

$$G(\lambda) = \begin{vmatrix} -(2\alpha_S - 1)r - \lambda & -(2\alpha_S - 1)r - \beta \\ -2(1 - \alpha_S)r & -2(1 - \alpha_S)r + \beta - \lambda \end{vmatrix} = \lambda^2 + \lambda(r - \beta) - \beta r \quad (5.51)$$

equal to zero. The zeros of $G(\lambda)$ are the eigenvalues, given by

$$\lambda = -\frac{1}{2}(r - \beta) \pm \frac{1}{2}|r + \beta| = -r, \beta, \quad (5.52)$$

and the result follows. \square

Proposition 5.9 (Equilibrium point $(0, 1)$ when $\beta_S - \beta_F \neq 0$).

Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels, and history functions which satisfy (5.4)–(5.10). For all $\beta_S, \beta_F \in \mathbb{R}$, $(0, 1)$ is locally stable when $\beta_S - \beta_F > 0$ and locally unstable when $\beta_S - \beta_F < 0$.

Proof. For $(S^*, F^*) = (0, 1)$, Equations (5.19) and (5.20) become

$$\frac{dx(t)}{dt} = 2(1 - \alpha_F)(-x(t) - y(t)) - \beta x(t), \quad (5.53)$$

$$\frac{dy(t)}{dt} = (2\alpha_F - 1)(-x(t) - y(t)) + \beta x(t). \quad (5.54)$$

Equations (5.53) and (5.54) have a solution of the form in Equation (5.23), so substitution gives

$$e^{\lambda t} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -2(1 - \alpha_F) - \beta & -2(1 - \alpha_F) \\ -(2\alpha_F - 1) + \beta & -(2\alpha_F - 1) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\lambda t}. \quad (5.55)$$

To ensure that $(c_1, c_2)^\top \neq 0$ we must have the characteristic equation

$$G(\lambda) = \begin{vmatrix} -2(1 - \alpha_F) - \beta - \lambda & -2(1 - \alpha_F) \\ -(2\alpha_F - 1) + \beta & -(2\alpha_F - 1) - \lambda \end{vmatrix} = \lambda^2 + \lambda(1 + \beta) + \beta \quad (5.56)$$

equal to zero. The zeros of $G(\lambda)$ are the eigenvalues, given by

$$\lambda = -\frac{1}{2}(1 + \beta) \pm \frac{1}{2}|1 - \beta| = -1, -\beta, \quad (5.57)$$

and the result follows. \square

Proposition 5.10 (Equilibrium points $(u, 1 - u)$ for $u \in [0, 1]$ when $\beta_S = \beta_F$).

Consider the system of delay differential equations (5.2) and (5.3) with parameters, delay kernels,

and history functions which satisfy (5.4)–(5.10). For all $\beta_S, \beta_F \in \mathbb{R}$ such that $\beta_S = \beta_F$ and for all $u \in [0, 1]$ the equilibrium point $(u, 1 - u)$ is locally stable.

Proof. For $(S^*, F^*) = (u, 1 - u)$ with $u \in \mathbb{R}$ and $\beta = 0$, Equations (5.19) and (5.20) become

$$\frac{dx(t)}{dt} = ((2\alpha_S - 1)ru + 2(1 - \alpha_F)(1 - u))(-x(t) - y(t)), \quad (5.58)$$

$$\frac{dy(t)}{dt} = (2(1 - \alpha_S)ru + (2\alpha_F - 1)(1 - u))(-x(t) - y(t)). \quad (5.59)$$

Equations (5.58) and (5.59) have a solution of the form in Equation (5.23), so substitution gives

$$e^{\lambda t} \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} -(2\alpha_S - 1)ru & -(2\alpha_S - 1)ru \\ -2(1 - \alpha_F)(1 - u) & -2(1 - \alpha_F)(1 - u) \\ -2(1 - \alpha_S)ru & -2(1 - \alpha_S)ru \\ -(2\alpha_F - 1)(1 - u) & -(2\alpha_F - 1)(1 - u) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} e^{\lambda t}. \quad (5.60)$$

To ensure that $(c_1, c_2)^T \neq 0$ we must have the characteristic equation

$$G(\lambda) = \begin{vmatrix} -(2\alpha_S - 1)ru - \lambda & -(2\alpha_S - 1)ru \\ -2(1 - \alpha_F)(1 - u) & -2(1 - \alpha_F)(1 - u) \\ -2(1 - \alpha_S)ru & -2(1 - \alpha_S)ru - \lambda \\ -(2\alpha_F - 1)(1 - u) & -(2\alpha_F - 1)(1 - u) \end{vmatrix} = \lambda^2 - \lambda(u(1 - r) - 1) \quad (5.61)$$

equal to zero. The zeros of $G(\lambda)$ are the eigenvalues, given by

$$\lambda = 0, u(1 - r) - 1. \quad (5.62)$$

Since $r \in (0, 1]$, for $u \in [0, 1]$ we have $u(1 - r) - 1 < 0$, therefore $(u, 1 - u)$ is locally stable. \square

5.5 Supporting numerical simulations

We obtain numerical solutions of Equations (5.2) and (5.3) using the forward Euler method, for which the temporal domain, $[0, 1000]$ h, is uniformly discretised with a time step of duration $\Delta t = 0.1$ h. To approximate the distributed delays we use the trapezoidal rule with uniform discretisation of the integration interval into 500 subintervals. The distributed delays depend on past values of the functions $S(t)$ and $F(t)$, which are obtained by interpolating between the previously estimated values for $S(t)$ and $F(t)$. The interpolation is achieved using piecewise

cubic Hermite interpolating polynomials, which are shape preserving. The sizes of the time step and the integration subintervals ensure grid-independence for our results. Examples of the simulations are shown in Figure 5.3.

The delay kernels in our model are set as probability density functions of a right-truncated Erlang distribution (Electronic Supplementary Material), shown in Figure 5.3(a). For slow-proliferating cells the Erlang density has shape $k = 12000$ and rate $\lambda = 20 \text{ h}^{-1}$ with mean 600 h, and is truncated at $U_S = 700 \text{ h}$. For fast-proliferating cells the Erlang density has shape $k = 20$ and rate $\lambda = 1 \text{ h}^{-1}$ with mean 20 h, and is truncated at $U_F = 100 \text{ h}$. All simulations use the parameters $K = 500$ and $r_S = r_F = 0.1 \text{ h}^{-1}$. The parameters α_S , α_F , β_S , and β_F are varied for the different simulations, as indicated in Figure 5.3(b)–(h).

There are many options for the functional form of the history functions. One simple option is to use constant functions, however it is reasonable to assume that the cells grew exponentially in the past, so we use exponential functions with growth rates equal to the intrinsic growth rates of the slow- and fast-proliferating cells. The history functions are $\phi_S(t) = 100e^{r_S t}$ for $t \in [-700, 0]$ and $\phi_F(t) = 100e^{r_F t}$ for $t \in [-100, 0]$, for Figure 5.3(b)–(h), $\phi_S(t) = 10^{-4}e^{r_S t}$ for $t \in [-700, 0]$ and $\phi_F(t) = 100e^{r_F t}$ for $t \in [-100, 0]$, for Figure 5.3(i), and $\phi_S(t) = 100e^{r_S t}$ for $t \in [-700, 0]$ and $\phi_F(t) = 10^{-4}e^{r_F t}$ for $t \in [-100, 0]$, for Figure 5.3(j). Since the state space is the function space C in Equation (5.11), choosing a different history function in C results in a different solution. When $\beta_S \neq \beta_F$, different history functions may produce different transient dynamics, whereas the solutions have the same long-term behaviour. When $\beta_S = \beta_F$, however, there are infinitely many equilibrium points so different history functions can produce solutions with different long-term behaviour.

Figure 5.3(b)–(d) shows simulations with $\beta_S = \beta_F = 0 \text{ h}^{-1}$, so no induced switching between slow and fast proliferation. By Theorem 5.5 there are infinitely many locally-stable equilibrium points corresponding to the line segment between $(K, 0)$ and $(0, K)$. The different equilibrium states are obtained by varying the levels of asymmetric division through α_S and α_F , or using different history functions.

In Figure 5.3(e)–(f) we show simulations with $\beta_S = 0 \text{ h}^{-1}$ and $\beta_F = 0.008 \text{ h}^{-1}$, therefore induced switching only from fast to slow proliferation. By Theorem 5.5 the equilibrium point $(K, 0)$ is locally stable and the equilibrium point $(0, K)$ is locally unstable. In Figure 5.3(g)–(h) we show simulations with $\beta_S = 0.008 \text{ h}^{-1}$ and $\beta_F = 0 \text{ h}^{-1}$, so induced switching only from slow to fast proliferation. By Theorem 5.5 the equilibrium point $(0, K)$ is locally stable and the equilibrium point $(K, 0)$ is locally unstable. These simulations illustrate that induced switching determines the long-term behaviour of the solutions, while asymmetric division only influences

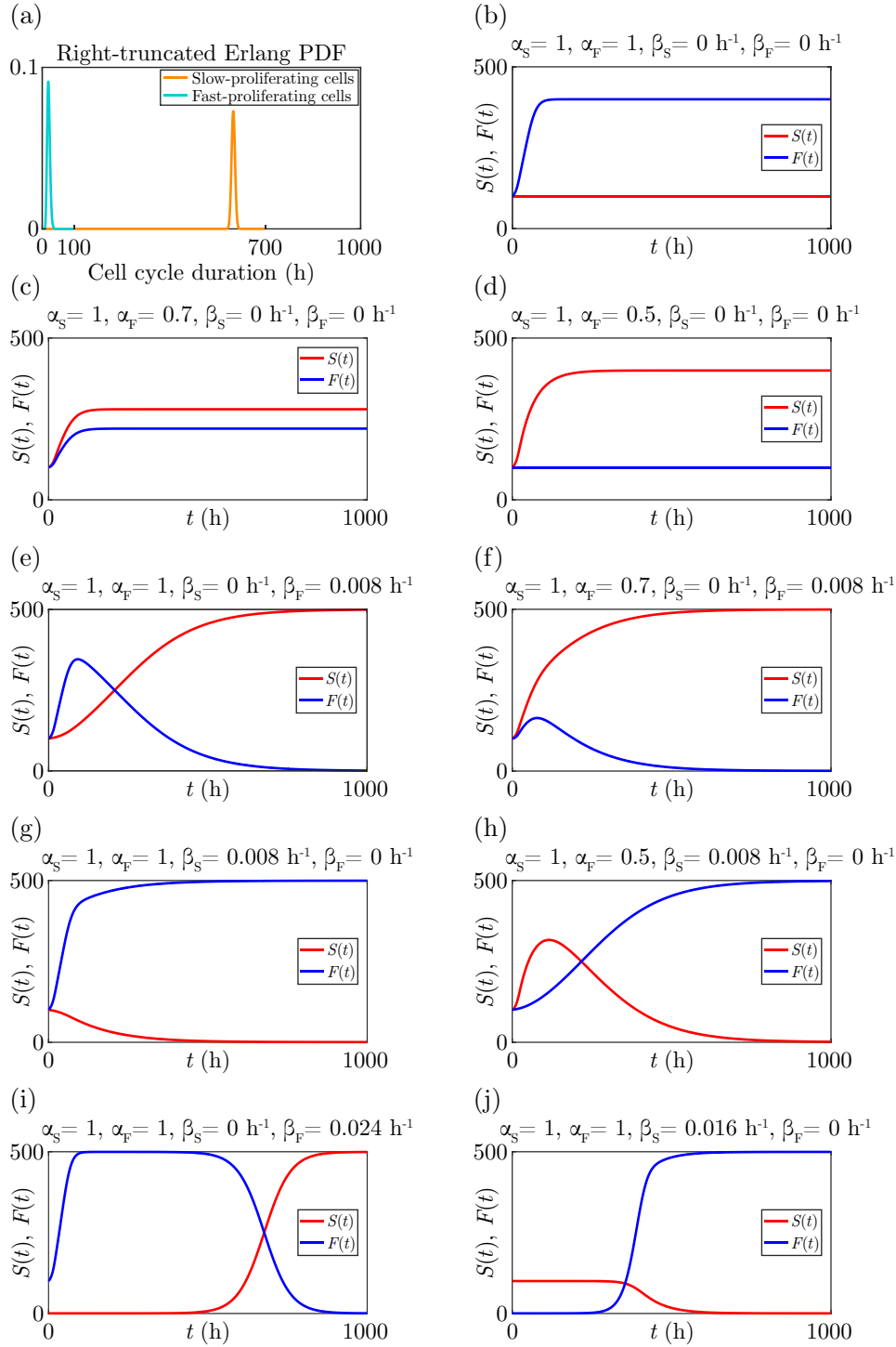


Figure 5.3: Numerical simulations of our model in Equations (5.2) and (5.3). (a) Each delay kernel is the probability density function (PDF) of a right-truncated Erlang distribution (Electronic Supplementary Material). For slow-proliferating cells the Erlang density function has shape $k = 12000$ and rate $\lambda = 20 \text{ h}^{-1}$ with mean 600 h, and is truncated at $U_S = 700 \text{ h}$. For fast-proliferating cells the Erlang density function has shape $k = 20$ and rate $\lambda = 1 \text{ h}^{-1}$ with mean 20 h, and is truncated at $U_F = 100 \text{ h}$. (b)–(j) The simulations all use the parameters $K = 500$ and $r_S = r_F = 0.1 \text{ h}^{-1}$. For (b)–(h) the history functions are $\phi_S(t) = 100e^{r_S t}$ and $\phi_F(t) = 100e^{r_F t}$, for (i) the history functions are $\phi_S(t) = 10^{-4}e^{r_S t}$ and $\phi_F(t) = 100e^{r_F t}$, and for (j) the history functions are $\phi_S(t) = 100e^{r_S t}$ and $\phi_F(t) = 10^{-4}e^{r_F t}$. Parameters specific to each simulation, namely α_S , α_F , β_S , and β_F , are indicated on the figure

the transient dynamics.

In Figure 5.3(i)–(j) we set one of the history functions ϕ_S or ϕ_F close to zero over its domain, illustrating how a very small subpopulation can become the main subpopulation through induced switching, possibly requiring a long time period. It is particularly interesting that, in Figure 5.3(i), the density of the fast-proliferating cells is very close to carrying capacity and appears to be at equilibrium for a long time, however through induced switching the slow-proliferating cells eventually become the main subpopulation.

5.6 Discussion and outlook

Proliferative heterogeneity in cancer cell populations constitutes a crucial challenge for cancer therapy, as slow-proliferating cells tend to be highly aggressive, have increased resistance to cytotoxic drugs, and can replenish the fast-proliferating subpopulation [27, 29, 66, 157]. The dynamics underlying tumour heterogeneity are not well understood, so improving cancer therapy depends on furthering this understanding [13, 66, 177]. Theoretical approaches are well-placed to assist in elucidating the transient dynamics of intratumoural heterogeneity.

In this article we present a delay differential equation model for heterogeneous cell proliferation in which the total population consists of a subpopulation of slow-proliferating cells and a subpopulation of fast-proliferating cells. Our model incorporates the two cellular processes of asymmetric cell division and induced switching between proliferative states, which are important contributors to the dynamic heterogeneity of a cancer cell population [159, 160, 163, 164]. The model is designed for investigating the transient dynamics of intratumoural heterogeneity with respect to cell proliferation. We employ delay differential equations in our model rather than ordinary differential equations in order to obtain transient dynamics consistent with the dynamics in a tumour. While the equilibrium states for our model are the same as those for the corresponding ordinary differential equations, the transient dynamics are very different, and model parameterisation with biologically-realistic values requires a model that incorporates realistic cell cycle durations for the slow- and fast-proliferating subpopulations.

Because the transient dynamics of a tumour are of primary interest, and the local stability analysis of our model provides only long-term behaviour near to the equilibrium, we must numerically simulate our model to explore the transient dynamics. We provide some examples of numerical simulations in Figure 5.3, where we specify the delay kernels to be right-truncated Erlang distributions (Section 2 of Electronic Supplementary Material), and vary the parameters to demonstrate some of the possible dynamics within a tumour cell population. To exemplify some of the experimental scenarios to which our model is applicable, we consider a tumour that

is treated with a cytotoxic drug which may induce cellular stress, causing the fast-proliferating cells to switch to the drug-resistant slow-proliferating phenotype [66] through the mechanisms of asymmetric cell division and induced switching.

We show simulations where there is no induced switching between the slow- and fast-proliferating subpopulations in Figure 5.3(b)–(d). If tumour cells experience no microenvironmental stress, then all cell divisions may be symmetric, corresponding to the situation in Figure 5.3(b) where the fast-proliferating cells rapidly populate the tumour until the total cell density reaches carrying capacity. If a drug is introduced into the tumour microenvironment then the fast-proliferating cells may experience cellular stress, inducing the fast-proliferating cells into asymmetric cell division as a survival strategy, as the slow-proliferating phenotype is drug resistant. Figure 5.3(c)–(d) illustrates this behaviour, first for an intermediate level of asymmetric division in (c), and then for the maximum level of asymmetric division in (d).

Figure 5.3(e) shows a simulation where fast-proliferating cells are under stress due to the presence of a drug, and are induced to switch to slow proliferation through signals from slow-proliferating cells as a survival strategy. Alternatively, the induced signalling may arise from the highly invasive slow-proliferating cells [66,178] influencing the less invasive fast-proliferating cells to switch to the more invasive slow-proliferating phenotype. In our simulation the total cell population appears to reach the carrying capacity density at around 200 hours, however the dynamics of induced switching of cells from fast to slow proliferation continues until all cells are slow proliferating. This is important because, while tumour growth has effectively ceased, the tumour is becoming increasingly drug resistant and invasive over time until the whole tumour is composed of drug resistant and invasive cells. Therefore, effective early treatment of the tumour is required in order to prevent the tumour from becoming more aggressive and treatment resistant. If the per capita interaction strength of slow-proliferating cells to induce fast-proliferating cells to switch to slow proliferation is obtained experimentally, then our model could be used to predict the transient change in the proportion of slow-proliferating cells in the population, and therefore the changes in invasiveness and drug resistance of the tumour.

Now consider a tumour composed of mostly fast-proliferating cells and a very small proportion of slow-proliferating cells, as in Figure 5.3(i). The fast-proliferating cells undergo induced switching to slow proliferation, perhaps due to stress from an introduced drug or to increase invasiveness. Initially the fast-proliferating subpopulation grows to reach near the carry capacity density, and the system appears to be in equilibrium for an extended period of time. As the tumour is almost completely composed of fast-proliferating cells, it is in the least invasive and most drug sensitive state. Without knowledge of the presence of induced switching, ex-

perimental investigations may not reveal that the tumour is in the process of becoming highly invasive and drug resistant. Indeed, once the density of the slow-proliferating cells has reached a sufficient but still very low level, the tumour rapidly becomes populated by slow-proliferating cells through induced switching of the fast-proliferating cells.

Finally, consider a small tumour comprised mostly of slow-proliferating cells and a very small proportion of fast-proliferating cells, as in Figure 5.3(j). Signals from the fast-proliferating cells induce the slow-proliferating cells to switch to fast proliferation. Experimentally, this could correspond to a tumour that has been treated with a drug which caused the death of most of the fast-proliferating cells. For an extended period of time the tumour grows very little, until the density of fast-proliferating cells is high enough that the induced switching from slow- to fast-proliferation rapidly grows the tumour to the maximum sustainable size. Our model is therefore able to provide an estimate of tumour growth over time following drug treatment, when the cells can undergo induced switching.

There are numerous possibilities for future work. Induced switching between proliferative states could take many forms. In tumours the slow-proliferating state may continually arise and disappear [28], so it would be interesting to accommodate time-dependent induced switching into the model, which could be either periodic or aperiodic. We could also consider the induced switching to have an explicit dependence on density, so that no switching occurs from a particular proliferative state when the density of cells from the other proliferative state is above a certain value. A similar explicit density dependence could be implemented for asymmetric cell division, which occurs at constant proportions in our current model. These explicit dependences on density may be relevant for slow-proliferating subpopulations in tumours that appear to persist over time and maintain the relative size of the subpopulation [27, 29]. Our model could also be extended to include the additional process of spontaneous switching between proliferative states, which is independent of other cells and may be stochastic.

While our model has implicit spatial structure, since the dependent variables are cell densities, we could include spatial structure explicitly. We could then explicitly model cell migration with a diffusive term [85]. Further, induced switching could be modelled as a more localised process where the rate of a cell switching proliferative states is determined by the density of cells in a different proliferative state within a given radius of the cell, where the interaction strength decreases with distance from the cell.

We could also extend our model to more than two dependent variables. For example, we could consider fast-, slow-, and very-slow-proliferating subpopulations. Another possible extension is to include apoptosis. Much of our analysis in this article is likely to be easily generalised to

an extended version of our model. The more challenging aspect could be the analysis of the corresponding transcendental characteristic equations, however taking a more abstract approach for an extended model with an arbitrary n dependent variables could simplify the problem.

Code availability

The code for the algorithm to replicate the numerical simulations in this work is available on GitHub at <https://github.com/DrSeanTVittadello/Vittadello2020>.

5.7 Supplementary Material

5.7.1 Experimental

We briefly describe the materials and methods employed to obtain our experimental data for the durations of cells in G1 and S/G2/M phases, shown in Figure 5.4. Further experimental details are given in [146]. Each histogram is constructed using data from 50 individual cells.

Our experimental data consist of microscopy time-series images of two-dimensional cell proliferation assays using the three human melanoma cell lines C8161 (kindly provided by Mary Hendrix, Chicago, IL, USA), WM983C and 1205Lu (both kindly provided by Meenhard Herlyn, Philadelphia, PA, USA), which have cell cycle durations of approximately 21, 23 and 37 h, respectively [13]. The cell lines were genotypically characterised [101–104], grown as described in [106], and authenticated by STR fingerprinting (QIMR Berghofer Medical Research Institute, Herston, Australia).

We maintain the cell cultures to prevent any induced synchronisation from cell cycle arrest in G1 phase, by passaging the cells every three days, and on the day prior to setting up an experiment, to maintain a subconfluent cell density and a fresh growth medium.

Experimental investigation of the progression of the cell cycle is visually enabled with fluorescent ubiquitination-based cell cycle indicator (FUCCI) technology [30]. FUCCI consists of two genetically-encoded reporters that enable visualisation of the cell cycle of individual live cells: when the cell is in G1 phase the nucleus fluoresces red, and when the cell is in S/G2/M phase the nucleus fluoresces green (Figure 5.1 in the main document). During the transition from G1 to S phase, called early S, both reporters fluoresce and the nucleus appears yellow. FUCCI is utilised in experimental studies of the cycling dynamics of cells in tumours [13], and reveals the differential cycling of the cell population.

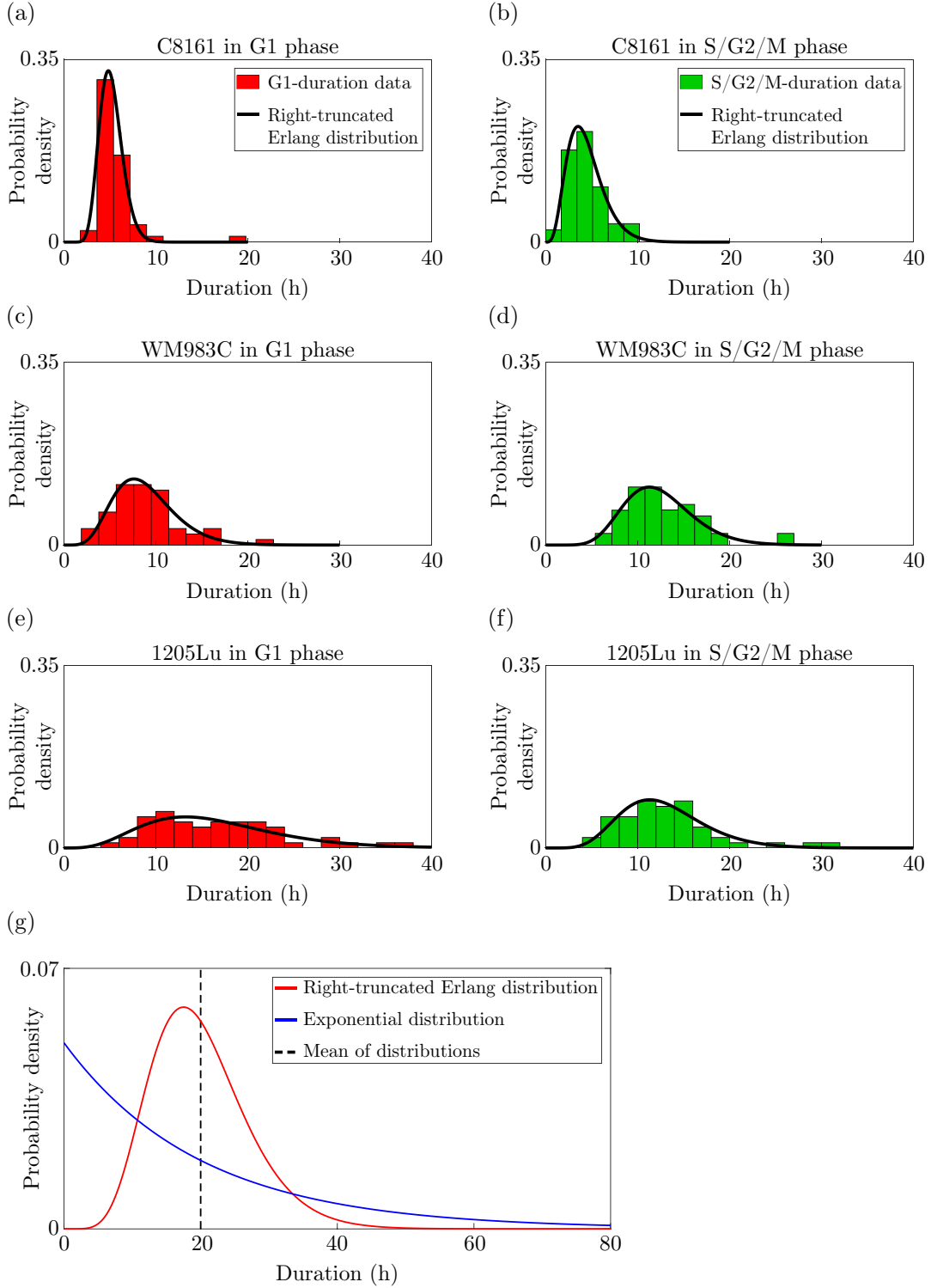


Figure 5.4: Histograms of G1- and S/G2/M-phase duration data with best fits of the right-truncated Erlang distribution in Equation (5.66), and a comparison of the probability density functions of the right-truncated Erlang and exponential distributions. Each histogram corresponds to 50 cells. The best-fit parameters for the right-truncated Erlang distribution are: (a) RTE(17, 3.3 h⁻¹, 20 h); (b) RTE(5, 1.1 h⁻¹, 20 h); (c) RTE(7, 0.79 h⁻¹, 30 h); (d) RTE(11, 0.89 h⁻¹, 30 h); (e) RTE(5, 0.30 h⁻¹, 40 h); (f) RTE(8, 0.62 h⁻¹, 40 h). (g) Probability density functions of the right-truncated Erlang and exponential distributions, for RTE(8, 0.4 h⁻¹, 80 h) and Exp(0.05 h⁻¹). The exponential distribution has mean 20 h, and the right-truncated Erlang distribution has mean 20.0000 h (to four decimal places)

5.7.2 Experimental data and models for cell cycle durations

Standard deterministic mathematical models of cell proliferation, such as exponential (5.63) and logistic (5.64) growth models,

$$\frac{dP(t)}{dt} = rP(t), \quad (5.63)$$

$$\frac{dP(t)}{dt} = rP(t) \left(1 - \frac{P(t)}{K} \right), \quad (5.64)$$

where $P(t)$ is the population density at time t , r is the intrinsic growth rate, and K is the carrying capacity density, are based on cell cycle durations with an exponential distribution $\text{Exp}(\theta)$ for rate θ . Realising Equations (5.63) and (5.64) as stochastic pure birth processes yields continuous-time homogeneous Markov chains with exponentially-distributed durations of cell cycle residence [179]. Experimental investigations, however, suggest that the duration of the cell cycle, and in particular each cell cycle phase, is not exponentially distributed [76,77,82,83,86]. Rather, it is often found that the hypoexponential distribution, characterised as the sum of k independent exponential random variables with distinct rate parameters λ_i , for $i = 1, \dots, k$, is a reasonable distribution for cell cycle duration [76,77,82,83,86]. The hypoexponential distribution generalises the Erlang distribution for which the exponential random variables have the same rate parameter λ . Since the exponential distribution allows for a relatively large probability of arbitrarily small cell cycle durations, ordinary differential equations such as (5.63) and (5.64) tend to overestimate the population growth rate, particularly for slow-proliferating cells.

We consider only bounded distributions for their greater biological realism. A particular example of a bounded delay kernel that may be relevant when modelling the cell cycle is the probability density for the right-truncated Erlang distribution. The probability density for the Erlang distribution is

$$g(z) = \frac{\lambda^k z^{k-1} e^{-\lambda z}}{(k-1)!}, \text{ for } z \in [0, \infty), \quad (5.65)$$

where k is the shape parameter, λ is the rate parameter, and the mean is k/λ . Restricting the Erlang distribution to the bounded interval $[0, U]$, where $U \in (0, \infty)$, gives the right-truncated Erlang distribution which has probability density

$$g_*(z) = \frac{g(z)}{\int_0^U g(w) dw}, \text{ for } z \in [0, U]. \quad (5.66)$$

We also show in Figure 5.4 the best fit to the data for the right-truncated Erlang distribution

$\text{RTE}(k, \lambda, U)$, where the shape k and rate λ correspond to the Erlang distribution, and U is the right-truncation point of the Erlang distribution. Each best fit is obtained with the MATLAB nonlinear least-squares solver `lsqnonlin` [122] with the trust-region-reflective algorithm [123]. These data demonstrate that the durations of G1 and S/G2/M, and therefore of the complete cell cycle, are well approximated by a right-truncated Erlang distribution. In Figure 5.4(g) we compare the probability density functions of the exponential and right-truncated Erlang distributions, which clearly demonstrates that these cell cycle durations are not exponentially distributed. Using the `lsqnonlin` solver we find that for each data set the norm of the residual from fitting the right-truncated Erlang distribution is no greater than the norm of the residual from fitting the Erlang distribution. This outcome is expected as the cumulative density of the Erlang distribution tends to one exponentially after the maximum data value. We could also left-truncate the Erlang distribution to ensure a positive lower bound for the distribution, depending on whether a better fit would be achieved with particular data.

5.7.3 Complete proof of Theorem 5.2

It suffices to prove that $S(t) \geq 0$ and $F(t) \geq 0$ for all $t > 0$, for then it follows from Lemma 5.3 that $S(t) \leq K$ and $F(t) \leq K$ for all $t > 0$. Define t_1 and t_2 by $t_1 = \inf\{t > 0 \mid S(t) < 0\}$ and $t_2 = \inf\{t > 0 \mid F(t) < 0\}$. We consider the infima in the extended real numbers, so that $t_1, t_2 \in (0, +\infty]$, where either infimum is equal to $+\infty$ if the corresponding set is empty. The proof consists of four separate cases.

Case 1: Let $\beta_S - \beta_F \geq 0$ and suppose $S(t) < 0$ for some $t > 0$.

Note that $t_1 \in \mathbb{R}$. If $t_1 < t_2$ then choose $t_3 \in (t_1, t_2)$ such that $S(t_3) < 0$, $dS(t)/dt|_{t=t_3} < 0$, and the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3 and since $F(t_3) \geq 0$, Equation (5.2) gives $dS(t)/dt|_{t=t_3} \geq 0$, a contradiction.

If $t_1 \geq t_2$ then, since $F(0) > 0$ and $F(t_2) = 0$, there exists $t_3 \in (0, t_2)$ such that $dF(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.3) gives $dF(t)/dt|_{t=t_3} \geq 0$, a contradiction. We conclude that $S(t) \geq 0$ for all $t > 0$.

Case 2: Let $\beta_S - \beta_F < 0$ and suppose $S(t) < 0$ for some $t > 0$.

Note that $t_1 \in \mathbb{R}$. If $t_1 \leq t_2$ then, since $S(0) > 0$ and $S(t_1) = 0$, there exists $t_3 \in (0, t_1)$ such that $dS(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.2) gives $dS(t)/dt|_{t=t_3} \geq 0$, a contradiction.

Suppose now that $t_1 > t_2$, and choose $t_3 \in (t_2, t_1)$ such that $F(t_3) < 0$, $dF(t)/dt|_{t=t_3} < 0$, and the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, and since $S(t_3) \geq 0$, Equation (5.2) gives $dF(t)/dt|_{t=t_3} \geq 0$, a contradiction.

Suppose now that $t_1 = t_2$. Since $S(0) > 0$ and $S(t_1) = 0$, there exists $t_3 \in (0, t_1)$ such that $dS(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.3) gives $dS(t)/dt|_{t=t_3} \geq 0$, a contradiction. We conclude that $S(t) \geq 0$ for all $t > 0$.

Case 3: Let $\beta_S - \beta_F \geq 0$ and suppose $F(t) < 0$ for some $t > 0$.

Note that $t_2 \in \mathbb{R}$. If $t_2 < t_1$ then, since $F(0) > 0$ and $F(t_2) = 0$, there exists $t_3 \in (0, t_2)$ such that $dF(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.3) gives $dF(t)/dt|_{t=t_3} \geq 0$, a contradiction.

Suppose now that $t_2 > t_1$, and choose $t_3 \in (t_1, t_2)$ such that $S(t_3) < 0$, $dS(t)/dt|_{t=t_3} < 0$, and the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, and since $F(t_3) \geq 0$, Equation (5.2) gives $dS(t)/dt|_{t=t_3} \geq 0$, a contradiction.

Suppose now that $t_1 = t_2$. Since $F(0) > 0$ and $F(t_2) = 0$, there exists $t_3 \in (0, t_2)$ such that

$dF(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.3) gives $dF(t)/dt|_{t=t_3} \geq 0$, a contradiction. We conclude that $F(t) \geq 0$ for all $t > 0$.

Case 4: Let $\beta_S - \beta_F < 0$ and suppose $F(t) < 0$ for some $t > 0$.

Note that $t_2 \in \mathbb{R}$. If $t_2 < t_1$ then choose $t_3 \in (t_2, t_1)$ such that $F(t_3) < 0$, $dF(t)/dt|_{t=t_3} < 0$, and the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3 and since $S(t_3) \geq 0$, Equation (5.3) gives $dF(t)/dt|_{t=t_3} \geq 0$, a contradiction.

Suppose now that $t_2 \geq t_1$. Since $S(0) > 0$ and $S(t_1) = 0$ there exists $t_3 \in (0, t_1)$ such that $dS(t)/dt|_{t=t_3} < 0$. Then, since $P(t_3) \leq K$ by Lemma 5.3, since the delays satisfy $\bar{S}(t_3) \geq 0$ and $\bar{F}(t_3) \geq 0$, and since $S(t_3), F(t_3) \geq 0$, Equation (5.2) gives $dS(t)/dt|_{t=t_3} \geq 0$, a contradiction.

We conclude that $F(t) \geq 0$ for all $t > 0$. □

5.7.4 Complete proof of Theorem 5.4

We first show that f defined in Equation (5.14) satisfies the following Lipschitz condition on every bounded subset of C : for all $M > 0$ there exists $L > 0$ such that for $\rho, \psi \in C([- \widehat{U}, 0], \mathbb{R}^2)$ with $\|\rho\|, \|\psi\| \leq M$ we have $\|f(\rho) - f(\psi)\|_2 \leq L\|\rho - \psi\|$.

To further simplify the notation in Equation (5.14) we define $\kappa_1 = (2\alpha_S - 1)r_S$, $\kappa_2 = 2(1 - \alpha_F)r_F$, $\kappa_3 = 2(1 - \alpha_S)r_S$, $\kappa_4 = (2\alpha_F - 1)r_F$, and $\kappa_5 = (\beta_S - \beta_F)/K$. Now,

$$\begin{aligned}
 f(\rho) - f(\psi) &= \begin{bmatrix} \left(\kappa_1 \overline{\rho_S} + \kappa_2 \overline{\rho_F} \right) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) - \kappa_5 \rho_S(0) \rho_F(0) \\ - \left(\kappa_1 \overline{\psi_S} + \kappa_2 \overline{\psi_F} \right) \left(1 - \frac{(\psi_S + \psi_F)(0)}{K} \right) + \kappa_5 \psi_S(0) \psi_F(0) \\ \left(\kappa_3 \overline{\rho_S} + \kappa_4 \overline{\rho_F} \right) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) + \kappa_5 \rho_S(0) \rho_F(0) \\ - \left(\kappa_3 \overline{\psi_S} + \kappa_4 \overline{\psi_F} \right) \left(1 - \frac{(\psi_S + \psi_F)(0)}{K} \right) - \kappa_5 \psi_S(0) \psi_F(0) \end{bmatrix} \\
 &= \begin{bmatrix} \left(\kappa_1 (\overline{\rho_S} - \overline{\psi_S}) + \kappa_2 (\overline{\rho_F} - \overline{\psi_F}) \right) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \\ + \left(\kappa_1 \overline{\psi_S} + \kappa_2 \overline{\psi_F} \right) \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K} \right) \\ + \kappa_5 (\psi_S - \rho_S)(0) \psi_F(0) + \kappa_5 (\psi_F - \rho_F)(0) \rho_S(0) \\ \left(\kappa_3 (\overline{\rho_S} - \overline{\psi_S}) + \kappa_4 (\overline{\rho_F} - \overline{\psi_F}) \right) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \\ + \left(\kappa_3 \overline{\psi_S} + \kappa_4 \overline{\psi_F} \right) \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K} \right) \\ + \kappa_5 (\rho_S - \psi_S)(0) \psi_F(0) + \kappa_5 (\rho_F - \psi_F)(0) \rho_S(0) \end{bmatrix}
 \end{aligned}$$

so, using the triangle inequality, we obtain

$$\|f(\rho) - f(\psi)\|_2 \leq \left\| \begin{bmatrix} \kappa_1 (\overline{\rho_S} - \overline{\psi_S}) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \\ \kappa_3 (\overline{\rho_S} - \overline{\psi_S}) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \end{bmatrix} \right\|_2$$

$$\begin{aligned}
& + \left\| \begin{bmatrix} \kappa_2(\overline{\rho_F} - \overline{\psi_F}) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \\ \kappa_4(\overline{\rho_F} - \overline{\psi_F}) \left(1 - \frac{(\rho_S + \rho_F)(0)}{K} \right) \end{bmatrix} \right\|_2 \\
& + \left\| \begin{bmatrix} \kappa_1 \overline{\psi_S} \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K} \right) \\ \kappa_3 \overline{\psi_S} \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K} \right) \end{bmatrix} \right\|_2 \\
& + \left\| \begin{bmatrix} \kappa_2 \overline{\psi_F} \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K} \right) \\ \kappa_4 \overline{\psi_F} \left(\frac{(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)}{K} \right) \end{bmatrix} \right\|_2 \\
& + \left\| \begin{bmatrix} \kappa_5(\psi_S - \rho_S)(0) \psi_F(0) \\ \kappa_5(\rho_S - \psi_S)(0) \psi_F(0) \end{bmatrix} \right\|_2 + \left\| \begin{bmatrix} \kappa_5(\psi_F - \rho_F)(0) \rho_S(0) \\ \kappa_5(\rho_F - \psi_F)(0) \rho_S(0) \end{bmatrix} \right\|_2 \\
& \leq \sqrt{\kappa_1^2 + \kappa_3^2} \left(1 + \frac{1}{K} |\rho_S(0)| + \frac{1}{K} |\rho_F(0)| \right) |(\overline{\rho_S} - \overline{\psi_S})| \\
& + \sqrt{\kappa_2^2 + \kappa_4^2} \left(1 + \frac{1}{K} |\rho_S(0)| + \frac{1}{K} |\rho_F(0)| \right) |(\overline{\rho_F} - \overline{\psi_F})| \\
& + \frac{1}{K} \sqrt{\kappa_1^2 + \kappa_3^2} |\overline{\psi_S}| |(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)| \\
& + \frac{1}{K} \sqrt{\kappa_2^2 + \kappa_4^2} |\overline{\psi_F}| |(\psi_S - \rho_S)(0) + (\psi_F - \rho_F)(0)| \\
& + \sqrt{2} |\kappa_5| |(\rho_S - \psi_S)(0)| |\psi_F(0)| + \sqrt{2} |\kappa_5| |(\rho_F - \psi_F)(0)| |\rho_S(0)| \\
& \leq \sqrt{\kappa_1^2 + \kappa_3^2} \left(1 + \frac{1}{K} \|\rho_S\| + \frac{1}{K} \|\rho_F\| \right) \|\rho_S - \psi_S\| \\
& + \sqrt{\kappa_2^2 + \kappa_4^2} \left(1 + \frac{1}{K} \|\rho_S\| + \frac{1}{K} \|\rho_F\| \right) \|\rho_F - \psi_F\| \\
& + \frac{1}{K} \sqrt{\kappa_1^2 + \kappa_3^2} \|\psi_S\| (\|\psi_S - \rho_S\| + \|\psi_F - \rho_F\|) \\
& + \frac{1}{K} \sqrt{\kappa_2^2 + \kappa_4^2} \|\psi_F\| (\|\psi_S - \rho_S\| + \|\psi_F - \rho_F\|)
\end{aligned}$$

$$\begin{aligned}
& + \sqrt{2}|\kappa_5| \|\rho_S - \psi_S\| \|\psi_F\| + \sqrt{2}|\kappa_5| \|\rho_F - \psi_F\| \|\rho_S\| \\
& \leq \left(\sqrt{\kappa_1^2 + \kappa_3^2} \left(1 + \frac{2M}{K}\right) + \sqrt{\kappa_2^2 + \kappa_4^2} \left(1 + \frac{2M}{K}\right) + \frac{2M}{K} \sqrt{\kappa_1^2 + \kappa_3^2} \right. \\
& \quad \left. + \frac{2M}{K} \sqrt{\kappa_2^2 + \kappa_4^2} + 2\sqrt{2}|\kappa_5|M \right) \|\rho - \psi\| \\
& = \left(\left(\sqrt{\kappa_1^2 + \kappa_3^2} + \sqrt{\kappa_2^2 + \kappa_4^2} \right) \left(1 + \frac{4M}{K}\right) + 2\sqrt{2}|\kappa_5|M \right) \|\rho - \psi\|,
\end{aligned}$$

so we can set L to be

$$L = \left(\sqrt{\kappa_1^2 + \kappa_3^2} + \sqrt{\kappa_2^2 + \kappa_4^2} \right) \left(1 + \frac{4M}{K}\right) + 2\sqrt{2}|\kappa_5|M$$

and then f satisfies the Lipschitz condition. Then [173, Page 32, Theorem 3.7] provides local existence and uniqueness of solutions for the system (5.2) and (5.3). Since our solutions of interest are bounded by Theorem 5.2, it follows from [173, Page 37, Proposition 3.10] that the solutions are continuable to all positive time. \square

5.7.5 Alternative proof of Proposition 5.6

The functions S and F are non-negative by Theorem 5.2, so the following inequalities show that the equilibrium point $(S^*, F^*) = (0, 0) \in C([- \widehat{U}, \infty), \mathbb{R}_{\geq 0}^2)$ for the system (5.2) and (5.3) is stable if and only if the equilibrium point $P^* = 0 \in C([- \widehat{U}, \infty), \mathbb{R}_{\geq 0})$ for Equation (5.12) is stable:

$$\|(S, F)\| \leq \|P\| \leq \sqrt{2} \|(S, F)\|. \quad (5.67)$$

The history function ϕ for the system (5.2) and (5.3) is strictly positive by Equation (5.9), and $\phi_S + \phi_F < K$ by Equation (5.10). It follows by continuity that $S(t) > 0$, $F(t) > 0$, and $S(t) + F(t) < K$ in a neighbourhood of $(0, 0)$. So, by Equation (5.12), $dP(t)/dt > 0$ in a neighbourhood of $P^* = 0$, which is therefore an unstable equilibrium point. We conclude that $(S^*, F^*) = (0, 0)$ is also an unstable equilibrium point by Equation (5.67). \square

5.7.6 Graphical illustration of Cauchy's argument principle

In Figure 5.5 we graphically illustrate our application of Cauchy's argument principle. We let Γ be the closed rectangular contour in Figure 5.5(a). Due to the integrals in the transcendental characteristic equation $G(\lambda)$ in Equation (5.25), a very large number of numerical integrations are required to calculate $G(\lambda)$ along a contour. So, we instead use the discrete delay version of $G(\lambda)$, denoted by $G_\delta(\lambda)$, which gives similar qualitative behaviour:

$$G_\delta(\lambda) = \lambda^2 - \lambda \left((2\alpha_S - 1)re^{-\lambda r_F \tau_S} + (2\alpha_F - 1)e^{-\lambda r_F \tau_F} \right) + (2\alpha_S + 2\alpha_F - 3)re^{-\lambda r_F (\tau_S + \tau_F)}, \quad (5.68)$$

obtained from $G(\lambda)$ with the Dirac kernels $g_S(z) = \delta(z - \tau_S)$ and $g_F(z) = \delta(z - \tau_F)$ for discrete delays τ_S and τ_F . Figures 5.5(b), (d), and (f) show the images $G_\delta(\Gamma)$ for three different sets of parameters for $G_\delta(\lambda)$, and Figures 5.5(c), (e), and (g) show the respective close-up views around the origin. Note that each coloured segment of $G_\delta(\Gamma)$ in (b)–(f) is the image of the same-coloured segment of the rectangular contour Γ in (a) under $G_\delta(\lambda)$. To calculate the winding number of $G_\delta(\Gamma)$ with respect to the origin we count the net number of times that $G_\delta(\Gamma)$ winds counter-clockwise around the origin, assigning +1 for each time $G_\delta(\Gamma)$ winds around the origin in a counter-clockwise direction, and -1 for each time $G_\delta(\Gamma)$ winds around the origin in a clockwise direction. Figure 5.5 illustrates that the behaviour of $G_\delta(\Gamma)$ can be complicated near the origin, so when calculating the winding number of $G_\delta(\Gamma)$ with respect to the origin we need to ensure that we also account for the possibility of $G_\delta(\Gamma)$ winding around the origin in a clockwise direction.

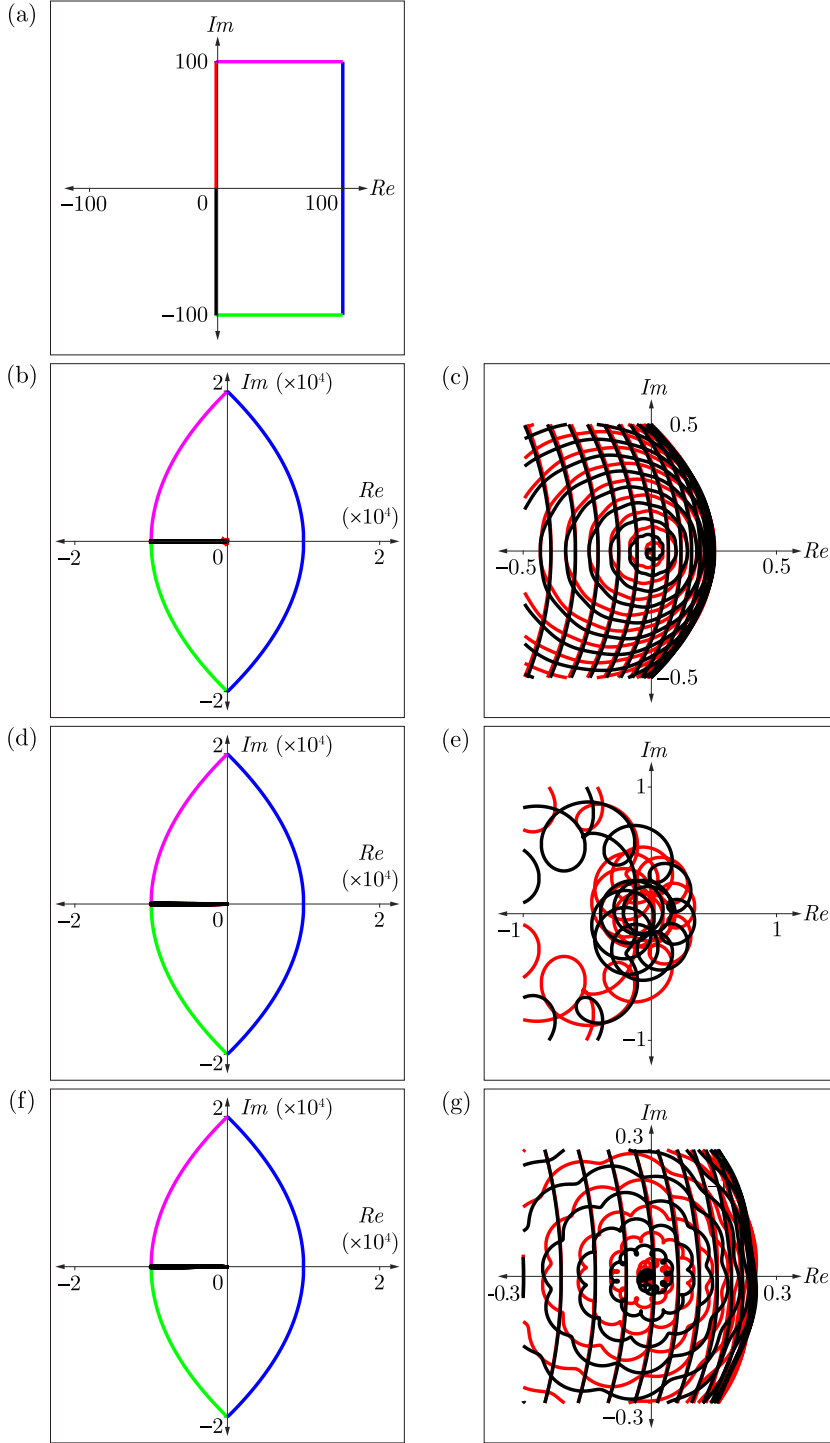


Figure 5.5: Graphical illustration of our application of Cauchy's argument principle for $G_\delta(\lambda)$ in Equation (5.68), when $2\alpha_S + 2\alpha_F - 3 > 0$. (a) A rectangular contour Γ in the right half-plane does not intersect the imaginary axis. (b) The image contour $G_\delta(\Gamma)$ when G_δ has the parameters $\alpha_S = 0.6$, $\alpha_F = 1$, $r_F = 1$, and $r = 0.01$, with a close-up view of the origin in (c). (d) The image contour $G_\delta(\gamma)$ when G_δ has the parameters $\alpha_S = 0.6$, $\alpha_F = 1$, $r_F = 0.1$, and $r = 1$, with a close-up view of the origin in (e). (f) The image contour $G_\delta(\Gamma)$ when G_δ has the parameters $\alpha_S = 1$, $\alpha_F = 1$, $r_F = 1$, and $r = 0.01$, with a close-up view of the origin in (g). Each coloured segment of $G_\delta(\Gamma)$ in (b)–(f) is the image of the same-coloured segment of the rectangular contour Γ in (a) under $G_\delta(\lambda)$

Conclusions and future work

6.1 Summary

In this thesis we employ experimental investigation, mathematical modelling, and mathematical analysis to explore how the detailed information about the cell cycle obtained with fluorescent cell cycle indicators can inform the development of novel mathematical models of cell migration and proliferation, address many key assumptions underlying mathematical models in the literature, and allow for more rigorous mathematical analysis of experimental observations.

The objectives of this thesis were to:

1. Incorporate cell cycle dynamics into a continuum model of cell migration based on the information provided by FUCCI technology in relation to the cell cycle phase.
2. Investigate whether cell proliferation assays prepared using standard cell-culture procedures are asynchronous, utilising FUCCI technology.
3. Rigorously examine the go-or-grow hypothesis using FUCCI technology and cell-cycle-inhibiting drugs.
4. Develop a continuum model of heterogeneous cell proliferation motivated by data from new proliferation experiments with FUCCI-expressing cells.

These four objectives are successfully addressed in the four articles presented in Chapters 2–5 of this thesis, which we summarise here.

In Chapter 2 we commence our study by developing a novel continuum mathematical model that can be used to quantitatively describe 2-D cell migration experiments employing FUCCI technology. We consider the total cell population as structured, consisting of distinct subpopulations of cells in a particular phase of the cell cycle. The fundamental model is a system of two coupled partial differential equations that describes the two subpopulations of cells in either G1 or S/G2/M phase, as directly labelled by FUCCI. Our model can describe the spatial and temporal distributions of the two subpopulations, and in particular the dynamics of scratch assays performed with cells highlighted by FUCCI, which we demonstrate by showing that the predictions of the model compare well with our new experimental data from scratch assay experiments for three cell lines with distinct cell-cycle characteristics.

Scratch assays are routinely employed to study cell migration, for example in the context of malignant invasion [107] and wound healing [46]. A typical scratch assay may require several days to grow the cells, set up the experiment, and acquire images of the cell population. Our model can simulate such an experiment on a single desktop computer in a few seconds. Further, we can easily vary the parameters in our model to simulate experiments over any period of time, any scratch width, any geometry, and any cell line provided that we know the durations of each cell cycle phase, which are typically available. Therefore, this computational modelling tool can provide valuable information to assist in the design and interpretation of these kinds of experiments conducted with FUCCI-expressing cells.

Under appropriate conditions, numerical solutions of our mathematical model demonstrate the existence of travelling wave solutions that are a combination of coupled wavefronts and pulses. Travelling wave solutions are of great practical interest as cell migration tends to exhibit travelling wave characteristics [46]. We derive an analytical expression for the minimum wave speed of the travelling waves, assuming that the diffusion coefficients are equal for cells in G1 and S/G2/M phases, which is biologically reasonable as we are considering subpopulations of the same cell type which differ only with respect to cell cycle phase. An implication of our work in Chapter 4 is the experimental justification of the equality of the diffusion coefficients for cells in G1 and S/G2/M phases. Our expression for the minimum wave speed depends on the cell cycle dynamics of the particular cells under consideration, namely the ratio of the duration of S/G2/M phase to the duration of G1 phase. The analytical expression we derive for the minimum wave speed is of practical interest, as a moving front of cells can be thought of as acting like a travelling wave solution, so our expression can provide a prediction for the speed of the moving front in experimental studies.

Following our work in Chapter 2 we started to investigate the go-or-grow hypothesis, per-

forming some preliminary proliferation experiments with three cell-cycle-inhibiting drugs to determine the minimum doses that induce arrest in each of the three FUCCI-expressing cell lines while minimising any other effects of the drugs. While analysing the time-series microscopy images for the corresponding control assays we observed that the total cell population appeared to grow exponentially, as expected, however the two subpopulations of cells in G1 and S/G2/M phases exhibited oscillations in their growth. These oscillations, which represent partial cell-cycle synchronisation of the cell population, were completely unexpected as we prepare our proliferation assays using standard methods thought to produce asynchronous cell populations. We therefore continued to investigate this cell synchronisation phenomenon, as it could have implications for our examination of the go-or-grow hypothesis. Further, many different mathematical models of proliferation have been proposed [50–56], and a key assumption often implicit in such models is that the cell population is asynchronous, so such models cannot explain the partial synchronisation that we observe experimentally.

The investigation of the partial synchronisation in the proliferation assays is reported in Chapter 3, where we present our new experimental data showing that cell proliferation assays, prepared using standard methods thought to produce asynchronous cell populations, persistently exhibit partial synchronisation which we call *inherent* synchronisation. Our experiments use FUCCI-expressing cells to highlight the oscillatory subpopulations within the total cell population, thereby revealing the normally-hidden cell synchronisation. These oscillatory subpopulations would never be observed without these cell cycle indicators. In contrast, our experimental data show that the total cell population appears to grow exponentially, as in an asynchronous population. These observations seem inconsistent, however we were able to reconcile these seemingly inconsistent observations by employing a multi-stage mathematical model of cell proliferation that can replicate the oscillations in the growth of the cell subpopulations. Our study has important implications for understanding and improving experimental reproducibility, as inherent synchronisation may affect the experimental reproducibility of studies aiming to investigate cell-cycle-dependent mechanisms, including changes in migration and drug response. Without quantitative experimental techniques like FUCCI to highlight the cell cycle, synchronisation and its effects on experimental outcomes and reproducibility may remain hidden.

Following our investigation of inherent synchronisation, we return in Chapter 4 to our examination of the go-or-grow hypothesis. One aim was to develop a rigorous procedure for testing whether individual cells in a population are subject to go-or-grow. Previous experimental studies of the go-or-grow hypothesis are conflicting, as some studies support the hypothesis [57, 62, 63] while others refute it [13, 64, 65]. The claims of many of these studies are based on qualitative

observation rather than quantitative analysis, so the conclusions are not well supported. Further, the correct measure of cell motility is cell diffusivity, which is typically not estimated in previous studies of go-or-grow.

Our examination of the go-or-grow hypothesis employs 2-D *in vitro* proliferation assays using three FUCCI-expressing melanoma cell lines, as melanoma is an ideal candidate for studying go-or-grow [57, 59, 68]. FUCCI allows us to study cell motility in G1 separately from cell motility in S/G2/M [13, 69, 70]. Specifically, we investigate cycling cells for differences in motility when the cells are in G1 compared with S/G2/M, since if cells are subject to go-or-grow then a cycling cell may become less motile as it progresses through the cell cycle and nears cell division. Further, given the potential for an arrested cell to become more motile, as it is not proliferating, we use a cell cycle-inhibiting drug to effect G1 arrest and compare the motility of the arrested cells with cycling cells.

Since the concept of go-or-grow applies to a single cell, we determine the trajectories of cells in the experimental time-series images using single-cell tracking. From the resulting trajectory data we then determine that the cell migration is globally isotropic and that the directional persistence of the cells dissipates within 1 h. We may therefore regard the cell migration as free diffusion for lag times of at least 1 h. We then calculate mean diffusivities for cells in G1 and S/G2/M phases and in cell-cycle arrest. In summary, our analysis of cell migration in 2-D assays using three melanoma cell lines does not support the go-or-grow hypothesis. We find that cell motility is independent of the cell-cycle phase, so that the implication from go-or-grow that cells are more motile in G1 than in S/G2/M when they are nearing cell division is not supported by our data. Notably, there is no change in cell motility when we effect drug-induced G1 arrest in the cells, again displaying a lack of support for the go-or-grow hypothesis.

In Chapter 5 we use data from 2-D *in vitro* proliferation assays with three FUCCI-expressing melanoma cell lines to illustrate that cell cycle durations are well approximated by an Erlang distribution and not by an exponential distribution. Typically, mathematical models of cell proliferation in the literature have an implicit exponential distribution of cell cycle durations [50–54, 71–74], resulting in growth kinetics that are too fast.

Our aim in Chapter 5 was to develop a new mathematical model of cell proliferation with proliferative heterogeneity. To incorporate realistic cell cycle durations into our model we employed a delay differential equation model for cell proliferation in which the delays are distributed. The cell population consists of a slow-proliferating subpopulation and a fast-proliferating subpopulation, as tumours consist of a small subpopulation of slow-proliferating cells that tends to survive anti-cancer drug treatment and can maintain the tumour by repopulating the fast-proliferating

subpopulation [27,29]. In our model the cells can switch between the two subpopulations through two cellular processes, namely asymmetric cell division and induced switching of proliferative states by surrounding cells. While there are mathematical models in the literature that consider proliferative heterogeneity, they generally lack the flexibility of the proliferative-state dynamics that we consider: some models account for one proliferating subpopulation only, which undergoes asymmetric division, while the other subpopulations are quiescent or differentiated [87,88]; other models simply consider subpopulations with different proliferative states without any cells switching between the subpopulations [89].

The results from our model are very interesting, particularly when comparing the results with the proliferative heterogeneity observed in tumours. Our model accounts for several characteristics of tumours [27,29,156,157]: the cancer cells in tumours can self-renew through asymmetric division, as can the cells in our model; tumours consist of both a slow-proliferating subpopulation and a fast-proliferating subpopulation, which is accounted for in our model; both slow- and fast-proliferating cells in a tumour have a distribution of cell cycle durations, captured by the time delays in our model. Overall, our model shows promise for providing insight into proliferative heterogeneity in cell populations, particularly with regard to tumours.

6.2 Main theme

The main theme of the work in this thesis, detailed in Chapters 2–5, is the utilisation of FUCCI in cell migration and proliferation experiments to inform the development of novel mathematical models and to allow for more detailed mathematical analysis of experimental data. Experiments using FUCCI cells provide for the simultaneous study of cell migration and the cell cycle, motivating the development of novel mathematical models of cell migration and proliferation as in Chapters 2 and 5. Further, the detailed information on cell subpopulations from experiments with FUCCI cells may reveal normally-hidden features of cell populations, such as the partial synchronisation discussed in Chapter 3, and can be used to investigate cell behaviour as a function of the cell cycle, such as our examination of go-or-grow in Chapter 4. Overall, modern biological techniques such as FUCCI can be used to improve the mathematical understanding of cell biology, from which new biological insights can assist in explaining experimental observations.

6.3 Future work

Here we discuss some of the possibilities for future research corresponding to each of the four articles presented in this thesis.

6.3.1 A quantitative investigation of the effects of anti-cancer drugs on cell migration and proliferation

In Chapter 2 we present a mathematical model for cell migration and proliferation, which can describe scratch assay experiments. It would be of substantial interest to employ our model to quantitatively investigate how the migration and proliferation of melanoma cells are affected by treatment with anti-cancer drugs. It is known that these drugs often inhibit the cell cycle, causing proliferation to cease, however many important questions remain unanswered regarding the effects of the drugs on melanoma cells [13].

6.3.2 Extension of our cell migration model to additional cell cycle phases

In Chapter 2 we develop two models corresponding to the cell cycle phases as indicated by FUCCI: the fundamental model for the two phases G1 and S/G2/M; and the extended model for the three phases G1, eS, and S/G2/M. With the development of improved cell-cycle indicators that can highlight additional phases of the cell cycle, we could easily extend our model to incorporate as many phases as required. Indeed, a recent extension of FUCCI, referred to as FUCCI4 [113], highlights all four phases of the cell cycle: G1, S, G2, and M. By extending our model to four coupled partial differential equations, our modelling framework could be used to quantitatively describe cell migration where individual cells are highlighted using FUCCI4. Modelling all four phases of the cell cycle, rather than combining three phases into S/G2/M as required for FUCCI data, may improve the accuracy of the model predictions compared with our current model. Further, a model incorporating the four phases G1, S, G2, and M would allow for the investigation of the three phases S, G2, and M individually rather than combined in S/G2/M, and would be of particular utility for studying the effects of anti-cancer drugs that specifically inhibit cell cycle progression of melanoma cells in one of S, G2, or M [33].

6.3.3 Formal proof of existence of travelling wave solutions

In Chapter 2 we demonstrated numerically that our models have travelling wave solutions. For the fundamental model, making the biological reasonable assumption that the diffusivities are equal for cells in G1 and S/G2/M phases, we derived an expression for the minimum wave

speed. It would be of interest, at least from a mathematical perspective, to formally prove the existence of the travelling wave solutions. For the fundamental model, phase space is 4-D. While it is generally straightforward to prove the existence of travelling wave solutions in 2-D, in higher dimensions the problem becomes much more difficult due to the increasingly complex topology of the spaces. We attempted to prove the existence of travelling wave solutions for our fundamental model with an approach based on the principle of Ważewski [180, Chapter 2, Page 24, Theorem 2.3], which recognises in an explicit manner that flows define homotopies, so that homotopy theory can be used for the study of differential equations. Due to the 4-D nature of our phase space, there were substantial difficulties in extending the local behaviour of solutions to global behaviour. Another complication in phase space was the presence of a pulse solution rather than just a wavefront solution. Due to the difficulties involved, there was no time to continue with this project.

For phase space of three dimensions and higher an approach based on the principle of Ważewski seems likely to be difficult except for special cases in which phase space is relatively simple. Possibly a better approach is to use another topological index such as the Conley index [180–183] which is well-suited to higher dimensions.

6.3.4 Investigating the causes of inherent synchronisation

We employed a multi-stage model in Chapter 3 to replicate the inherent synchronisation observed in our cell proliferation experiments. We haven't, however, identified the causes of the synchronisation, of which there are a number of possibilities. While our experimental procedures ensured that there was no induced synchronisation, our cell lines are adherent cells so require attachment to a substrate to progress through the cell cycle. During the experiment set up, the cells are detached from the culture flask and placed into suspension while a cell count is performed prior to seeding the cells onto a plate. During the period the cells are in suspension, which could be up to an hour, some of the cells may die as a result of an inability to progress through the cell cycle. The cells most likely to die are those close to mitosis, therefore in S/G2/M phase which is highlighted as green by FUCCI. Once the suspended cells are seeded on a plate, the distribution of red to green cells could now be shifted towards the red cells, thereby producing partial synchronisation during the beginning of the experiment. Another possibility is that the synchronisation arises from the initial synchronisation of daughter cells from the same parent, as they both start in G1. Another further possibility is that spontaneous synchronisation is occurring by intercellular signalling, whereby cells influence the cell cycle progression of surrounding cells.

One approach to test for these possibilities, and others, is to simulate a proliferation assay with an agent-based model using the multi-stage version of the Gillespie algorithm [184].

6.3.5 Investigating go-or-grow in 3-D

In Chapter 4 we developed a rigorous methodology for examining the go-or-grow hypothesis, and then used this methodology to test the hypothesis with melanoma cells in 2-D assays. The natural extension is to utilise this methodology to study go-or-grow in 3-D. While the methodology is completely applicable, there are some difficulties to be overcome when analysing images of spheroids in 3-D: cell tracking can be more complicated as cells can move behind other cells, however good tracking software would minimise this issue; cells can appear to be migrating while suppressing proliferation, however this may be due to other reasons such as nuclear damage resulting from the cells moving through constricted pores.

6.3.6 Mathematical model of heterogeneous cell proliferation with time-dependent induced switching

In tumours the slow-proliferating subpopulation can continually arise and disappear (Roesch et al., 2010). It would therefore be interesting to accommodate this process into our model with a time-dependent switching term, which could be either periodic or aperiodic.

6.3.7 Mathematical model of heterogeneous cell proliferation with an explicit density-dependence of proliferative-state switching

We could consider the induced switching to have an explicit density dependence, so that no switching occurs from a particular proliferative state when the density of cells from the other proliferative state is above a certain value. A similar explicit density dependence could be implemented for asymmetric cell division, which occurs at constant proportions in our current model. These explicit dependences on density may be relevant for slow-proliferating subpopulations in tumours that appear to persist over time and maintain the relative size of the subpopulation (Perego et al., 2018; Vallette et al., 2019).

6.3.8 Mathematical model of heterogeneous cell proliferation with spontaneous switching

Another process for switching of proliferative states that would be interesting to study is spontaneous switching between subpopulations, which is independent of other cells.

6.3.9 Mathematical model of heterogeneous cell proliferation with spatial dimensions

While our model has implicit spatiality, since the dependent variables are cell densities, we could include spatial dimensions explicitly to model cell migration with diffusive terms. Further, induced switching could be modelled as a more localised process, where the chance of a cell switching proliferative states is determined by the density of cells in a different proliferative state within a given radius of the cell, where the interaction strength decreases with distance from the cell.

6.4 Additional co-authored publications and preprints motivated by this thesis

Over the course of the work in this thesis the author has been involved in three collaborative projects, at both institutional and international levels, that were motivated by this work. The resulting publications and preprints are noted here:

- Simpson MJ, Jin W, **Vittadello ST**, Tambyah TA, Ryan JM, Gunasingh G, Haass NK, McCue SW. Stochastic models of cell invasion with fluorescent cell cycle indicators. *Physica A: Statistical Mechanics and its Applications*. 2018;510:375–386. doi:10.1016/j.physa.2018.06.128.
- Simpson MJ, Baker RE, **Vittadello ST**, Maclaren OJ. Practical parameter identifiability for spatio-temporal models of cell invasion. *Journal of the Royal Society Interface*. 2020;17:20200055. doi:10.1098/rsif.2020.0055.
- Gavagnin E, **Vittadello ST**, Gunasingh G, Haass NK, Simpson MJ, Rogers T, Yates CA. Synchronised oscillations in growing cell populations are explained by demographic noise. bioRxiv 2020.03.13.987032. 2020. doi:10.1101/2020.03.13.987032.

6.5 Final remarks

In this thesis we have considered some examples of how the detailed cell-cycle information provided by FUCCI can be used to develop new mathematical models of cell migration and proliferation, and how the application of mathematical models and techniques can help to interpret experimental data. While we have addressed some particular questions in this thesis, even more have arisen during the course of this research, highlighting that there is still much to learn about cell migration and proliferation. The interdisciplinary approach to research has proved to be very productive in this thesis. Given the enormous amount of data now produced

from experimental biology, it is of great benefit that mathematicians and biologists collaborate to analyse and interpret the data, and improve experiment design. With such collaboration, the fields of applied mathematics and biology both have much to gain.

Bibliography

- [1] Dekoninck S, Blanpain C. Stem cell dynamics, migration and plasticity during wound healing. *Nature Cell Biology*. 2019;21:18–24.
- [2] Friedl P, Gilmour D. Collective cell migration in morphogenesis, regeneration and cancer. *Nature Reviews Molecular Cell Biology*. 2009;10:445–457.
- [3] Friedl P, Weigelin B. Interstitial leukocyte migration and immune function. *Nature Immunology*. 2008;9:960–969.
- [4] Ridley AJ, Schwartz MA, Burridge K, Firtel RA, Ginsberg MH, Borisy G, et al. Cell migration: integrating signals from front to back. *Science*. 2003;302:1704–1709.
- [5] Werner S, Grose R. Regulation of wound healing by growth factors and cytokines. *Physiological Reviews*. 2003;83:835–870.
- [6] Lämmermann T, Bader BL, Monkley SJ, Worbs T, Wedlich-Söldner R, Hirsch K, et al. Rapid leukocyte migration by integrin-independent flowing and squeezing. *Nature*. 2008;453:51–55.
- [7] Etienne-Manneville S. Polarity proteins in migration and invasion. *Oncogene*. 2008;27:6970–6980.
- [8] Jones MC, Zha J, Humphries MJ. Connections between the cell cycle, cell adhesion and the cytoskeleton. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2019;374:20180227.
- [9] Zhu J, Thompson CB. Metabolic regulation of cell growth and proliferation. *Nature Reviews Molecular Cell Biology*. 2019;20:436–450.

- [10] Matson JP, Cook JG. Cell cycle proliferation decisions: the impact of single cell analyses. *The FEBS Journal*. 2017;284:362–375.
- [11] Duronio RJ, Xiong Y. Signaling pathways that control cell proliferation. *Cold Spring Harbor Perspectives in Biology*. 2013;5:a008904.
- [12] Ju RJ, Stehbens SJ, Haass NK. The role of melanoma cell-stroma interaction in cell motility, invasion, and metastasis. *Frontiers in Medicine*. 2018;5:307.
- [13] Haass NK, Beaumont KA, Hill DS, Anfosso A, Mrass P, Munoz MA, et al. Real-time cell cycle imaging during melanoma growth, invasion, and drug response. *Pigment Cell & Melanoma Research*. 2014;27:764–776.
- [14] Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. *Cell*. 2011;144:646–674.
- [15] Friedl P, Wolf K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nature Reviews Cancer*. 2003;3:362–374.
- [16] Coussens LM, Werb Z. Inflammation and cancer. *Nature*. 2002;420:860–867.
- [17] Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000;100:57–70.
- [18] Edmondson R, Broglie JJ, Adcock AF, Yang L. Three-dimensional cell culture systems and their applications in drug discovery and cell-based biosensors. *ASSAY and Drug Development Technologies*. 2014;12:207–218.
- [19] Yamada KM, Sixt M. Mechanisms of 3D cell migration. *Nature Reviews Molecular Cell Biology*. 2019;20:738–752.
- [20] Ascione F, Vasaturo A, Caserta S, D’Esposito V, Formisano P, Guido S. Comparison between fibroblast wound healing and cell random migration assays *in vitro*. *Experimental Cell Research*. 2016;347:123–132.
- [21] Beaumont KA, Mohana-Kumaran N, Haass NK. Modeling melanoma *in vitro* and *in vivo*. *Healthcare*. 2014;2:27–46.
- [22] Liang CC, Park AY, Guan JL. *In vitro* scratch assay: a convenient and inexpensive method for analysis of cell migration *in vitro*. *Nature Protocols*. 2007;2:329–333.
- [23] Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D, Darnell J. *Molecular Cell Biology*. 4th ed. W. H. Freeman; 2000.

-
- [24] Vermeulen K, Van Bockstaele DR, Berneman ZN. The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer. *Cell Proliferation*. 2003;36:131–149.
- [25] McClatchey AI, Yap AS. Contact inhibition (of proliferation) redux. *Current Opinion in Cell Biology*. 2012;24:685–694.
- [26] Puliafito A, Hufnagel L, Neveu P, Streichan S, Sigal A, Fygenson DK, et al. Collective and single cell behavior in epithelial contact inhibition. *Proceedings of the National Academy of Sciences of the United States of America*. 2012;109:739–744.
- [27] Perego M, Maurer M, Wang JX, Shaffer S, Müller AC, Parapatics K, et al. A slow-cycling subpopulation of melanoma cells with highly invasive properties. *Oncogene*. 2018;37:302–312.
- [28] Roesch A, Fukunaga-Kalabis M, Schmidt EC, Zabierowski SE, Brafford PA, Vultur A, et al. A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell*. 2010;141:583–594.
- [29] Vallette FM, Olivier C, Lézot F, Oliver L, Cochonneau D, Lalier L, et al. Dormant, quiescent, tolerant and persister cells: Four synonyms for the same target in cancer. *Biochemical Pharmacology*. 2019;162:169–176.
- [30] Sakaue-Sawano A, Kurokawa H, Morimura T, Hanyu A, Hama H, Osawa H, et al. Visualizing spatiotemporal dynamics of multicellular cell-cycle progression. *Cell*. 2008;132:487–498.
- [31] Shand FHW, Ueha S, Otsuji M, Koid SS, Shichino S, Tsukui T, et al. Tracking of inter-tissue migration reveals the origins of tumor-infiltrating monocytes. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111:7771–7776.
- [32] Beaumont KA, Anfosso A, Ahmed F, Weninger W, Haass NK. Imaging- and flow cytometry-based analysis of cell position and the cell cycle in 3D melanoma spheroids. *Journal of Visualized Experiments*. 2015;106:e53486.
- [33] Beaumont KA, Hill DS, Daignault SM, Lui GYL, Sharp DM, Gabrielli B, et al. Cell cycle phase-specific drug resistance as an escape mechanism of melanoma cells. *Journal of Investigative Dermatology*. 2016;136:1479–1489.
- [34] Dowling MR, Kan A, Heinzl S, Zhou JHS, Marchingo JM, Wellard CJ, et al. Stretched cell cycle model for proliferating lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111:6377–6382.

- [35] Sandler O, Mizrahi SP, Weiss N, Agam O, Simon I, Balaban NQ. Lineage correlations of single cell division time as a probe of cell-cycle dynamics. *Nature*. 2015;519:468–471.
- [36] Pauklin S, Vallier L. The cell-cycle state of stem cells determines cell fate propensity. *Cell*. 2013;155:135–147.
- [37] Roccio M, Schmitter D, Knobloch M, Okawa Y, Sage D, Lutolf MP. Predicting stem cell fate changes by differential cell cycle progression patterns. *Development*. 2013;140:459–470.
- [38] Murray JD. *Mathematical Biology: 1. An Introduction*. Third edition ed. New York: Springer; 2002.
- [39] Fisher RA. The wave of advance of advantageous genes. *Annals of Eugenics*. 1937;7:355–369.
- [40] Alarcón T, Byrne HM, Maini PK. A cellular automaton model for tumour growth in inhomogeneous environment. *Journal of Theoretical Biology*. 2003;225:257–274.
- [41] Anderson ARA, Chaplain MAJ. Continuous and discrete mathematical models of tumor-induced angiogenesis. *Bulletin of Mathematical Biology*. 1998;60:857–899.
- [42] Johnston ST, Shah ET, Chopin LK, McElwain DLS, Simpson MJ. Estimating cell diffusivity and cell proliferation rate by interpreting IncuCyte ZOOM™ assay data using the Fisher-Kolmogorov model. *BMC Systems Biology*. 2015;9:38.
- [43] Tremel A, Cai A, Tirtaatmadja N, Hughes BD, Stevens GW, Landman KA, et al. Cell migration and proliferation during monolayer formation and wound healing. *Chemical Engineering Science*. 2009;64:247–253.
- [44] Simpson MJ, Zhang DC, Mariani M, Landman KA, Newgreen DF. Cell proliferation drives neural crest cell invasion of the intestine. *Developments in Biologicals*. 2007;302:553–568.
- [45] Sengers BG, Please CP, Oreffo ROC. Experimental characterization and computational modelling of two-dimensional cell spreading for skeletal regeneration. *Journal of the Royal Society Interface*. 2007;4:1107–1117.
- [46] Maini PK, McElwain DLS, Leavesley D. Travelling waves in a wound healing assay. *Applied Mathematics Letters*. 2004;17:575–580.
- [47] Landman KA, Cai AQ, Hughes BD. Travelling waves of attached and detached cells in a wound-healing cell migration assay. *Bulletin of Mathematical Biology*. 2007;69:2119–2138.

-
- [48] Canosa J. On a nonlinear diffusion equation describing population growth. IBM Journal of Research and Development. 1973;17:307–313.
- [49] Simpson MJ, Landman KA, Hughes BD, Newgreen DF. Looking inside an invasion wave of cells using continuum models: Proliferation is the key. Journal of Theoretical Biology. 2006;243:343–360.
- [50] Sherratt JA, Murray JD. Models of epidermal wound healing. Proceedings of the Royal Society B. 1990;241:29–36.
- [51] Swanson KR, Bridge C, Murray JD, Alvord EC. Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. Journal of the Neurological Sciences. 2003;216:1–10.
- [52] Maini PK, McElwain DLS, Leavesley DI. Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells. Tissue Engineering. 2004;10:475–482.
- [53] Scott JG, Basanta D, Anderson ARA, Gerlee P. A mathematical model of tumour self-seeding reveals secondary metastatic deposits as drivers of primary tumour growth. Journal of the Royal Society Interface. 2013;10:20130011.
- [54] Sarapata EA, de Pillis LG. A comparison and catalog of intrinsic tumor growth models. Bulletin of Mathematical Biology. 2014;76:2010–2024.
- [55] Böttcher MA, Dingli D, Werner B, Traulsen A. Replicative cellular age distributions in compartmentalized tissues. Journal of the Royal Society Interface. 2018;15:20180272.
- [56] Treloar KK, Simpson MJ, McElwain DLS, Baker RE. Are *in vitro* estimates of cell diffusivity and cell proliferation rate sensitive to assay geometry? Journal of Theoretical Biology. 2014;356:71–84.
- [57] Hoek KS, Eichhoff OM, Schlegel NC, Döbbeling U, Kobert N, Schaerer L, et al. *In vivo* switching of human melanoma cells between proliferative and invasive states. Cancer Research. 2008;68:650–656.
- [58] Zhang J, Goliwas KF, Wang W, Taufalele PV, Bordeleau F, Reinhart-King CA. Energetic regulation of coordinated leader-follower dynamics during collective invasion of breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America. 2019;116:7867–7872.

- [59] Zipser MC, Eichhoff OM, Widmer DS, Schlegel NC, Schoenewolf NL, Stuart D, et al. A proliferative melanoma cell phenotype is responsive to RAF/MEK inhibition independent of BRAF mutation status. *Pigment Cell & Melanoma Research*. 2011;24:326–333.
- [60] Giese A, Bjerkvig R, Berens ME, Westphal M. Cost of migration: invasion of malignant gliomas and implications for treatment. *Journal of Clinical Oncology*. 2003;21:1624–1636.
- [61] Czirók A, Schlett K, Madarász E, Vicsek T. Exponential Distribution of Locomotion Activity in Cell Cultures. *Physical Review Letters*. 1998;81:3038–3041.
- [62] Schultz C, Lemke N, Ge S, Golembieski WA, Rempel SA. Secreted protein acidic and rich in cysteine promotes glioma invasion and delays tumor growth *in vivo*. *Cancer Research*. 2002;62:6270–6277.
- [63] Giese A, Loo MA, Tran N, Haskett D, Coons SW, Berens ME. Dichotomy of astrocytoma migration and proliferation. *International Journal of Cancer*. 1996;67:275–282.
- [64] Garay T, Juhász É, Molnár E, Eisenbauer M, Czirók A, Dekan B, et al. Cell migration or cytokinesis and proliferation? — Revisiting the “go or grow” hypothesis in cancer cells in vitro. *Experimental Cell Research*. 2013;319:3094–3103.
- [65] Corcoran A, Del Maestro RF. Testing the “Go or Grow” hypothesis in human medulloblastoma cell lines in two and three dimensions. *Neurosurgery*. 2003;53:174–185.
- [66] Ahmed F, Haass NK. Microenvironment-driven dynamic heterogeneity and phenotypic plasticity as a mechanism of melanoma therapy resistance. *Frontiers in Oncology*. 2018;8:173.
- [67] Grzywa TM, Paskal W, Włodarski PK. Intratumor and intertumor heterogeneity in melanoma. *Translational Oncology*. 2017;10:956–975.
- [68] Hoek KS, Goding CR. Cancer stem cells versus phenotype-switching in melanoma. *Pigment Cell & Melanoma Research*. 2010;23:746–759.
- [69] Chu TLH, Connell M, Zhou L, He Z, Won J, Chen H, et al. Cell cycle-dependent tumor engraftment and migration are enabled by Aurora-A. *Molecular Cancer Research*. 2018;16:16–31.
- [70] Kagawa Y, Matsumoto S, Kamioka Y, Mimori K, Naito Y, Ishii T, et al. Cell cycle-dependent Rho GTPase activity dynamically regulates cancer cell motility and invasion in vivo. *PLoS ONE*. 2013;8:e83629.

-
- [71] Webb GF. A model of proliferating cell populations with inherited cycle length. *Journal of Mathematical Biology*. 1986;23:269–282.
- [72] Lebowitz JL, Rubinow SI. A theory for the age and generation time distribution of a microbial population. *Journal of Mathematical Biology*. 1974;1:17–36.
- [73] Byrne H, Drasdo D. Individual-based and continuum models of growing cell populations: a comparison. *Journal of Mathematical Biology*. 2009;58:657–687.
- [74] Cai AQ, Landman KA, Hughes BD. Multi-scale modeling of a wound-healing cell migration assay. *Journal of Theoretical Biology*. 2007;245:576–594.
- [75] Pavel M, Renna M, Park SJ, Menzies FM, Ricketts T, Füllgrabe J, et al. Contact inhibition controls cell survival and proliferation via YAP/TAZ-autophagy axis. *Nature Communications*. 2018;9:2961.
- [76] Chao HX, Fakhreddin RI, Shimerov HK, Kedziora KM, Kumar RJ, Perez J, et al. Evidence that the human cell cycle is a series of uncoupled, memoryless phases. *Molecular Systems Biology*. 2019;15:e8604.
- [77] Weber TS, Jaehnert I, Schichor C, Or-Guil M, Carneiro J. Quantifying the length and variance of the eukaryotic cell cycle phases by a stochastic model and dual nucleoside pulse labelling. *PLoS Computational Biology*. 2014;10:e1003616.
- [78] Villasana M, Radunskaya A. A delay differential equation model for tumor growth. *Journal of Mathematical Biology*. 2003;47:270–294.
- [79] Baker CTH, Bocharov GA, Paul CAH, Rihan FA. Modelling and analysis of time-lags in some basic patterns of cell proliferation. *Journal of Mathematical Biology*. 1998;37:341–371.
- [80] Baker CTH, Bocharov GA, Paul CAH. Mathematical modelling of the interleukin-2 T-Cell system: a comparative study of approaches based on ordinary and delay differential equation. *Journal of Theoretical Medicine*. 1997;1:117–128.
- [81] Mackey MC, Rudnicki R. Global stability in a delayed partial differential equation describing cellular replication. *Journal of Mathematical Biology*. 1994;33:89–109.
- [82] Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models incorporating a multi-stage cell cycle replicate normally-hidden inherent synchronization in cell proliferation. *Journal of the Royal Society Interface*. 2019;16:20190382.

- [83] Gavagnin E, Ford MJ, Mort RL, Rogers T, Yates CA. The invasion speed of cell migration models with realistic cell cycle time distributions. *Journal of Theoretical Biology*. 2019;481:91–99.
- [84] Simpson MJ, Jin W, Vittadello ST, Tambyah TA, Ryan JM, Gunasingh G, et al. Stochastic models of cell invasion with fluorescent cell cycle indicators. *Physica A: Statistical Mechanics and its Applications*. 2018;510:375–386.
- [85] Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Mathematical models for cell migration with real-time cell cycle dynamics. *Biophysical Journal*. 2018;114:1241–1253.
- [86] Yates CA, Ford MJ, Mort RL. A multi-stage representation of cell proliferation as a Markov process. *Bulletin of Mathematical Biology*. 2017;79:2905–2928.
- [87] Greene JM, Levy D, Fung KL, Souza PS, Gottesman MM, Lavi O. Modeling intrinsic heterogeneity and growth of cancer cells. *Journal of Theoretical Biology*. 2015;367:262–277.
- [88] Arino O, Kimmel M. Asymptotic behavior of a nonlinear functional-integral equation of cell kinetics with unequal division. *Journal of Mathematical Biology*. 1989;27:341–354.
- [89] Jin W, McCue SW, Simpson MJ. Extended logistic growth model for heterogeneous populations. *Journal of Theoretical Biology*. 2018;445:51–61.
- [90] Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, et al. *Molecular Biology of the Cell*. 6th ed. Connecticut: Taylor & Francis Ltd.; 2014.
- [91] Sadeghi HM, Seitz B, Hayashi S, LaBree L, McDonnell PJ. In vitro effects of Mitomycin-C on human keratocytes. *Journal of Refractive Surgery*. 1998;14:534–40.
- [92] Simpson MJ, Treloar KK, Binder BJ, Haridas P, Manton KJ, Leavesley DI, et al. Quantifying the roles of cell motility and cell proliferation in a circular barrier assay. *Journal of the Royal Society Interface*. 2013;10.
- [93] Holmes W, Mata M, Edelstein-Keshet L. Local perturbation analysis: A computational tool for biophysical reaction–diffusion models. *Biophysical Journal*. 2015;108:230–236.
- [94] Simpson MJ, Landman KA. Characterizing and minimizing the operator split error for Fisher’s equation. *Applied Mathematics Letters*. 2006;19:604–612.

-
- [95] Marel AK, Zorn M, Klingner C, Wedlich-Söldner R, Frey E, Rädler J. Flow and diffusion in channel-guided cell migration. *Biophysical Journal*. 2014;107:1054–1064.
- [96] Holder AB, Rodrigo MR, Herrero MA. A model for acid-mediated tumour growth with nonlinear acid production term. *Applied Mathematics and Computation*. 2014;227:176–198.
- [97] Kimpton LS, Whiteley JP, Waters SL, King JR, Oliver JM. Multiple travelling-wave solutions in a minimal model for cell motility. *Mathematical Medicine and Biology*. 2013;30:241–272.
- [98] Marchant BP, Norbury J, Byrne HM. Biphasic behaviour in malignant invasion. *Mathematical Medicine and Biology*. 2006;23:173–196.
- [99] Curtis CW, Bortz DM. Propagation of fronts in the Fisher-Kolmogorov equation with spatially varying diffusion. *Physical Review E*. 2012;86:066108.
- [100] Denman PK, McElwain DLS, Norbury J. Analysis of travelling waves associated with the modelling of aerosolised skin grafts. *Bulletin of Mathematical Biology*. 2007;69:495–523.
- [101] Davies MA, Stemke-Hale K, Lin E, Tellez C, Deng W, Gopal YN, et al. Integrated molecular and clinical analysis of AKT activation in metastatic melanoma. *Clinical Cancer Research*. 2009;15:7538–7546.
- [102] Hoek KS, Schlegel NC, Brafford P, Sucker A, Ugurel S, Kumar R, et al. Metastatic potential of melanomas defined by specific gene expression profiles with no BRAF signature. *Pigment Cell Research*. 2006;19:290–302.
- [103] Smalley KSM, Contractor R, Haass NK, Kulp AN, Atilla-Gokcumen GE, Williams DS, et al. An organometallic protein kinase inhibitor pharmacologically activates p53 and induces apoptosis in human melanoma cells. *Cancer Research*. 2007;67:209–217.
- [104] Smalley KSM, Contractor R, Haass NK, Lee JT, Nathanson KL, Medina CA, et al. Ki67 expression levels are a better marker of reduced melanoma growth following MEK inhibitor treatment than phospho-ERK levels. *British Journal of Cancer*. 2007;96:445–449.
- [105] Smalley KSM, Brafford P, Haass NK, Brandner JM, Brown E, Herlyn M. Up-regulated expression of zonula occludens protein-1 in human melanoma associates with N-cadherin and contributes to invasion and adhesion. *American Journal of Pathology*. 2005;166:1541–1554.

- [106] Spoerri L, Beaumont KA, Anfosso A, Haass NK. Real-time cell cycle imaging in a 3D cell culture model of melanoma. *Methods in Molecular Biology*. 2017;1612:401–416.
- [107] Jin W, Shah ET, Penington CJ, McCue SW, Chopin LK, Simpson MJ. Reproducibility of scratch assays is affected by the initial degree of confluence: Experiments, modelling and model selection. *Journal of Theoretical Biology*. 2016;390:136–145.
- [108] Warne DJ, Baker RE, Simpson MJ. Optimal Quantification of Contact Inhibition in Cell Populations. *Biophysical Journal*. 2017;113:1920–1924.
- [109] Yohem K, Slymen D, Bregman M, Meyskens F. Radiation survival of murine and human melanoma cells utilizing two assay systems: Monolayer and soft agar. *British Journal of Cancer*. 1988;57:64–69.
- [110] Treloar KK, Simpson MJ, Haridas P, Manton KJ, Leavesley DI, McElwain DLS, et al. Multiple types of data are required to identify the mechanisms influencing the spatial expansion of melanoma cell colonies. *BMC Systems Biology*. 2013;7:137.
- [111] Haridas P, Penington CJ, McGovern JA, McElwain DLS, Simpson MJ. Quantifying rates of cell migration and cell proliferation in co-culture barrier assays reveals how skin and melanoma cells interact during melanoma spreading and invasion. *Journal of Theoretical Biology*. 2017;423:13–25.
- [112] Abramowitz M, Stegun IA, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. United States Department of Commerce, National Bureau of Standards (NBS); 1964.
- [113] Bajar BT, Lam AJ, Badiee RK, Oh YH, Chu J, Zhou XX, et al. Fluorescent indicators for simultaneous reporting of all four cell cycle phases. *Nature Methods*. 2016;13:993–996.
- [114] Morton KW, Mayers DF. *Numerical Solution of Partial Differential Equations*. 2nd ed. Cambridge: Cambridge University Press; 2005.
- [115] Tsoularis A, Wallace J. Analysis of logistic growth models. *Mathematical Biosciences*. 2002;179:21–55.
- [116] Otani K, Naito Y, Sakaguchi Y, Seo Y, Takahashi Y, Kikuta J, et al. Cell-cycle-controlled radiation therapy was effective for treating a murine malignant melanoma cell line in vitro and in vivo. *Scientific Reports*. 2016;6:30689.

-
- [117] Bae H, Go YH, Kwon T, Sung BJ, Cha HJ. A theoretical model for the cell cycle and drug induced cell cycle arrest of FUCCI systems with cell-to-cell variation during mitosis. *Pharmaceutical Research*. 2019;36:57.
- [118] Taylor IW, Hodson PJ. Cell cycle regulation by environmental pH. *Journal of Cellular Physiology*. 1984;121:517–525.
- [119] Davis PK, Ho A, Dowdy SF. Biological methods for cell-cycle synchronization of mammalian cells. *BioTechniques*. 2001;30:1322–1331.
- [120] Haass NK, Gabrielli B. Cell cycle-tailored targeting of metastatic melanoma: Challenges and opportunities. *Experimental Dermatology*. 2017;26:649–655.
- [121] MATLAB fit. Fit curve or surface to data (R2019b). Accessed February 2020.; 2019. Available from: <https://au.mathworks.com/help/curvefit/fit.html>.
- [122] MATLAB lsqnonlin. Solve nonlinear least-squares (nonlinear data-fitting) problems (R2019b). Accessed February 2020.; 2019. Available from: <https://mathworks.com/help/optim/ug/lsqnonlin.html>.
- [123] Coleman TF, Li Y. An interior trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*. 1996;6:418–445.
- [124] MATLAB fft. Fast Fourier transform (R2019b). Accessed February 2020.; 2019. Available from: <https://au.mathworks.com/help/matlab/ref/fft.html>.
- [125] Pfeifer CR, Xia Y, Zhu K, Liu D, Irianto J, García VMM, et al. Constricted migration increases DNA damage and independently represses cell cycle. *Molecular Biology of the Cell*. 2018;29:1948–1962.
- [126] Smal I, Loog M, Niessen W, Meijering E. Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Transactions on Medical Imaging*. 2010;29:282–301.
- [127] Schiffmann DA, Dikovskaya D, Appleton PL, Newton IP, Creager DA, Allan C, et al. Open Microscopy Environment and FindSpots: integrating image informatics with quantitative multidimensional image analysis. *BioTechniques*. 2006;41:199–208.
- [128] Tinevez JY, Perry N, Schindelin J, Hoopes GM, Reynolds GD, Laplantine E, et al. TrackMate: An open and extensible platform for single-particle tracking. *Methods*. 2017;115:80–90.

- [129] Banigan EJ, Harris TH, Christian DA, Hunter CA, Liu AJ. Heterogeneous CD8⁺ T cell migration in the lymph node in the absence of inflammation revealed by quantitative migration analysis. *PLoS Computational Biology*. 2015;11:e1004058.
- [130] Harris TH, Banigan EJ, Christian DA, Konradt C, Wojno EDT, Norose K, et al. Generalized Lévy walks and the role of chemokines in migration of effector CD8⁺ T cells. *Nature*. 2012;486:545–548.
- [131] Rudnick J, Gaspari G. The asphericity of random walks. *Journal of Physics A: Mathematical and General*. 1986;19:L191–L193.
- [132] Guisoni N, Mazzitello KI, Diambra L. Modeling active cell movement with the Potts model. *Frontiers in Physics*. 2018;6.
- [133] Wu PH, Giri A, Sun SX, Wirtz D. Three-dimensional cell migration does not follow a random walk. *Proceedings of the National Academy of Sciences of the United States of America*. 2014;111:3949–3954.
- [134] Selmeçzi D, Mosler S, Hagedorn PH, Larsen NB, Flyvbjerg H. Cell motility as persistent random motion: theories from experiments. *Biophysical Journal*. 2005;89:912–931.
- [135] Campos D, Méndez V, Llopis I. Persistent random motion: uncovering cell migration dynamics. *Journal of Theoretical Biology*. 2010;267:526–534.
- [136] Holmström K, Petersson J. A review of the parameter estimation problem of fitting positive exponential sums to empirical data. *Applied Mathematics and Computation*. 2002;126:31–61.
- [137] Stokes CL, Lauffenburger DA, Williams SK. Migration of individual microvessel endothelial cells: stochastic model and parameter measurement. *Journal of Cell Science*. 1991;99:419–430.
- [138] Gail MH, Boone CW. The locomotion of mouse fibroblasts in tissue culture. *Biophysical Journal*. 1970;10:980–993.
- [139] Takagi H, Sato MJ, Yanagida T, Ueda M. Functional analysis of spontaneous cell movement under different physiological conditions. *PLoS ONE*. 2008;3:e2648.
- [140] Michalet X. Mean square displacement analysis of single-particle trajectories with localization error: Brownian motion in an isotropic medium. *Physical Review E*. 2010;82:041914.

-
- [141] Qian H, Sheetz MP, Elson EL. Single particle tracking: analysis of diffusion and flow in two-dimensional systems. *Biophysical Journal*. 1991;60:910–921.
- [142] Saxton MJ. Single-particle tracking: the distribution of diffusion coefficients. *Biophysical Journal*. 1997;72:1744–1753.
- [143] MATLAB fitlm. Fit linear regression model (R2019b). Accessed February 2020.; 2019. Available from: <https://au.mathworks.com/help/stats/fitlm.html>.
- [144] Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers EJ, Berk R, et al. Redefine statistical significance. *Nat Hum Behav*. 2018;2:6–10.
- [145] Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Reviews*. 2007;82:591–605.
- [146] Vittadello ST, McCue SW, Gunasingh G, Haass NK, Simpson MJ. Examining go-or-grow using fluorescent cell-cycle indicators and cell-cycle-inhibiting drugs. *Biophysical Journal*. 2020;118:1243–1247.
- [147] Byrne HM. The effect of time delays on the dynamics of avascular tumor growth. *Mathematical Biosciences*. 1997;144:83–117.
- [148] Getto P, Waurick M. A differential equation with state-dependent delay from cell population biology. *Journal of Differential Equations*. 2016;260:6176–6200.
- [149] Getto P, Gyllenberg M, Nakata Y, Scarabel F. Stability analysis of a state-dependent delay differential equation for cell maturation: analytical and numerical methods. *Journal of Mathematical Biology*. 2019;79:281–328.
- [150] Cassidy T, Humphries AR. A mathematical model of viral oncology as an immunoncology instigator. *Mathematical Medicine and Biology*. 2020;37:117–151.
- [151] Arino O. A survey of structured cell population dynamics. *Acta Biotheoretica*. 1995;43:3–25.
- [152] Gabriel P, Garbett SP, Quaranta V, Tyson DR, Webb GF. The contribution of age structure to cell population responses to targeted therapeutics. *Journal of Theoretical Biology*. 2012;311:19–27.
- [153] Billy F, Clairambault J, Fercoq O, Gaubert S, Lepoutre T, Ouillon T, et al. Synchronisation and control of proliferation in cycling cell population models with age structure. *Mathematics and Computers in Simulation*. 2014;96:66–94.

- [154] Clairambault J, Fercoq O. Physiologically Structured Cell Population Dynamic Models with Applications to Combined Drug Delivery Optimisation in Oncology. *Mathematical Modelling of Natural Phenomena*. 2016;11:45–70.
- [155] Cassidy T, Craig M, Humphries AR. Equivalences between age structured models and state dependent distributed delay differential equations. *Mathematical Biosciences and Engineering*. 2019;16:5419–5450.
- [156] Moore N, Lyle S. Quiescent, slow-cycling stem cell populations in cancer: A review of the evidence and discussion of significance. *Journal of Oncology*. 2011;2011:396076.
- [157] Ahn A, Chatterjee A, Eccles MR. The slow cycling phenotype: A growing problem for treatment resistance in melanoma. *Molecular Cancer Therapeutics*. 2017;16:1002–1009.
- [158] Gallaher JA, Brown JS, Anderson ARA. The impact of proliferation-migration tradeoffs on phenotypic evolution in cancer. *Scientific Reports*. 2019;9:2425.
- [159] Bajaj J, Zimdahl B, Reya T. Fearful symmetry: subversion of asymmetric division in cancer development and progression. *Cancer Research*. 2015;75:792–797.
- [160] Dey-Guha I, Alves CP, Yeh AC, Salony, Sole X, Darp R, et al. A mechanism for asymmetric cell division resulting in proliferative asynchronicity. *Molecular Cancer Research*. 2015;13:223–230.
- [161] Smalley KSM, Herlyn M. Integrating tumor-initiating cells into the paradigm for melanoma targeted therapy. *International Journal of Cancer*. 2009;124:1245–1250.
- [162] Dey-Guha I, Wolfer A, Yeh AC, Albeck JG, Darp R, Leon E, et al. Asymmetric cancer cell division regulated by AKT. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;108:12845–12850.
- [163] Nelson CM, Chen CS. Cell–cell signaling by direct contact increases cell proliferation via a PI3K-dependent signal. *FEBS Letters*. 2002;514:238–242.
- [164] West J, Newton PK. Cellular interactions constrain tumor growth. *Proceedings of the National Academy of Sciences of the United States of America*. 2019;116:1918–1923.
- [165] Lu L. Numerical stability of the Θ -methods for systems of differential equations with several delay terms. *Journal of Computational and Applied Mathematics*. 1991;34:291–304.

-
- [166] Engelborghs K, Luzyanina T, Roose D. Numerical bifurcation analysis of delay differential equations. *Journal of Computational and Applied Mathematics*. 2000;125:265–275.
- [167] Sun L. Stability analysis for delay differential equations with multidelays and numerical examples. *Mathematics of Computation*. 2006;75:151–165.
- [168] McCluskey CC. Global stability of an *SIR* epidemic model with delay and general non-linear incidence. *Mathematical Biosciences and Engineering*. 2010;7:837–850.
- [169] Khasawneh FA, Mann BP. Stability of delay integro-differential equations using a spectral element method. *Mathematical and Computer Modelling*. 2011;54:2493–2503.
- [170] Huang C, Cao J, Wen F, Yang X. Stability analysis of SIR model with distributed delay on complex networks. *PLOS ONE*. 2016;11:e0158813.
- [171] Kaslik E, Neamtu M. Stability and Hopf bifurcation analysis for the hypothalamic-pituitary-adrenal axis model with memory. *Mathematical Medicine and Biology*. 2018;35:49–78.
- [172] Kuang Y. *Delay Differential Equations: With Applications in Population Dynamics*. Academic Press; 1993.
- [173] Smith H. *An Introduction to Delay Differential Equations with Applications to the Life Sciences*. Springer-Verlag GmbH; 2011.
- [174] Diekmann O, van Gils SA, Verduyn Lunel SM, Walther HO. *Delay Equations*. Springer New York; 1995.
- [175] Ahlfors LV. *Complex Analysis*. 3rd ed. McGraw-Hill; 1979.
- [176] Rudin W. *Real and Complex Analysis*. 3rd ed. McGraw-Hill Education; 1986.
- [177] Haass NK. Dynamic tumor heterogeneity in melanoma therapy: how do we address this in a novel model system? *Melanoma Management*. 2015;2:93–95.
- [178] Chapman A, Fernandez del Ama L, Ferguson J, Kamarashev J, Wellbrock C, Hurlstone A. Heterogeneous tumor subpopulations cooperate to drive invasion. *Cell Reports*. 2014;8:688–695.
- [179] Allen LJS. *An Introduction to Stochastic Processes with Applications to Biology*. 2nd ed. Taylor & Francis Ltd.; 2010.

- [180] Conley C. Isolated Invariant Sets and the Morse Index. vol. 38 of CBMS Regional Conference Series in Mathematics. American Mathematical Society; 1978.
- [181] Gardner R, Smoller J. The existence of periodic travelling waves for singularly perturbed predator-prey equations via the Conley Index. *Journal of Differential Equations*. 1983;47:133–161.
- [182] Arnold L, Jones CKRT, Mischaikow K, Raugel G. *Dynamical Systems*. Johnson R, editor. Springer Berlin Heidelberg; 1995.
- [183] Kaźmierczak B. Existence of travelling wave solutions for reaction-diffusion-convection systems via the Conley index theory. *Topological Methods in Nonlinear Analysis*. 2001;17:359–403.
- [184] Gillespie DT. Exact stochastic simulation of coupled chemical reactions. *Journal of Physical Chemistry*. 1977;81:2340–2361.