Mathematical Modelling of Cancer Stem Cells

by

Colin Turner

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The traditional view of cancer asserts that a malignant tumour is composed of a population of cells, all of which share the ability to divide without limit. Within the last decade, however, this notion has lost ground to the emerging *cancer stem cell hypothesis*, which counters that only a (typically small) sub-population of so-called 'cancer stem cells' has the capacity to proliferate indefinitely, and hence to drive and maintain tumour growth. Cancer stem cells have been putatively identified in leukemias and, more recently, in a variety of solid tumours including those of the breast and brain. The cancer stem cell hypothesis helps to explain certain clinically-observed phenomena, including the apparent inability of conventional anti-cancer therapies to eradicate the disease despite (transient) reduction of overall tumour bulk – presumably these treatments fail to kill the underlying cancer stem cells. Herein, we develop stochastic and deterministic temporal models of tumour growth based on the cancer stem cell hypothesis, and apply these models to discussions of the treatment of glioblastoma multiforme, a common type of brain cancer believed to be maintained by cancer stem cells, and to the phenomenon of the epithelial-mesenchymal transition, a process thought to be important in generating cancer cells capable of metastasis.

Acknowledgments

I'd like to thank my supervisors, Mohammad Kohandel and Siv Sivaloganathan, for their support, in addition to the opportunities they have given me and the lessons they have taught me. I would like to thank my fellow graduate students, particularly Rudy Gunawan and Gibin Powathil for their apparently infinite patience in dealing with my more or less constant volley of questions and concerns, and Eddie Dupont for firing back with a more or less constant barrage of very offensive jokes and one-liners. And finally, I'd like to thank my parents and sisters for the unwavering and universal support they've given me over the last twenty-five or so years.

Contents

List of Tables **ix**

List of Figures **x**

1 Introduction **1**

2 Background **6**

2.1 A brief histoy of cancer stem cell biology 6

2.1.1 Till and McCulloch and the advent of stem cell biology . . . 7

2.1.2 Subsequent characterisation of hematopoietic stem cells . . . 9

2.1.3 Recognition of heterogeneity within tumours 10

2.1.4 Technology 11

2.1.5 Cancer stem cells in leukemias and in solid tumours 12

2.1.6 The origins of cancer stem cells 14

2.1.7 Criticism 14

2.2 Mathematical Modelling of (Cancer) Stem Cells 18

2.2.1 Deterministic Modelling of CSCs in Leukemias and the Brain 18
2.2.2 Modelling the Role of Stem Cells in Colorectal Cancer . . . 21
2.2.3 Stochastic Modelling of the Origins of Cancer 23
2.2.4 Modelling the Growth of (Cancer) Stem Cells as a Birth-and-Death Process 24

3 Modelling Cancer Stem Cells 27

3.1 A simple model 27
3.1.1 Division properties of cancer stem cells 28
3.1.2 Properties of Non-Stem Cells 29
3.1.3 On the stochasticity of cell division 30
3.2 Markov Processes 32
3.2.1 Random Variables 32
3.2.2 Stochastic Processes 33
3.2.3 The Markov Property 33
3.2.4 The Chapman-Kolmogorov Equation 34
3.2.5 The Master Equation 35
3.3 Formulating the Model 39
3.3.1 Averages 40
3.3.2 Fraction of CSCs 44
3.3.3 Survival Rate 46
3.4 Gillespie's Algorithm 52

3.5 Comparison of Stochastic and Deterministic Models 57
3.6 A General Model 59
3.6.1 The role of progenitor cells 59
3.6.2 A simplified approach 68
3.6.3 Fraction of CSCs 71
3.6.4 Comparison of the Full and Condensed Models 73

4 Application: Treatment of Brain Cancer 76

4.1 Biological Orientation 77
4.1.1 Neural Stem Cells in Homeostasis 77
4.1.2 Brain Tumour Treatment and BTSCs 78
4.2 Modelling Brain Tumour Treatment 79

5 Sphere-Forming Assays, (Cancer) Stem Cells, and the Epithelial-Mesenchymal Transition 85

5.1 Sphere-Forming Assays 86
5.1.1 The Mammosphere Assay 89
5.1.2 Parameter Values 90
5.1.3 Simulation of the Mammosphere Assay 94
5.2 CSCs and the Epithelial-Mesenchymal Transition 104
5.2.1 Scenario One 106
5.2.2 Scenario Two 111
5.2.3 Discussion and Comparison of Scenarios 116

6 Conclusion and Future Work **121**

References **125**

List of Tables

5.1 Parameter values, $N = 4$ 96

5.2 Parameter values, $N = 5$ 96

5.3 Parameter values, $N = 6$ 97

5.4 Parameter values, $N = 7$ 98

List of Figures

2.1 The hematopoietic stem cell hierarchy 10

3.1 Three types of cancer stem cell division 29

3.2 Number of cells and fraction of CSCs for a given set of parameters . 44

3.3 Fraction of CSCs versus time 46

3.4 Survival rate versus time 53

3.5 Sample realizations of the stochastic process for different values of n_S^0 60

3.6 Illustration of cellular hierarchy for BTSCs 62

3.7 Number and fraction of cells versus time, $N = 5$ 65

3.8 Number and fraction of cells versus time, $N = 7$ 66

3.9 Comparison of full and condensed models, $N = 5$ 74

3.10 Comparison of full and condensed models, $N = 7$ 75

4.1 Theoretical treatment of brain tumour stem cells 83

5.1 Image of a mammosphere 87

5.2 Supplemental data of Mani *et al.* (Cell 2008) 93

5.3 10 000 simulations of the mammosphere assay, $N = 4$ 99

5.4 10 000 simulations of the mammosphere assay, $N = 5$ 100

5.5 10 000 simulations of the mammosphere assay, $N = 6$ 101

5.6 10 000 simulations of the mammosphere assay, $N = 7$ 102

5.7 Production of CSCs via EMT: Scenarios One and Two 106

5.8 EMT Scenario 1, $N = 4$ 108

5.9 EMT Scenario 1, $N = 5$ 108

5.10 EMT Scenario 1, $N = 6$ 109

5.11 EMT Scenario 1, $N = 7$ 109

5.12 EMT Scenario 1, $N = 5$, with time delay 110

5.13 EMT Scenario 2, $N = 4$ 112

5.14 EMT Scenario 2, $N = 5$ 113

5.15 EMT Scenario 2, $N = 6$ 114

5.16 EMT Scenario 2, $N = 7$ 115

5.17 EMT Scenario 2, $N = 5$, with time delay 117

5.18 Comparison of EMT Scenarios 1 and 2, $N = 5$ 119

Chapter 1

Introduction

The subject of cancer needs little introduction, given the extent to which it intrudes on human lives. As an elementary definition, a cancer is any one of a class of diseases characterized by uncontrolled cell growth and invasion into surrounding tissues. Over the course of the last decade, however, the classical notion of a cancer as being composed of malignant cells that are homogeneous with respect to their potential to divide without limit has been increasingly challenged. Instead, some scientists are viewing certain cancers as consisting of hierarchies of sub-populations of cells with varying abilities to proliferate. At the base of these hierarchies are cancer stem cells (CSCs), so called because of the properties they share with normal stem cells. This view is known as the cancer stem cell hypothesis, with cancer stem cells first being identified in various leukemias and more recently in a variety of solid tumours including those of the brain, breast, colon, and prostate; mathematical modelling of cancer growth based on the cancer stem cell hypothesis is the subject of this thesis.

One particularly devastating malignancy belonging to the group classified as

CSC disorders is glioblastoma multiforme (GBM), a disease that is at once the most common and the most aggressive primary brain tumour. GBM patients typically succumb to the disease within a year of diagnosis, even with aggressive treatment consisting of surgical resection followed by radiotherapy, chemotherapy or both – this poor prognosis has not been significantly improved upon through the course of decades of research.

It has been suggested that the failure of standard treatment protocols to eradicate GBM may be due to the presence of a sub-population of GBM cells that exercise resistance to conventional cytotoxic treatments, and it has further been postulated that cells from this same sub-population are uniquely responsible for initiating and maintaining tumour growth. These cells have been termed CSCs, with the designation deriving from the fact that their defining properties – namely, the abilities to self-renew (that is, to divide such that at least one daughter cell retains the CSC properties of its mother CSC) and to produce mature cells of various lineages – are also the defining properties of normal adult stem cells. Recently, human GBM cells isolated on the basis of expression of a particular membrane-associated protein were shown to have these cardinal stem cell properties, and hence were putatively identified as GBM stem cells. While the cancer stem cell hypothesis is entrenched in the theory of certain blood cancers, its extension to the field of solid-tumour biology is nascent, with putative identification and isolation of CSCs from breast and brain cancers occurring in 2003. Challenging the full-scale acceptance of the CSC hypothesis in solid-tumour biology are several objections raised by critics, including questioning of the relevance of the "gold-standard" assay in identifying CSCs – the implantation of human cancer cells into immunodeficient mice, a model system which may not necessarily test for the same growth properties that characterize a tumourigenic cell in a human host. Continued research by experimentalists aims

to address the validity of the CSC model in various human malignancies, and it is speculated by some that certain tumours will be found to adhere to the CSC hypothesis while others do not. Regardless, the additional scrutiny of cancers due to their being viewed in the light of the CSC hypothesis is likely to contribute to an enhanced understanding of the individual diseases.

The CSC hypothesis has some important implications for the study and treatment of cancers. Philosophically, it implies that a cancer is a form of aberrant organ growth, with malignant cells being caricatures of their normal counterparts, in which a hierarchy of cells is maintained – imparting the tumour with an order and contrasting the common view of cancer as the epitome of disarray. More practically, it suggests that current cytotoxic treatment strageties such as chemotherapy and radiotherapy that aim to indiscriminately kill dividing cells may be futile because they may spare CSCs. Another ramification of the CSC hypothesis is that the dissemination of a CSC from the primary tumour and its re-establishment in a secondary location is necessary for metastatic growth to occur. An interesting, albeit alarming, recent proposal is that the same mechanism which gives certain cancer cells the ability to disseminate (the "epithelial-mesenchymal transition", or EMT) may also endow these cells with the CSC phenotype.

This thesis, devoted to the mathematical modelling of the growth of cancer cells based on the CSC hypothesis, is organized into six chapters. Chapter Two is dedicated to previous work, and is divided into two sections. The first section presents a brief historical overview of the scientific developments leading from the initiation of the field of (normal) stem cell biology to the formulation of the CSC hypothesis for solid tumours; this section concludes with a review of some of the criticisms and unanswered questions associated with the cancer stem cell hypothesis. The second section of Chapter Two reflects on previous mathematical modelling – while math-

ematical modelling has made valuable contributions to many aspects of the study of cancers, we restrict the discussion herein to previous models in cancer biology and in normal stem cell biology that are pertinent to our consideration of the CSC hypothesis.

In Chapter Three, we simultaneously develop a discrete, stochastic model and a continuous, ordinary differential equation model of the growth of populations of cancer cells categorized into sub-populations of cells defined based on their positions within a CSC hierarchy. The stochastic model treats cell division as a random event, and is of particular relevance when the number of cells under consideration is small, while the continuous model represents a tractable simplification of the stochastic approach that is appropriate when large numbers of cells are being considered, as they often are in discussions of cancer. Here, we begin with a simple case in which a system is composed of two sub-populations of cells (CSCs and non-stem cancer cells), and extend this to the general case in which arbitrarily many sub-populations of cells are included.

Chapter Four is the first of two chapters in which the models of Chapter Three are borrowed for clinical or experimental applications. In Chapter Four, we discuss the treatment (and treatment failures) of GBM. In particular, the recent observations that GBM CSCs may be preferentially resistant to radiotherapy and that particular proteins (bone morphogenetic proteins) may induce GBM CSCs to shed their CSC phenotype via differentiation are considered. In Chapter Five, we examine the aforementioned link between the EMT and the stem cell phenotype, particularly in the context of cells closely related to breast cancer cells, with the purpose of attempting to distinguish between two proposed mechanisms by which experimentally-induced EMT may enrich the sub-population of mammary stem cells. To facilitate this, we also consider the mammosphere assay, an experimental

technique used in the culturing of both normal mammary stem cells and breast CSCs. Finally, in Chapter Six, we discuss conclusions and opportunities for future work.

Chapter 2

Background

2.1 A brief histoy of cancer stem cell biology

Here, we present a brief chronological history of the developments leading from the initiation of the field of stem cell biology in the mid-20th century to the formulation of the cancer stem cell hypothesis in solid tumours, beginning with (and focusing special attention on) the pioneering work of Till and McCulloch on hematopoietic stem cells. While this history is by no means complete, we attempt to highlight important discoveries and technological advancements of relevance to the chapters to come. It is useful at this point to make clear the following definition: a (cancer) stem cell is a cell that has the abilities to *self-renew* (by which we mean divide such that at least one daughter cell has the same stem cell properties as its mother cell) and to produce differentiated (mature) cells of a variety of lineages.

2.1.1 Till and McCulloch and the advent of stem cell biology

The field of stem cell biology began in earnest following the end of the Second World War, as scientists sought to understand the toxic effects of radiation that caused human death, both over short and long time periods, necessitating studies on mice. It was found that the symptoms of mice exposed to whole-body X-irradiation mimicked the human symptoms of radiation poisoning – the minimal lethal dose resulted in death due to failure of the hematopoietic (that is, blood-forming) system [Jacobson et al., 1949]. Interestingly, mice receiving the same irradiation dose recovered if either the spleen or a single large bone (both of which are sites of blood production) was shielded [Jacobson et al., 1949], suggesting that these sites harbour the capacity to replenish depleted populations of hematopoeitic cells.

The hematopoietic system is particularly vulnerable to irradiation because of the high frequency of division that its cells undergo, as the effects of irradiation establish themselves as toxic during the process of cell division. The utility of radiation in killing cells of another (albeit aberrant) system characterized by rapidly-proliferating cells – cancer – did not go unrecognized; hence, it became necessary to quantify the sensitivities of normal cell types to radiation in order to find a balance between the elimination of malignant cells and the sparing of well-behaved ones. To this end, biophysicist James Till and biologist/physician Ernest McCulloch conducted seminal research in which they measured the capacity of irradiated murine bone marrow cells to reconstitute the hematopoietic systems of irradiated mice [Till and McCulloch, 1961]. In this manner they intended to investigate the radiation sensitivity of proliferating cells – serendipitously, they also suggested the existence

of a hematopoietic stem cell.

Their experiments relied on a spleen colony-forming assay: it was known that injection of a sufficient number of bone marrow cells into sufficiently-irradiated mice led to the formation of nodules in the spleens of the recipient animals. Till and McCulloch (1961) counted the number of spleen nodules formed as a function of the number of irradiated bone marrow cells received by the transplanted mice (for various irradiation doses), and found that the relationship between these quantities was approximately linear, with the line-of-best-fit obtained via least-squares approximately passing through the origin – suggesting, although not proving, that a single bone marrow cell could give rise to a spleen nodule.

Proof that the spleen colonies were clonal in origin (that is, that a single spleen colony was generated in its entirety by a single hematopoietic cell) did follow soon after, again from the laboratory of Till and McCulloch (1963). Murine bone marrow cells were injected into recipient mice that had received a small pre-injection dose of x-irradiation; shortly following the injection, the recipient mice were again irradiated, so that in total, they received sufficient irradiation to suppress their hematopoietic systems, thus rendering the animals' hematopoietic systems susceptible to reconsitution by the transplanted cells. The injected bone marrow cells received only a part of this total irradiation, bestowing on the surviving cells unique karyotypic signatures (chromosomal abnormalities) as a consequence of randomly-induced mutations. Eleven days later, the mice were euthanized and their spleen nodules dissociated. Examination under the microscope revealed that, in spleen colonies exhibiting distinguishable configurations of chromosomes, the "overwhelming majority" of cells in a given colony exhibited the characteristic chromosomal abnormality [Becker et al., 1963]. This provided direct support for the conclusion that an entire spleen colony, typically consisting of over one million cells after a pe-

riod of ten days, could originate entirely from one cell. This "colony-forming cell" reflected the definition of a stem cell in other ways besides its obvious potential for sustained proliferation – it exhibited self-renewal (since other colony-forming cells were found to be present in spleen colonies) and the capacity to produce differentiated cells of various lineages (myelocytes and erythrocytes of various levels of maturity were observed [Till and McCulloch, 1961]). These cells were also found to be quite rare, with a frequency of about one colony-forming cell per 10^4 bone marrow cells [Till and McCulloch, 1961].

2.1.2 Subsequent characterisation of hematopoietic stem cells

The seminal work of Till and McCulloch in the early 1960s demonstrated the existence of hematopoietic stem cells (HSCs) and afforded them a functional definition, but it did not provide a method of isolating these cells so that they could be studied directly [Weissman, 2002]. Over the next three decades, a succession of studies by various researchers gradually refined the scientific community's knowledge of the HSC phenotype [Weissman, 2002], leading to the purification of mouse HSCs (on the basis of expression of a variety of cell-surface-associated proteins) by Irving Weissman's group in 1988 [Spangrude et al., 1988]. Building on this work on murine stem cells, subsequent research has steadily been conducted in order to elucidate the identity and function of human hematopoietic stem cells. Figure 2.1 illustrates the current state-of-knowledge of the hierarchy of the hematopoietic system, with HSCs at the base.

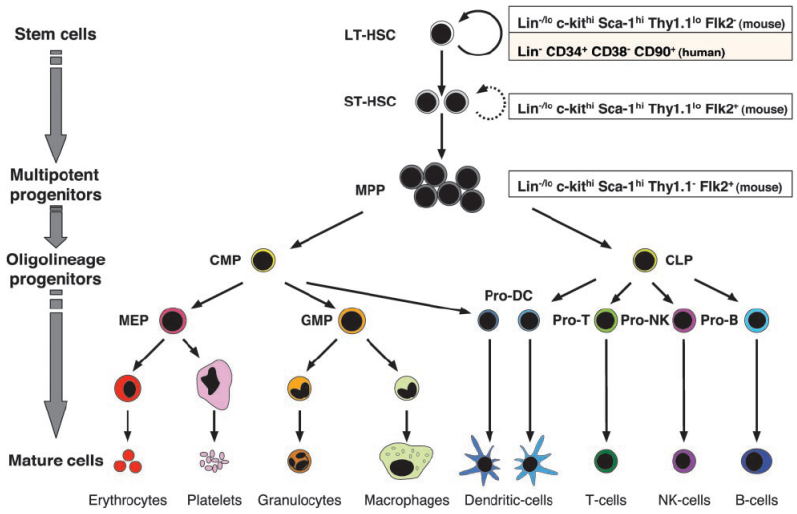


Figure 2.1: An illustration of the hierarchy of cells originating from the hematopoietic stem cell, used with permission from [Passegué et al., 2003]; copyright 2003 National Academy of Sciences, U.S.A.

2.1.3 Recognition of heterogeneity within tumours

Concurrent with the burgeoning of the field of stem cell biology, various independent researchers began to note the apparent heterogeneity of cancer cells with respect to their potential to initiate new tumours. In 1961, prompted by the observation that disseminated tumour cells are commonly found in the blood and lymph fluid of cancer patients yet actual metastatic growths are relatively rare, Southam and Brunschwig (1961) reported results of an experiment in which they auto-transplanted cancer cells into various locations under the skin of terminal cancer patients [Southam and Brunschwig, 1961]. (The experiments of Southam and Brunschwig (1961) are interesting today not only for the results they yielded but also for their methodology, which would be unjustifiable by contemporary ethical standards.) The authors found that at least 10^6 (unsorted) cells were required to

initiate a secondary tumour, suggesting that not all cells of the original tumour were capable of initiating tumour growth. On the heels of this work, Bruce and van der Gaag [Bruce and van der Gaag H., 1963] demonstrated in 1963 that only a fraction (1-4%) of murine lymphoma cells were capable of forming spleen colonies [similar to those described by Till and McCulloch (1961)] in recipient mice.

While these reports by Southam and Brunschwig (1961) and Bruce and van der Gaag (1963) only suggested the probable existence of cancer stem cells, McCulloch and his colleagues were writing freely of "tumor stem cells" by the 1970s, with McCulloch *et al.* describing in 1971 their observation that only about 1% of murine myeloma cells formed colonies *in vitro* [Park et al., 1971]. By the late 1970s, Hamburger and Salmon had developed their own *in vitro* colony-forming assay; using this experimental system, they extended the demonstration of proliferative heterogeneity of cancer cells from lymphomas and leukaemias to solid tumours, showing that for various solid tumours, only 1 in 1000 to 1 in 5000 tumour cells was capable of forming colonies [Hamburger and Salmon, 1977].

2.1.4 Technology

A handful of important technological advances were instrumental in facilitating the refinement of the phenotypic definition of normal and cancer stem cells, both in mice and men. One of these was the development, by Köhler and Milstein in 1975, of a process by which monoclonal antibodies specific to any cell surface protein could be produced (monoclonal antibodies are proteins engineered to be produced by a colony of immune-system cells) [Kohler and Milstein, 1975]. This gave experimentalists the ability to 'tag' sub-populations of cells with certain cell-surface protein expression profiles. These tagged cells could then be sorted through fluorescence-

activated cell sorting (FACS), a process then newly-introduced by Herzenberg and colleagues [Bonner et al., 1972] allowing for the sorting and separation of single, viable cells.

As direct studies of human hematopoietic repopulation by putative human hematopoietic stem cells after irradiation were (and still are) impractical for a number of (mostly ethical) reasons, studies of human stem cells necessitated the engineering of strains of mice able to tolerate the introduction of foreign cells. This was accomplished with the development of severe combined immunodeficient (SCID) and non-obese diabetic severe combined immunodeficient (NOD/SCID) mice as model systems in the late 1980s and early 1990s, respectively (see, e.g., [Kamel-Reid and Dick, 1988]).

2.1.5 Cancer stem cells in leukemias and in solid tumours

While the works listed above (and others like them) indicated that different cancer cells from the same tumours varied in their abilities to maintain continued proliferation, they did not specify *which* cells were cancer stem cells. The first definitive demonstration of a cancer stem cell came in the mid-1990s, with the identification of the human acute myeloid leukemia (AML) stem cell by John Dick and colleagues, who experimented with NOD/SCID mice [Lapidot et al., 1994]. These cells, termed more specifically SCID leukemia-initiating cells (SL-ICs), were isolated from human leukemias on the basis of their expressing the cell-surface protein CD34 while *not* expressing the cell-surface protein CD38 (a phenotype usually denoted $CD34^+/CD38^-$). It was found that injection of as few as 5000 $CD34^+/CD38^-$ human AML cells could initiate leukemia in NOD/SCID mice, whereas injection of up to 100 times as many $CD34^-/CD38^+$ cells failed to do so. Thus, it was determined

that the SL-ICs lie in the $CD34^+/CD38^-$ fraction, which represented anywhere from 0.02-1% of the total fraction of leukemia cells in various (human) patients. Crucial to the characterization of the SL-IC as a "cancer stem cell", these cells were shown to have marked capacity to differentiate (producing a myriad of cell types *in vivo* that were representative of the constitution of the original leukemia), and demonstrated the ability to self-renew (forming secondary leukemias after extraction from the first murine host and introduction into new mice). Over the next several years, the cancer stem cell hypothesis became firmly intrenched in the accepted understanding of acute myeloid leukemia as the phenotype and functional role of the SL-IC was refined. The full blossoming of the cancer stem cell field, however, would require the discovery of cancer stem cells in malignancies beyond those of the blood.

In 2003, the first report of cancer stem cells in a solid tumour came from the laboratory of Michael Clarke [Al-Hajj et al., 2003], whose team prospectively identified tumourigenic breast cancer cells, again on the basis of a particular phenotype (these cells express the cell-surface protein CD44, while lacking or having low expression of CD24 and no expression of any lineage marker proteins indicating cellular maturity – hence, they are denoted $CD44^+/CD24^{-/low}Lineage^-$). As in the case of the the SL-IC, Clarke's group employed a NOD/SCID mouse model, into which they xenografted human breast cancer cells. Within a year, a human brain tumour stem cell was also identified, this time by the group of Peter Dirks [Singh et al., 2004]. Subsequently, cancer stem cells have been putatively identified in many types of solid tumours, including prostate [Collins et al., 2005], colon [O'Brien et al., 2007], pancreatic [Li et al., 2007], head-and-neck [Prince et al., 2007] and lung [Eramo et al., 2008] cancers.

2.1.6 The origins of cancer stem cells

The connotations of the term "cancer stem cell" have led to some confusion regarding the role of CSCs in oncogenesis (the development of cancer). We point out that CSCs are distinct from the "cell of origin" of a cancer, with the latter term referring specifically to the cell that is the subject of the first mutation on the genetic road to malignancy [Visvader and Lindeman, 2008].

The question, "from which type of cell does a CSC arise?" is another continuing source of debate. Those same functional parallels between normal stem cells and CSCs that lend CSCs their name make it attractive to suppose that CSCs arise by transformation of normal stem cells. Normal stem cells, by virtue of the property of self-renewal, have the longevity that may be necessary to acquire a series of mutations that leads to oncogenic transformation, whereas more mature (and hence relatively short-lived) cells may not. Furthermore, it is logical to suppose that fewer changes (i.e. fewer mutations) would need to occur in a normal stem cell for that cell to gain the functionality of a CSC, than for a more mature cell that is not already in command of the powers of self-renewal and multilineage differentiation to produce the same CSC [Dirks, 2008]. While these arguments exist, they still remain largely philosophical and, indeed, evidence exists (particularly in leukemias) suggesting that many CSCs arise from the transformation of more mature progenitor cells rather than from stem cells [Krivtsov et al., 2006, Lobo et al., 2007].

2.1.7 Criticism

While the CSC hypothesis seems to be generally accepted for certain leukemias [Kern and Shibata, 2007], its relevance to solid tumour growth remains controversial. Critics are hesitant to adopt the CSC hypothesis on the basis of several

arguments. The most superficial of these pertains to nomenclature: the "cancer stem cell" moniker is objectionable to some on the grounds that it can be confused with the "cell of origin" [Kennedy et al., 2007], as discussed above; others point out that, as the full extent of the mechanistic analogy between normal and cancer stem cells remains to be elucidated, terms such as "tumour-initiating cell" may be more appropriate [Hill and Perris, 2007].

Concerns not grounded in semantics exist, as well. Virtually all demonstrations of CSCs in solid tumours have adhered to the experimental framework of sorting human tumour cells based on phenotypic markers, and then implanting sub-populations of cells into NOD/SCID mice. Critics fear (and validly so) that such xenograft assays test not a sub-population of cells' true tumourigenic potential, but rather the ability of cells to adapt to the (foreign) murine environment [Shipitsin and Polyak, 2008]. The NOD/SCID xenograft model, then, may underestimate the fraction of tumour cells that truly have the ability to initiate tumour growth. Such criticism is suggestive of supposition of an alternative model for tumour growth, namely the "stochastic model", which asserts that the majority of tumour cells have the ability to self-renew and initiate tumours, but due to various external factors each cell has only a small probability of demonstrating this in any given assay [Adams and Strasser, 2008, Reya et al., 2001].

The stochastic model can be considered as a feature of the clonal evolution model [Visvader and Lindeman, 2008], under which tumour heterogeneity arises as a consequence of different clones (harbouring different genetic mutations) being at different stages of malignant transformation [Adams and Strasser, 2008]. Mounting evidence suggests that some cancers may strictly adhere to the clonal evolution/stochastic model while others strictly follow the cancer stem cell model, with other cancers representing various hybrids of the two viewpoints [Adams and

Strasser, 2008]. For example, in such a hybrid model, the genetic instability that is fundamental to the clonal evolution model may have the effect of forcing clonal evolution on CSCs themselves [Visvader and Lindeman, 2008].

Another uncertainty with the experimental results supporting the CSC hypothesis is an apparent inconsistency in the reported numbers of CSCs required to form tumours in mice, as pointed out by Richard Hill [Hill and Perris, 2007]. For example, in their putative identification of a colon cancer stem cell on the basis of expression of the marker protein CD133, O'Brien *et al.* found that injection of as few as 200 $CD133^+$ CSCs into NOD/SCID mice resulted in tumour initiation, whereas about 60 000 cells from the unsorted population (comprising a mixture of both $CD133^+$ and $CD133^-$ cells) were required to do the same [O'Brien et al., 2007]. With roughly 12% of cells in the unsorted population expressing CD133 this means that in the latter case, about 7000 putative CSCs were required for tumour initiation – a value which contrasts that of 200 CSCs from the sorted population. While this apparent disparity may have a biological explanation that does not violate the CSC hypothesis – it has been postulated that the presence of $CD133^-$ cells downregulates the tumourigenicity of $CD133^+$ CSCs, or alternatively that a small number of cells (e.g. 200) can escape immune detection when injected into mice whereas larger numbers elicit an immune response [Hill and Perris, 2007] – it is still a concern.

It has generally been presumed that CSCs constitute a rare sub-population of tumour cells, and one of the implications of the CSC hypothesis is that it is this rare subpopulation that needs to be understood and targeted therapeutically. Recently, however, experiments by the research group of Sean Morrison have cast doubt on the presupposed rarity of cells with tumour-initiating ability in melanoma [Quintana et al., 2008]. Melanoma is a disease reported to be driven by CSCs, based on

the observation that only roughly one per million human melanoma cells was found able to initiate tumour growth in NOD/SCID mice [Schatton et al., 2008]. Quintana *et al.* (2008) xenotransplanted human melanoma cells into a modified strain of NOD/SCID mice that also lacked natural-killer cell activity (an important class of cells of the immune system), and furthermore co-injected the melanoma cells with Matrigel (a protein mixture resembling typical extracellular environments). In this modified xenotransplantation system, melanoma cells exhibited a marked increase in the ability to initiate tumour growth – as many as one in four melanoma cells formed tumours, even when implanted as single cells. Tumour-initiating cells also varied widely with respect to phenotypic characteristics such as cell-surface protein expression. These results demonstrate that the identification and observed rarity of melanoma stem cells by Schatton *et al.* (2008) are artifacts of the NOD/SCID xenotransplantation assay, and that a new assay yields an interpretation more consistent with the stochastic model of tumour growth. Thus, this seems to argue against the existence of CSCs in at least a particular class of melanomas, or otherwise at least diminish the significance of CSCs in this particular disease – although it has been pointed out that rarity is not necessarily a defining property of CSCs [Baker, 2008], if all or most cancer cells have tumourigenic ability then the traditional strategy of focusing on therapies that simply eradicate the bulk of tumour cells is not threatened by the CSC hypothesis.

Quintana *et al.* (2008) have highlighted the importance of xenotransplantation techniques in dicating the growth properties of xenotransplanted cells. It will be interesting to see if the use of novel assays leads to the observation of similar tumour-initiating ubiquity in other cancers previously hypothesized to be driven by (rare) CSCs. Various investigators, including Quintana *et al.* (2008), suggest it is likely that some cancers follow a CSC model, while others do not. Taken together,

these criticisms and unresolved questions illustrate some of the complexities of contemporary cancer research and highlight the need for continued investigation.

2.2 Mathematical Modelling of (Cancer) Stem Cells

Mathematical modelling has made significant contributions to the collective understanding of cancer biology since the pioneering work of Nordling [Nordling, 1953] and Armitage and Doll [Armitage and Doll, 1954] in deciphering cancer age-incidence data to conclude that multiple mutations are required for full-blown malignancy to occur. The spectrum of cancer-related topics addressed mathematically by various groups over the course of the last five decades is too broad for consideration in this thesis. Instead, we focus on mathematical modelling of relevance to the cancer stem cell hypothesis; even with this restriction we inevitably and regrettably neglect the efforts of some. The following is a brief description of some of the mathematical modelling related to the contents of the coming chapters.

2.2.1 Deterministic Modelling of CSCs in Leukemias and the Brain

As the architecture of the (normal) hematopoietic system became increasingly elucidated through the decades following the seminal work of Till and McCulloch (1961), interdisciplinary researchers began employing mathematical models to interpret, and extrapolate, their results. Throughout the late 1970s and early 1980s, H.-Erich Wichmann and Markus Loeffler, in particular, developed a compartmental

model of hematopoietic stem cell regulation which they then subjected to various biologically-relevant insults in order to predict results [Wichmann and Loeffler, 1985]. Their model was a deterministic, compartment-based model featuring six comparments of hematopoietic cells (pluripotent stem cells, early and late erythropoeitic progenitor cells, erythropoietic precursor cells, granulopoietic progenitor cells, and granulopoietic precursor cells); these compartments were interrelated through various feedback mechanisms, with the population of cells in each compartment being governed by an ordinary differential equation (spatial variations were not considered).

While the work of Wichmann and Loeffler (1985) is noteworthy in hematolagy for its depth and scope, its significance in the context of this thesis lies in its uptake two decades later by Ganguly and Puri, who adapted it to a discussion of cancer stem cells [Ganguly and Puri, 2006]. Ganguly and Puri considered four compartments of normal cells – stem cells (SC), early progenitor cells (EP), late progenitor cells (LP) and mature cells (M) – with the division properties of each of the first three compartments governed in part by feedback from other compartments. Three additional compartments of abnormal cells (SC_A, EP_A, and abnormal precursors AP) were introduced, with cells being transferred from the SC compartment to the SC_A compartment and from the EP compartment to the EP_A compartment, each with some mutation rate – these abnormal cells were assumed to be regulated according to the same mechanisms as normal cells, and their associated parameters were taken to have the same values as their normal counterparts. Among the authors' conclusions were that mutations in the SC compartment lead to faster growth of abnormal progeny than do mutations in the EP compartment, and that under the assumed regulatory system, an insult that depletes the number of normal mature cells (e.g. radiotherapy) ultimately leads to an increase in the production

of abnormal progeny. In a later paper, Ganguly and Puri included the effects of chemotherapy [Ganguly and Puri, 2007], concluding that while targeting of only the AP sub-population leads to an immediate arrest of AP cell growth, targeting of only the SC_A sub-population ultimately results in lower cell numbers although the time period before these effects are observable is greater than in the former case. We point out that Ganguly and Puri [Ganguly and Puri, 2006, 2007] have assumed the same hierarchy and regulatory networks as in the hematopoietic system in their consideration of non-specific cancer growth based on the CSC hypothesis, while the precise hierarchies and mechanisms of regulation of most solid tumours implicated in the CSC hypothesis (e.g. those of the brain) remain to be elucidated.

ODE modelling has been used recently to investigate the role of leukemic stem cells (LSCs) in the treatment of chronic myeloid leukemia (CML) by Michor *et al.* [Michor et al., 2005a], who divided each of the three populations of normal hematopoietic, leukemic, and drug-resistant leukemic cells into four compartments (stem, progenitor, differentiated and terminally-differentiated cells) governed by ODEs. Comparison of the model results to clinically-derived measurements of CML cell numbers during the course of treatment with the anti-CML drug imatinib led Michor *et al.* [Michor et al., 2005a] to conclude that imatinib, while effective in depleting leukemic progenitor cells and differentiated leukemic cells, does not deplete the sub-population of LSCs.

An alternative interpretation of the imatinib-treated CML data was offered by Komarova and Wodarz, who considered only sub-populations of actively-dividing LSCs and quiescent LSCs (a "quiescent" cell is one that has temporarily exited the cell cycle and hence is insensitive to anti-cancer therapies such as imatinib that target actively-dividing cells) [Komarova and Wodarz, 2007]. Applying ODE modelling to these compartments, Komarova and Wodarz (2007) suggested that

imatinib *does* deplete the LSC sub-population and that extended use of the drug may eventually lead to eradication of the malignancy as quiescent LSCs are drawn back into the actively-cycling sub-population.

CML progresses through three clinically-defined phases: the chronic phase, the accelerated phase, and blast-crisis [Michor, 2008]. It is thought that the accumulation of additional mutations facilitates the progression from one phase to the next, and the question of which cellular compartment is the target of the blast-crisis-initiating mutation was addressed by Michor [Michor, 2007]. Again using ODE modelling to describe the growth of two sub-populations of CML cells – LSCs and leukemic progenitor cells – Michor concluded through comparison of her model-predicted rates of blast-crisis initiation with clinically-measured rates that the blast-crisis precursor cell likely lies in the leukemic progenitor cell sub-population, rather than the LSC sub-population.

2.2.2 Modelling the Role of Stem Cells in Colorectal Cancer

Deterministic compartmental models have also been used in the context of colorectal cancer (CRC). Although a colon cancer stem cell was not identified until 2007 [O'Brien et al., 2007, Ricci-Vitiani et al., 2007], colorectal cancer has long been thought of as a stem cell disorder, due to the long-standing knowledge that the lining of the colon is arranged into many "crypts", each of which is maintained by normal stem cells. This makes the case of colon cancer somewhat unique among solid tumours, in that it was considered a disease involving cancerous stem cells before the entrenchment of the CSC hypothesis in the solid tumour literature. A significant body of mathematical work has been devoted to understanding colonic

crypt stem cell dynamics in homeostasis; such work is reviewed in [van Leeuwen et al., 2006] and [Johnston et al., 2007b], for example. We here turn our attention to models relevant specifically to colorectal cancer.

Johnston *et al.* developed a compartmental model of colonic cell populations in homeostasis and cancer that kept track of three interrelated sub-populations – stem cells, transit (i.e. progenitor) cells, and fully differentiated (i.e. mature) cells [Johnston et al., 2007a]. Their model was age-structured, in the sense that in addition to counting the numbers of cells in each compartment, the authors kept track of the age distributions of the sub-populations. To complement their analysis of the age-structured model, Johnston *et al.* (2007) simultaneously developed a continuous ODE model, describing the same three cellular sub-populations. By imposing various forms of theoretical regulation on the divisions of stem and progenitor cells, the authors sought conditions (i.e. parameter changes) that led to breakdown of regulation and hence unbounded growth (interpreted as cancer). The work of Johnston *et al.* (2007) built on earlier efforts by Tomlinson and Bodmer [Tomlinson and Bodmer, 1995], who had considered a related compartmental model to predict that the uncontrolled growth observed in colon cancer could be attributed to the failure of stem cells to differentiate or undergo apoptosis (programmed cell death).

It is of note that the CSC hypothesis, while not being explicitly discussed, has been taken as implicit in much of this colorectal cancer modelling, in the sense that CRC is viewed as resulting from, and being maintained by, a dysregulated population of stem cells. The first modelling of colorectal cancer done explicitly in the context of the CSC hypothesis, evidently, was by Boman et al. [Boman et al., 2007]. Boman and colleagues presented a four-compartment ODE model in order to determine what changes could account for the dramatic expansion of the numbers of stem cells from the healthy colonic crypt to the pathological case of colorectal

cancer. They considered stem cells, intermediate proliferating cells (analogous to progenitor cells), non-proliferating cells (analogous to fully mature cells or cells that are in a state of irreversible quiescence), and eradicated (dead) cells. Performing a computational analysis on the ODEs they formulated, Boman *et al.* concluded that an increase in the rate of symmetric self-renewal of stem cells is key in facilitating the aforementioned expansion of stem cells.

2.2.3 Stochastic Modelling of the Origins of Cancer

In recent years, a large body of work has come forth from Harvard University's Program for Evolutionary Dynamics, headed by Martin Nowak, and affiliates. Much of this group's research has employed stochastic modelling of stem cell dynamics to investigate the role of mutations in various cell compartments in the initiation and progression of cancer. (While the class of research directed at understanding the origins of cancer from a normal stem cell hierarchy may not be cancer stem cell modelling *per se*, we devote some attention to it here because of the frequency with which it arises in a search for "cancer stem cell" modelling literature, and because of its possible relevance to future modelling of CSC dynamics that includes such features as clonal evolution.)

A Moran process is a stochastic process in which individual objects (e.g. stem cells) are randomly chosen to reproduce and are randomly chosen to be removed from the system such that the total number of individuals at any given time remains *constant* [Michor et al., 2006]. While, clearly, the restriction that the number of cells remains constant makes this model inappropriate for consideration of a growing tumour, the population of normal stem cells under homeostasis can be considered as a Moran process . Indeed, Michor *et al.* have made this assumption

in addressing the origins of CML, a disease to which the *BCR-ABL* fusion mutation (the "Philadelphia chromosome") is infamously linked [Michor et al., 2006]. Assuming that hematopoietic stem cells divide according to a Moran process, but that at each division there is a probability that a cell carrying the Philadelphia chromosome will arise and that the acquisition of the Philadelphia chromosome increases the reproductive "fitness" of a stem cell, Michor *et al.* [Michor et al., 2006] proposed that a single mutation (that is, the Philadelphia chromosome mutation) is sufficient to match the age-incidence data of CML diagnoses. This prediction is significant because many cancers, including CML, are often thought to require several mutations before being characterized as malignant [Michor, 2008]. We note that Michor and colleagues have used similar models to investigate the dynamics of genetic events involved in colorectal cancer initiation, as summarized in the review articles [Michor et al., 2004] and [Michor et al., 2005b].

2.2.4 Modelling the Growth of (Cancer) Stem Cells as a Birth-and-Death Process

Till and McCulloch quickly followed their proof of the clonal nature of spleen colonies [Becker et al., 1963] with a mathematical study inquiring into the nature of regulation of the newly-discovered HSC [Till et al., 1964]. The observation that the number of spleen colony-forming cells per colony varied markedly from colony to colony suggested that control of stem cell division may be somewhat slack, leading Till *et al.* to propose that hematopoietic stem cells may be mathematically modelled by a "birth-and-death" process, a continuous-time stochastic process in which the quantity of interest (in this case, the number of HSCs) may during a single stochastic event vary only from the (integer) state n to $n + 1$ (a "birth")

or from n to $n-1$ (a "death" – birth-and-death processes will be considered in greater detail in the following chapter). In this context, a "birth" is a symmetric self-renewing HSC division in which one HSC divides to give two daughter HSCs, while a "death" is the differentiation of a HSC (asymmetric divisions in which an HSC divides to yield one daughter HSC and one differentiated daughter cell were not taken into account); the birth (death) of an individual HSC was assumed by Till *et al.* to occur with constant probability p_2 (p_0). Comparison of simulation of this birth-and-death process with experimental measurements of the numbers of colony-forming cells per colony suggested that the birth-and-death process does indeed serve as an appropriate model of HSC proliferation, giving credence to the idea of stem cell divisions involving randomness.

Sachs *et al.* used the birth-and-death process as a model in cancer biology to study the effects of dose timing during tumour radiotherapy [Sachs et al., 1996]. Here, CSC terminology was not explicitly used, but rather the number of "clonogens" – cells capable of regenerating the tumour – was modelled during the late stages of radiotherapy, at which time it is assumed that the number of clonogens is small and hence stochastic fluctuations may be important. Assuming that, independent of irradiation, clonogens have constant birth and death rates b and d respectively, and that irradiation with (not necessarily constant) dose rate $\dot{D}(t)$ is added to the death rate during the course of radiotherapy, Sachs *et al.* (1996) solved for the tumour control probability (TCP, defined as the probability of a tumour having zero clonogens as time $t \to \infty$). It was determined that the TCP is higher when the irradiation dosage is front-loaded (i.e. made higher at the beginning of the treatment schedule).

In research motivated by (but not directly related to) cancer, the group of Philip Jones used a variation of the birth-and-death process to put the hierarchy of cells

comprising the skin of a mouse's tail under scrutiny [Clayton et al., 2007]. The conventional model of the maintenance of murine epidermal homeostasis insists that a sub-population of stem cells self-renews asymmetrically while producing a sub-population of progenitor cells, which in turn divide a small number of times before differentiating to give mature cells that eventually migrate out of the basal layer of the epidermis [Clayton et al., 2007]. Clayton *et al.* (2007) proposed instead a model in which only one sub-population of epidermal cells is actively dividing in the basal layer: a sub-population of self-renewing epidermal "progenitor" cells (these cells have the characteristics of stem cells and thus may be called epidermal stem cells – we caution that the term "progenitor cell" is often used ambiguously in the literature, adding confusion to an already-confusing subject). These epidermal progenitor cells give rise to a sub-population of non-proliferating mature cells. Consideration of a stochastic process in which progenitor cells can self-renew symmetrically to produce two daughter progenitor cells with constant probability r, divide symmetrically to yield two mature cells with probability r, or self-renew asymmetrically to produce a progenitor cell and a mature cell with probability $(1-2r)$ led Clayton *et al.* (2007) to conclude that this model featuring a single class of progenitor cell provided a better fit to their data than did the existing stem cell-to-progenitor cell-to-mature cell model. Clayton *et al.* (2007) enforced homeostasis by requiring that the probability of symmetric progenitor cell self-renewal equalled the probability of symmetric progenitor cell differentiation. In the following, we develop a similar model without this restriction, allowing for analysis of tumour growth based on the cancer stem cell hypothesis.

Chapter 3

Modelling Cancer Stem Cells

Herein, we present a general model describing the growth of cancer cells based on the cancer stem cell hypothesis. In the following chapters, this model will be adapted to a handful of applications. Our initial interest in the CSC hypothesis was piqued by the recognition of the role of CSCs in brain tumours, and by subsequent discussions with experimentalists and clinicians working in this field. However, the model framework is intended to be relevant to CSC dynamics in any solid tumour. Portions of this chapter are related to work presented in [Turner et al., In press].

3.1 A simple model

Our goal is to develop a mathematical model describing the number of cells in a solid tumour, the growth and maintenance of which are governed according to the cancer stem cell hypothesis. In the simplest case, we may consider two sub-populations of cancer cells: a sub-population of cancer stem cells (CSCs), denoted by S, and a sub-population of non-stem cancer cells, which we denote by P for

reasons which will later become apparent. At the simplest level, we have a CSC hierarchy of the form

$$S \longrightarrow P,$$

wherein CSCs give rise to non-stem cancer cells. We now consider the details of this hierarchy.

3.1.1 Division properties of cancer stem cells

The unique ability of CSCs to both drive and maintain tumour growth appears to be a function of these cells' capacity for two distinct types of self-renewal [Morrison and Kimble, 2006]. CSCs can undergo symmetric self-renewal, a cell division in which both daughters possess the stem cell characteristics of the mother CSC, resulting in expansion of the CSC population, or asymmetric self-renewal, in which one CSC and one non-stem cancer cell are produced [Morrison and Kimble, 2006]. In addition to self-renewal, CSCs may permit symmetric division yielding two non-stem daughter cells; we shall refer to this event as "symmetric commitment" [Morrison and Kimble, 2006]. Even the simplest model of cancer cell population dynamics based on the CSC hypothesis should include these three types of division, which are illustrated in Figure 3.1. (We note that normal stem cells also employ these three types of division in the maintenance of homeostasis [Morrison and Kimble, 2006].)

We assume that CSCs divide with some division rate ρ_S, measured in units of day^{-1}. That is, a CSC undergoes any one of the three aforementioned types of division with rate ρ_S. We further assume that the "decision" to undergo symmetric self-renewal is made with probability r_1, while asymmetric self-renewal and symmetric commitment are undergone with probabilities r_2 and r_3, respectively. We

(a) (b) (c)

Figure 3.1: The three types of CSC division (CSCs shown in blue; non-stem cancer cells in gold): (a) symmetric self-renewal (b) asymmetric self-renewal (c) symmetric commitment.

normalize these three probabilities:

$$r_1 + r_2 + r_3 = 1.$$

The overall CSC division rate $\rho_S = \rho_S(r_1+r_2+r_3)$ can, if one prefers, be abandoned as a concept in favour of considering the rates $\rho_S r_1$, $\rho_S r_2$, and $\rho_S r_3$ for symmetric self-renewal, asymmetric self-renewal, and symmetric commitment, respectively.

3.1.2 Properties of Non-Stem Cells

Non-stem cancer cells, P, are produced by CSCs via asymmetric self-renewal and symmetric commitment. (The notation P reflects the fact that the cells downstream of stem cells in a stem cell hierarchy are generally referred to as "progenitor cells" – we will consider progenitor cells more explicitly later.) At the present time (i.e. in the two-compartment model that is currently under development), we do not consider the divisions of these non-stem cancer cells, as under the CSC hypothesis only CSCs are capable of initiating and maintaining tumour growth. In general, the progeny P of CSCs are thought to divide only a limited number of times before terminally differentiating – this effect will be considered later. For now, P cells may

die with rate Γ (units of day^{-1}). Note that we are considering CSCs to be immortal in the sense that they do not die but rather are lost only to differentiation.

To summarize thus far, we are (for the time being) interested in a two-compartment cancer stem cell hierarchy in which the following division pathways are active:

$$\begin{aligned} S &\xrightarrow{\rho_S r_1} S + S \\ S &\xrightarrow{\rho_S r_2} S + P \\ S &\xrightarrow{\rho_S r_3} P + P \\ P &\xrightarrow{\Gamma} \emptyset. \end{aligned} \tag{3.1}$$

3.1.3 On the stochasticity of cell division

Here, we develop a discrete stochastic model of cancer stem cell dynamics in order to shed light on some of the mechanisms driving tumour growth as well as the implications of these mechanisms on treatment strategies. Such an approach allows us to incorporate the inherent stochasticity of biological phenomena and, equally importantly, allows for meaningful examination of small numbers of cells, where a deterministic method fails. A similar technique was employed, for example, by Clayton *et al.* in modelling cell populations in normal murine epidermal homeostasis [Clayton et al., 2007]. From this stochastic model, we will derive (deterministic) equations for the average numbers of cells in particular compartments; these deterministic equations will be useful later, when the behaviour of large numbers of cells is studied.

Before progressing, we should comment on our assumptions regarding the randomness of cell divisions. First, we have assumed that the choice of type of stem cell division (symmetric self-renewal, symmetric commitment, or asymmetric division)

is a random, rather than a deterministic, process. This is tantamount to assuming that cell fate is assigned randomly. In many cases, such as in the development of the nervous system of *Drosophila melanogaster* (the common fruit fly), cell fate appears to be assigned in a pre-programmed, deterministic manner [Losick and Desplan, 2008]. In other systems, however, cell fate appears to have a stochastic component [Losick and Desplan, 2008]. Decades ago, Till and McCulloch [Till et al., 1964] found evidence that hematopoietic stem cells undergo either self-renewal or differentiation randomly, and conjectured that perhaps control mechanisms act by biasing the associated probabilites; more recently it has been proposed that stochastic fluctuations may be crucial in dictating the fate of hematopoietic progenitor cells [Laslo et al., 2006]. Recent work tracking cell divisions in the colonic crypts of mice suggests that colon stem cells may exhibit stochasticity in undergoing symmetric versus asymmetric divisions [Yatabe et al., 2001]. While it is still possible that the decision to differentiate or self-renew is essentially a deterministic process that depends on a multitude of extra- and intracellular variables, neither the nature nor the existence of such regulation has yet been established (although, for example, it has recently been reported that the elasticity of the surrounding microenvironment may be important in dictating the fate of the progeny of mesenchymal stem cells [Engler et al., 2006]). Accordingly, we adopt a stochastic approach to stem cell division, which is reflective of the current state-of-knowledge. We speculate that such an assumption may be particularly valid in cancers or *in vitro*, the former of which is characterized by its apparent lack of regulation and the latter also being an environment in which homeostatic conditions are presumably not maintained. It is further assumed that the process of cell division is a Markov process, as is discussed in the following section.

3.2 Markov Processes

This section is concerned with formulating the process of interest, namely the process of growth of a population of cancer cells under the CSC hypothesis, in such a way that it is amenable to some sort of mathematical treatment. To this end, we frame the process as a Markov process, and use this as the basis for our analysis. We begin with some preliminary concepts.

3.2.1 Random Variables

A *random variable* X is a quantity which (i) assumes a value x from a set of possible values called the "state space" (or alternatively, the "sample space", "range", "set of states", or "phase space" [Van Kampen, 2007]), and which (ii) has an associated probability distribution function

$$F(x) = P(X \leq x),$$

which is interpreted in the usual way as reading, "the probability that the random variable X takes a value less than or equal to x." The state space may be continuous or discrete; when it is continuous, we typically speak of the probability density function $f(x)$ associated with the probability distribution function,

$$f(x) = \frac{dF}{dx}.$$

When the state space is discrete, as it will be in our investigations, it is convenient to instead consider the probability mass function

$$p_n = P(X = x_n).$$

Furthermore, a random variable X may be multidimensional. When this is the case, X is often treated as a vector $\vec{X} = (X_1, \ldots, X_N)$ where N is the number

of components of the multidimensional random variable. In the case of a multidimensional random variable, the concepts of probability distribution/density/mass functions still apply and are adjusted accordingly.

3.2.2 Stochastic Processes

Intuitively speaking, a *stochastic process* is the time-evolution of a family of time-dependent random variables $X(t)$ defined for $t \in T$ (the parameter t is usually time, although it may be something else). By "time-dependent random variable" we mean random variable $X(t)$ such that for a fixed choice of t, $X(t)$ is a random variable in the sense described above.

3.2.3 The Markov Property

A Markov process is a stochastic process for which

$$f(x_k, t_k | x_{k-1}, \ldots, x_1; t_{k-1}, \ldots, t_1) = f(x_k, t_k | x_{k-1}, t_{k-1}), \tag{3.2}$$

where $t_1 \leq t_2 \leq \ldots \leq t_k$, and $f(A|B)$ denotes the conditional probability of the event A occurring *given* that the event B occurred, as in the usual case. In the case of a discrete state space, we replace the x_i's of Equation 3.2 with n_i's; the probability density function f may also be replaced with the probability mass function p. Equation 3.2 constitutes the Markov property: it says that the conditional probability density of the stochastic process at time t_k depends only on the state of the process at time t_{k-1} and not at earlier times. In other words, the conditional probability of the system moving to state x_k at the next time t_k given its particular trajectory from state x_1 at time t_1 to its present state x_{k-1} at time t_{k-1} is determined entirely by knowledge of its current state x_{k-1} and not by its previous

history. It is in this sense that a Markov process is said to have the property of memorylessness.

3.2.4 The Chapman-Kolmogorov Equation

In the following, we adopt the notation associated with discrete state space stochastic processes, with the understanding that the results can be applied to continuous state space processes simply by replacing probability mass functions with probability density functions, n_i's with x_i's, and sums with integrals. It is true for any stochastic process $X(t)$ that

$$p(n_1, n_2, n_3; t_1, t_2, t_3) = p(n_3, t_3|n_2, n_1; t_1, t_2)p(n_2, t_2|n_1, t_1)p(n_1, t_1), \tag{3.3}$$

where the left-hand side of Equation 3.3 is the joint probability of $X(t)$ taking the values n_1 at time t_1, n_2 at t_2, and n_3 at t_3 and the right-hand side of Equation 3.3 is the product of the probability of $X(t)$ taking the value n_1 at t_1, the conditional probability of $X(t)$ taking the value n_2 at time t_2 *given* that it took the value n_1 at time t_1, and so on. If we impose the time-ordering $t_1 \leq t_2 \leq t_3$ and assume the Markov property, then Equation 3.3 reduces to

$$p(n_1, n_2, n_3; t_1, t_2, t_3) = p(n_3, t_3|n_2, t_2)p(n_2, t_2|n_1, t_1)p(n_1, t_1). \tag{3.4}$$

Likewise, it is easy to see that

$$p(n_1, \ldots, n_k; t_1, \ldots, t_k) = p(n_k, t_k|n_{k-1}, t_{k-1}) \ldots p(n_2, t_2|n_1, t_1)p(n_1, t_1),$$

if the time-ordering suggested above is maintained, and thus the entire process is specified if we have knowledge of the probability $p(n_1, t_1)$ and the *transition probability* $p(n_2, t_2|n_1, t_1)$. Thus, we seek a relationship between these probabilities.

Summing Equation 3.4 over all possible values of n_2, we have

$$p(n_1, n_3; t_1, t_3) = p(n_1, t_1) \sum_{n_2=-\infty}^{\infty} p(n_3, t_3|n_2, t_2)p(n_2, t_2|n_1, t_1). \tag{3.5}$$

Dividing Equation 3.5 by $p(n_1, t_1)$ and using the fact that

$$\frac{p(n_1, n_3; t_1, t_3)}{p(n_1, t_1)} = p(n_3, t_3|n_1, t_1),$$

we find that

$$p(n_3, t_3|n_1, t_1) = \sum_{n_2=-\infty}^{\infty} p(n_3, t_3|n_2, t_2)p(n_2, t_2|n_1, t_1), \tag{3.6}$$

which is the Chapman-Kolmogorov Equation.

3.2.5 The Master Equation

The Chapman-Kolmogorov Equation 3.6 is a condition on the transition probabilities of a Markov process, and, as a functional relation, it is difficult to handle. Instead of working directly with the Chapman-Kolmogorov Equation, we seek an equivalent ordinary differential equation that is obtained by taking the time difference $\tau' = t_3 - t_2$ to be infinitesimal. Hence, we must find the behaviour of the transition probability $p(n_3, t_3|n_2, t_2)$ as $\tau' \to 0$.

To this end, we define $q(n, t; \tau)$ as the probability that, given the process is in state n at time t, the process makes a transition away from state n at some time between t and $t + \tau$, and assume that for infinitesimally small time intervals τ, $q(n, t; \tau)$ takes the form $a(n, t)\tau$. As $\tau \to 0$, $a(n, t)\tau$ is then the probability that any transition occurs at some time in $[t, t + \tau)$, and the probability of two or more transitions in the interval $(t, t+\tau)$ is $o(\tau)$ by which we mean as $\tau \to 0$, $o(\tau)/\tau \to 0$. Following Gillespie [Gillespie], we also define $w(n^*|n, t)$ as the probability that the stochastic process, upon leaving state n at time t, transitions to state n^*.

Thus, the probability of the process, being in state n_1 at time t_1, making one transition at some time $t^* \in [t_1, t_1 + \tau)$ *and* of this transition being to the state n_2 is the product $a(n_1, t_1)\tau w(n_2, n_1|t^*)$. On the other hand, the probability that *no* transition occurs from state n_1 in the interval $[t_1, t_1 + \tau)$ is $1 - a(n_1, t_1)\tau$; if no transition occurs then the probability of the process being in state n_2 at time $t_1 + \tau$ is simply δ_{n_2,n_1}, where

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

is the Kronecker delta function. Thus, the transition probability $p(n_2, t_2|n_1, t_1)$ is

$$p(n_2, t_2|n_1, t_1) = [1 - a(n_1, t_1)\tau]\delta_{n_2,n_1} + a(n_1, t_1)\tau w(n_2|n_1, t^*) + o(\tau), \tag{3.7}$$

with $\tau = t_2 - t_1$. If we further assume that $w(n^*|n, t)$ is a smooth function of t, then we can replace the time t^* with the infinitesimally close t_1 in Equation 3.7:

$$p(n_2, t_2|n_1, t_1) = [1 - a(n_1, t_1)\tau]\delta_{n_2,n_1} + a(n_1, t_1)\tau w(n_2|n_1, t_1) + o(\tau). \tag{3.8}$$

To match the more common notation of, for example, van Kampen [Van Kampen, 2007], we define $W(n^*|n, t) \equiv a(n, t)w(n^*|n, t)$ as the (joint) *transition probability per unit time* that the process, being in state n at time t, makes the transition to state n^*. We notice that

$$\sum_{n^*=-\infty}^{\infty} W(n^*|n, t) = a(n, t) \sum_{n^*=-\infty}^{\infty} w(n^*|n, t),$$

and that

$$\sum_{n^*=-\infty}^{\infty} w(n^*|n, t) = 1,$$

so that

$$w(n^*|n, t) = \frac{W(n^*|n, t)}{\sum_{n^*=-\infty}^{\infty} W(n^*|n, t)}.$$

Using these relations in Equation 3.8, we have

$$p(n_2, t_2|n_1, t_1) = \left[1 - \sum_{n_2=-\infty}^{\infty} W(n_2|n_1, t_1)\tau\right] \delta_{n_2,n_1} + W(n_2|n_1, t)\tau + o(\tau). \quad (3.9)$$

We define a Markov process to be *temporally homogeneous* if $a(n, t) = a(n)$ and $w(n^*|n, t) = w(n^*|n)$, in which case it is clear that $W(n^*|n, t) = W(n^*|n)$. Thus, for a temporally homogeneous Markov process, the transition probability on the left-hand side of Equation 3.9 evidently does not depend on the times t_1 and t_2 but only on the time difference $\tau = t_2 - t_1$. Adopting the notation of van Kampen [Van Kampen, 2007], we write $T_\tau(n_2|n_1) \equiv p(n_2, t_2|n_1, t_1)$ for such a Markov process. For brevity and for consistency with van Kampen, we also write

$$a_0(n_1) = \sum_{n_2=-\infty}^{\infty} W(n_2|n_1);$$

now, we may rewrite Equation 3.9 as

$$T_\tau(n_2|n_1) = (1 - a_0(n_1)\tau)\delta_{n_2,n_1} + W(n_2|n_1)\tau + o(\tau). \quad (3.10)$$

Having established the behaviour of the transition probability T_τ for a temporally homogeneous Markov process as the time difference τ becomes infinitesimally small, we return to the Chapman Kolmogorov Equation 3.6, taking $\tau = t_2 - t_1$, $\tau' = t_3 - t_2$, and replacing $T_{\tau'}$ with the expression in Equation 3.10 as is done by

van Kampen [Van Kampen, 2007]:

$$\begin{aligned}
T_{\tau+\tau'}(n_3|n_1) &= \sum_{n_2=-\infty}^{\infty} T_{\tau'}(n_3|n_2)T_{\tau}(n_2|n_1) \\
&= \sum_{n_2=-\infty}^{\infty} [(1-a_0(n_2)\tau')\delta_{n_3,n_2} + W(n_3|n_2)\tau' + o(\tau')]T_{\tau}(n_2|n_1) \\
&= \sum_{n_2=-\infty}^{\infty} (1-a_0(n_2)\tau')\delta_{n_3,n_2}T_{\tau}(n_2|n_1) + \tau' \sum_{n_2=-\infty}^{\infty} W(n_3|n_2)T_{\tau}(n_2|n_1) \\
&+ o(\tau') \sum_{n_2=-\infty}^{\infty} T_{\tau}(n_2|n_1) \\
&= (1-a_0(n_3)\tau')T_{\tau}(n_3|n_1) + \tau' \sum_{n_2=-\infty}^{\infty} W(n_3|n_2)T_{\tau}(n_2|n_1) \\
&+ o(\tau') \sum_{n_2=-\infty}^{\infty} T_{\tau}(n_2|n_1). \qquad (3.11)
\end{aligned}$$

Using in Equation 3.11 the definition of a_0 from above after rearranging and dividing by τ', we have

$$\begin{aligned}
\frac{T_{\tau+\tau'}(n_3|n_1) - T_{\tau}(n_3|n_1)}{\tau'} &= -\sum_{n_2=-\infty}^{\infty} W(n_2|n_3)T_{\tau}(n_3|n_1) + \sum_{n_2=-\infty}^{\infty} W(n_3|n_2)T_{\tau}(n_2|n_1) \\
&+ \frac{o(\tau)}{\tau'} \sum_{n_2=-\infty}^{\infty} T_{\tau}(n_2|n_1) \\
&= \sum_{n_2=-\infty}^{\infty} \{W(n_3|n_2)T_{\tau}(n_2|n_1) - W(n_2|n_3)T_{\tau}(n_3|n_1)\} \\
&+ \frac{o(\tau')}{\tau'} \sum_{n_2=-\infty}^{\infty} T_{\tau}(n_2|n_1).
\end{aligned}$$

Now taking the limit as $\tau' \to 0$ we have

$$\frac{d}{d\tau}T_{\tau}(n_3|n_1) = \sum_{n_2=-\infty}^{\infty} \{W(n_3|n_2)T_{\tau}(n_2|n_1) - W(n_2|n_3)T_{\tau}(n_3|n_1)\}. \qquad (3.12)$$

Usually this is written as

$$\frac{d}{dt}p_n(t) = \sum_m \{w_{nm}p_m(t) - w_{mn}p_n(t)\}, \qquad (3.13)$$

where the index m runs over all possible states of the system, and with the understanding that $p_n(t)$ is indeed the *conditional* probability of the system being in state n at time t *given* that is was in state n_1 at time t_1.

3.3 Formulating the Model

The stochastic process currently being considered is the two-dimensional stochastic process $(N_S(t), N_P(t))$, where the time-dependent random variables $N_S(t)$ and $N_P(t)$ represent the number of CSCs and the number of non-stem cancer cells, respectively, observed at time t. The sample space of this process is the set of all two-dimensional vectors whose components are non-negative integers. We assume that this process is a Markov process, an assumption that has been used successfully in the past by, for example, Till *et al.* [Till et al., 1964] and Clayton *et al.* [Clayton et al., 2007]. In this context, we interpret the rates depicted in Equation 3.1 as probabilites per unit time of an individual cell undergoing a particular division. For example, $\rho_S r_1$ is the probability per unit time of an individual CSC undergoing symmetric self-renewal. Thus, the (temporally homogeneous) transition probability per unit time of the process going from state (n_S, n_P) to state (n_{S+1}, n_P) is $\rho_S r_1 n_S$.

From Equation 3.13, we can write the master equation governing the time-evolution of the (conditional) probability

$$p_{n_S,n_P}(t) = P(N_S(t) = n_S, N_P(t) = n_P | N_S(0) = n_S^0, N_P(0) = n_P^0)$$

as

$$\begin{aligned}
\frac{d}{dt}p_{n_S,n_P}(t) &= \rho_S r_1(n_S - 1)p_{n_S-1,n_P}(t) + \rho_S r_2 n_S p_{n_S,n_P-1}(t) \\
&+ \rho_S r_3(n_S + 1)p_{n_S+1,n_P-2}(t) - \rho_S(r_1 + r_2 + r_3)n_S p_{n_S,n_P}(t) \\
&+ \Gamma(n_P + 1)p_{n_S,n_P+1}(t) - \Gamma n_P p_{n_S,n_P}(t). \qquad (3.14)
\end{aligned}$$

Equation 3.14 is valid for $n_S \geq 1, n_P \geq 2$, as the process does not permit negative numbers of cells or births of stem cells from state $(n_S, n_P) = (0, n_P)$. In the cases $n_S = 0$ or $n_P = 0, 1$, Equation 3.14 is modified accordingly – for example, we have

$$\begin{aligned} \frac{d}{dt}p_{0,n_P}(t) &= \rho_S r_3 p_{1,n_P-2}(t) + \Gamma(n_P+1)p_{0,n_P+1}(t) \\ &- \Gamma n_P p_{0,n_P}(t), \end{aligned}$$

for $n_P \geq 2$, and

$$\frac{d}{dt}p_{0,1}(t) = 2\Gamma p_{0,2}(t) - \Gamma p_{0,1}(t).$$

Equation 3.14 is subject to the initial condition

$$p_{n_S,n_P}(0) = \delta_{n_S,n_S^0}\delta_{n_P,n_P^0}.$$

While a master equation such as Equation 3.14 is more convenient to work with than the associated Chapman-Kolmogorov Equation, solving it can be a daunting, if not impossible, task. Fortunately, some information can be extracted from Equation 3.14 without actually solving it. In the following, we derive equations for the average numbers of the stochastic variables considered in Equation 3.14.

3.3.1 Averages

Recall that the expected value, or average, of a function $f(X)$ of a (discrete) random variable X is

$$\langle f(X) \rangle = \sum_i f(x_i)p(x_i),$$

where $p(x_i)$ is the probability that x takes the value x_i, and the index i extends over all values of x in the sample space. We exploit this, and the form of Equation 3.14, to find an ordinary differential equation describing the time evolution of the

average numbers of cells $\langle N_S \rangle$ and $\langle N_P \rangle$. In the following, for brevity, we begin suppressing the indices of $p_{n_S,n_P}(t)$ that do not differ from n_S, n_P [for example, we write $p(t) = p_{n_S,n_P}(t)$ and $p_{n_S-1}(t) = p_{n_S-1,n_P}(t)$]. To proceed, we multiply Equation 3.14 by n_S and sum both sides over all $\{\mathrm{n}\}$, where $\{\mathrm{n}\}$ represents the set of all possible states (n_S, n_P):

$$\begin{aligned}
\sum_{\{n\}} n_S \frac{dp(t)}{dt} &= \sum_{\{n\}} n_S \rho_S r_1 (n_S - 1) p_{n_S-1}(t) + \sum_{\{n\}} n_S \rho_S r_2 n_S p_{n_P-1}(t) \\
&+ \sum_{\{n\}} n_S \rho_S r_3 (n_S + 1) p_{n_S+1,n_P-2}(t) - \sum_{\{n\}} n_S \rho_S (r_1 + r_2 + r_3) n_S p(t) \\
&+ \sum_{\{n\}} n_S \Gamma (n_P + 1) p_{n_P+1}(t) - \sum_{\{n\}} n_S \Gamma n_P p(t). \qquad (3.15)
\end{aligned}$$

Next, we re-label the indices in each of the sums on the right-hand-side of Equation 3.15 so that each of the probabilities is in the form $p(t) = p_{n_S,n_P}(t)$:

$$\begin{aligned}
\sum_{\{n\}} n_S \frac{dp(t)}{dt} &= \sum_{\{n\}} (n_S + 1) \rho_S r_1 n_S p(t) + \sum_{\{n\}} n_S \rho_S r_2 n_S p(t) \\
&+ \sum_{\{n\}} (n_S - 1) \rho_S r_3 n_S p(t) - \sum_{\{n\}} n_S \rho_S (r_1 + r_2 + r_3) n_S p(t) \\
&+ \sum_{\{n\}} n_S \Gamma n_P p(t) - \sum_{\{n\}} n_S \Gamma n_P p(t). \qquad (3.16)
\end{aligned}$$

Expanding the summands of Equation 3.16 leads to

$$\begin{aligned}
\sum_{\{n\}} n_S \frac{dp(t)}{dt} &= \rho_S r_1 \sum_{\{n\}} n_S^2 p(t) + \rho_S r_1 \sum_{\{n\}} n_S p(t) \\
&+ \rho_S r_2 \sum_{\{n\}} n_S^2 p(t) + \rho_S r_3 \sum_{\{n\}} n_S^2 p(t) - \rho_S r_3 \sum_{\{n\}} n_S p(t) \\
&- \rho_S r_1 \sum_{\{n\}} n_S^2 p(t) - \rho_S r_2 \sum_{\{n\}} n_S^2 p(t) - \rho_S r_3 \sum_{\{n\}} n_S^2 p(t) \\
&+ \Gamma \sum_{\{n\}} n_S n_P p(t) - \Gamma \sum_{\{n\}} n_S n_P p(t), \qquad (3.17)
\end{aligned}$$

and noting that

$$\langle N_S \rangle = \sum_{\{n\}} n_S p(t),$$

and

$$\frac{d\langle N_S \rangle}{dt} = \frac{d}{dt}\left(\sum_{\{n\}} n_S p(t)\right) = \sum_{\{n\}}\left(n_S \frac{dp(t)}{dt}\right),$$

Equation 3.17 can be further simplified to

$$\frac{d\langle N_S \rangle}{dt} = \rho_S(r_1 - r_3)\langle N_S \rangle. \tag{3.18}$$

Following the same procedure as above, except multiplying Equation 3.14 by n_P instead of n_S in the first step, the following ODE governing the average number of non-stem cancer cells $\langle N_P \rangle$ is obtained:

$$\frac{d\langle N_P \rangle}{dt} = \rho_S(r_2 + 2r_3)\langle N_S \rangle - \Gamma\langle N_P \rangle. \tag{3.19}$$

We now establish the convention of writing $S(t) = \langle N_S \rangle$ and $P(t) = \langle N_P \rangle$. The purpose of this is dual: to explicitly illustrate the time dependence of the averages, and to free ourselves of a cumbersome notation. Furthermore, we define the net CSC symmetric division probability $r = r_1 - r_3$. Using this notation, the pair of ODEs describing the average numbers of cancer stem and non-stem cells can be written as

$$\begin{aligned} \frac{dS(t)}{dt} &= \rho_S r S(t) \\ \frac{dP(t)}{dt} &= \rho_S(1-r)S(t) - \Gamma P(t), \end{aligned} \tag{3.20}$$

subject to the initial condition $S(0) = S_0 = n_S^0$, $P(0) = P_0 = n_P^0$.

Integrating the first line of Equation 3.20 yields

$$S(t) = S_0 e^{\rho_S r t}, \tag{3.21}$$

after we have applied the initial condition from Equation 3.20. Substituting $S(t)$ from Equation 3.21 into the second line of Equation 3.20, we have

$$\frac{dP(t)}{dt} + \Gamma P(t) = \rho_S(1-r)S_0 e^{\rho_S r t}. \tag{3.22}$$

The above is a linear, inhomogeneous ODE for $P(t)$; to solve it, we can use the method of integrating factors [Boyce and DiPrima, 1992]. Our integrating factor is

$$\mu(t) = e^{\int \Gamma dt} = e^{\Gamma t};$$

multiplying Equation 3.22 by this integrating factor, we have

$$\frac{d(P(t)e^{\Gamma t})}{dt} = \rho_S(1-r)S_0 e^{\rho_S r t + \Gamma t}. \tag{3.23}$$

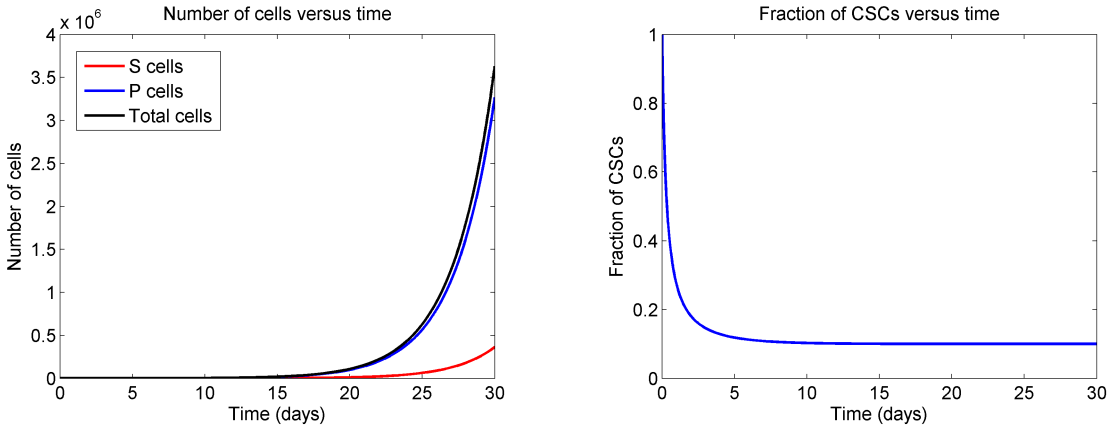
Integrating Equation 3.23, we have

$$e^{\Gamma t}P(t) = \frac{\rho_S(1-r)S_0}{\rho_S r + \Gamma} e^{(\rho_S r + \Gamma)t} + C, \text{ for } \rho_S r + \Gamma \neq 0,$$

where C is an integrative constant to be determined by the initial condition on Equation 3.20. Applying this condition, and solving for C, we have

$$P(t) = \left(\frac{\rho_S(1-r)S_0}{\rho_S r + \Gamma}\right)\left(e^{\rho_S r t} - e^{-\Gamma t}\right) + P_0 e^{-\Gamma t}. \tag{3.24}$$

Provided $S_0 > 0$, the quantities $S(t)$ and $P(t)$ given in Equations 3.21 and 3.24 clearly exhibit exponential growth when $\rho_S r > 0$ (that is, when $r_1 > r_3$). When $r_1 = r_3$, the number of CSCs remains at its intial value S_0.If $r_1 < r_3$ then the CSC sub-population (and hence the P-cell sub-population) dies out. The case $r_1 > r_3$ (that is, $r > 0$) is thus most relevant to a growing tumour. While $S(t)$ and $P(t)$ grow without bound, the fraction of CSCs $\frac{S(t)}{S(t)+P(t)}$ instead reaches a steady-state value, as is illustrated in Figure 3.2 and analyzed in the following subsection.



(a) (b)

Figure 3.2: (a) Number of CSCs (S) and non-stem cancer cells (P) versus time and (b) fraction of CSCs versus time, for $\rho_S = 3.5$, $r = 0.1$, $\Gamma = 0$ and initial condition $S_0 = 1$, $P_0 = 0$.

3.3.2 Fraction of CSCs

A quantity of significant interest is the fraction of CSCs in the total population of cancer cells. This could be derived via the brute-force manipulation of Equations 3.21 and 3.24; however, we find it preferable to bypass that calculation in favour of a slightly different approach. To this end, we define

$$X(t) = \frac{S(t)}{S(t) + P(t)},$$

and note that

$$X'(t) = \frac{S'(t)[S(t) + P(t)] - S(t)[S'(t) + P'(t)]}{[S(t) + P(t)]^2}.$$

Substituting expressions for $S'(t)$ and $P'(t)$ from Equation 3.20 into the above, we obtain

$$X'(t) = \frac{(\rho_S r - \rho_S)S^2(t)}{[S(t) + P(t)]^2} + \frac{(\rho_S r + \Gamma)S(t)}{[S(t) + P(t)]} \frac{P(t)}{[S(t) + P(t)]}. \tag{3.25}$$

Using the identity $\frac{P(t)}{S(t)+P(t)} = 1 - \frac{S(t)}{S(t)+P(t)}$ in Equation 3.25, and then employing our definition of $X(t)$, we find that

$$X'(t) = (\rho_S r - \rho_S)X^2(t) + (\rho_S r + \Gamma)X(t)[1 - X(t)]. \tag{3.26}$$

Rearranging Equation 3.26,

$$X'(t) = (\rho_S r + \Gamma)X(t)\left[1 - \frac{(\rho_S + \Gamma)X(t)}{\rho_S r + \Gamma}\right]. \tag{3.27}$$

Equation 3.27 is recognized as having the form of a logistic growth equation, for which the "carrying capacity" of the fraction of CSCs [Murray, 2002] is

$$K = \frac{\rho_S r + \Gamma}{\rho_S + \Gamma}.$$

The time-dependent solution of Equation 3.27 is [Murray, 2002]

$$X(t) = \frac{X_0 K e^{(\rho_S r+\Gamma)t}}{K + X_0(e^{(\rho_S r+\Gamma)t} - 1)}, \tag{3.28}$$

where $X_0 = S_0/(S_0 + P_0)$. Equation 3.27 (with the initial condition $X(0) = X_0$) has a non-trivial steady-state solution $X = K$, valid for $(\rho_S r + \Gamma) > 0$ and $X_0 > 0$. In particular,

$$X(t) \to \frac{\rho_S r + \Gamma}{\rho_S + \Gamma} \text{ as } t \to \infty.$$

Figure 3.3 illustrates the tendency of the CSC fraction toward the steady-state for different initial fractions of CSCs.

So, the model predicts that the fraction of CSCs reaches a steady-state, or plateau phase, that is determined by its growth characteristics. Evidently (and intuitively), the difference $r = r_1 - r_3$ of the rates of symmetric self-renewal and symmetric commitment divisions by CSCs plays an important role in dictating the fraction of CSCs in the plateau phase. When Γ is small compared to $\rho_S r$, then the steady-state fraction of CSCs is $\frac{\rho_S r+\Gamma}{\rho_S+\Gamma} \approx r$.

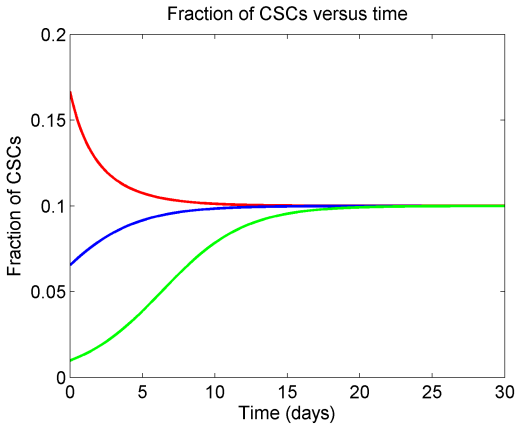


Figure 3.3: Fraction of CSCs versus time for $X_0 = 0.17$ (red), $X_0 = 0.07$ (blue), and $X_0 = 0.01$ (green), using the same parameter values as in Figure 3.2.

3.3.3 Survival Rate

Equation 3.21 states that the average number of CSCs grows exponentially for any $S_0 > 0$ when $r = r_1 - r_3 > 0$. However, an interesting feature of the stochastic process under consideration is that the sub-population of CSCs may become extinct, even when the initial condition and parameter r predict exponential growth of the mean. This is significant, because under the CSC hypothesis, extinction of the CSC sub-population implies eventual extinction of the entire tumour (provided the death rate of non-stem cells is $\Gamma > 0$). To find the probability of the sub-population becoming extinct at some time t, we consider the single-component stochastic process $N_S(t)$. This stochastic process is a continuous-time Markov process for which the sample space consists of the non-negative integers and for which only transitions between adjacent states are permitted; such a process is called a "one-step" or "birth-and-death" process.

In particular, an individual CSC may symmetrically self-renew with probability

per unit time $\rho_S r_1$ (corresponding to an increase of the population by one) or may symmetrically commit with probability per unit time $\rho_S r_3$ (corresponding to a decrease of the population by one). (In this light, it is not hard to see why the "birth-and-death" moniker is appropriate.) Note that the event of asymmetric self-renewal does not change the number of CSCs. Concentrating on CSCs only, the birth-and-death process $N_S(t)$ contains the following types of divisions with associated rates:

$$\begin{aligned} S &\xrightarrow{\rho_S r_1} S + S \\ S &\xrightarrow{\rho_S r_3} \emptyset. \end{aligned} \tag{3.29}$$

The corresponding master equation is

$$\begin{aligned} \frac{dp_{n_S}(t)}{dt} &= \rho_S r_1 (n_S - 1) p_{n_S - 1}(t) + \rho_S r_3 (n_S + 1) p_{n_S + 1}(t) \\ &- \rho_S (r_1 + r_3) n_S p_{n_S}(t), \end{aligned} \tag{3.30}$$

for $n_S \geq 1$, with

$$\frac{dp_0(t)}{dt} = \rho_S r_3 p_1(t),$$

where $p_{n_S}(t)$ is the conditional probability of the random variable $N_S(t)$ taking the value n_S at time t given $N_S(0) = n_S^0$ at time $t = 0$ (that is, $p_{n_S}(0) = \delta_{n_S, n_S^0}$).

Following the method used by Bailey [Bailey, 1964], we introduce the probability generating function $P(x, t)$, defined by

$$P(x, t) = \sum_{n=0}^{\infty} x^n p_n(t). \tag{3.31}$$

We are interested in the probability $p_0(t)$ that a population of n_S^0 CSCs at time $t_0 = 0$ has gone extinct at some later time t; to this end, we notice that

$$P(0, t) = p_0(t) + \sum_{n=1}^{\infty} 0^n p_n(t) = p_0(t), \tag{3.32}$$

and proceed to cast the master equation 3.30 into an equation for the probability generating function. Multiplying Equation 3.30 by x^{n_S} and summing over all n_S, we find

$$\begin{aligned}\sum_{n_S=0}^{\infty} x^{n_S}\frac{dp_{n_S}(t)}{dt} &= \sum_{n_S=1}^{\infty} x^{n_S}\rho_S r_1(n_S-1)p_{n_S-1}(t) + \sum_{n_S=0}^{\infty} x^{n_S}\rho_S r_3(n_S+1)p_{n_S+1}(t) \\ &- \sum_{n_S=0}^{\infty} x^{n_S}\rho_S(r_1+r_3)n_S p_{n_S}(t). \end{aligned} \tag{3.33}$$

Re-labeling indices in Equation 3.33,

$$\begin{aligned}\sum_{n_S=0}^{\infty} x^{n_S}\frac{dp_{n_S}(t)}{dt} &= \sum_{n_S=0}^{\infty} x^{n_S+1}\rho_S r_1 n_S p_{n_S}(t) + \sum_{n_S=1}^{\infty} x^{n_S-1}\rho_S r_3 n_S p_{n_S}(t) \\ &- \sum_{n_S=0}^{\infty} x^{n_S}\rho_S(r_1+r_3)n_S p_{n_S}(t) \\ &= \rho_S r_1 x^2 \sum_{n_S=1}^{\infty} n_S x^{n_S-1} p_{n_S}(t) + \rho_S r_3 \sum_{n_S=1}^{\infty} n_S x^{n_S-1} p_{n_S}(t) \\ &- \rho_S(r_1+r_3)x \sum_{n_S=1}^{\infty} n_S x^{n_S-1} p_{n_S}(t). \end{aligned} \tag{3.34}$$

The observations that

$$\sum_{n_S=0}^{\infty} x^{n_S}\frac{dp_{n_S}(t)}{dt} = \frac{\partial P(x,t)}{\partial t}$$

and

$$\sum_{n_S=1}^{\infty} n_S x^{n_S-1} p_{n_S}(t) = \frac{\partial P(x,t)}{\partial x}$$

allow us to write Equation 3.34 in the form

$$\frac{\partial P(x,t)}{\partial t} = (\rho_S r_1 x^2 - \rho_S(r_1+r_3)x + \rho_S r_3)\frac{\partial P(x,t)}{\partial x}, \tag{3.35}$$

with

$$P(x,0) = x^{n_S^0}.$$

This PDE can be solved by the method of characteristics [Garabedian, 1964]: introducting the parameter s, the characteristic equations are

$$\begin{aligned} \frac{dt}{ds} &= 1 \\ \frac{dx}{ds} &= -(\rho_S r_1 x^2 - \rho_S(r_1 + r_3)x + \rho_S r_3) \\ \frac{dP}{ds} &= 0. \end{aligned} \tag{3.36}$$

The first line of Equation 3.36 gives $t = s$, where we have taken $t(0) = 0$. The second line of Equation 3.36 is a Riccati equation [Boyce and DiPrima, 1992]. Factoring, and using $s = t$ from above, we have that the characteristic curves are given by

$$dt = \frac{dx}{\rho_S(1-x)(r_1 x - r_3)}, \tag{3.37}$$

which makes two things apparent: (i) in the special case $r_1 = r_3$, Equation 3.37 becomes

$$-dt = \frac{dx}{\rho_S r_1 (x-1)^2}, \tag{3.38}$$

and (ii) a particular solution of the Riccati Equation 3.37 is $x = 1$. Dealing with Equation 3.38 first, we make the change-of-variables $\nu = x - 1$ so that Equation 3.38 becomes

$$-dt = \frac{d\nu}{\rho_S r_1 \nu^2},$$

which upon integration gives

$$\nu = \frac{1}{\rho_S r_1 t + k},$$

where k is some integrative constant. This gives

$$x = 1 + \frac{1}{\rho_S r_1 t + k}. \tag{3.39}$$

Taking $x(0) = x_0$ and $t(0) = 0$, we find that

$$k = \frac{1}{x_0 - 1},$$

and hence

$$x_0 = 1 + \frac{1}{\frac{1}{x-1} - \rho_S r_1 t}. \tag{3.40}$$

The third line of Equation 3.36 says that the probability generating function $P(x,t)$ is constant along the characteristic curves. Thus, $P(x,t) = P(x_0,0)$. The initial condition $P(x,0) = x^{n_S^0}$ gives, in conjunction with Equation 3.40,

$$P(x,t) = \left(1 - \frac{1}{\frac{1}{1-x} + \rho_S r_1 t}\right)^{n_S^0}. \tag{3.41}$$

The probability of extinction $P(0,t)$ is thus

$$P(0,t) = p_0(t) = \left(1 - \frac{1}{1 + \rho_S r_1 t}\right)^{n_S^0},$$

for $r_1 = r_3$, which can be written more simply as

$$p_0(t) = \left(\frac{\rho_S r_1 t}{\rho_S r_1 t + 1}\right)^{n_S^0}, r_1 = r_3. \tag{3.42}$$

We now return to Equation 3.37 for the more general situation $r_1 \neq r_3$. A Riccati equation such as this can be reduced to a linear ODE by making the substitution

$$x = x_1 + \frac{1}{\nu},$$

where x_1 is a particular solution of the Riccati equation. We have observed by inspection that $x = 1$ is a particular solution of Equation 3.37 and thus take

$$x = 1 + \frac{1}{\nu}.$$

This substitution transforms Equation 3.37 into

$$dt = \frac{d\nu}{\rho_S(r_1 - r_3)\nu + \rho_S r_1}, \tag{3.43}$$

which (using the method of integrating factors) gives

$$\nu = \frac{\frac{-r_1}{r_1 - r_3} e^{-\rho_S(r_1 - r_3)t} + k}{e^{-\rho_S(r_1 - r_3)t}},$$

whereupon imposing $\nu(0) = \nu_0$ and $t(0) = 0$ the integrative constant is found to be

$$k = \nu_0 + \frac{r_1}{r_1 - r_3}.$$

Thus, we have

$$\nu = \frac{\frac{-r_1}{r_1 - r_3} e^{-\rho_S (r_1 - r_3)t} + \nu_0 + \frac{r_1}{r_1 - r_3}}{e^{-\rho_S (r_1 - r_3)t}}. \tag{3.44}$$

Recalling that

$$x = 1 + \frac{1}{\nu},$$

we find

$$x = 1 + \frac{e^{-\rho_S (r_1 - r_3)t}}{\frac{1}{x_0 - 1} + \frac{r_1}{r_1 - r_3} - \frac{r_1}{r_1 - r_3} e^{-\rho_S (r_1 - r_3)t}}, \tag{3.45}$$

where we have used

$$x_0 = 1 + \frac{1}{v_0}.$$

Thus,

$$x_0 = 1 - \frac{1}{\frac{e^{-\rho_S (r_1 - r_3)t}}{1 - x} - \frac{r_1}{r_1 - r_3} \left(e^{-\rho_S (r_1 - r_3)t} - 1 \right)}, \tag{3.46}$$

whereupon $P(x,t) = P(x_0, 0) = x_0^{n_S^0}$ gives

$$P(x,t) = \left(1 - \frac{1}{\frac{e^{-\rho_S (r_1 - r_3)t}}{1 - x} - \frac{r_1}{r_1 - r_3} \left(e^{-\rho_S (r_1 - r_3)t} - 1 \right)} \right)^{n_S^0}. \tag{3.47}$$

Not forgetting our original motivation, we set $x = 0$ to find $P(0,t) = p_0(t)$:

$$p_0(t) = \left(\frac{r_3 \left(e^{\rho_S (r_1 - r_3)t} - 1 \right)}{r_1 e^{\rho_S (r_1 - r_3)t} - r_3} \right)^{n_S^0}, r_1 \neq r_3, \tag{3.48}$$

where some simplifying steps have been omitted. Collecting Equations 3.42 and 3.48 together, we have the probability of extinction of CSCs as

$$p_0(t) = \begin{cases} \left(\frac{r_3 \left(e^{\rho_S (r_1 - r_3)t} - 1 \right)}{r_1 e^{\rho_S (r_1 - r_3)t} - r_3} \right)^{n_S^0}, & \text{if } r_1 \neq r_3 \\ \left(\frac{\rho_S r_1}{\rho_S r_1 + 1} \right)^{n_S^0}, & \text{if } r_1 = r_3. \end{cases} \tag{3.49}$$

We may be interested in the long-term behaviour of the extinction probabilities given in Equation 3.49. Taking the limit as $t \to \infty$, we find that

$$p_0(t) \to \begin{cases} \left(\frac{r_3}{r_1}\right)^{n_S^0}, \text{ if } r_1 > r_3 \\ 1, \text{ if } r_1 \leq r_3. \end{cases} \tag{3.50}$$

Thus, the model makes the somewhat encouraging prediction that the occurrence of a single CSC will not necessarily result in a tumour, even if the probability of self-renewal is greater than that of differentiation. As a numerical example, consider the case in which oncogenic transformation or metastasis results in the presence of a single CSC with intrinsic division probabilities $r_1 = 0.4$, $r_2 = 0.3$, and $r_3 = 0.3$. This cell has only a 25% chance of forming a lasting colony (i.e. a tumour), an interpretation that is in contrast to the exponential growth predicted by the average equations. Figure 3.4 plots the CSC sub-population survival rate $r_{surv} = 1 - p_0(t)$ against time, based on Equation 3.50, along with estimates of r_{surv} obtained via the stochastic simulatinon algorithm described by Gillespie, which is described in the following section.

3.4 Gillespie's Algorithm

The difficulties associated with analytically solving the master equation can be circumvented by computational simulation of the stochastic process. To this end, we introduce Gillespie's algorithm, a stochastic simulation algorithm originally developed with the intention of simulating the time evolution of systems of reacting chemical species [Gillespie, 1977]. Rather than deal with the master equation, Gillespie's algorithm uses random number generation to numerically simulate realizations of the corresponding Markov process. To begin, we are interested in

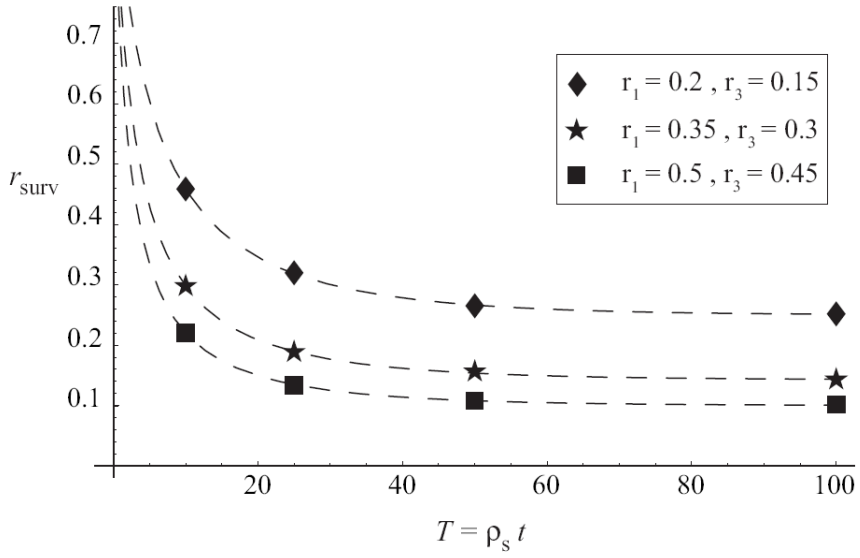


Figure 3.4: Survival rate (1 - $p_0(t)$) versus time, for various values of r_1, r_2, and r_3, with $r = r_1 - r_3$ fixed for the purpose of comparison. The initial number of CSCs is one. Dashed lines are obtained from Equation 3.49. As T grows large, the curves tend to $1 - \frac{r_3}{r_1}$, as expected based on Equation 3.50. Data points are from 100 000 realizations of Gillespie's algorithm, encoded by Adam R. Stinchcombe and subsequently modified appropriately by the author. Image from [Turner et al., In press].

the probability $p(\tau, n^*|n,t)d\tau$, which Gillespie defines as the joint probability that, given the process is in state n at time t, the next transition will occur between times $t+\tau$ and $t+\tau+d\tau$ and will be to the state n^* [Gillespie]; Gillespie alternately calls $p(\tau, n^*|n,t)$ the "next-jump density function" or the "reaction probability density function".

Here we use the notation introduced in subsection 3.2.5, so that $q(n,t;\tau)$ is the probability that, given the process is in state n at time t, the process makes a transition away from state n at some time between t and $t+\tau$ (for infinitesimal τ we have $q(n,t;\tau) = a(n,t)\tau$, as before), and $w(n^*|n,t)$ is the probability that the process makes the transition to state n^* upon leaving state n at time t. The "next-jump density function" $p(\tau, n^*|n,t)$ is evidently the probability that (i) *no* transition occurs between times t and $t+\tau$ *and* that (ii) a transition occurs in the infinitesimal time interval $(t+\tau, t+\tau+d\tau)$ *and* that (iii) this transition is to state n^*. This can be expressed as

$$p(\tau, n^*|n,t)d\tau = [1-q(n,t;\tau)][a(n,t+\tau)d\tau][w(n^*|n,t^*)], \tag{3.51}$$

where the first factor of the right-hand-side of Equation 3.51 is the probability of event (i), the second factor is the probability of event (ii), and the third factor is the probability of event (iii) (occurring at time $t^* \in (t+\tau, t+\tau+d\tau)$). It is convenient to consider $q^*(n,t;\tau) = 1-q(n,t;\tau)$, as Gillespie does [Gillespie]. Then,

$$\begin{aligned} q^*(n,t;\tau+d\tau) &= q^*(n,t;\tau)[q^*(n,t+\tau;d\tau)] \\ &= q^*(n,t;\tau)[1-q(n,t+\tau;d\tau)] \\ &= q^*(n,t;\tau)[1-a(n,t+\tau)d\tau]. \end{aligned} \tag{3.52}$$

Rearranging Equation 3.52, we find

$$q^*(n,t;\tau+d\tau) - q^*(n,t;\tau) = -q^*(n,t;\tau)a(n,t+\tau)d\tau. \tag{3.53}$$

Writing $dq^*(n,t;\tau) = q^*(n,t;\tau+d\tau) - q^*(n,t;\tau)$, we have

$$\frac{dq^*(n,t;\tau)}{d\tau} = -q^*(n,t;\tau)a(n,t+\tau). \tag{3.54}$$

Thus, integrating, we have

$$q^*(n,t;\tau) = e^{-\int_0^\tau a(n,t+\tau')d\tau'},$$

giving

$$q(n,t;\tau) = 1 - e^{-\int_0^\tau a(n,t+\tau')d\tau'}. \tag{3.55}$$

Now we make the assumption that the Markov process is temporally homogeneous, so that $a(n,t) = a(n)$ and $w(n^*|n,t) = w(n^*|n)$. Under this assumption, Equation 3.55 reduces to

$$q(n,t;\tau) = 1 - e^{-a(n)\tau},$$

which, upon substitution into Equation 3.51 yields

$$p(\tau, n^*|n,t) = e^{-a(n)\tau}a(n)w(n^*|n), \tag{3.56}$$

where we have cancelled the factors $d\tau$ that appeared on both sides of the equation. It is convenient to consider

$$p_1(\tau|n,t) = a(n)e^{-a(n)\tau}$$

as the joint probability that, given the process is in state n at time t, the next transition occurs in $(t+\tau, t+\tau+d\tau)$. Thus, $p_1(\tau|n,t)$ is interpreted as the distribution of the "waiting times" between transitions of the Markov process. Hence, we have

$$p(\tau, n^*|n,t) = p_1(\tau|n,t)w(n^*|n).$$

Gillespie's Algorithm, then, beginning with a Markov process in state n at time t, proceeds by first choosing (via a random number generator) a waiting time τ distributed according to

$$p_1(\tau|n,t) = a(n)e^{-a(n)\tau}.$$

The next transition of the Markov process occurs at time $t + \tau$; the *state* to which the process moves at time $t + \tau$ from state n is then determined, again by a random number generator. In practice, namely in the context of the stochastic processes of the present chapter, this proceeds as follows [Gillespie, 1977].

1. Set $t = t_0$ and $n = n_0$ (the state n of the Markov process is, in general, a vector quantity).

2. Choose the waiting time τ until the next transition. To do so, calculate $a(n)$ according to
$$a(n) = \sum_{i=1}^{M} c_i n_i,$$
where the index i runs over the M distinct division pathways of the process, c_i is the probability per unit time that an individual cell will undergo division i, and n_i is the number of cells that may undergo division i. Now, take
$$\tau = [1/a(n)] \ln(1/\sigma_1),$$
where σ_1 is a random number drawn from the unit uniform distribution [Gillespie, 1977].

3. Choose the state n^* to which the process will transition at time $t + \tau$. To do so, consider $a_k = c_k n_k$ for $k = 1, \ldots, M$. Division k brings the process from state n to some state n_k, with $w(n_k|n) = a_k/a(n)$ for $k = 1, \ldots, M$. Choose a number σ_2 from the unit uniform distribution and take k to be the integer for which [Gillespie, 1977]
$$\sum_{i=1}^{k-1} a_i < \sigma_2 a(n) \leq \sum_{i=1}^{k} a_k.$$

4. Update the system according to $t = t + \tau$, $n = n_k$.

5. Return to step 2.

3.5 Comparison of Stochastic and Deterministic Models

Herein we have concurrently developed stochastic and deterministic models of the growth of cancer cell populations based on the CSC hypothesis, with the intention that the stochastic model be used when small numbers of cells are under consideration, and the deterministic model when phenomena involving large numbers of cells are being investigated. The example following the derivation of the chance of extinction of CSCs above emphasizes the importance of fluctuations on the scale of small numbers of cells, and illustrates a particular difference between the stochastic model and the deterministic model based on averages. One may be interested in further understanding how these two models compare.

In a previous section, we used the master equation 3.14 to find an ODE for the average number $S(t) = \langle n_S \rangle$ of CSCs at time t. In the same manner, i.e. by mutliplying Equation 3.14 by n_S^2 and summing over all n_S, we can obtain the following equation for the second moment $S_2(t) = \langle n_S^2 \rangle$:

$$\frac{dS_2(t)}{dt} = 2\rho_S(r_1 - r_3)S_2(t) + \rho_S(r_1 + r_3)S(t). \tag{3.57}$$

Taking $S(t)$ from Equation 3.21 and substituting into Equation 3.57 gives

$$\frac{dS_2(t)}{dt} = 2\rho_S(r_1 - r_3)S_2(t) + \rho_S(r_1 + r_3)n_S^0 e^{\rho_S(r_1 - r_3)t}, \tag{3.58}$$

which is solved by the method of integrating factors to yield

$$S_2(t) = n_S^0 \left(\frac{r_1 + r_3}{r_1 - r_3}\right) e^{\rho_S(r_1 - r_3)t} \left(e^{\rho_S(r_1 - r_3)t} - 1\right) + (n_S^0)^2 e^{2\rho_S(r_1 - r_3)t}, \tag{3.59}$$

where we require that $r_1 \neq r_3$. If $r_1 = r_3$, then Equation 3.58 simplifies to

$$\frac{dS_2(t)}{dt} = 2\rho_S r_1 n_S^0, \tag{3.60}$$

which is solved via integration to give

$$S_2(t) = 2\rho_S r_1 n_S^0 t + (n_S^0)^2. \tag{3.61}$$

Using Equations 3.59 and 3.61 in the formula for the standard deviation,

$$\sigma_S(t) = \sqrt{\langle n_S^2 \rangle - \langle n_S \rangle^2} = \sqrt{S_2(t) - S^2(t)},$$

we thus have

$$\sigma_S(t) = \begin{cases} \sqrt{n_S^0 \left(\frac{r_1+r_3}{r_1-r_3}\right)} e^{\frac{1}{2}\rho_S(r_1-r_3)t} \sqrt{e^{\rho_S(r_1-r_3)t} - 1}, \text{ if } r_1 \neq r_3, \\ \sqrt{2\rho_S r_1 n_S^0 t}, \text{ if } r_1 = r_3. \end{cases} \tag{3.62}$$

Evidently, the standard deviation of the stochastic realizations about the mean number of CSCs grows exponentially for $r_1 > r_3$ (which is the case for a growing tumour). The relative standard deviation $\sigma_S(t)/S(t)$, however, obeys

$$\frac{\sigma_S(t)}{S(t)} = \frac{\sqrt{\langle n_S^2 \rangle - \langle n_S \rangle^2}}{\langle n_S \rangle} = \begin{cases} \frac{1}{\sqrt{n_S^0}} \sqrt{\frac{r_1+r_3}{r_1-r_3}} e^{-\frac{1}{2}\rho_S(r_1-r_3)t} \sqrt{e^{\rho_S(r_1-r_3)t} - 1}, \text{ if } r_1 \neq r_3, \\ \frac{1}{\sqrt{n_S^0}} \sqrt{2\rho_S r_1 t}, \text{ if } r_1 = r_3. \end{cases} \tag{3.63}$$

From Equation 3.63, the relative standard deviation corresponding to $r_1 = r_3$ is proportional to $t^{1/2}$, as expected for a birth-death process in which the rates of birth and death are equal [Bailey, 1964]. For very small times, i.e. as $t \to 0$, the relative standard deviation given in Equation 3.63 vanishes. For the more general case $r_1 \neq r_3$, taking the limit as $t \to \infty$, we see that

$$\frac{\sigma_S(t)}{S(t)} = \frac{\sqrt{\langle n_S^2 \rangle - \langle n_S \rangle^2}}{\langle n_S \rangle} \to \frac{1}{\sqrt{n_S^0}} \sqrt{\frac{r_1 + r_3}{r_1 - r_3}}, \text{ if } r_1 \neq r_3. \tag{3.64}$$

In Figure 3.5, realizations of the stochastic model obtained via Gillespie's algorithm are plotted for various values of the initial number n_S^0 of CSCs. We see

that, when $n_S^0 = 1$, some of the sample realizations show behaviour that differs markedly from the behaviour of the mean. Thus, we suggest that when a single CSC (or a small number of CSCs) is present in a patient at some time (perhaps at the very onset of a primary cancer, at the beginning stages of a metastatic growth, or following treatment that narrowly failed to eradicate the CSC sub-population), little can be said about the growth dynamics that will ensue. The CSC (or CSCs) may differentiate such that the CSC sub-population dies out. Alternatively, the CSC population may quickly enter the phase of exponential growth due to early symmetric self-renewing divisions, or else may exhibit some significant lag period before entering the exponential growth phase at a later time. This consideration may help in understanding the widespread variation in times to cancer detection or relapse.

3.6 A General Model

To this point, we have considered a model of cancer cell growth based on the CSC hypothesis with only two subpopulations of cancer cells: CSCs and non-stem cancer cells, the latter of which we have denoted P-type cells to reflect the fact that the cells "downstream" of stem cells in a stem cell hierarchy are usually called progenitor cells. Here, we elaborate on the role of progenitor cells and generalize the model developed above to include additional subpopulations of cells.

3.6.1 The role of progenitor cells

Progenitor cells, the non-stem cell progeny of stem cells, differ from stem cells in that they have limited proliferative potential and limited ability to differentiate.

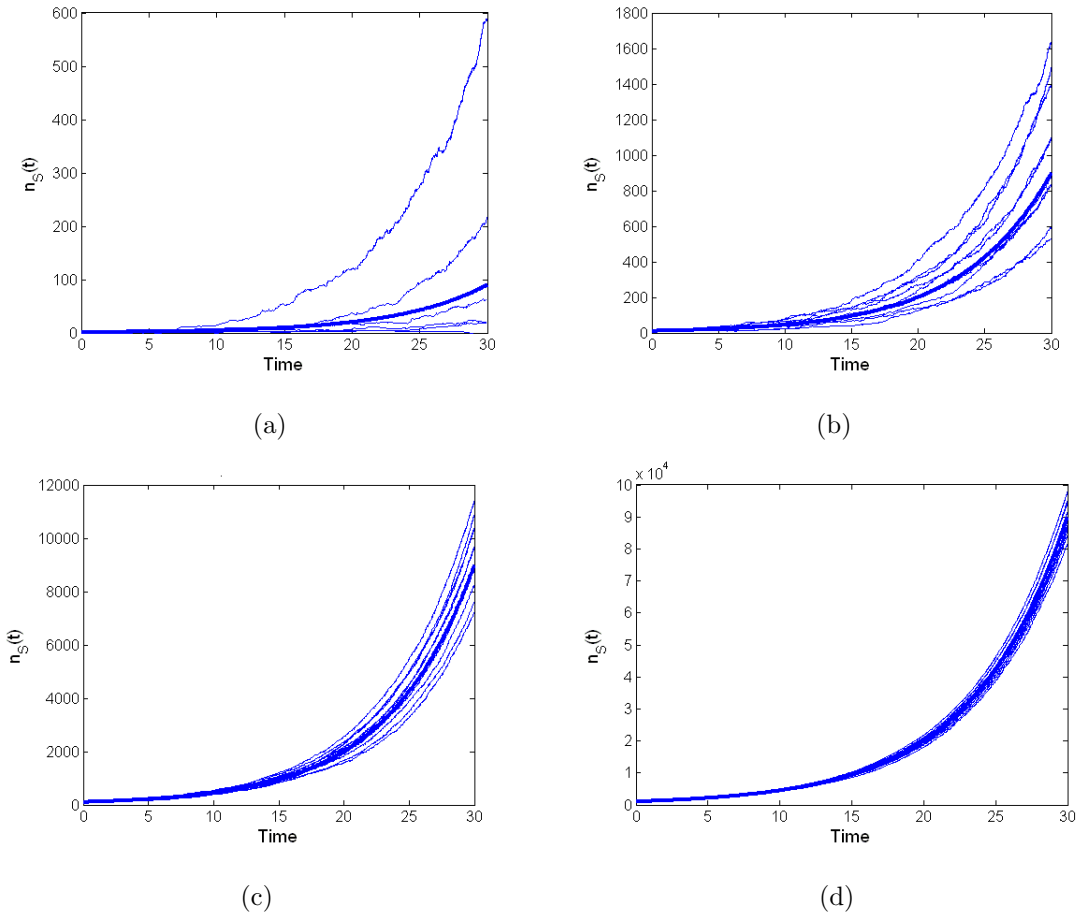


Figure 3.5: 10 sample realizations of the stochastic model, with (a) $n_S^0 = 1$, (b) $n_S^0 = 10$, (c) $n_S^0 = 100$ and (d) $n_S^0 = 1000$. Average values of the number of CSCs, as calculated from Equation 3.21 are shown in bold.

Typically, an early (relatively immature) progenitor will divide into later (more mature) progenitors, undergoing only several rounds of self-renewing cell division before terminally differentiating. While the precise mechanisms of this process are likely more complicated (and are largely unknown at this time), the important effect is ultimately an amplification of the number of mature cells (denoted M) (these progenitor cells are hence sometimes termed "transit amplifying" cells [Clarke and Fuller, 2006]). In general, we assume that a CSC hierarchy contains N generations of progenitor cells, with N fixed. We will assume that first-generation progenitor cells, denoted P_1, divide with rate ρ_{P_1} to produce two second-generation progenitor cells (P_2). Likewise, i^{th}-generation progenitor cells (P_i) divide with rate ρ_{P_i} to produce two P_{i+1} cells. Finally, N^{th}-generation progenitor cells (P_N) divide with rate ρ_{P_N} to produce two mature cells. Generally it will be assumed that the progenitor cell division rates ρ_{P_i} are the same for each of the generations $i = 1, \ldots, N$. When this is the case, we will write $\rho_{P_i} = \rho_P$. Note that we have assumed each progenitor cell division to be symmetric – in particular, of the form $P_i \to P_{i+1} + P_{i+1}$ (for $i = 1, \ldots, N-1$). Other types of progenitor cell division may be envisaged; for example, divisions of the form $P_i \to P_{i+1} + M$ for any generation $i = 1, \ldots, N-1$ of progenitor cells. While the biological prevalence of such divisions is not yet known, we comment that such division pathways could be incorporated into the model as experimentalists further elucidate the nature of the progenitor cell sub-populations in various physiological and pathological systems.

The mature (fully differentiated) cells of the tumour are considered as caricatures of their normal counterparts, and, unlike stem and progenitor cells, lack any potential for division. They do, however, have the ability to die with rate Γ. We assume that the stem and progenitor cell populations do not have associated death rates. Figure 3.6 illustrates a simplified version of the CSC hierarchy for brain

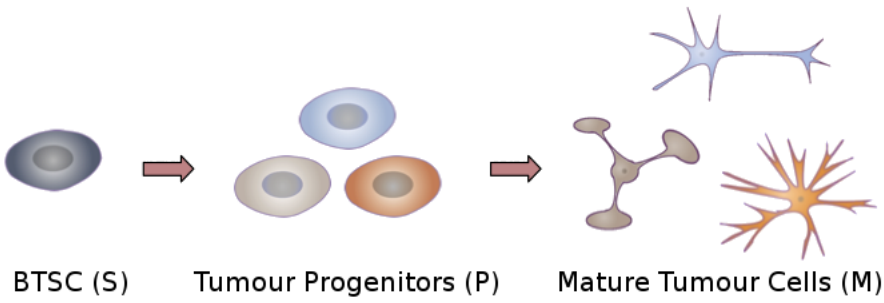


Figure 3.6: Illustration of cellular hierarchy for BTSCs

tumours. As specifics including the number of generations of progenitor cells are not yet well-understood, details are omitted here.

To summarize thus far, we permit the following types of cell division, with associated rates shown above the corresponding arrows:

$$\begin{aligned}
S &\xrightarrow{\rho_S r_1} S + S \\
S &\xrightarrow{\rho_S r_2} S + P_1 \\
S &\xrightarrow{\rho_S r_3} P_1 + P_1 \\
P_1 &\xrightarrow{\rho_{P_1}} P_2 + P_2 \\
&\vdots \\
P_i &\xrightarrow{\rho_{P_i}} P_{i+1} + P_{i+1} \\
&\vdots \\
P_N &\xrightarrow{\rho_{P_N}} M + M \\
M &\xrightarrow{\Gamma} \emptyset.
\end{aligned} \tag{3.65}$$

We define $p_{n_S,n_{P_1},\ldots,n_{P_N},n_M}(t)$ as the (conditional) probability of the system having n_S stem cells, n_{P_i} ith-generation progenitor cells, and n_M mature cells at time

t, given that n_S^0 stem cells, $n_{P_i}^0$ ith-generation progenitor cells, and n_M^0 mature cells were present at time t_0. $p_{n_S,n_{P_1},\ldots,n_{P_N},n_M}(t)$ is governed by the following master equation:

$$\begin{aligned}
\frac{dp(t)}{dt} &= \rho_S r_1(n_S-1)p_{n_S-1}(t) + \rho_S r_2 n_S p_{n_P-1}(t) \\
&+ \rho_S r_3(n_S+1)p_{n_S+1,n_P-2}(t) - \rho_S(r_1+r_2+r_3)n_S p(t) \\
&+ \sum_{i=1}^{N-1}\left\{\rho_{P_i}(n_{P_i}+1)p_{n_{P_i}+1,n_{P_{i+1}}-2}(t) - \rho_{P_i}n_{P_i}p(t)\right\} \\
&+ \rho_{P_N}(n_{P_N}+1)p_{n_{P_N}+1,n_M-2}(t) - \rho_{P_N}n_{P_N}p(t) \\
&+ \Gamma(n_M+1)p_{n_M+1}(t) - \Gamma n_M p(t), \qquad (3.66)
\end{aligned}$$

where for brevity, we have suppressed writing in full the indices of $p_{n_S,\ldots,n_M}(t)$ that do not differ from $n_S, n_{P_1}, \ldots, n_N, n_M$. (For example, for $N=4$, $p_{n_S,n_{P_1},n_{P_2},n_{P_3}+1,n_{P_4}-2,n_M}(t)$ is written as $p_{n_{P_3}+1,n_4-2}(t)$.) As in Equation 3.14, Equation 3.66 is valid for $n_S \geq 1$, $n_{P_i} \geq 2$, $i = 1,\ldots,N$, $n_M \geq 2$, and is adjusted accordingly otherwise. The initial condition for Equation 3.66 is

$$p_{n_S,n_{P_1},\ldots,n_{P_N},n_M}(t_0) = \delta_{n_S,n_S^0}\delta_{n_{P_1},n_{P_1}^0}\ldots\delta_{n_{P_N},n_{P_N}^0}\delta_{n_M,n_M^0}.$$

As we have done for the two-compartment model, we can use Equation 3.66 to find equations for the average numbers of S, P_i, and M cells. Doing so, we find

$$\begin{aligned}
\frac{d}{dt}S(t) &= \rho_S r S(t) \\
\frac{d}{dt}P_1(t) &= \rho_S(1-r)S(t) - \rho_{P_1}P_1(t) \\
\frac{d}{dt}P_i(t) &= 2\rho_{P_{i-1}}P_{i-1}(t) - \rho_{P_i}P_i(t), \text{ for } i = 2,\ldots,N \\
\frac{d}{dt}M(t) &= 2\rho_{P_N}P_N(t) - \Gamma M(t), \qquad (3.67)
\end{aligned}$$

with $S(0) = S_0 = n_S^0$, $P_i(0) = P_0^i = n_{P_i}^0$ for $i = 1,\ldots,N$, and $M(0) = M_0 = n_M^0$, and $r = r_1 - r_3$ as before.

Figures 3.7 and 3.8 show the behaviour of solutions to Equation 3.67 for two different sets of parameter values, corresponding to $N = 5$ and $N = 7$ generations of progenitor cells. Generally it is assumed that progenitor cells undergo several rounds of division before terminally differentiating; the particular values $N = 5$ and $N = 7$ were chosen for their relevance to the modelling of mammary stem cell dynamics to be presented in Chapter Five. In both cases, the system has $S_0 = 1$, $P_0^i = 0$ for $i = 1, \ldots, N$, and $M_0 = 0$, and in both cases we see a cascading increasing in the numbers of S, $P_1, \ldots, P_N$ and M cells, respectively. In Figure 3.7 (b), it is interesting to note that the sub-populations of 1^{st}- and 2^{nd}-generation progenitor cells remain smaller than the sub-population of CSCs; likewise, Figure 3.8 (b) shows that the sub-populations of 1^{st}-, 2^{nd}-, 3^{rd}- and 4^{th}-generation progenitor cells remain smaller than the CSC sub-population. It is generally assumed in the literature that the number of progenitor cells is significantly greater than the number of CSCs; indeed, we see in both Figures 3.7 and 3.8 that the total number of progenitor cells becomes much greater than the number of CSCs. However, the prediction that early-generation progenitor cells are fewer in number than CSCs (depending on parameter values) may not have been obvious.

The fractions of CSCs, total progenitors (that is, the sum of the fractions of each of the N generations of progenitor cells) and mature cells are shown in Figures 3.7 (c) and 3.8 (c). We see that, as in the case of the two-compartment model discussed above, the *fractions* of cells plateau while the absolute numbers of cells grow exponentially. In the plateau phase, for the parameter values selected, CSCs constitute the smallest fraction of the total cells (when all of the N generations of progenitor cells are grouped together into the progenitor cell fraction), while mature cells constitute the largest fraction.

Recall that it has been assumed that mature cells are terminally-differentiated,

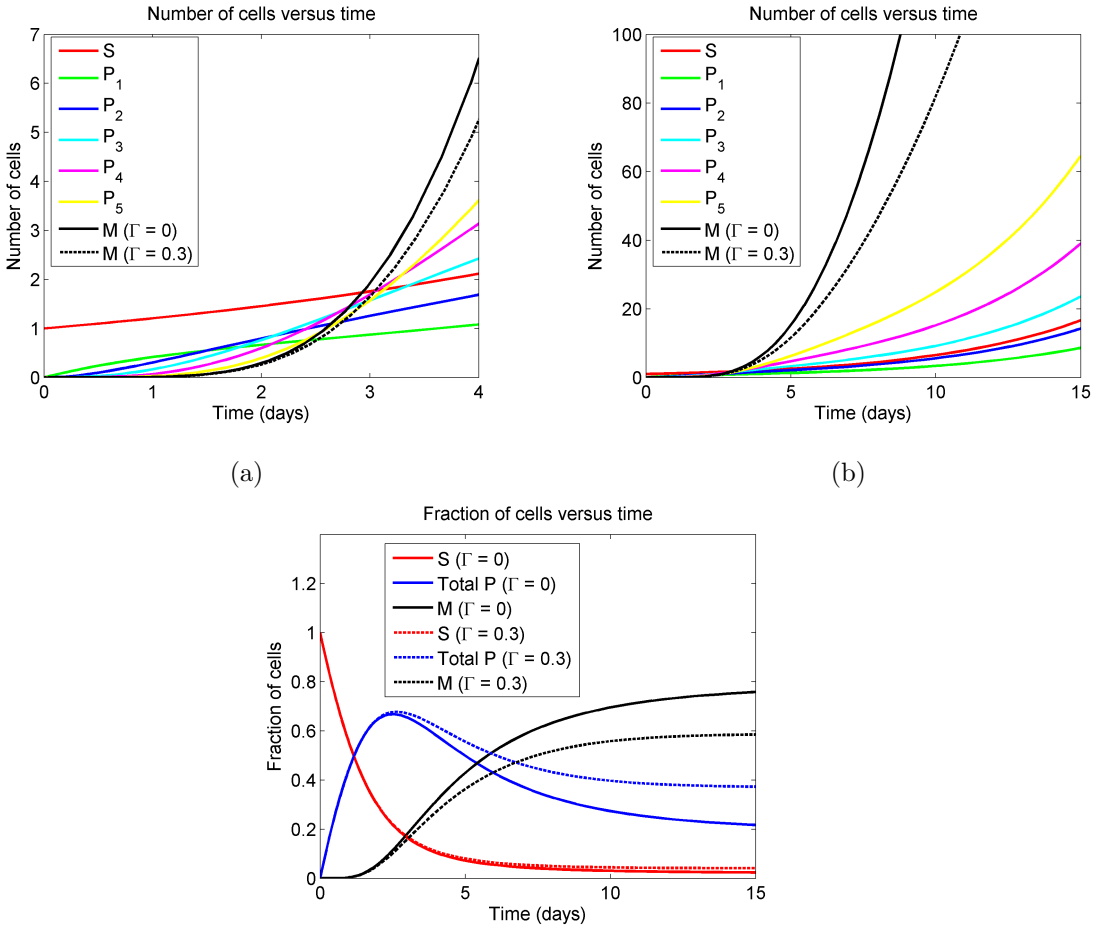
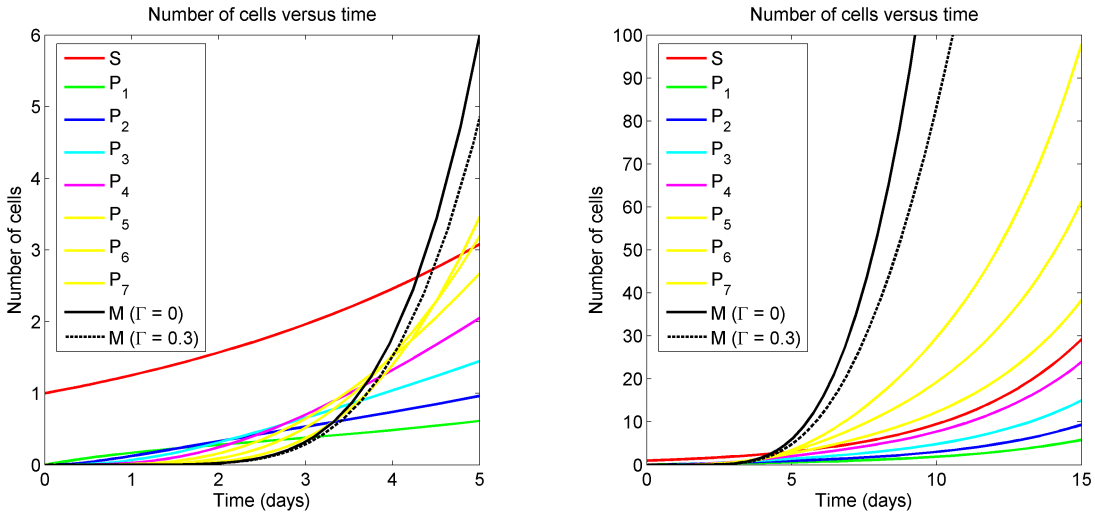
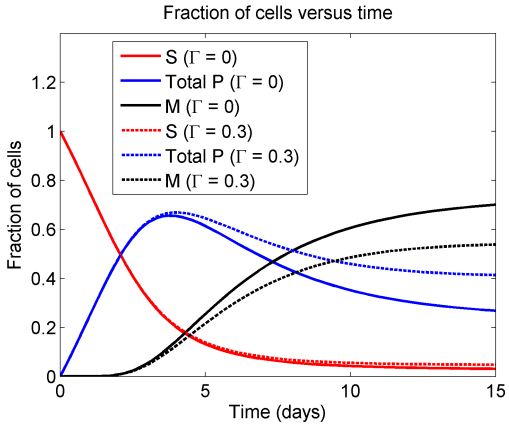


Figure 3.7: Number of cells (a), (b) and fraction of cells (c) versus time for a CSC system with 5 generations of progenitor cells. Parameter values are as follows: $\rho_S = 0.75$ day^{-1}, $\rho_P = 0.9$ day^{-1}, $r = 0.25$, $\Gamma = 0$ day^{-1} (solid curves), $\Gamma = 0.3$ day^{-1} (dashed curves).



(a)

(b)



(c)

Figure 3.8: Number of cells (a), (b) and fraction of cells (c) versus time for a CSC system with 7 generations of progenitor cells. Parameter values are as follows: $\rho_S = 0.45$ day^{-1}, $\rho_P = 0.9$ day^{-1}, $r = 0.5$, $\Gamma = 0$ day^{-1} (solid curves), $\Gamma = 0.3$ day^{-1} (dashed curves).

non-dividing cells. In the context of clinical cancer biology, the fraction of cells in the cell cycle (that is, actively dividing or preparing to divide) is known as the growth fraction [McKinnel et al., 2006]. Estimates of growth fractions of human tumours range from 0.06 to 0.9 [Hall, 1994].

Particular cells may not be part of the growth fraction because they are terminally-differentiated, non-proliferating cancer cells – i.e. cells of our type M. Alternatively, cancer cells may not be part of the growth fraction because they are quiescent, by which we mean that they have (temporarily) exited the cell cycle but may once again begin dividing when triggered to do so by some cue. In normal stem cell biology, quiescence of stem cells is thought to be a widespread phenomenon [Wicha et al., 2006]. In CSC biology, it is often hypothesized that CSCs may also at times be quiescent [Wicha et al., 2006]. This is currently not a feature of our model, although the introduction of quiescence may be classified in the category of future work. While we have not included quiescence, it is evident in Figures 3.7(c) and 3.8(c) that even if a fraction of the fraction of CSCs were considered to be quiescent, then the majority of non-proliferating cells would still be M-type cells. Denoting the fraction of M cells f_M and the growth fraction f_G, we have that $f_G \approx 1 - f_M$. From Figures 3.7(c) and 3.8(c), the growth fraction f_G evidently increases (decreases) as the mature cell death rate Γ increases (decreases) and other parameter values are held fixed. We point this out to illustrate that a wide range of clinically-relevant growth fractions can be obtained in our model simply by varying Γ. Of course, we expect that by varying other parameters such as ρ_S, r and ρ_P, the growth fraction can also be altered.

We may either numerically solve Equations 3.67, or else solve Equations 3.67 analytically. We refrain from doing the latter here because, despite the linear homogeneous nature of Equations 3.67, the solution of these $N + 2$ coupled ODEs

will become cumbersome to the point that its meaning is obscured. Instead, we propose a simplified, or "condensed", model that may be more amenable to such treatment and may yield some useful results.

3.6.2 A simplified approach

As mentioned above, the net effect of the N generations of progenitor cells is to amplify the number of mature cells, so that a single P_1 cell eventually gives rise to 2^N M cells, according to the progression

$$P_1 \overset{\rho_{P_1}}{\rightarrow} 2P_2 \overset{\rho_{P_2}}{\rightarrow} \ldots \overset{\rho_{P_{N-1}}}{\rightarrow} 2^{N-1}P_N \overset{\rho_{P_N}}{\rightarrow} 2M.$$

This effect can be captured more succinctly, by simply assuming that a P_1 cell divides to produce 2^N M cells with some division rate ρ_P^*. (This simplification embodies what we will refer to as the "condensed" model of cancer stem cell growth, while the model in Equation 3.66 will be referred to as the "full" model.) To find ρ_P^*, we note that each of the growth rates ρ_{P_i} is related to the reciprocal of the cell cycle time τ_i of P_i cells by $\rho_{P_i} = \ln 2/\tau_i$. Thus, the effective growth rate of progenitor cells in the condensed model is

$$\rho_P^* = \frac{\ln 2}{\sum_{i=1}^{N} \tau_i} = \frac{\ln 2}{\sum_{i=1}^{N} \frac{\ln 2}{\rho_{P_i}}} = \frac{1}{\sum_{i=1}^{N} \frac{1}{\rho_{P_i}}}, \tag{3.68}$$

which reduces to $\rho_P^* = \rho_P/N$ in the case that $\rho_{P_i} = \rho_P$ for $i = 1, \ldots, N$. We will generally assume that this is the case (i.e. that the division rates are the same for each generation of progenitor cell); we simply point out that the situation may be generalized to include cases in which the division rates of different progenitor cells are not the same. Subsituting the notation P for the lone distinct class of

progenitor cells P_1, the types of cell division permitted in the condensed model are then

$$\begin{aligned} S &\xrightarrow{\rho_S r_1} S+S \\ S &\xrightarrow{\rho_S r_2} S+P \\ S &\xrightarrow{\rho_S r_3} P+P \\ P &\xrightarrow{\rho_P^*} 2^N M \\ M &\xrightarrow{\Gamma} \emptyset. \end{aligned} \tag{3.69}$$

The condensed model of CSC growth has the advantage of simplicity over the full model; in particular, it allows for adjustment of the number of generations of progenitor cells N without altering the structure of the model. Such a model has been used, for example, in modelling the treatment dynamics of chronic myeloid leukemia [Michor et al., 2005a]. We will consider a comparison of the full and condensed models later, but first we proceed to write the master and average equations.

For the condensed model depicted in Equation 3.70, we define $p_{n_S,n_P,n_M}(t)$ as the (conditional) probability of the system containing n_S stem cells, n_P progenitor cells, and n_M mature cells at time t, given that n_S^0 stem cells, n_P^0 progenitor cells, and n_M^0 mature cells were present at time t_0. This conditional probability obeys the master equation

$$\begin{aligned} \frac{d}{dt}p(t) &= \rho_S r_1(n_S-1)p_{n_S-1}(t) + \rho_S r_2 n_S p_{n_P-1}(t) \\ &+ \rho_S r_3(n_S+1)p_{n_S+1,n_P-2}(t) - \rho_S(r_1+r_2+r_3)n_S p(t) \\ &+ \rho_P^*(n_P+1)p_{n_P+1,n_M-2^N}(t) - \rho_P^* n_P p(t) \\ &+ \Gamma(n_M+1)p_{n_M+1}(t) - \Gamma n_M p(t). \end{aligned} \tag{3.70}$$

As in Equation 3.66, we have suppressed writing in full the indices of $p_{n_S,n_P,n_M}(t)$ that do not differ from n_S, n_P, n_M. Equation 3.66 is valid for $n_S \geq 1$, $n_P \geq 2$, $n_M \geq 2^N$, and is adjusted accordingly otherwise. The initial condition for Equation 3.70 is

$$p_{n_S,n_P,n_M}(t_0) = \delta_{n_S,n_S^0}\delta_{n_P,n_P^0}\delta n_M, n_M^0.$$

Using Equation 3.70 to find equations for the average numbers of S, P, and M cells, we find

$$\begin{aligned}
\frac{d}{dt}S(t) &= \rho_S r S(t) \\
\frac{d}{dt}P(t) &= \rho_S(1-r)S(t) - \rho_P^* P(t) \\
\frac{d}{dt}M(t) &= 2^N \rho_P^* P(t) - \Gamma M(t),
\end{aligned} \tag{3.71}$$

subject to $S(0) = S_0 = n_S^0$, $P(0) = P_0 = n_P^0$, and $M(0) = M_0 = n_M^0$, with $r = r_1 - r_3$, as before.

Since we have not bestowed upon our CSCs any additional abilities (in the form of division pathways), we see that the expression for $S(t)$ in Equation 3.71 is the same as in Equation 3.18. Solving the second equation using the solution from the first (given in Equation 3.21) and likewise using the solution of the second equation in solving the third equation, we have

$$P(t) = \left(\frac{\rho_S(1-r)S_0}{\rho_S r + \rho_P^*}\right)\left(e^{\rho_S r t} - e^{-\rho_P^* t}\right) + P_0 e^{-\rho_P^* t}, \tag{3.72}$$

$$\begin{aligned}
M(t) &= 2^N \rho_P^* \left(\frac{\rho_S(1-r)S_0}{\rho_S r + \rho_P^*}\right)\left(\frac{e^{\rho_S r t}}{\rho_S r + \Gamma} - \frac{e^{-\rho_P^* t}}{\Gamma - \rho_P^*}\right) + \left(\frac{2^N \rho_P^* P_0}{\Gamma - \rho_P^*}\right) e^{-\rho_P^* t} \\
&+ \left[M_0 - 2^N \rho_P^* \left(\frac{\rho_S(1-r)S_0}{\rho_S r + \rho_P^*}\right)\left(\frac{1}{\rho_S r + \Gamma} - \frac{1}{\Gamma - \rho_P^*}\right) - \frac{2^N \rho_P^* P_0}{\Gamma - \rho_P^*}\right] e^{-\Gamma t}.
\end{aligned} \tag{3.73}$$

3.6.3 Fraction of CSCs

With the intention of finding the fraction of CSCs in this three-compartment model, we define

$$X(t) = \frac{S(t) + P(t)}{S(t)} = 1 + \frac{P(t)}{S(t)}$$

and

$$Y(t) = \frac{S(t) + P(t) + M(t)}{S(t)} = X(t) + \frac{M(t)}{S(t)}.$$

(The reciprocal of $Y(t)$ is thus the desired fraction of cancer stem cells.) Using Equation 3.71, it is found that

$$\begin{aligned}
\frac{d}{dt}X(t) &= \frac{S(t)\frac{d}{dt}P(t) - P(t)\frac{d}{dt}S(t)}{S^2(t)} \\
&= \frac{S(t)(\rho_S(1-r)S(t) - \rho_P^* P(t))}{S^2(t)} - \frac{P(t)\rho_S r S(t)}{S^2(t)} \\
&= \rho_S(1-r) - (\rho_P^* + \rho_S r)\frac{P(t)}{S(t)} \\
&= \rho_S(1-r) + (\rho_P^* + \rho_S r) - (\rho_P^* + \rho_S r) - (\rho_P^* + \rho_S r)\frac{P(t)}{S(t)} \\
&= (\rho_S + \rho_P^*) - (\rho_P^* + \rho_S r)X(t). \qquad (3.74)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\frac{d}{dt}Y(t) &= \frac{d}{dt}X(t) + \frac{S(t)\frac{d}{dt}M(t) - M(t)\frac{d}{dt}S(t)}{S^2(t)} \\
&= \frac{d}{dt}X(t) + \frac{S(t)(2^N\rho_P^* P(t) - \Gamma M(t))}{S^2(t)} - \frac{M(t)\rho_S r S(t)}{S^2(t)} \\
&= \frac{d}{dt}X(t) + 2^N\rho_P^*\frac{P(t)}{S(t)} - (\rho_S r + \Gamma)\frac{M(t)}{S(t)} \\
&= \frac{d}{dt}X(t) + 2^N\rho_P^*\frac{P(t)}{S(t)} - (\rho_S r + \Gamma)\frac{M(t)}{S(t)} + 2^N\rho_P^* - 2^N\rho_P^* \\
&\quad +(\rho_S r + \Gamma)\left(1 - 1 + \frac{P(t)}{S(t)} - \frac{P(t)}{S(t)}\right) \\
&= -2^N\rho_P^* + \frac{d}{dt}X(t) + (2^N\rho_P^* + \rho_S r + \Gamma)\left(1 + \frac{P(t)}{S(t)}\right) \\
&\quad -(\rho_S r + \Gamma)\left(1 + \frac{P(t)}{S(t)} + \frac{M(t)}{S(t)}\right) + (\rho_S r + \Gamma)\left(1 + \frac{P(t)}{S(t)}\right) \\
&= -2^N\rho_P^* + \frac{d}{dt}X(t) + (2^N\rho_P^* + \rho_S r + \Gamma)X(t) \\
&\quad -(\rho_S r + \Gamma)Y(t) \\
&= -2^N\rho_P^* + (\rho_S + \rho_P^*) - (\rho_P^* + \rho_S r)X(t) \\
&+ (2^N\rho_P^* + \rho_S r + \Gamma)X(t) - (\rho_S r + \Gamma)Y(t). \qquad (3.75)
\end{aligned}$$

To find the steady-state values $\overline{X}$ and $\overline{Y}$ of $X(t)$ and $Y(t)$ respectively (and hence of the fraction of cancer stem cells $1/Y(t)$, we simply set $X'(t) = Y'(t) = 0$. From Equation 3.74, we find that

$$\overline{X} = \frac{\rho_S + \rho_P^*}{(\rho_P^* + \rho_S r)}. \qquad (3.76)$$

Substituting $X'(t) = Y'(t) = 0$ into Equation 3.75 along with $X(t) = \overline{X}$ as given in Equation 3.76, we solve for $\overline{Y}$ to yield

$$\overline{Y} = \frac{2^N\rho_P^*\rho_S(1-r) + (\rho_S r + \Gamma)(\rho_S + \rho_P^*)}{(\rho_S r + \Gamma)(\rho_S r + \rho_P^*)}. \qquad (3.77)$$

Upon taking the reciprocal of Equation 3.77, we see that

$$\frac{S(t)}{S(t) + P(t) + M(t)} \to \frac{(\rho_S r + \Gamma)(\rho_S r + \rho_P^*)}{2^N\rho_P^*\rho_S(1-r) + (\rho_S r + \Gamma)(\rho_S + \rho_P^*)}, \text{ as } t \to \infty. \qquad (3.78)$$

3.6.4 Comparison of the Full and Condensed Models

A comparison of the full and condensed models is presented in Figures 3.9 and 3.10, corresponding to stem cell hierarchies featuring $N = 5$ and $N = 7$ generations of progenitor cells, respectively. Figures 3.9 (b) and 3.10 (b) show that the desired plateauing of a system in which only progenitor cells exists occurs for both the full and condensed models. In Figures 3.9 (a), (c) and 3.10 (a), (c) we see that while the behaviour of the condensed model mimics that of the full model, the full model lags slightly behind the condensed model, with the relative difference appearing greatest in the early stages of growth. Figures 3.9 (d) and 3.10 (d) serve to illustrate that the fraction of CSCs exhibits very similar behaviour for both models after a transient period of disagreement, suggesting that Equation 3.78 serves as a good approximation to the long-term fraction of CSCs in the full model as well as being exact for the condensed model. A comparison of Figures 3.9 and 3.10 suggests that the quantitative disagreement between the two models is exaggerated as the number of generations N of progenitor cells is increased.

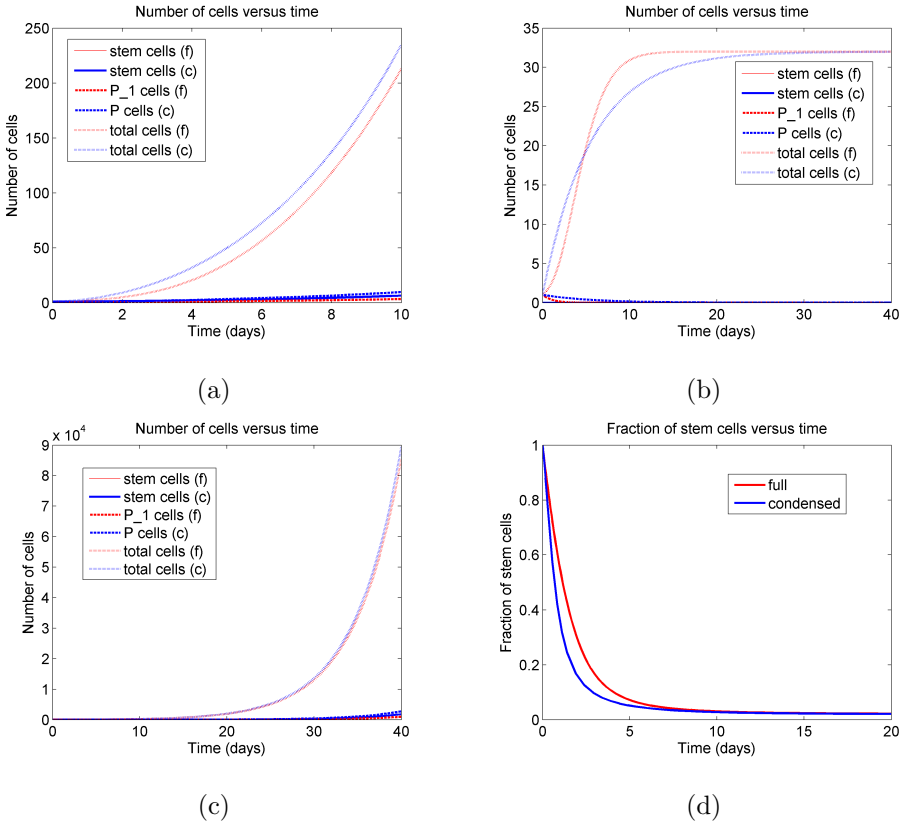


Figure 3.9: Comparison of the full (f) and condensed (c) models, with one stem cell at time $t = 0$ (a), (c), (d) and one P_1 (P) cell at time $t = 0$ (b). Parameter values are as follows: $N = 5$, $\rho_S = 0.75$ day^{-1}, $\rho_P = 0.9$ day^{-1}, $r = 0.25$, $\Gamma = 0$.

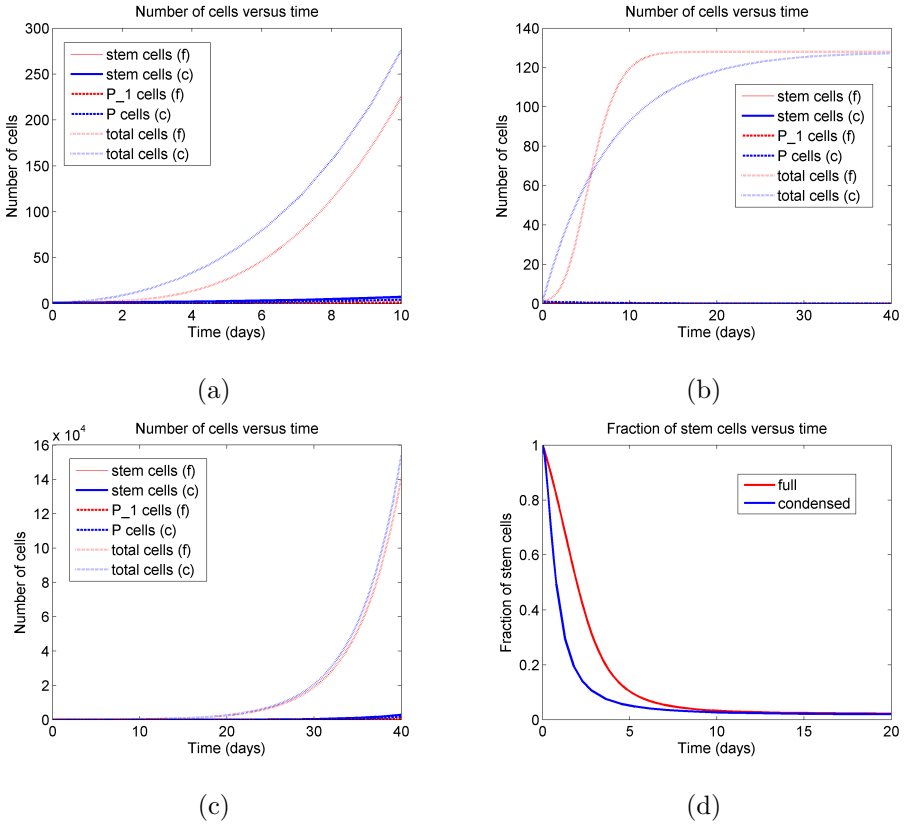


Figure 3.10: Comparison of the full (f) and condensed (c) models, with one stem cell at time $t = 0$ (a), (c), (d) and one P_1 (P) cell at time $t = 0$ (b). Parameter values are as follows: $N = 7$, $\rho_S = 0.4$ day^{-1}, $\rho_P = 1.05$ day^{-1}, $r = 0.5$, $\Gamma = 0$ day^{-1}.

Chapter 4

Application: Treatment of Brain Cancer

While the prognoses for many types of cancers have improved in recent years as diligent research has led to better diagnostic and treatment techniques, brain tumours remain consistently devastating in both adults and children. Cancers of the brain and spinal cord are the second most common cause of cancer mortality in children (National Cancer Institute of Canada data: http://www.ncic.cancer.ca). In adults, the median patient survival time following diagnosis with the most prevalent type of brain cancer, glioblastoma multiforme (GBM), is a dismal 6 to 12 months, a prognosis not significantly improved upon over the last few decades [DeAngelis, 2005]. The failure of standard treatment strategies consisting of surgical resection followed by radiation and/or chemotherapy to substantially improve patient outcomes reflects the fact that the mechanisms driving human brain tumour growth, as well as the interactions of cancer cells with their microenvironment and with therapeutics, are not yet well understood. In order to explain clinical and experimental results,

including the shortcomings of conventional treatments, much recent research has focused on the study of brain tumour development and growth in terms of the cancer stem cell hypothesis. Here, we use an adaptation of the simple two-compartment model developed in Chapter Three to examine recent developments in the theory of the treatment of brain tumours; much of this chapter is reproduced from [Turner et al., In press].

4.1 Biological Orientation

4.1.1 Neural Stem Cells in Homeostasis

The role of stem cells in the maintenance of the adult human central nervous system is only beginning to be understood [Uchida et al., 2000], although significantly more is known about embryonic brain development. The concept that neurogenesis (the production of neurons) continues throughout adulthood has only recently overturned the long-held doctrine of "no new neurons" after birth, with the demonstration that new neurons are indeed produced in certain areas of the brain until death [Eriksson et al., 1998]. Widespread recognition of adult mammalian neurogenesis is predated by initial work aimed to isolate mammalian neural stem cells – in particular, Reynolds and Weiss showed that cells expressing the protein Nestin could be extracted from the brains of adult mice and induced to proliferate and differentiate into neurons and astrocytes [Reynolds and Weiss, 1992].

The search for stem cells of the adult human central nervous system narrowed significantly in 2000, when Uchida *et al.* reported the isolation of adult human neural stem cells on the basis of the cell surface protein CD133 [Uchida et al., 2000]. CD133, also known as AC133 or human Prominin-1, is an 865-amino acid

long glycosylated protein embedded in the plasma membrane, consisting of five transmembrane domains including two prominent extracellular loops [Shmelkov et al., 2005, Neuzil et al., 2007]. Although its biological function has yet to be established [Shmelkov et al., 2005, Neuzil et al., 2007], it was first identified as a marker of hematopoietic stem cells and since its use by Uchida *et al.* (2000), it has been implicated as a marker for various putative cancer stem cells.

Despite the relative nascency of the field of neural stem cell biology, the cancer stem cell hypothesis has recently been extended to brain cancers [Ignatova et al., 2002, Hemmati et al., 2003, Singh et al., 2003, 2004, Galli et al., 2004]. In 2003, brain tumour cells expressing CD133 were identified as brain tumour stem cells (BTSCs) based on their exclusive ability to commence and support tumour growth [Singh et al., 2004]. These $CD133^+$ cells, isolated from human brain tumours, were able to generate tumours with the phenotypic signature of the original human malignancy when transplanted in small numbers (as low as 100) into the brains of non-obese diabetic/severe combined immunodeficient (NOD/SCID) mice. Importantly, $CD133^-$ cells were unable to initiate tumourigenesis in mice, even when transplanted in numbers on the order of tens of thousands.

4.1.2 Brain Tumour Treatment and BTSCs

An implication of the cancer stem cell hypothesis is that treatments aiming to indiscriminately destroy tumour cells in bulk may fail to consistently provide a cure because they spare a sub-population of cancer stem cells. Consistent with this proposition is the observation that human $CD133^+$ glioma cells exhibit radioresistance due to preferential activation of the DNA damage checkpoint response – in particular, the fraction of $CD133^+$ glioma cells has recently been found to be en-

riched following treatment with ionizing radiation, both *in vitro* and in the brains of immunocompromised mice [Bao et al., 2006].

Another treatment possibility may soon emerge, as Piccirillo *et al.* have recently demonstrated that certain bone morphogenetic proteins (BMPs) are capable of inducing $CD133^+$ human GBM cells to differentiate and adopt a $CD133^-$ cell phenotype, both in culture and, more importantly, in the brains of mice [Piccirillo et al., 2006]. Thus, pharmacological application of BMPs to brain tumours may direct BTSCs to differentiate into cells that are more vulnerable to traditional anti-cancer treatments (i.e. radiotherapy and chemotherapy). It is becoming increasingly clear that, under the brain cancer stem cell hypothesis, any potentially curative therapy must target BTSCs. The depletion of the cancer stem cell pool via induced differentiation represents one promising strategy.

4.2 Modelling Brain Tumour Treatment

Here, we focus on the application of radiotherapy as considered in [Bao et al., 2006] and theoretical BMP treatment suggested by [Piccirillo et al., 2006]. Previously (in Chapter 3), we considered conditions of exponential growth, which are appropriate *in vitro* or during the early stages of tumour development. To account for the *in vivo* effects of competition for space and nutrient limitations, we incorporate logistic growth by replacing the formerly-constant overall BTSC division rate ρ_S by

$$\widetilde{\rho_S}(S,P) = \rho_S \left(1 - \frac{S}{S_{\text{lim}}} - \frac{P}{P_{\text{lim}}}\right), \tag{4.1}$$

where S_{lim} and P_{lim} are the limiting populations of stem and progenitor cells, respectively.

The use of a logistic growth model now allows for discussion of various *in vivo* treatment strategies. In the following, we consider only the dynamics of BTSCs and progenitor cells, as these are the proliferating cells that are the targets of therapy.

We first consider the application of radiotherapy by using the exponential decay model described by Kohandel *et al.* [Kohandel et al., 2007]. The model is incorporated by adding treatment terms to Equations 3.20 for the numbers of stem and progenitor cells respectively, so that they read

$$\begin{aligned} \frac{dS(t)}{dt} &= \widetilde{\rho_S} r S(t) - \alpha_S S(t) \sum_i d_i f\left(\frac{t-t_i}{\tau_S}\right) \\ \frac{dP(t)}{dt} &= \widetilde{\rho_S}(1-r) S(t) - \alpha_P P \sum_i d_i f\left(\frac{t-t_i}{\tau_P}\right). \end{aligned} \tag{4.2}$$

Here, α_S and α_P represent the radiosensitivities of the CSCs and progenitor cells (respectively), in units of Gy^{-1}. The ith acute dosage (in Gy per day) is denoted d_i, which is applied at time t_i. The radiation clearance time (order of doubling time) of CSCs is τ_S (day), and τ_P (day) is that of the $\mathrm{CD133}^-$ progenitor cells. Finally, we define the function $f(x)$ by

$$f(x) = \begin{cases} e^{-x}, \text{ when } x \geq 0 \\ 0 \text{ otherwise.} \end{cases}$$

It has recently been observed in the laboratory that $\mathrm{CD133}^+$ GBM cells exhibit greater radioresistance than do cells lacking CD133. In particular, following treatment of cultures of cells isolated from primary human glioblastomas or from human glioblastoma xenografts grown in murine hosts with 2-5 Gy ionizing radiation, the fraction of $\mathrm{CD133}^+$ cells was found to have increased four- to five-fold [Bao et al., 2006]. Similar results were obtained *in vivo*, with murine subjects bearing

xenograft tumours. These results indicate that the radiosensitivity of BTSCs may be significantly smaller than that of GBM progenitor cells, and thus in our model [Equation 4.2] we should choose $\alpha_S < \alpha_P$. Based on data of Bao *et al.* (2006), α_S was estimated to be 0.2 Gy^{-1}, a value consistent with a previous estimate of stem cell radiosensitivity given by Sachs and Brenner [Sachs and Brenner, 2005]. We estimate α_P to be threefold greater (0.6 Gy^{-1}). Although not considered herein, Bao *et al.* (2006) suggest that administration of an inhibitor of the Chk1 and Chk2 checkpoint kinases (specifically, debromohymenialdisine) involved in cell cycle regulation concurrent with ionizing radiation renders CD133^+ cells more vulnerable, thus acting to increase the value of α_S.

While recognizing the radioresistance of BTSCs is an important if somewhat grim realization, encouraging news comes from recent experiments by Piccirillo and colleagues supporting the notion that BMPs may induce CD133^+ GBM cells to differentiate into cells with decreased tumourigenic potential [Piccirillo et al., 2006]. BMPs are a subgroup of the transforming growth factor β (TGFβ) family of cell-regulating proteins [Chen et al., 2004]. While BMPs play various roles throughout the body, in neural development they typically induce differentiation into astroglial cells [Piccirillo et al., 2006]. *In vitro*, treatment of glioblastoma-derived cells with BMPs resulted in significantly reduced (in the range of 50%) CD133^+ populations. *In vivo*, immunodeficient mice that received gradual administration of BMP4 via beads implanted into their brains either concurrent with or following xenograft of glioma cells lived longer than control mice. The precise mechanisms through which BMPs reduce the tumourigenicity of CD133^+ GBM cells remain unclear [Piccirillo et al., 2006]; mathematically, in our model we interpret the effects of BMP4 as decreasing the net symmetric division rate r while leaving r_2 fixed. Based on the work of Piccirillo *et al.* (2006) we estimate that, starting from a pre-treatment

value of $r = 0.1$, the effect of treatment with BMP4 is to reduce r to $r = -0.1$. Note that we have previously defined $r = r_1 - r_3$, so that the change of r to a negative value represents a simultaneous increase in the proportion of symmetric differentiation divisions and decrease in the proportion of symmetric self-renewing divisions.

In the laboratory, the minimum number of $CD133^+$ GBM cells required for tumour formation upon injection into immunocompromised mice has been reported as approximately 100, while xenograft of up to 10^6 $CD133^-$ cells lacked the capacity to be tumourigenic [Singh et al., 2004, Bao et al., 2006]. For our simulations, we take $S_0 = 5000$ and $P_0 = 10^5$. This is roughly equivalent to implanting a tumour of 2-3mm in diameter, which is initially about five percent stem cells by composition. The doubling time of the $CD133^+$ subpopulation is estimated to be about two days [Kohandel et al., 2007], resulting in $\rho_S r \approx 0.35\text{day}^{-1}$. We take S_{lim} and P_{lim} to be 10^7 and 10^8, respectively.

Solving Equations 4.2 numerically, we can consider the effects of various treatment strategies on GBM cell populations, as shown in Figure 4.1. A feature of our results is the observed enrichment of the $CD133^+$ population following treatment with ionizing radiation, consistent with the experimental results of Bao *et al.* (2006) In our model, the greater radiosensitivity of $CD133^-$ cells dictates that the fraction of $CD133^+$ cells increases. Furthermore, the logistic growth condition leads to an increase in the growth rate as $CD133^-$ cells are destroyed; this allows $CD133^+$ cells to repopulate. Consequently, we observe a slight increase in the number of $CD133^+$ cells relative to the control case once radiotherapy has ended and the number of cells has plateaued [Figure 4.1 (a)].

Our numerical results indicate that a BMP-type therapy is effective in decreasing $CD133^+$ cell numbers at the expense of a slight increase in the number of $CD133^-$

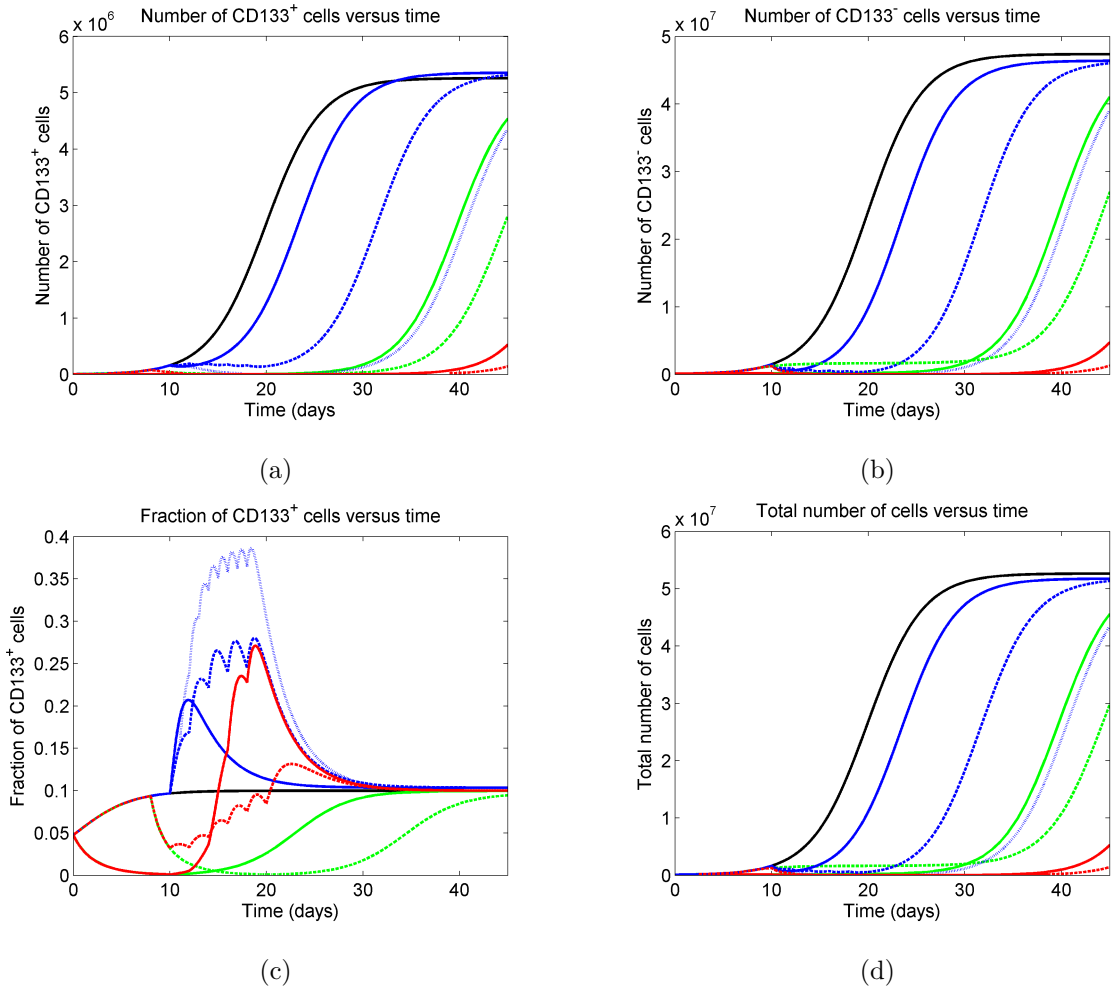


Figure 4.1: Number of CD133$^+$ cells (a), number of CD133$^-$ cells (b), fraction of CD133$^+$ cells (c) and total number of cells (d) following various combinations of treatments. The legend is as follows: black, solid (no radiation or BMPs); blue, solid (3Gy ionizing radiation (IR) administered at day 10); blue, dashed (10Gy IR administered in 2Gy doses on days 10, 12, 14, 16, 18); blue, dotted (18Gy IR administered in 2Gy doses on each of days 10-18); green, solid (10 days BMP4 administered from days 0-10); green, dashed (BMP4 administered from days 8-20); red, solid (BMP4 administered from days 0-10 followed by 10Gy radiation administered in 2Gy doses on days 10, 12, 14, 16, 18); red, dashed (BMP4 administered from days 8-20 with 10Gy radiation administered in 2Gy doses on days 10, 12, 14, 16, 18). Image from [Turner et al., In press].

cells and hence in the total number of cells. This net increase is a consequence of our assumption that each $CD133^+$ cell produces two $CD133^-$ cells when it differentiates. If, rather, a $CD133^+$ cell transitions directly into a $CD133^-$ cell (that is, an event of the form $S \rightarrow P$), then such an increase in overall tumour bulk is not to be expected. In either case, our results suggest that radiotherapy may be more effective when combined with BMP (or another type of differentiation-inducing) therapy, as is evidenced by the length of time for which cell number is suppressed below saturation in the case of radio- and BMP combination therapy relative to other strategies. It should be stressed, however, that at the present time clinical therapy with BMPs is not feasible due to the many questions that remain regarding the actions and consequences of these proteins – the theoretical results merely provide additional motivation to investigate such differentiation-inducing factors. A further prediction of the model is that, following the discontinuation of a treatment regime composed of any combination of radiotherapy and BMP-type therapy, the GBM cell population will recover to its original constitution (unless the $CD133^+$ population has been rendered extinct). That is, the period of change in the percentage of $CD133^+$ GBM cells that begins with the onset of treatment is only transient, and the tumour will eventually recover its original phenotype, as shown in Figure 4.1 (c).

Chapter 5

Sphere-Forming Assays, (Cancer) Stem Cells, and the Epithelial-Mesenchymal Transition

In recent years the epithelial-mesenchymal transition (EMT), a cellular differentiation process wherein epithelial cells lose expression of genes that facilitate the ordered inter-cellular junctions that are characteristic of epithelial surfaces and instead become mobile, has been implicated in endowing certain cancer cells with the ability to metastasize – that is, to disseminate from the primary tumour, enter circulation via blood vessels, and establish themselves in new locations throughout the body, potentially leading to secondary tumours. It is also an implication of the CSC hypothesis that, in order for successful metastatic growth to occur, the metastasized cell must be a CSC. A link between EMT and CSCs has been pro-

posed – particularly, it has been postulated that the EMT may foster metastasis by producing cells that are not only motile but are indeed CSCs, with recent experimental support for this notion in breast cancer and breast-cancer-related cell lines [Mani et al., 2008, Morel et al., 2008]. Here, we apply the models of Chapter Three to investigate the nature of the mechanism by which EMT induces the experimentally-observed enrichment of cells expressing stem-like characteristics in populations of mammary epithelial cells [Mani et al., 2008, Morel et al., 2008]. To aid in this investigation (particularly in determining appropriate values of model parameters), we first consider modelling the mammosphere assay, a technique used in the investigation and enrichment of mammary stem cells.

5.1 Sphere-Forming Assays

In the study of stem cells and CSCs, it is crucial to design laboratory protocols that serve to functionally identify, and enrich, populations of these cells. To this end, investigators have developed sphere-forming assays, wherein cells are dissociated and placed into culture conditions such that they may develop into free-floating, sphere-shaped clusters of cells [Singec et al., 2006]. Such assays were originally developed for the study of neural stem cells [Reynolds and Weiss, 1992] – in this context, the spheres of cells are known as "neurospheres". More recently, an analogous "mammosphere" assay has been established for the study of mammary stem cells [Dontu et al., 2003]. Because sphere-forming assays are relevant techniques in cancer biology and normal stem cell biology, we will frequently use the abbreviation (C)SC to indicate that the cells that are the subject of discussion may be either stem cells or cancer stem cells.

Mammosphere and neurosphere assays have been used to estimate the numbers

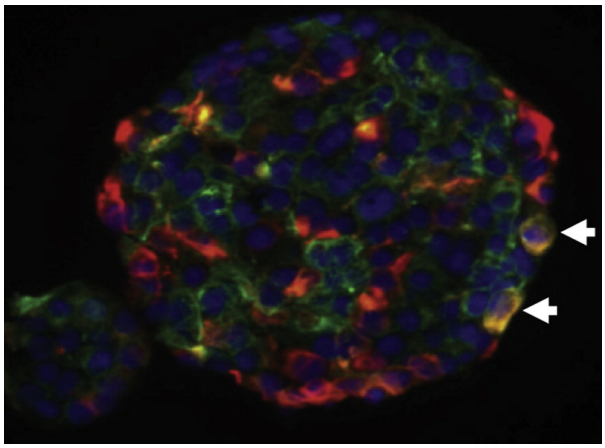


Figure 5.1: An image of a mammosphere; yellow staining represents the presence of markers indicating cells with mulitpotentiality, suggestive of mammary stem cells. Adapted from [Mani et al., 2008].

of (C)SCs in cell populations [Reynolds and Rietze, 2005]; often in such cases it is assumed that the fraction of cells plated that actually form spheres is representative of the frequency of (C)SCs in that population [Reynolds and Rietze, 2005] (here, by "plating" a cell, we mean introducing that cell to the sphere-forming assay). Interpretation of data from sphere-forming assays in this manner requires the making of some assumptions:

- each (C)SC plated forms a sphere
- each sphere is clonal in origin (i.e. is derived from a single (C)SC)
- non-stem cells (or non-CSCs) do not form spheres.

Evidence challenging the second assumption has been presented by Singec *et al.* [Singec et al., 2006], who demonstrated, using video microscopy, that neurospheres

have a tendency to migrate toward each other and fuse, contradicting the assumption of clonality. The complicating effects of sphere migration and fusion are likely dependent on the density at which cells are plated – that is, when many cells are plated per unit volume, fusion is likely to be more significant than when only a few cells are plated per unit volume. (In a non-spatial model of cell growth such as ours, the phenomena of cell migration and fusion cannot be directly addressed. We simply mention that cell migration and fusion *may* contribute to polyclonality of spheres, and suggest that in the future, a spatial model may incorporate these effects.)

Additionally, it is maintained by some that non-stem cells (namely, progenitor cells) *do* sometimes show the ability to form spheres [Reynolds and Rietze, 2005]. A handful of examples supporting this claim exist in the literature [Seaberg and van der Kooy, 2002, Tropepe et al., 2000, Zhang et al., 1998]. However, here we do not consider progenitor cells as having the ability to self-renew and so it is a feature of our model that progenitor cells will not form spheres in sphere-forming assays.

Regarding the first assumption (i.e. that each (C)SC plated forms a sphere), we point out that it is a feature of a stochastic model of (C)SC division that each (C)SC does *not necessarily* form a sphere, as will become evident shortly (the first division of a (C)SC plated in a sphere-forming assay may be a symmetric commitment-type division, for example). Thus, a stochastic model of (C)SC division predicts that the number of spheres formed in a sphere-forming assay will be an under-representation of the actual number of (C)SCs in the original population.

5.1.1 The Mammosphere Assay

Our particular interest in sphere-forming assays is a result of the use of the mammosphere assay in recent experiments by the group of Robert Weinberg suggesting the role of the EMT in generating cells with properties of stem cells [Mani et al., 2008], as will be discussed later in this chapter. Now, we focus our attention on the mammosphere assay, as performed by Mani *et al.* [Mani et al., 2008]. We do so for two reasons: to improve our understanding of the nature of the mammosphere-forming assay, and to aid in the search for parameter values to be used in the following section, which investigates possible mechanisms by which mammary epithelial cells may gain the properties of mammary stem cells.

Mani *et al.* [Mani et al., 2008] defined a mammosphere to be an adherent cluster of cells with diameter d at least 75μm after 7-10 days in culture, (presumably) arising from a single cell. We note that a mammosphere should be at least approximately spherical, and that its volume is thus given by

$$V \approx \frac{4}{3}\pi r^3 = \frac{\pi}{6}d^3.$$

Thus, a mammosphere of diameter 75μm has volume $V \approx 2.2\times 10^5 \mu\text{m}^3$. We assume that an individual mammary epithelial/mesenchymal cell is roughly spherical with diameter 11μm, lending it a volume of roughly $700\mu\text{m}^3$. Kepler's Conjecture asserts that the packing of spheres (in three-dimensional space) results in at least just over 25% 'wasted' volume [Casti, 2001]; we suppose 30% 'wasted space' within a mammosphere (presumably occupied by extracellular matrix secreted by cells) so that $0.7 \times 2.2 \times 10^5 \mu\text{m}^3 = 1.54 \times 10^5 \mu\text{m}^3$ is available for occupation by cells. Dividing this available space by the estimated volume of a single cell gives $1.54 \times 10^5 \mu\text{m}^3/700\mu\text{m}^3 = 220$ as the estimated number of cells representing the threshold of classification of a culture of cells as a mammosphere. Thus, we assume that a

mammosphere has at least 220 cells after 7-10 days of growth.

To simulate this mammosphere assay, we perform numerous runs of Gillespie's algorithm using the division pathways illustrated in either Equation 3.66 or Equation 3.70 of Chapter Three. Each run begins with a single stem cell and lasts for a period of $t = 8.5$ days. A mammosphere is considered to have formed if the total number of cells is at least 220, corresponding to the above estimate.

5.1.2 Parameter Values

We would like to assign numerical values to the seven parameters in Equation 3.70: ρ_S, r_1, r_2, r_3, ρ_P, N, and Γ. The problem of determining growth and death rates for cancer modelling is not new; however, in the present context of CSC biology, this problem is aggravated by the presence of sub-population-specific growth rates ρ_S and ρ_P and division-type probabilities r_1, r_2 and r_3. In treating a cancer as a collection of sub-populations of cells having distinct growth characteristics, we cannot readily make use of experimentally-acquired measurements of cell populations as wholes. Some recent experimental work shows promise toward directly elucidating the growth characteristics of individual cell types. The group of Tannishtha Reya imaged hematopoietic stem cells in real time by time-lapse microscopy, measuring frequencies of symmetric self-renewal, asymmetric self-renewal and symmetric commitment of these cells [Wu et al., 2007]. (Interestingly and importantly, they also showed that these division frequencies can be altered by certain cancer-associated proteins [Wu et al., 2007].) Prior to this, the group of Eric Jervis at the University of Waterloo began using high-resolution videomicroscopy to track cell division and morphological properties of hematopoietic stem cells [Dykstra et al., 2006]. Advances such as these may soon lead to the direct measurement of (C)SC and

progenitor cell properties, but in the meantime we propose an indirect method of estimating these quantities, based on the numerical simulation of sphere-forming assays over ranges of candidate parameter values and the comparison of results to a set of experimentally-derived selection criteria. We now proceed in this direction.

For simplicity, and because we lack quantitative information to the contrary, it is assumed here that the probability per unit time Γ of an individual mature cell dying is zero. This assumption, coupled with the normalization $r_1 + r_2 + r_3 = 1$ imposed on the probabilities of each of the three types of stem cell division, leaves us with five free parameters. To narrow the range of possible values of the remaining parameters, we look to extract additional conditions from the data of Mani *et al.* [Mani et al., 2008].

Mani *et al.* [Mani et al., 2008] selected cells for the mammosphere assay from a population of nontumourigenic immortalized mammary epithelial cells (HMLECs). HMLECs are cells that have been engineered to overcome the typically-observed eventual senescence that occurs when normal human mammary epithelial cells are cultured [Elenbaas et al., 2001]. Figure 5.2, from the Supplementary Data of [Mani et al., 2008], indicates that over a twelve-day period, the fraction of cells expressing high levels of the cell surface protein CD44 and low or undetectable levels of the cell surface protein CD24 ($CD44^{high}/CD24^{low}$ cells – this phenotype serves as a marker of mammary epithelial stem cells and of breast cancer stem cells [Mani et al., 2008]) fluctuated around a value of about 2%. We interpret this value as the steady-state fraction of $CD44^{high}/CD24^{low}$ cells in the population. Furthermore, we assume that the $CD44^{high}/CD24^{low}$ phenotype is a perfect marker of mammary epithelial stem cells, in the sense that the terms "stem cell" and "$CD44^{high}/CD24^{low}$ cell" can be used interchangeably. Thus, we require that the steady-state fraction of stem cells F_S be approximately 0.02 – this serves as one additional criterion that our set of

parameter values should satisfy.

We note that, while the $CD44^{high}/CD24^{low}$ phenotype is used in laboratory settings as a marker of mammary epithelial stem cells and breast cancer stem cells, its relevance as such has been questioned [Fillmore and Kuperwasser, 2007], and the assumption that it is a perfect marker does not reflect the possibility that some non-stem cells (e.g. early-generation progenitor cells) may also share this cell surface protein expression profile. Recent reports suggest that expression of the enzyme aldehyde dehydrogenase (ALDH) [Ginestier et al., 2007] or of epithelial-specific antigen (ESA) in addition to the $CD44^{high}/CD24^{low}$ phenotype [Fillmore and Kuperwasser, 2008] may serve as a better marker. However, as data relating the numbers of cells defined by these phenotypes to the EMT are currently unavailable, we reserve the refinement of this modelling assumption for future work.

It was reported that roughly $4/1000 = 0.4\%$ of HMLECs plated in the mammosphere assay formed mammospheres [Mani et al., 2008] when these cells were selected from an unsorted population (that is, from a mixture of $CD44^{high}/CD24^{low}$ cells and cells without this phenotype). We interpret the formation of a mammosphere in this way as two independent events occurring in succession: first, a stem cell is chosen from the unsorted population of HMLECs. Next, that stem cell proliferates in such a way that it forms a mammosphere. If we assume that cells are taken randomly from a well-mixed population of HMLECs with a steady-state composition in which $(100 \times F_S)\%$ of cells are stem cells, then the first event (the selection of a stem cell) occurs with a probability of F_S. The second event (the formation of a mammosphere by a selected stem cell) occurs with some relative frequency f_m; the independence of these two events means that

$$MFE = F_S f_m,$$

	Percentage of CD44^{hi}/CD24^{low} cells following Tamoxifen treatment				
	Days 0	3	6	9	12
Vector	1.38	1.44	2.34	1.99	2.36
Snail-ER	0.15	0.62	1.80	14.30	13.20
Twist-ER	0.16	0.10	2.20	21.60	21.50

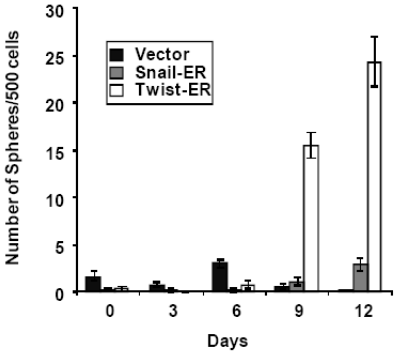


Figure 5.2: Supplemental Figures S2C (left) and S2D (right), from [Mani et al., 2008]. The left-hand-side gives the percentage of HMLECs expressing the CD44^{high}/CD24^{low} breast/breast cancer stem cell phenotype as measured at times following exposure of cells to tamoxifen, while the right-hand-side quantifies mammosphere-forming ability of the same cells. "Vector" HMLECs are control cells expressing neither Snail nor Twist.

where we have defined the mammosphere-forming efficiency (MFE) of a population of cells as the relative frequency of a cell from the population forming a mammosphere. For consistency with Mani *et al.* [Mani et al., 2008] (Figure 5.2) we require that the MFE be roughly 0.004 (or in the range of 0.003-0.005).

The average mammosphere size was also measured by Mani *et al.* [Mani et al., 2008] to be about 300 cells. In summary, we have three numerical criteria that we would like our set of parameter values to satisfy:

- the steady-state fraction of stem cells F_S is about 0.02 (0.02 ± 0.002)
- the $MFE = F_S f_m$ is about 0.004 (0.004 ± 0.001)
- the average size of a mammosphere (AMS) is about 300 cells (300 ± 30).

In the list above, we have used the qualifier "about" to reflect the fact that a degree of uncertainty in these values exists. This uncertainty, exacerbated by the number of free parameters being greater than the number of quantitative selection criteria, in turn lends a degree of arbitrariness to our selection of parameter values.

5.1.3 Simulation of the Mammosphere Assay

To proceed, we perform mammosphere simulations using Gillespie's algorithm, looping over a range of values of each of the free parameters ρ_S, ρ_P, N, r_1 and r_3. In particular, we take candidate values of $\rho_S \in \{0.25, 0.3, \ldots, 1.55, 1.6\}$ (with units of day^{-1}), and candidate values of ρ_P from the same set. We restrict N to $N \in \{3, 4, 5, 6, 7, 8\}$. For each of r_1, r_2 and r_3, we consider values in $\{0.05, 0.1, \ldots, 0.85, 0.9\}$ subject to $r_1 + r_2 + r_3 = 1$.

In Chapter Three, we introduced two forms of a general model of growth for cell populations with N generations of progenitor cells: a "full model", with division pathways given in Equation 3.66, and a "condensed model", with division pathways given in Equation 3.70. The condensed model has the advantage of simplicity, as it has only three distinct sub-populations of cells S, P and M, regardless of the number of generations N of progenitor cells that it models – this simplicity translates to a significant advantage versus the full model in terms of computational time required to perform stochastic simulations of the model via Gillespie's algorithm. The full model, on the other hand, has the advantage of physicality – distinct subpopulations of progenitor cells divide, and instantaneous jumps from a single first-generation progenitor cell to 2^N mature cells are not permitted in the full model while they are in the condensed model. This distinction may be relevant when considering simulation of the mammosphere assay, in which relatively small numbers of cells divide over a relatively short period of time.

Weighing these considerations, we use the condensed model in our initial round of stochastic simulations – this initial round serves to narrow the set of consistent parameter values, using the criteria listed above as parameter selection criteria. With a much smaller set of parameter values from which to choose (i.e. over which to loop in simulations of the mammosphere assay), the inefficiency associated with simulating the larger full model becomes less of a drawback.

Simulating the mammosphere-forming assay 1000 times per set of parameter values, we find that several distinct sets of parameter values meet the selection criteria, as recorded in Tables 5.1 - 5.4. In the cases of $N = 4$ and $N = 5$, only the parameter ρ_P shows significany variability in the sense that numerous values of ρ_P satisfy the selection criteria when the other parameters are fixed. For $N = 6$ and $N = 7$, more sets of consistent parameter values are obtained, with the trend that

ρ_S (day^{-1})	ρ_P (day^{-1})	r	r_2	AMS	f_m	MFE	F_S
0.8	1.6	0.2	0.7	308	0.149	0.0031	0.0204
0.85	1.45	0.2	0.7	326	0.150	0.0033	0.0213
0.85	1.5	0.2	0.7	322	0.155	0.0033	0.0211
0.85	1.55	0.2	0.7	325	0.181	0.0039	0.0209
0.85	1.65	0.2	0.7	323	0.184	0.0039	0.0207

Table 5.1: Consistent sets of parameter values for N=4, using the full model.

ρ_S (day^{-1})	ρ_P (day^{-1})	r	r_2	AMS	f_m	MFE	F_S
0.7	1.05	0.25	0.65	314	0.189	0.0035	0.0186
0.75	0.85	0.25	0.65	300	0.149	0.0031	0.0219
0.75	0.9	0.25	0.65	318	0.208	0.0042	0.0211
0.75	0.95	0.25	0.65	325	0.188	0.0037	0.0205

Table 5.2: Consistent sets of parameter values for N=5, using the full model.

as ρ_S increases, the difference $r = r_1 - r_3$ decreases. In all cases $\rho_S < \rho_P$, which is consistent with the often-stated supposition that stem cells divide more slowly than do progenitor cells. We caution that the values of the parameters obtained, besides being inherently dependent on the structure of our model, also depend on the assumed cell size and the neglecting of cell death.

In Figures 5.3 - 5.6, we plot the distribution of stem cells across all "colonies" for three representative sets of parameter values, where each colony corresponds to a realization of the mammosphere assay, for 10 000 realizations. Mammospheres are colonies of size greater than or equal to 220, where by "colony size" we mean the sum of all sub-populations of cells at time 8.5 days for that particular realization of the mammosphere assay. Each point in Figures 5.3 - 5.6 (a) corresponds to a realization of the assay with the colony size given by the position along the horizontal axis and the number of stem cells in that particular colony given by the position along the vertical axis. Likewise, in Figures 5.3 - 5.6 (b), each point corresponds to a realization of the mammosphere assay with colony size given as

ρ_S (day^{-1})	ρ_P (day^{-1})	r	r_2	AMS	f_m	MFE	F_S
0.45	1.05	0.4	0.5	306	0.146	0.0032	0.0220
0.45	1.1	0.4	0.4	317	0.149	0.0032	0.0213
0.5	1.0	0.35	0.45	323	0.169	0.0031	0.0181
0.55	0.85	0.35	0.55	328	0.137	0.0030	0.0219
0.55	0.85	0.35	0.35	316	0.162	0.0036	0.0219
0.55	0.9	0.35	0.55	319	0.166	0.0035	0.0209
0.55	0.9	0.35	0.45	314	0.162	0.0034	0.0209
0.55	0.9	0.35	0.35	307	0.151	0.0032	0.0209
0.55	0.95	0.35	0.55	320	0.206	0.0041	0.0201
0.55	0.95	0.35	0.45	328	0.222	0.0045	0.0201
0.55	1.0	0.35	0.55	320	0.214	0.0041	0.0193
0.55	1.0	0.35	0.45	328	0.210	0.0041	0.0193
0.6	0.95	0.35	0.45	330	0.232	0.0050	0.0215
0.65	0.75	0.3	0.6	321	0.153	0.0030	0.0198
0.7	0.75	0.3	0.6	326	0.207	0.0044	0.0212

Table 5.3: Consistent sets of parameter values for N=6, using the full model.

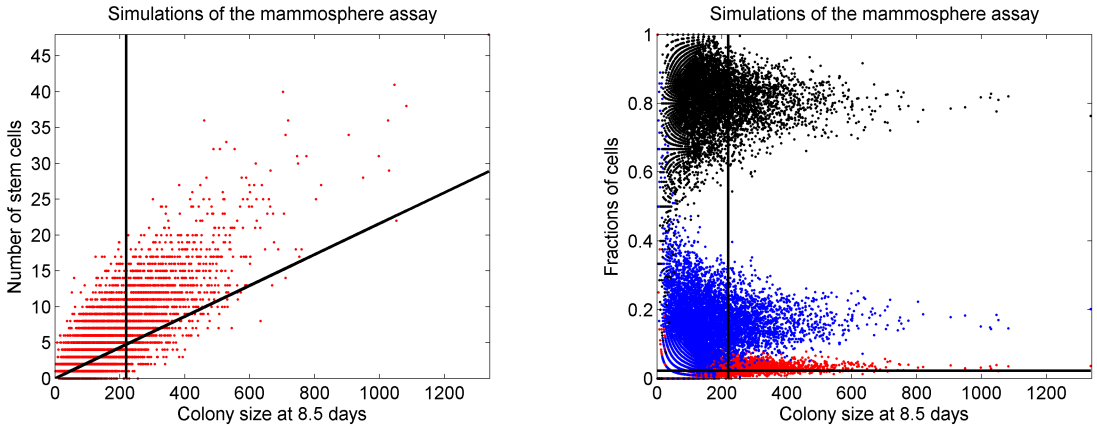
before and the fraction of stem cells, progenitor cells (of any generation) and mature cells in that particular colony shown in red, blue and black, respectively.

Figures 5.3 - 5.6 (a) show that, as a general trend, the number of stem cells per colony increases as the colony size increases, although this trend is weakest in the case of $N = 7$. We also observe, however, that the number of stem cells can vary widely even among colonies of the same size (as expected, in accordance with Equation 3.62 for the initial condition $n_S^0 = 1$). Comparing to Figures 5.3 - 5.6 (b), we see that the variability in the fraction of stem cells among colonies of the same size is much smaller than the variability in absolute numbers of stem cells seen in Figures 5.3 - 5.6 (a). For colonies at or above the mammosphere threshold of 220, the fraction of stem cells appears to be relatively constant – that is, it is independent of colony size.

It is often assumed that sphere-forming assays serve to enrich for stem cells. Observing Figures 5.3 - 5.6, we see that the fraction of stem cells among mammospheres indeed most often lies above the steady-state stem cell fraction of a

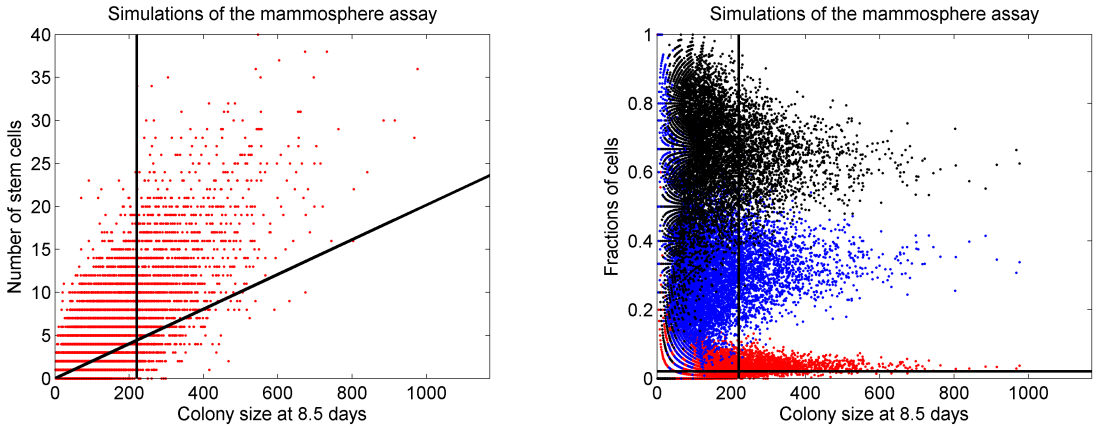
ρ_S (day^{-1})	ρ_P (day^{-1})	r	r_2	AMS	f_m	MFE	F_S
0.25	1.4	0.6	0.2	288	0.147	0.0031	0.0209
0.25	1.55	0.6	0.2	292	0.155	0.0031	0.0198
0.25	1.6	0.6	0.3	302	0.166	0.0032	0.0195
0.25	1.6	0.6	0.2	301	0.177	0.0035	0.0195
0.3	1.2	0.55	0.25	298	0.153	0.0031	0.0199
0.3	1.25	0.55	0.35	305	0.155	0.0030	0.0194
0.3	1.25	0.55	0.25	299	0.185	0.0031	0.0194
0.3	1.25	0.55	0.15	298	0.155	0.0030	0.0194
0.3	1.3	0.55	0.25	304	0.175	0.0033	0.0189
0.3	1.3	0.55	0.15	292	0.178	0.0034	0.0189
0.3	1.35	0.55	0.35	306	0.180	0.0033	0.0185
0.3	1.35	0.55	0.25	310	0.203	0.0038	0.0185
0.3	1.35	0.55	0.15	304	0.210	0.0039	0.0185
0.3	1.4	0.55	0.25	309	0.200	0.0036	0.0181
0.3	1.4	0.55	0.15	305	0.232	0.0042	0.0181
0.3	1.55	0.6	0.2	328	0.223	0.0049	0.0219
0.3	1.6	0.6	0.2	327	0.227	0.0049	0.0215
0.35	1.1	0.5	0.4	315	0.183	0.0033	0.0182
0.35	1.1	0.5	0.1	305	0.167	0.0030	0.0182
0.35	1.25	0.55	0.35	321	0.222	0.0048	0.0216
0.35	1.25	0.55	0.15	320	0.231	0.0050	0.0216
0.35	1.3	0.55	0.25	322	0.238	0.0050	0.0210
0.4	1.05	0.5	0.3	324	0.213	0.0045	0.0211
0.4	1.05	0.5	0.2	317	0.230	0.0048	0.0211
0.4	1.05	0.5	0.1	327	0.217	0.0046	0.0211
0.4	1.1	0.5	0.2	316	0.221	0.0045	0.0203
0.45	0.85	0.45	0.45	311	0.148	0.0032	0.0214
0.45	0.85	0.45	0.35	322	0.154	0.0033	0.0214
0.45	0.85	0.45	0.15	317	0.160	0.0034	0.0214
0.45	0.9	0.45	0.45	323	0.191	0.0039	0.0202
0.45	0.9	0.45	0.35	320	0.190	0.0038	0.0202
0.45	0.9	0.45	0.25	323	0.179	0.0036	0.0202
0.45	0.9	0.45	0.15	315	0.184	0.0037	0.0202
0.45	0.9	0.45	0.05	329	0.162	0.0033	0.0202
0.45	0.95	0.45	0.45	328	0.193	0.0037	0.0192
0.45	0.95	0.45	0.35	327	0.224	0.0043	0.0192
0.45	1.0	0.45	0.15	327	0.235	0.0043	0.0183
0.5	0.8	0.4	0.5	308	0.167	0.0031	0.0185
0.6	0.7	0.35	0.45	320	0.168	0.0031	0.0184
0.6	0.7	0.35	0.35	311	0.177	0.0033	0.0184
0.6	0.7	0.35	0.15	316	0.166	0.0031	0.0184
0.65	0.7	0.35	0.35	315	0.230	0.0047	0.0203

Table 5.4: Consistent sets of parameter values for N=7, using the full model.



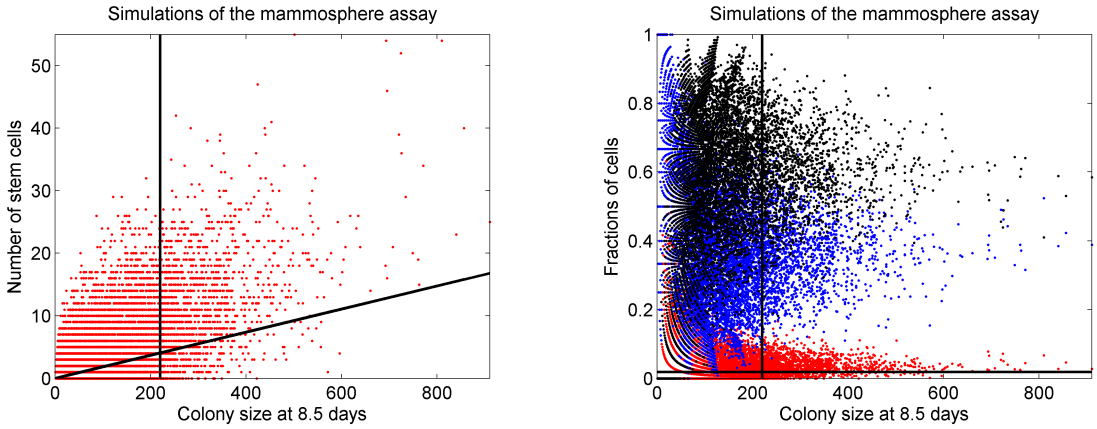
(a) (b)

Figure 5.3: Number of stem cells per colony (a) and fractions of stem (red), progenitor (blue), and mature (black) cells per colony (b), for 10 000 realizations of the mammosphere assay via Gillespie's algorithm, with $N = 4$, $\rho_S = 0.85$ day^{-1}, $\rho_P = 1.5$ day^{-1}, $r_1 = 0.25$, $r_2 = 0.7$, $r_3 = 0.05$. Mammosphere threshold (220) is shown as a vertical line; non-vertical lines represent steady-state numbers of stem cells corresponding to total colony size (a) and steady-state fraction of stem cells (b).



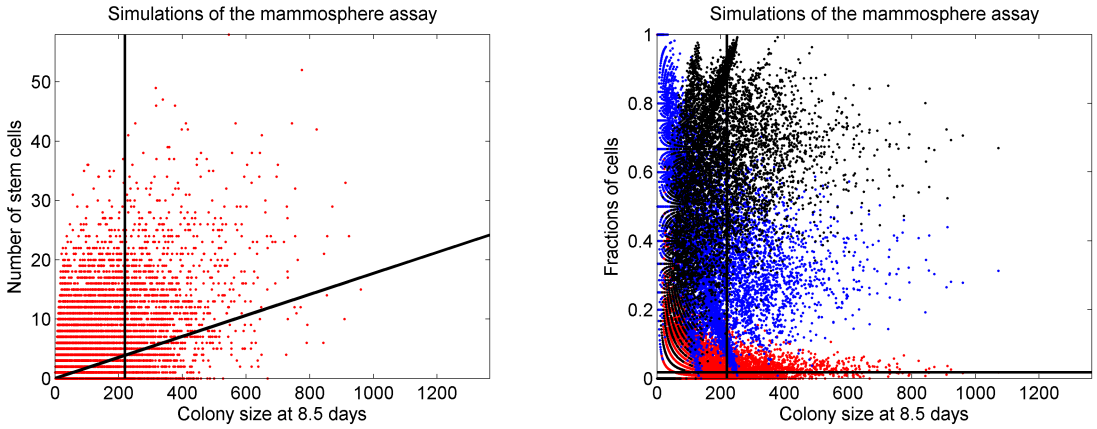
(a) (b)

Figure 5.4: Number of stem cells per colony (a) and fractions of stem (red), progenitor (blue), and mature (black) cells per colony (b), for 10 000 realizations of the mammosphere assay via Gillespie's algorithm, with $N = 5$, $\rho_S = 0.75$ day^{-1}, $\rho_P = 0.9$ day^{-1}, $r_1 = 0.3$, $r_2 = 0.65$, $r_3 = 0.05$. Mammosphere threshold (220) is shown as a vertical line; non-vertical lines represent steady-state numbers of stem cells corresponding to total colony size (a) and steady-state fraction of stem cells (b).



(a) (b)

Figure 5.5: Number of stem cells per colony (a) and fractions of stem (red), progenitor (blue), and mature (black) cells per colony (b), for 10 000 realizations of the mammosphere assay via Gillespie's algorithm, with $N = 6$, $\rho_S = 0.55$ day^{-1}, $\rho_P = 0.9$ day^{-1}, $r_1 = 0.45$, $r_2 = 0.45$, $r_3 = 0.1$. Mammosphere threshold (220) is shown as a vertical line; non-vertical lines represent steady-state numbers of stem cells corresponding to total colony size (a) and steady-state fraction of stem cells (b).



(a) (b)

Figure 5.6: Number of stem cells per colony (a) and fractions of stem (red), progenitor (blue), and mature (black) cells per colony (b), for 10 000 realizations of the mammosphere assay via Gillespie's algorithm, with $N = 7$, $\rho_S = 0.4$ day^{-1}, $\rho_P = 1.05$ day^{-1}, $r_1 = 0.65$, $r_2 = 0.2$, $r_3 = 0.15$. Mammosphere threshold (220) is shown as a vertical line; non-vertical lines represent steady-state numbers of stem cells corresponding to total colony size (a) and steady-state fraction of stem cells (b).

culture (indicated in Figures 5.3 - 5.6 (b) by a solid horizontal line); however, *some* mammospheres (particularly those with total cell numbers only just above the mammosphere definition threshold) contain relatively *fewer* stem cells than the steady-state population and thus represent a dilution in terms of stem cell content.

A feature of the progression from Figure 5.3 (b) to Figure 5.6 (b) is that, as N increases, the fractions of progenitor and mature cells begin to approach each other. We suggest that, in experimental practice, it would be helpful to dissociate mammospheres following the mammosphere assay and sort according to markers of fully-differentiated (i.e. mature) cells, stem cells, and "others" (i.e. progenitor cells). Comparison of distributions obtained via stochastic simulation may then help to elucidate numerical values of the growth parameters of these cells – for example, in distinguishing between the cases of the parameter values used in Figures 5.3 - 5.6. Indeed, very recent analysis of mammospheres produced by normal human mammary cells (that is, non-immortalized human mammary epithelial cells) indicated that about 70% of cells within mammospheres stained positive for CK18, a membrane-associated protein indicative of (differentiated) luminal epithelial cells, while 7% stained positive for CK14, a marker of (differentiated) myoepithelial cells, and 30% stained positive for CK19, a marker of early progenitor and myoepithelial cells [Dey et al., 2009]. While these statistics are likely confounded by overlap among marker proteins, a composition of about 70-80% mature/late-generation progenitor cells with the remaining 20-30% being stem cells or earlier-generation progenitor cells seems roughly consistent with Figures 5.3 - 5.6 (b).

More detailed experimental results may also suggest that some sort of regulatory feedback is at work. For example, one could imagine that a stem cell might be more likely to undergo symmetric self-renewal if few other stem cells are present, and less likely to undergo symmetric self-renewal if many stem cells are present. Such

feedback is implied by the fact that organs achieve homeostasis, although it is not clear that effects present *in vivo* are necessarily present *in vitro*. We suggest that in the future such cell number-dependent regulation of stem cell division could be included in the model, if required. For now, however, we take the parameter values used in generating Figures 5.3 - 5.6 and proceed to the next section.

5.2 CSCs and the Epithelial-Mesenchymal Transition

Epithelial-mesenchymal transition (EMT) is a cellular differentiation process in which epithelial cells lose representative phenotypic characteristics such as expression of E-cadherin (a transmembrane protein important in cell adhesion) and instead adopt features of mesenchymal cells, including increased cellular mobility. This process plays essential roles in normal development – for example, it enables formation of the secondary palate and heart valves in embryonal development [Yang and Weinberg, 2008]. However, in recent years the study of EMT has transcended developmental biology as cancer biologists have noted the similarities between developmental EMT and the acquisition of invasive properties by metastatic cancer cells [Yang and Weinberg, 2008]. The research groups of Weinberg [Mani et al., 2008] and Puisieux [Morel et al., 2008] have implicated EMT in cancer stem cell biology, suggesting that EMT generates cells with properties of stem cells.

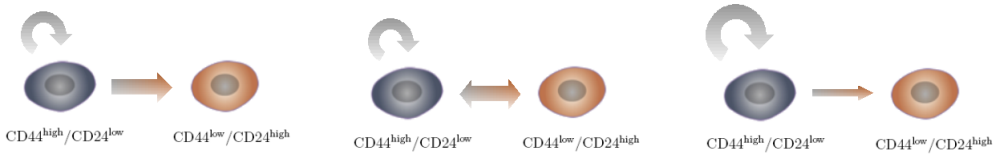
In brief, the work of Mani *et al.* [Mani et al., 2008] involved inducing EMT in HMLECs through several different means, and observing the phenotypes and functionalities in comparison to control cells. One method of inducing EMT in HMLECs was the forced, constitutive expression of either the protein Snail, which

directly inhibits transcription of the E-cadherin gene, or the protein Twist, which indirectly facilitates similar E-cadherin inhibition [Yang and Weinberg, 2008]; another method of EMT induction was exposure to TGFβ1, a protein implicated in signalling the initiation of EMT [Yang and Weinberg, 2008]. A third method was the coupling of either Snail or Twist to a modified estrogen receptor. By exposure to the estrogen receptor ligand tamoxifen, Snail or Twist expression could be triggered – this expression was reversible, in that Snail or Twist expression could be stopped via withdrawal of tamoxifen. We examine data derived from this third method of EMT induction in our analysis, as only these data are presented in a way that illustrates their time-dependence (Figure 5.2).

The effects of EMT induction on HMLECs were judged according to observed changes in phenotype and function. Phenotypic changes included loss or gain of expression of various genes associated with mesenchymal cells or epithelial cells, but the characteristic of most note was increased expression of the $CD44^{high}/CD24^{low}$ cell-surface protein expression profile, the phenotype attributed to breast (and breast cancer) stem cells. The ability of HMLECs exposed to EMT-inducing stimuli to form mammospheres was also quantified, as this function is generally considered to be demonstrative of stem cells. $CD44^{high}/CD24^{low}$ expression and mammosphere-forming ability of HMLECs induced to undergo EMT for varying lengths of time as measured by Mani *et al.* [Mani et al., 2008] are recorded in Figure 5.2.

It is noted from Figure 5.2 that both the fraction of HMLECs expressing the stem cell phenotype and the number of mammospheres formed per 500 cells increase with increasing exposure to tamoxifen (and hence increasing EMT induction). Radisky and LaBarge [Radisky and LaBarge, 2008] propose two possible explanations of how EMT induction leads to these observed increases in stem cell appearance and function. The first, which we call "Scenario One", is that exist-

ing non-stem HMLECs (that is, $CD44^{low}/CD24^{high}$ cells – we herein assume, as in the previous section, that the $CD44^{high}/CD24^{low}$ phenotype is a perfect marker of mammary stem cells) were converted directly to stem cells ($CD44^{high}/CD24^{low}$ cells) via EMT. The second proposed possibility ("Scenario Two") is that the EMT inhibited differentiation/enhanced proliferation of existing stem cells, while possibly inhibiting proliferation of non-stem cells. These scenarios are illustrated in Figure 5.7.



(a) (b) (c)

Figure 5.7: Mechanisms of production of stem and non-stem cells: (a) control (normal) case, (b) Scenario One (production of stem cells from non-stem cells, and (c) Scenario Two (increased stem cell self-renewal/decreased differentiation of stem cells).

5.2.1 Scenario One

Under Scenario One, we define the rate constants α and β (units of day^{-1}) to be the rates at which non-stem cells are transferred from the progenitor cell and mature cell compartments, respectively, to the stem cell compartment, and modify

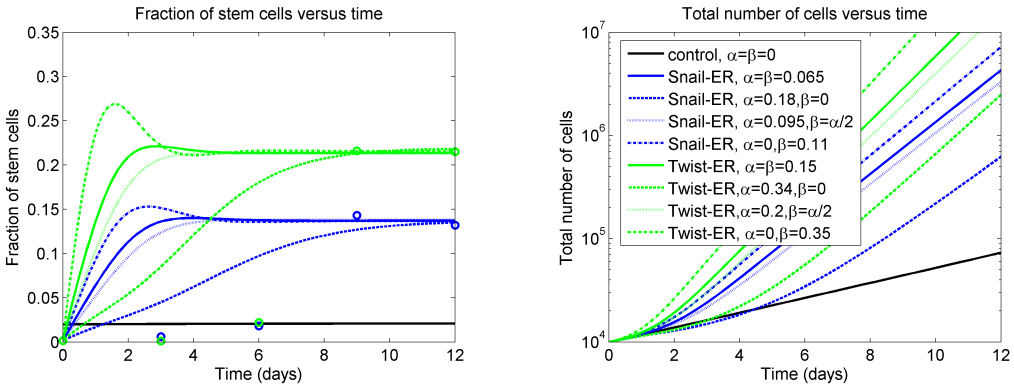
Equations 3.67 accordingly:

$$\begin{aligned}
\frac{d}{dt}S(t) &= \rho_S r S(t) + \alpha \sum_{i=1}^{N} P_i(t) + \beta M(t) \\
\frac{d}{dt}P_1(t) &= \rho_S (1-r) S(t) - \rho_{P_1} P_1(t) - \alpha P_1(t) \\
\frac{d}{dt}P_i(t) &= 2\rho_{P_{i-1}} P_{i-1}(t) - \rho_{P_i} P_i(t) - \alpha P_i(t), \text{ for } i = 2, \ldots, N \\
\frac{d}{dt}M(t) &= 2\rho_{P_N} P_N(t) - \Gamma M(t) - \beta M(t),
\end{aligned} \tag{5.1}$$

where we remind the reader that $r = r_1 - r_3$. We solve Equations 5.1 numerically for each of the sets of parameter values used in Figures 5.3 - 5.6, using initial conditions that correspond to 10^4 cells distributed such that the fraction of stem cells at time $t = 0$ is given by the data of Mani *et al.* [Mani et al., 2008] in Figure 5.2 and the fractions of mature cells and progenitor cells are approximately in steady-state. Results are shown in Figures 5.8 - 5.11, wherein the values of α and β that best fit the experimental data presented in Figure 5.2 have been determined by inspection.

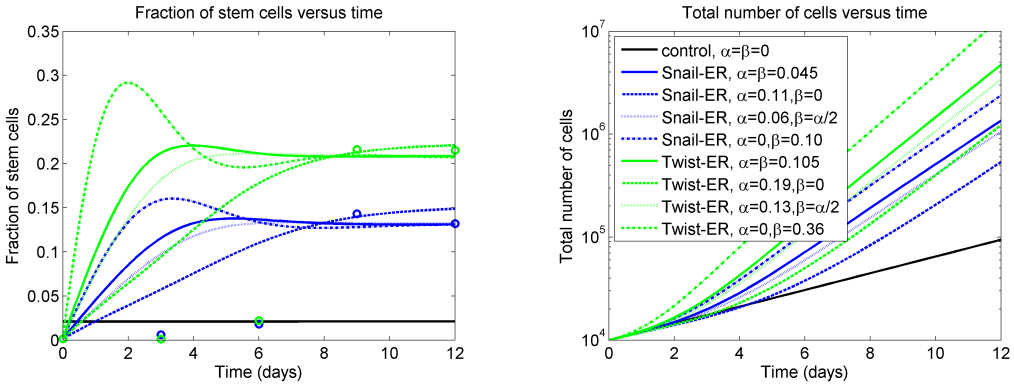
Several cases are considered in Figures 5.8 - 5.11, including those in which only one of the de-differentiation rates α and β is non-zero. Of these cases, the only one that can be said to capture the trend of the experimental data of Mani *et al.* [Mani et al., 2008] is that in which $\beta = 0$ and α is non-zero. This is interpreted biologically as the case in which only progenitor cells, and not non-proliferating mature cells, are susceptible to de-differentiation via the EMT. We point out that this case may be unrealistic, as we expect the mature cells, the bulk of which are the canonical epithelial cells, to respond to the EMT.

In the other cases, the theoretically-predicted rises in the fractions of $\text{CD44}^{\text{high}}/\text{CD24}^{\text{low}}$ Snail-ER and Twist-ER cells occur much more abruptly than the corresponding experimentall-observed rises. Experimentally, a delay of nearly six days passes before any appreciable increase in the $\text{CD44}^{\text{high}}/\text{CD24}^{\text{low}}$ fraction is observed. We



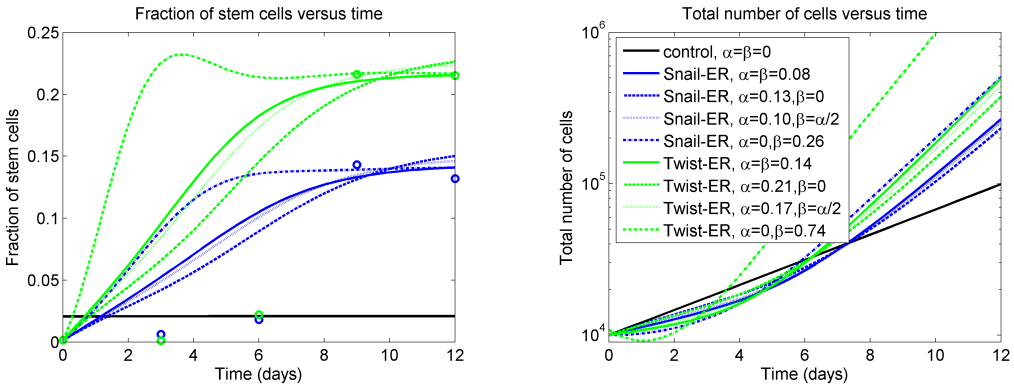
(a) (b)

Figure 5.8: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Parameter values not shown in the legend are as follows: $N = 4$, $\rho_S = 0.85$ days^{-1}, $\rho_P = 1.5$ days^{-1}, $r = r_1 - r_3 = 0.2$, $\Gamma = 0$.



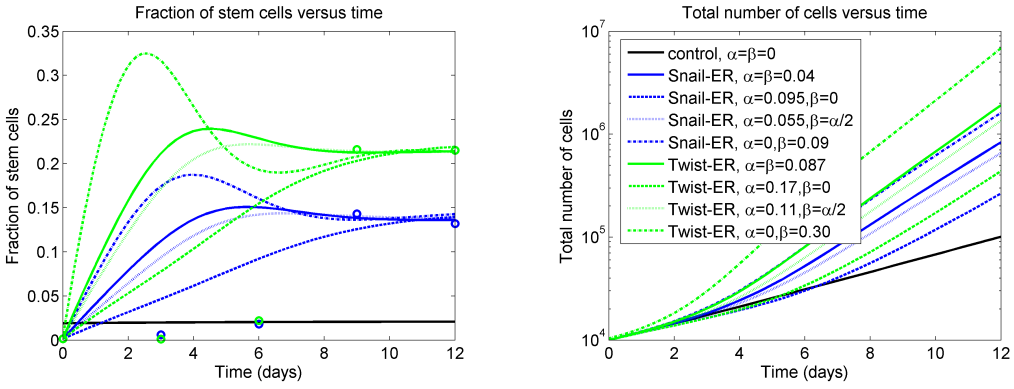
(a) (b)

Figure 5.9: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Parameter values not shown in the legend are as follows: $N = 5$, $\rho_S = 0.75$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.25$, $\Gamma = 0$.



(a) (b)

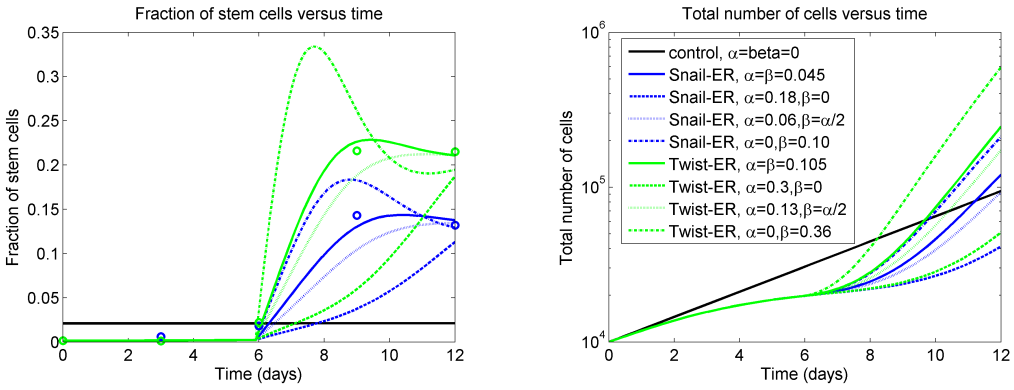
Figure 5.10: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Parameter values not shown in the legend are as follows: $N = 6$, $\rho_S = 0.55$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.35$, $\Gamma = 0$.



(a) (b)

Figure 5.11: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Parameter values not shown in the legend are as follows: $N = 7$, $\rho_S = 0.4$ days^{-1}, $\rho_P = 1.05$ days^{-1}, $r = r_1 - r_3 = 0.5$, $\Gamma = 0$.

propose instead that some initial delay period occurs following the addition of the EMT-inducing tamoxifen to the culture during which cells do not fully transition to the proliferating stem cell state. While the precise nature of such a delay is unknown at the present time, one possibility is that cells respond quickly (i.e. without delay) to the EMT-inducing stimulus and adopt a mesenchymal phenotype, but experience an extended delay before they gain the ability to self-renew and enter the functional stem cell compartment. Fitting this model to the experimental data, we take $\alpha = \beta = 0$ until $t = 5.9$ days, after which these rate constants assume their appropriate values, given in the legend of Figure 5.12. Note that the "all-or-nothing" form of the de-differentiation rates α and β assumed here is just one possible form that might capture the proposed delay – in the future, one might also consider time-dependent forms $\alpha(t)$ and $\beta(t)$ that increase smoothly with time.



(a) (b)

Figure 5.12: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. The de-differentiation rates α and β are set to zero until time $t = 5.9$ days in each case. Parameter values not shown in the legend are as follows: $N = 5$, $\rho_S = 0.75$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.25$, $\Gamma = 0$.

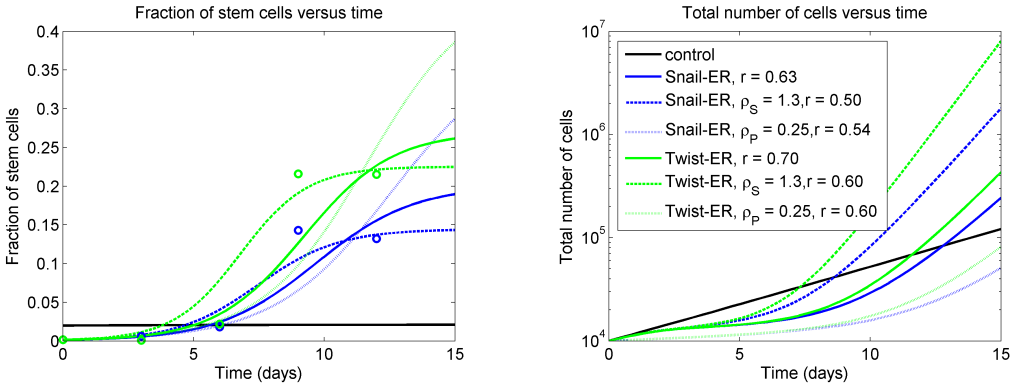
An interesting feature of the fractions of stem cells for Snail-ER and Twist-ER cells observed in Figure 5.12 (a) for the cases when $\beta > 0$ is that the fractions of stem cells increase quickly (after 5.9 days) – and increase *above* their respective plateau values before steadying. The experimentally-measured Snail-ER and Twist-ER data of Mani *et al.* [Mani et al., 2008] (Figure 5.2) both show similar behaviour – although we caution that the fractions of $CD44^{high}/CD24^{low}$ cells reported in Figure 5.2 are not given along with error estimates, and error in fluorescence-activated cell sorting of cells may be significant.

We see in Figure 5.12 (a) that the case $\alpha = \beta > 0$ offers a very good fit to the data when the aforementioned time delay is assumed. This corresponds biologically to the case in which progenitor and mature cells are equally susceptible to de-differentiation to the stem cell state via EMT.

5.2.2 Scenario Two

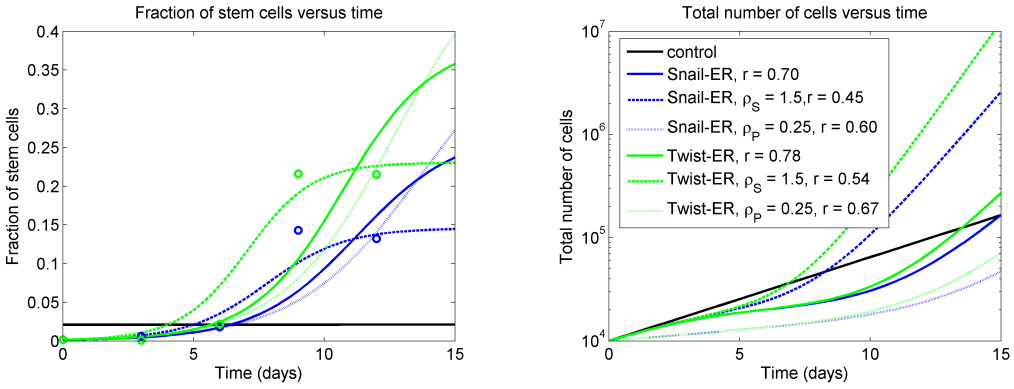
In Scenario Two, we do not assume the presence of any additional pathways: we use Equations 5.1 with $\alpha = \beta = 0$. Instead, we assume that the effect of EMT is to alter the values of the parameters r, ρ_S, and ρ_P associated with the division of stem and progenitor cells. Again choosing r, ρ_S and ρ_P by inspection via comparison with the data of Mani *et al.* [Mani et al., 2008], we plot the fraction of stem cells as functions of time under Scenario Two in Figures 5.13 to 5.16 for Snail-ER and Twist-ER cells.

In Figures 5.13 - 5.16, we have considered cases where (i) only the net proportion of self-renewing stem cell divisions r is affected by the EMT, (ii) both the stem cell division rate ρ_S and the net proportion of self-renewing stem cell divisions r are upregulated, and (iii) r is increased in concurrence with a decrease in the rate of



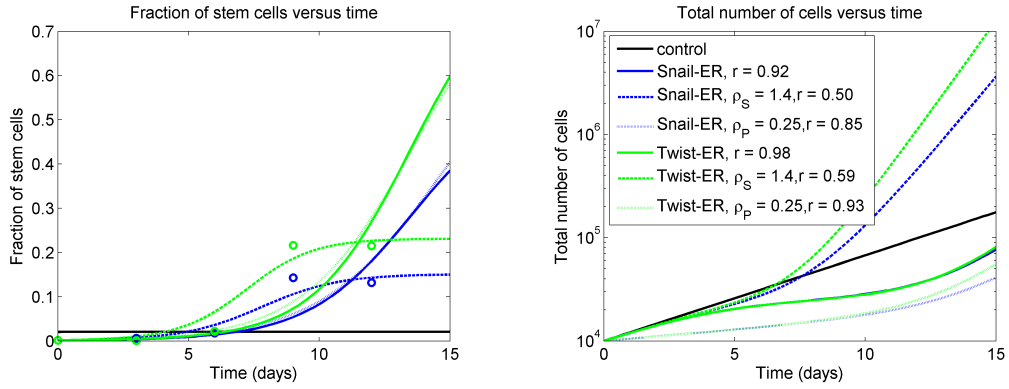
(a) (b)

Figure 5.13: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Parameter values are as follows: $N = 4$, $\rho_S = 0.85$ days^{-1}, $\rho_P = 1.5$ days^{-1}, $r = r_1 - r_3 = 0.2$, $\Gamma = 0$. Only values differing from control values are shown in the legend.



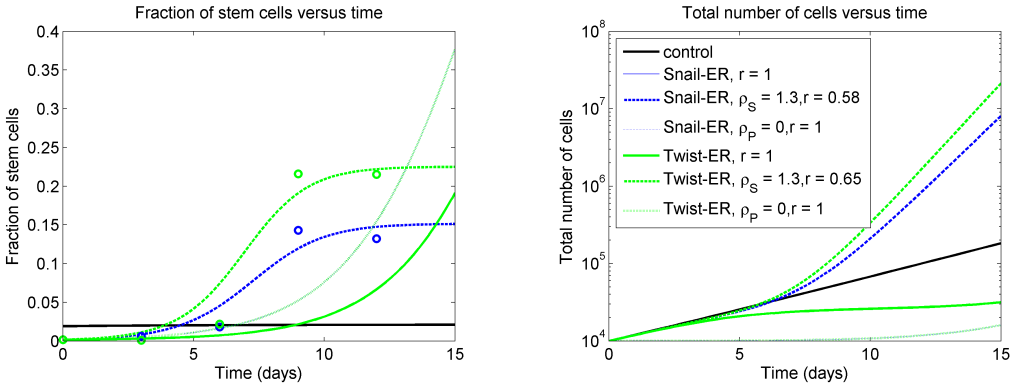
(a) (b)

Figure 5.14: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Control parameter values are as follows: $N = 5$, $\rho_S = 0.75$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.25$, $\Gamma = 0$. Only values differing from control values are shown in the legend.



(a) (b)

Figure 5.15: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Control parameter values are as follows: $N = 6$, $\rho_S = 0.55$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.35$, $\Gamma = 0$. Only values differing from control values are shown in the legend.



(a) (b)

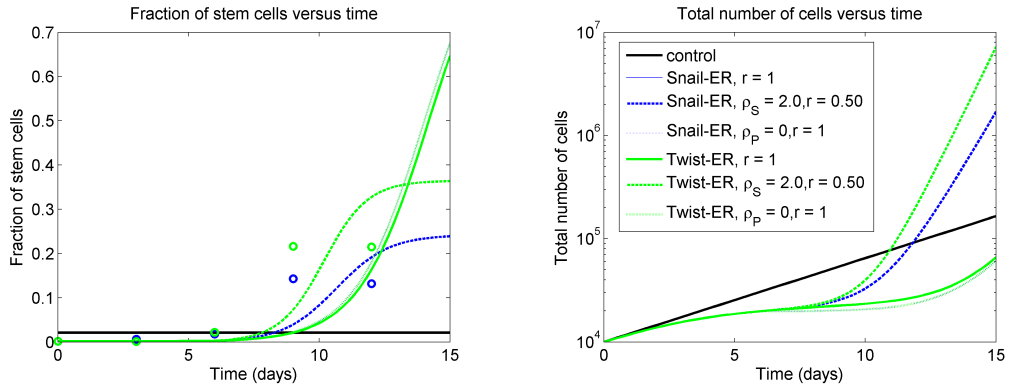
Figure 5.16: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Control parameter values are as follows: $N = 7$, $\rho_S = 0.4$ days^{-1}, $\rho_P = 1.05$ days^{-1}, $r = r_1 - r_3 = 0.5$, $\Gamma = 0$. Only values differing from control values are shown in the legend.

progenitor cell division. We see that only case (ii) provides a satisfactory fit to the data. When only r is assumed to be altered by the EMT, the fraction of stem cells tends to a significantly higher value than one might deduce from the experimental data; this is particularly evident in Figure 5.16, corresponding to $N = 7$. Similarly, case (iii) leads to a much higher long-term stem cell fraction than indicated by the data.

For the purpose of comparison with Figure 5.12, we have considered in Figure 5.17 Scenario Two with the 5.9-day time delay used above. Clearly, the assumption of a (significant) time delay before the onset of the effects of EMT in conjunction with Scenario Two provides a poor fit. We thus conclude that, if Scenario Two is the mechanism by which the EMT enhances the fraction of mammary epithelial stem cells, it does so by concurrently upregulating the stem cell division rate ρ_S and the net proportion of self-renewing stem cell divisions r without a significant time delay.

5.2.3 Discussion and Comparison of Scenarios

We have presented two distinct scenarios by which induction of the EMT may act to increase the proportion of stem cells in a population of mammary epithelial stem cells. Scenario One predicts that both progenitor cells and mature cells may be converted to stem cells, and has the attractive feature of capturing the peak in the fraction of stem cells observed at day 9 of EMT induction by Mani *et al.* [Mani et al., 2008]. Scenario Two, on the other hand, predicts that the EMT acts by increasing the division rate ρ_S of already-existing stem cells while also altering the net proportion of symmetric stem cell self-renewals r – we suggest that the value of r is important in determining the value toward which the fraction of stem cells



(a) (b)

Figure 5.17: Fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Control parameter values are as follows: $N = 5$, $\rho_S = 0.75$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.25$, $\Gamma = 0$. Only values differing from control values are shown in the legend; in each case, control values are maintained until time $t = 5.9$ days, at which point the values indicated in the legend are assumed.

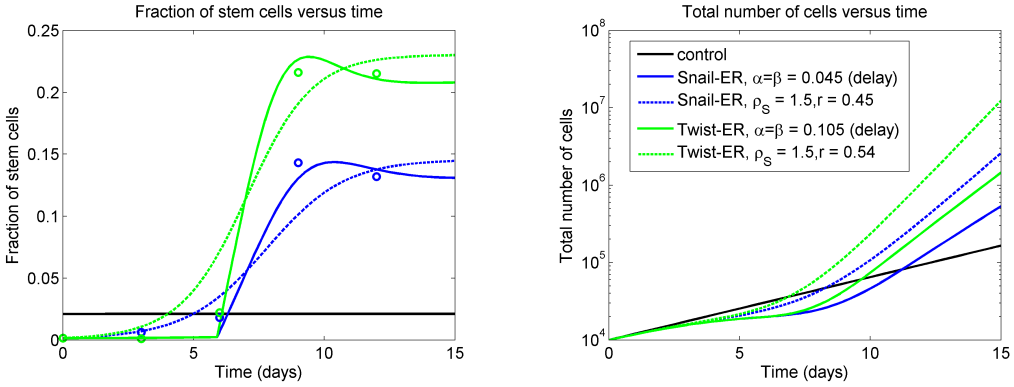
will tend, while the growth rate ρ_S is instrumental in determining how quickly this plateau phase is reached.

We point out that Mani *et al.* [Mani et al., 2008] have performed experiments in an attempt to rule out Scenario Two, in which cells were sorted by fluorescence-activated cell sorting for the $CD44^{low}/CD24^{high}$ phenotype and subjected to EMT-inducing stimuli [Mani et al., 2008]. It was observed that, following a period of EMT induction, the sorted $CD44^{low}/CD24^{high}$ sub-population *did* gain expression of the $CD44^{high}/CD24^{low}$ phenotype and the ability to form mammospheres – although small numbers of cells from the control $CD44^{low}/CD24^{high}$ sub-population (that is, cells not induced to undergo EMT) also showed the stem cell phenotype and the ability to form mammospheres after some time in culture. Thus, it is still possible that a small number of $CD44^{high}/CD24^{low}$ cells, despite fluorescence-activated cell sorting, were included in the sorted $CD44^{low}/CD24^{high}$ sub-population and that these cells were enriched via Scenario Two when induced to undergo EMT.

In attempting to distinguish between Scenarios One and Two (and answer the question, "by what mechanism does the EMT enhance the proportion of stem cells in a population of human mammary epithelial cells?"), we can suggest some considerations. With regard to Scenario Two, we point out that in order to achieve a qualitatively good fit to the data we have assumed that the stem cell proliferation rate is upregulated significantly through the EMT, and it is not intuitively clear that such a profound increase in proliferation is a biological possibility.

Figure 5.18 compares the dynamics of the fractions of stem cells and the total number of cells predicted by Scenario One ($\alpha = \beta$ with time delay) and Scenario Two (ρ_S, r upregulated) for $N = 5$. We suggest that increased temporal resolution in measuring fractions of stem cells may help to distinguish between the two Scenarios. Further, we observe in Figure 5.18 (b) that the total number of cells increases

more quickly under Scenario Two than under Scenario One. Mani *et al.* [Mani et al., 2008] do not report absolute numbers of cells; we suggest, however, that measurement of cell numbers during EMT induction may serve as one additional criterion by which the validities of Scenarios One and Two may be judged.



(a) (b)

Figure 5.18: Comparison of EMT Scenarios One and Two: fraction of stem cells (a) and total number of cells (b) versus time for control cells (black), cells induced to express Snail-ER during the twelve-day time period (blue), and cells induced to express Twist-ER during the twelve-day time period (green), with experimental data points shown as circles of the corresponding colour. Control parameter values are as follows: $N = 5$, $\rho_S = 0.75$ days^{-1}, $\rho_P = 0.9$ days^{-1}, $r = r_1 - r_3 = 0.25$, $\Gamma = 0$. Only values differing from control values are shown in the legend; in each case, control values are maintained until time $t = 5.9$ days, at which point the values indicated in the legend are assumed.

In Figure 5.2, it is seen that even as the fraction of CD44$^{\text{high}}$/CD24$^{\text{low}}$ cells reaches a plateau between days 9 and 12, the frequency of mammosphere inititation continues to increase for the Snail-ER and Twist-ER cells. The plateauing of the fraction of stem cells from day 9 to day 12 suggests that cells taken from the EMT-inducing culture and plated in the mammosphere assay at any time from day 9 to day 12 should have roughly the same ability to form mammospheres, *if* it is

further assumed that the values the growth parameters take upon introduction of the cells into the mammosphere culture do not depend on the length of time previously exposed to EMT-inducing stimuli. It would be interesting to see how the fractions of $CD44^{high}/CD24^{low}$ cells and mammosphere formation frequencies of Snail-ER and Twist-ER cells behave as exposure to EMT-inducing stimuli is extended beyond 12 days. It may be the case that the length of exposure to EMT affects the behaviour (i.e. growth parameter values) of stem cells following removal from EMT-inducing culture.

A third scenario not considered herein is that the EMT induces (C)SCs or non-stem cells to transform into cells of a secondary sub-population of (C)SCs – in particular, one that is responsible for the metastatic behaviour of the tumour. Distinct sub-populations of CSCs have indeed been identified in pancreatic tumours, with one sub-population driving primary tumour growth and another driving metastasis [Hermann et al., 2007]. This possible link between CSCs, metastasis and the EMT deserves further attention.

Chapter 6

Conclusion and Future Work

Herein, we have presented stochastic and deterministic models describing the growth of populations of cancer cells based on the cancer stem cell hypothesis. The deterministic models have been applied to the theoretical treatment of brain tumours and to investigate the generation of mammary stem cells via the epithelial-mesenchymal transition, while a stochastic model has been applied to simulate the mammosphere-forming assay. We suggest that other applications will become prudent as experimentalists identify additional questions regarding the dynamics of cancer stem cells.

Mathematical modelling of the cancer stem cell hypothesis is likely to prove useful in two somewhat distinct ways. First, in attempting to establish a mathematical framework that encapsulates such a complex biological process that is only beginning to be understood, important insights and questions may surface that will help to direct future research. Following these initial stages, mathematical modelling will become increasingly useful in predicting strategies for battling the tumour and its resilient cancer stem cells. It seems clear that a deeper understanding, combined

with quantitative modelling of cancer stem cells, is central, not only for the design of effective experimental studies, but also for the development of effective therapies that will target cancer stem cells.

In modelling brain cancer and mammary stem cells we have assumed that the $CD133^{+}$ and $CD44^{high}/CD24^{low}$ phenotypes, respectively, are perfect markers of brain cancer and mammary stem cells in the sense that we have assumed that all cells of these phenotypes are (C)SCs and that no cells of other phenotypes are (C)SCs. It is true that this may not necessarily be the case. Indeed, biologists already seem to doubt the specificity of these markers, although they do so largely on the grounds that not all cells expressing these markers behave as stem cells in functional assays (i.e. form spheres in sphere-forming assays or form tumours in tumour-forming assays). However, we point out that the assumption of stochasticity entails that not all (C)SCs will perform as such in tests for (C)SC function, and thus we expect that not even all cells expressing a truly perfect (C)SC marker will appear to be (C)SCs based on functional assays. We suggest that, as understanding of the extra- and intra-cellular mechanisms regulating (C)SC self-renewal develops, it may be reasonable to attempt to bias the stochastic behaviour of (C)SCs toward, for example, symmetric self-renewal so that the difficulties associated with assigning a phenotypic marker to a cell that is defined based on functionality yet displays stochasticity with respect to this functionality can be overcome. Further refinement of the (C)SC phenotype will likely come from imaging and tracking of the divisions of single (putative) stem cells, as discussed earlier; such individual cell analysis will also help in assigning numerical values to the parameters associated with growth and division. Analysis of microarray data may also help to unravel the differences in gene expression that lend cancer stem, progenitor, and fully differentiated cells their respective phenotypes.

Throughout, we have alluded to possible future work. This includes adapting the model to accomodate the feature of clonal evolution – in particular, we suggest that with some "mutation probability" a dividing cell (either (C)SC or progenitor cell) may mutate such that it possesses a new phenotype defined by growth parameters that differ from those of the original population. It may also be useful to include inter-cellular regulation of stem cell division – for example, the probability of a stem cell to undergo a particular type of division may depend explicitly on the numbers of stem, progenitor or mature cells in the system. Examination of the regulation of CSCs via non-stem cells may help in understanding the "numbers problem" described by Richard Hill and mentioned in Section 2.1.7, wherein the number of putative CSCs required to initiate tumourigenesis in immunocompromised mice is *greater* when the putative CSCs are co-injected with non-stem cancer cells than when the putative CSCs are injected alone.

Another possible adaptation is the inclusion of cell quiescence, wherein with some probability stem cells enter a state of temporarily-arrested growth – this may be particularly important in the context of treatment (and treatment resistance), and the aforementioned probability of quiescence may in general depend on environmental cues.

Ultimately, it may be worthwhile to develop a spatial model, although such a model would likely be of such complexity that it would be strictly computational. Using a spatial model, one would be able to investigate the role of CSCs in tumour angiogenesis (the formation of new blood vessels), as well as the effects of the tumour microenvironment (in particular, the distribution of oxygen concentration) on dictating response to treatment. One of the undeniable benefits of viewing cancer through the lens of the cancer stem cell hypothesis is the emphasis on recognizing the heterogeneity of individual cells (whether this heterogenity be inherent, as

suggested by the CSC hypothesis, or due to external factors, as suggested by the "stochastic" model of tumourigenesis), which will surely contribute to our collective understanding of the mechanisms that drive cancer and lend it its resilience.

References

J. M. Adams and A. Strasser. Is tumor growth sustained by rare cancer stem cells or dominant clones? *Cancer Res.*, 68:4018–4021, Jun 2008.

M. Al-Hajj, M. S. Wicha, A. Benito-Hernandez, S. J. Morrison, and M. F. Clarke. Prospective identification of tumorigenic breast cancer cells. *Proc. Natl. Acad. Sci. U.S.A.*, 100:3983–3988, Apr 2003.

P. Armitage and R. Doll. The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br. J. Cancer*, 8:1–12, Mar 1954.

N. T. J. Bailey. *The Elements of Stochastic Processes with Applications to the Natural Sciences*. John Wiley and Sons, Inc., New York, 1964.

Monya Baker. Cancer stem cells, becoming common. http://www.nature.com/stemcells/2008/0812/081203/full/stemcells.2008.153.html, December 2008.

S. Bao, Q. Wu, R. E. McLendon, Y. Hao, Q. Shi, A. B. Hjelmeland, M. W. Dewhirst, D. D. Bigner, and J. N. Rich. Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature*, 444:756–760, Dec 2006.

A. J. Becker, E. A. McCulloch, and J. E. Till. Cytological demonstration of the

clonal nature of spleen colonies derived from transplanted mouse marrow cells. *Nature*, 197:452–454, Feb 1963.

B. M. Boman, M. S. Wicha, J. Z. Fields, and O. A. Runquist. Symmetric division of cancer stem cells–a key mechanism in tumor growth that should be targeted in future therapeutic approaches. *Clin. Pharmacol. Ther.*, 81:893–898, Jun 2007.

W. A. Bonner, H. R. Hulett, R. G. Sweet, and L. A. Herzenberg. Fluorescence activated cell sorting. *Rev Sci Instrum*, 43:404–409, Mar 1972.

W. E. Boyce and R. C. DiPrima. *Elementary Differential Equations and Boundary Value Problems*. John Wiley and Sons, Inc., New York, 3rd edition, 1992.

W. R. Bruce and van der Gaag H. A quantitative assay for number of murine lymphoma cells capable of proliferation in vivo. *Nature*, 199(488):79–&, 1963. ISSN 0028-0836.

John L. Casti. *Mathematical mountaintops : the five most famous problems of all time / John Casti*. Oxford University Press, New York ; Oxford :, 2001. ISBN 0195141717.

D. Chen, M. Zhao, and G. R. Mundy. Bone morphogenetic proteins. *Growth Factors*, 22:233–241, Dec 2004.

M. F. Clarke and M. Fuller. Stem cells and cancer: two faces of eve. *Cell*, 124: 1111–1115, Mar 2006.

E. Clayton, D. P. Doup, A. M. Klein, D. J. Winton, B. D. Simons, and P. H. Jones. A single type of progenitor cell maintains normal epidermis. *Nature*, 446:185–189, Mar 2007.

A. T. Collins, P. A. Berry, C. Hyde, M. J. Stower, and N. J. Maitland. Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res.*, 65:10946–10951, Dec 2005.

L. M. DeAngelis. Chemotherapy for brain tumors–a new beginning. *N. Engl. J. Med.*, 352:1036–1038, Mar 2005.

D. Dey, M. Saxena, A. N. Paranjape, V. Krishnan, R. Giraddi, M. V. Kumar, G. Mukherjee, and A. Rangarajan. Phenotypic and functional characterization of human mammary stem/progenitor cells in long term culture. *PLoS ONE*, 4: e5329, 2009.

P. B. Dirks. Brain tumour stem cells: the undercurrents of human brain cancer and their relationship to neural stem cells. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.*, 363:139–152, Jan 2008.

G. Dontu, W. M. Abdallah, J. M. Foley, K. W. Jackson, M. F. Clarke, M. J. Kawamura, and M. S. Wicha. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.*, 17:1253–1270, May 2003.

B. Dykstra, J. Ramunas, D. Kent, L. McCaffrey, E. Szumsky, L. Kelly, K. Farn, A. Blaylock, C. Eaves, and E. Jervis. High-resolution video monitoring of hematopoietic stem cells cultured in single-cell arrays identifies new features of self-renewal. *Proc. Natl. Acad. Sci. U.S.A.*, 103:8185–8190, May 2006.

B. Elenbaas, L. Spirio, F. Koerner, M. D. Fleming, D. B. Zimonjic, J. L. Donaher, N. C. Popescu, W. C. Hahn, and R. A. Weinberg. Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev.*, 15:50–65, Jan 2001.

A. J. Engler, S. Sen, H. L. Sweeney, and D. E. Discher. Matrix elasticity directs stem cell lineage specification. *Cell*, 126:677–689, Aug 2006.

A. Eramo, F. Lotti, G. Sette, E. Pilozzi, M. Biffoni, A. Di Virgilio, C. Conticello, L. Ruco, C. Peschle, and R. De Maria. Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.*, 15:504–514, Mar 2008.

P. S. Eriksson, E. Perfilieva, T. Bjrk-Eriksson, A. M. Alborn, C. Nordborg, D. A. Peterson, and F. H. Gage. Neurogenesis in the adult human hippocampus. *Nat. Med.*, 4:1313–1317, Nov 1998.

C. Fillmore and C. Kuperwasser. Human breast cancer stem cell markers CD44 and CD24: enriching for cells with functional properties in mice or in man? *Breast Cancer Res.*, 9:303, 2007.

C. M. Fillmore and C. Kuperwasser. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res.*, 10:R25, 2008.

R. Galli, E. Binda, U. Orfanelli, B. Cipelletti, A. Gritti, S. De Vitis, R. Fiocco, C. Foroni, F. Dimeco, and A. Vescovi. Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cancer Res.*, 64: 7011–7021, Oct 2004.

R. Ganguly and I. K. Puri. Mathematical model for the cancer stem cell hypothesis. *Cell Prolif.*, 39:3–14, Feb 2006.

R. Ganguly and I. K. Puri. Mathematical model for chemotherapeutic drug efficacy in arresting tumour growth based on the cancer stem cell hypothesis. *Cell Prolif.*, 40:338–354, Jun 2007.

P. R. Garabedian. *Partial Differential Equations*. John Wiley and Sons, Inc., New York, 1964.

D. T. Gillespie. Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, 81(25):2340–2361, 1977.

Daniel T. Gillespie. *Markov Processes: An Introduction for Physical Scientists*. Academic Press, October . ISBN 0122839552.

C. Ginestier, M. H. Hur, E. Charafe-Jauffret, F. Monville, J. Dutcher, M. Brown, J. Jacquemier, P. Viens, C. G. Kleer, S. Liu, A. Schott, D. Hayes, D. Birnbaum, M. S. Wicha, and G. Dontu. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell*, 1: 555–567, Nov 2007.

E. J. Hall. *Radiobiology for the Radiologist*. J. B. Lippincott Company, Philadelphia, 4th edition, 1994.

A. W. Hamburger and S. E. Salmon. Primary bioassay of human tumor stem-cells. *Science*, 197(4302):461–463, 1977. ISSN 0036-8075.

H. D. Hemmati, I. Nakano, J. A. Lazareff, M. Masterman-Smith, D. H. Geschwind, M. Bronner-Fraser, and H. I. Kornblum. Cancerous stem cells can arise from pediatric brain tumors. *Proc. Natl. Acad. Sci. U.S.A.*, 100:15178–15183, Dec 2003.

P. C. Hermann, S. L. Huber, T. Herrler, A. Aicher, J. W. Ellwart, M. Guba, C. J. Bruns, and C. Heeschen. Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell*, 1:313–323, Sep 2007.

R. P. Hill and R. Perris. "Destemming" cancer stem cells. *J. Natl. Cancer Inst.*, 99:1435–1440, Oct 2007.

T. N. Ignatova, V. G. Kukekov, E. D. Laywell, O. N. Suslov, F. D. Vrionis, and D. A. Steindler. Human cortical glial tumors contain neural stem-like cells expressing astroglial and neuronal markers in vitro. *Glia*, 39:193–206, Sep 2002.

L. O. Jacobson, E. K. Marks, M. J. Robson, E. Gaston, and R. E. Zirkle. The effect of spleen protection on mortality following x-irradiation. *J. Lab. Clin. Med.*, 34 (11):1538–1543, 1949. ISSN 0022-2143.

M. D. Johnston, C. M. Edwards, W. F. Bodmer, P. K. Maini, and S. J. Chapman. Mathematical modeling of cell population dynamics in the colonic crypt and in colorectal cancer. *Proc. Natl. Acad. Sci. U.S.A.*, 104:4008–4013, Mar 2007a.

M. D. Johnston, C. M. Edwards, W. F. Bodmer, P. K. Maini, and S. J. Chapman. Examples of mathematical modeling: tales from the crypt. *Cell Cycle*, 6:2106–2112, Sep 2007b.

S. Kamel-Reid and J. E. Dick. Engraftment of immune-deficient mice with human hematopoietic stem cells. *Science*, 242:1706–1709, Dec 1988.

J. A. Kennedy, F. Barab, A. G. Poeppl, J. C. Wang, and J. E. Dick. Comment on "Tumor growth need not be driven by rare cancer stem cells". *Science*, 318: 1722; author reply 1722, Dec 2007.

S. E. Kern and D. Shibata. The fuzzy math of solid tumor stem cells: a perspective. *Cancer Res.*, 67:8985–8988, Oct 2007.

M. Kohandel, M. Kardar, M. Milosevic, and S. Sivaloganathan. Dynamics of tumor

growth and combination of anti-angiogenic and cytotoxic therapies. *Phys. Med. Biol.*, 52:3665–3677, Jul 2007.

G. Kohler and C. Milstein. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature*, 256:495–497, Aug 1975.

N. L. Komarova and D. Wodarz. Effect of cellular quiescence on the success of targeted CML therapy. *PLoS ONE*, 2:e990, 2007.

A. V. Krivtsov, D. Twomey, Z. Feng, M. C. Stubbs, Y. Wang, J. Faber, J. E. Levine, J. Wang, W. C. Hahn, D. G. Gilliland, T. R. Golub, and S. A. Armstrong. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature*, 442:818–822, Aug 2006.

T. Lapidot, C. Sirard, J. Vormoor, B. Murdoch, T. Hoang, J. Caceres-Cortes, M. Minden, B. Paterson, M. A. Caligiuri, and J. E. Dick. A cell initiating human acute myeloid leukaemia after transplantation into SCID mice. *Nature*, 367:645–648, Feb 1994.

P. Laslo, C. J. Spooner, A. Warmflash, D. W. Lancki, H. J. Lee, R. Sciammas, B. N. Gantner, A. R. Dinner, and H. Singh. Multilineage transcriptional priming and determination of alternate hematopoietic cell fates. *Cell*, 126:755–766, Aug 2006.

C. Li, D. G. Heidt, P. Dalerba, C. F. Burant, L. Zhang, V. Adsay, M. Wicha, M. F. Clarke, and D. M. Simeone. Identification of pancreatic cancer stem cells. *Cancer Res.*, 67:1030–1037, Feb 2007.

N. A. Lobo, Y. Shimono, D. Qian, and M. F. Clarke. The biology of cancer stem cells. *Annu. Rev. Cell Dev. Biol.*, 23:675–699, 2007.

R. Losick and C. Desplan. Stochasticity and cell fate. *Science*, 320:65–68, Apr 2008.

S. A. Mani, W. Guo, M.-J. Liao, E. N. Eaton, A. Ayyanan, A. Y. Zhou, M. Brooks, F. Reinhard, C. C. Zhang, M. Shipitsin, L. L. Campbell, K. Polyak, C. Brisken, J. Yang, and R. A. Weinberg. The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell*, 133(4):704–715, MAY 16 2008. ISSN 0092-8674. doi: 10.1016/j.cell.2008.03.027.

R. G. McKinnel, R. E. Parchment, A. O. Perantoni, Damjanov I., and Pierce G. B. *The Biological Basis of Cancer*. Cambridge University Press, New York, 2nd edition, 2006.

F. Michor. Mathematical models of cancer stem cells. *J. Clin. Oncol.*, 26:2854–2861, Jun 2008.

F. Michor. Chronic myeloid leukemia blast crisis arises from progenitors. *Stem Cells*, 25:1114–1118, May 2007.

F. Michor, Y. Iwasa, and M. A. Nowak. Dynamics of cancer progression. *Nat. Rev. Cancer*, 4:197–205, Mar 2004.

F. Michor, T. P. Hughes, Y. Iwasa, S. Branford, N. P. Shah, C. L. Sawyers, and M. A. Nowak. Dynamics of chronic myeloid leukaemia. *Nature*, 435:1267–1270, Jun 2005a.

F. Michor, Y. Iwasa, C. Lengauer, and M. A. Nowak. Dynamics of colorectal cancer. *Semin. Cancer Biol.*, 15:484–493, Dec 2005b.

F. Michor, Y. Iwasa, and M. A. Nowak. The age incidence of chronic myeloid

leukemia can be explained by a one-mutation model. *Proc. Natl. Acad. Sci. U.S.A.*, 103:14931–14934, Oct 2006.

A. P. Morel, M. Livre, C. Thomas, G. Hinkal, S. Ansieau, and A. Puisieux. Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS ONE*, 3:e2888, 2008.

S. J. Morrison and J. Kimble. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature*, 441:1068–1074, Jun 2006.

J. D. Murray. *Mathematical Biology I: An Introduction*. Springer, New York, 3rd edition, 2002.

J. Neuzil, M. Stantic, R. Zobalova, J. Chladova, X. Wang, L. Prochazka, L. Dong, L. Andera, and S. J. Ralph. Tumour-initiating cells vs. cancer 'stem' cells and CD133: what's in the name? *Biochem. Biophys. Res. Commun.*, 355:855–859, Apr 2007.

C. O. Nordling. A new theory on cancer-inducing mechanism. *Br. J. Cancer*, 7: 68–72, Mar 1953.

C. A. O'Brien, A. Pollett, S. Gallinger, and J. E. Dick. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature*, 445: 106–110, Jan 2007.

C. H. Park, Bergsagel D. E., and McCulloch E. A. Mouse myeloma tumor stem cells - primary cell culture assay. *J. Natl. Cancer Inst.*, 46(2):411–&, 1971. ISSN 0027-8874.

E. Passegué, C. H. M. Jamieson, L. E. Ailles, and I. L. Weissman. Normal and

leukemic hematopoiesis: Are leukemias a stem cell disorder or a reacquisition of stem cell characteristics? *Proc. Natl. Acad. Sci. U.S.A.*, 100:11842–11849, 2003.

S. G. Piccirillo, B. A. Reynolds, N. Zanetti, G. Lamorte, E. Binda, G. Broggi, H. Brem, A. Olivi, F. Dimeco, and A. L. Vescovi. Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature*, 444:761–765, Dec 2006.

M. E. Prince, R. Sivanandan, A. Kaczorowski, G. T. Wolf, M. J. Kaplan, P. Dalerba, I. L. Weissman, M. F. Clarke, and L. E. Ailles. Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc. Natl. Acad. Sci. U.S.A.*, 104:973–978, Jan 2007.

E. Quintana, M. Shackleton, M. S. Sabel, D. R. Fullen, T. M. Johnson, and S. J. Morrison. Efficient tumour formation by single human melanoma cells. *Nature*, 456:593–598, Dec 2008.

D. C. Radisky and M. A. LaBarge. Epithelial-mesenchymal transition and the stem cell phenotype. *Cell Stem Cell*, 2(6):511–512, Jun 2008. ISSN 1934-5909. doi: 10.1016/j.stem.2008.05.007.

T. Reya, S. J. Morrison, M. F. Clarke, and I. L. Weissman. Stem cells, cancer, and cancer stem cells. *Nature*, 414:105–111, Nov 2001.

B. A. Reynolds and R. L. Rietze. Neural stem cells and neurospheres–re-evaluating the relationship. *Nat. Methods*, 2:333–336, May 2005.

B. A. Reynolds and S. Weiss. Generation of neurons and astrocytes from isolated cells of the adult mammalian central nervous system. *Science*, 255:1707–1710, Mar 1992.

L. Ricci-Vitiani, D. G. Lombardi, E. Pilozzi, M. Biffoni, M. Todaro, C. Peschle, and R. De Maria. Identification and expansion of human colon-cancer-initiating cells. *Nature*, 445:111–115, Jan 2007.

R. K. Sachs and D. J. Brenner. Solid tumor risks after high doses of ionizing radiation. *Proc. Natl. Acad. Sci. U.S.A.*, 102:13040–13045, Sep 2005.

R. K. Sachs, W. F. Heidenreich, and D. J. Brenner. Dose timing in tumor radiotherapy: considerations of cell number stochasticity. *Math Biosci*, 138:131–146, Dec 1996.

T. Schatton, G. F. Murphy, N. Y. Frank, K. Yamaura, A. M. Waaga-Gasser, M. Gasser, Q. Zhan, S. Jordan, L. M. Duncan, C. Weishaupt, R. C. Fuhlbrigge, T. S. Kupper, M. H. Sayegh, and M. H. Frank. Identification of cells initiating human melanomas. *Nature*, 451:345–349, Jan 2008.

R. M. Seaberg and D. van der Kooy. Adult rodent neurogenic regions: the ventricular subependyma contains neural stem cells, but the dentate gyrus contains restricted progenitors. *J. Neurosci.*, 22:1784–1793, Mar 2002.

M. Shipitsin and K. Polyak. The cancer stem cell hypothesis: in search of definitions, markers, and relevance. *Lab. Invest.*, 88:459–463, May 2008.

S. V. Shmelkov, R. St Clair, D. Lyden, and S. Rafii. AC133/CD133/Prominin-1. *Int. J. Biochem. Cell Biol.*, 37:715–719, Apr 2005.

I. Singec, R. Knoth, R. P. Meyer, J. Maciaczyk, B. Volk, G. Nikkhah, M. Frotscher, and E. Y. Snyder. Defining the actual sensitivity and specificity of the neurosphere assay in stem cell biology. *Nat. Methods*, 3:801–806, Oct 2006.

S. K. Singh, I. D. Clarke, M. Terasaki, V. E. Bonn, C. Hawkins, J. Squire, and P. B. Dirks. Identification of a cancer stem cell in human brain tumors. *Cancer Res.*, 63:5821–5828, Sep 2003.

S. K. Singh, C. Hawkins, I. D. Clarke, J. A. Squire, J. Bayani, T. Hide, R. M. Henkelman, M. D. Cusimano, and P. B. Dirks. Identification of human brain tumour initiating cells. *Nature*, 432:396–401, Nov 2004.

C. M. Southam and A. Brunschwig. Quantitative studies of autotransplantation of human cancer - preliminary report. *Cancer*, 14(5):971–978, 1961. ISSN 0008-543X.

G. J. Spangrude, S. Heimfeld, and I. L. Weissman. Purification and characterization of mouse hematopoietic stem cells. *Science*, 241:58–62, Jul 1988.

J. E. Till and E. A. McCulloch. Direct measurement of radiation sensitivity of normal mouse bone marrow cells. *Radiation Research*, 14(2):213–222, 1961. ISSN 0033-7587.

J. E. Till, E. A. McCulloch, and L. Siminovitch. A Stochastic Model of Stem Cell Proliferation, Based on the Growth of Spleen Colony-Forming Cells. *Proc. Natl. Acad. Sci. U.S.A.*, 51:29–36, Jan 1964.

I. P. Tomlinson and W. F. Bodmer. Failure of programmed cell death and differentiation as causes of tumors: some simple mathematical models. *Proc. Natl. Acad. Sci. U.S.A.*, 92:11130–11134, Nov 1995.

V. Tropepe, B. L. Coles, B. J. Chiasson, D. J. Horsford, A. J. Elia, R. R. McInnes, and D. van der Kooy. Retinal stem cells in the adult mammalian eye. *Science*, 287:2032–2036, Mar 2000.

C. Turner, A. R. Stinchcombe, M. Kohandel, S. Singh, and S. Sivaloganathan. Characterization of brain cancer stem cells: a mathematical approach. *Cell Prolif.*, In press.

N. Uchida, D. W. Buck, D. He, M. J. Reitsma, M. Masek, T. V. Phan, A. S. Tsukamoto, F. H. Gage, and I. L. Weissman. Direct isolation of human central nervous system stem cells. *Proc. Natl. Acad. Sci. U.S.A.*, 97:14720–14725, Dec 2000.

N. G. Van Kampen. *Stochastic Processes in Physics and Chemistry, Third Edition.* North Holland, April 2007. ISBN 0444529659.

I. M. van Leeuwen, H. M. Byrne, O. E. Jensen, and J. R. King. Crypt dynamics and colorectal cancer: advances in mathematical modelling. *Cell Prolif.*, 39:157–181, Jun 2006.

J. E. Visvader and G. J. Lindeman. Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. *Nat. Rev. Cancer*, 8:755–768, Oct 2008.

I. L. Weissman. The road ended up at stem cells. *Immunol. Rev.*, 185:159–174, Jul 2002.

M. S. Wicha, S. Liu, and G. Dontu. Cancer stem cells: an old idea–a paradigm shift. *Cancer Res.*, 66:1883–1890, Feb 2006.

H. E. Wichmann and M. Loeffler. *Mathematical Modeling of Cell Proliferation: Stem Cell Regulation in Hemopoiesis*, volume 1. CRC Press, Boca Raton, FL, 1985.

M. Wu, H. Y. Kwon, F. Rattis, J. Blum, C. Zhao, R. Ashkenazi, T. L. Jackson, N. Gaiano, T. Oliver, and T. Reya. Imaging hematopoietic precursor division in real time. *Cell Stem Cell*, 1:541–554, Nov 2007.

J. Yang and R. A. Weinberg. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev. Cell*, 14:818–829, Jun 2008.

Y. Yatabe, S. Tavar, and D. Shibata. Investigating stem cells in human colon by using methylation patterns. *Proc. Natl. Acad. Sci. U.S.A.*, 98:10839–10844, Sep 2001.

S. C. Zhang, D. Lipsitz, and I. D. Duncan. Self-renewing canine oligodendroglial progenitor expanded as oligospheres. *J. Neurosci. Res.*, 54:181–190, Oct 1998.