The nonlinear nature of biology

Marthe Måløy

July 22, 2018

Abstract

In this thesis I explore the stability and the breakdown of stability of biological systems. The main examples are the blood system and invasion of cancer. However, the models presented in the thesis apply to several other examples.

Biological systems are characterised by both competition and cooperation. Cooperation is based on an unsolvable dilemma: Even though mutual cooperation leads to higher payoff than mutual defection, a defector has higher payoff than a co-operator when they meet. It is not possible to represent this dilemma with a linear and deterministic model. Hence, the dilemma of cooperation must have a nonlinear and/or stochastic representation.

More general, by using a linearised model to describe a biological system, one might lose dimensions inherent in the complexity of the system. In this thesis I illustrate that a nonlinear description of a biological system is potentially more accurate and might provide new information.

The thesis is made up of three papers. Paper 3 presents the most general model which considers a relative stable population that is invaded by an alternative strategy. That is, a new type of individual is in general not advantageous when it appears in stable population. The newcomers can grow in number due to stochasticity. However, they can only become advantageous if they manage to change the environment in such a way that they increase their fitness. The model presented in paper 3 is an extension of the Moran process that captures this dynamics.

Paper 2 proposes a model that links self-organisation with symmetric and asymmetric cell division. The model assumes that cell divisions are completely random events, and that the daughter cells resulting from asymmetric and symmetric divisions are, in general equal, and still, the tissue has the flexibility to self-renew, produce mature cells and regenerate, due to self-organisation.

Paper 1 presents a model that illustrates that if symmetric stem cell division is regulated by differentiated cells, then the fitness of the stem cells can be affected by modifying the death rate of the mature cells. This result is interesting because stem cells are less sensitive than mature cells to medical therapy, and our results imply that stem cells can be manipulated indirectly by medical treatments that target the mature cells.

Acknowledgements

First I would like to thank my supervisors, Per Jakobsen and Bjørn Brandsdal, who have given me the freedom to pursuit my ideas and at the same time being available for questions and feedback. In particular, the conversations with my main supervisor, Per Jakobsen, regarding nonlinear phenomena and how models in physics can be applied to biology, have been vitally important in the process of developing the models presented in this thesis.

I would also like to thanks Rafael Lahoz-Beltrá, Juan Carlos Nuño and Antonio Bru. I learned a lot about biology and the amazing culture of Spain during my stay in Madrid in 2010.

I owe a great thanks to my supervisor for my master thesis, Boris Kruglilov, who were the first to point out to me that there is no clear distinction between deterministic and random behavior, and to the teachers of *Examen facultatum* for interesting lectures and conversations that enabled me to see my work in a in a broader context. What I learned from Boris and the philosophy teachers is the backbone of this thesis.

I must also express my gratitude to my co-author and brilliant mathematician, Frode Måløy. This thesis would not be possible without you – such a cliche, but so true. You have helped me with everything: programming, figures, proofreading. The list goes on. I would lost in the code without you.

I would also like to thanks Kristian Meisingset and the rest of editorial of *Minerva* for giving me the opportunity to write to a broader audience, for their feedback on how I can improve my writing, and for their attempts to teach me the comma rules. Many of the ideas presented in my latest scientific paper springs from the *Minerva* texts.

I have learned a lot about how to improve the efficiency of algorithms from conversations with Arthur Schuchter. But most of all, you are a really great person who have meant a lot to my family and me. When this is all over, I hope we can celebrate together on *Kvalhornet*.

I would like to express gratitude to my family for doing fun stuff with my children when their mother is being really boring. Thanks for Christmas dinners, New Year parties, fishing trips and road trips. Especially thanks to Ingrid Måløy, Torgeir Johansen, Ragnhild Rørvik and Odd Høyem.

And finally, the most amazing people in the world, Liv and Trym, and the best parent ever, Frode. I am so lucky to have you in my life. I look forwards to spend more time with you.

List of publications

The three papers that make up the thesis:

Paper 1 Marte Rørvik Høyem, Frode Måløy, Per Jakobsen and Bjørn Brandsdal Stem cell regulation: Implications when differentiated cells regulate symmetric stem cell division. *Journal of Theoretical Biology* 380, 203–219 (2015).

Paper 2 Marthe Måløy, Frode Måløy, Per Jakobsen and Bjørn Brandsdal. Dynamic self-organisation of haematopoiesis and (a)symmetric cell division. *Journal of Theoretical Biology* 414, 147–164 (2017).

Paper 3 Marthe Måløy, Frode Måløy, Rafael Lahoz-Beltrá, Juan Carlos Nuño and Antonio Bru. Extended Moran process that captures the struggle for fitness

Other contributions that serve as background material of the thesis

- Marte Rørvik Høyem. Differential Invariants of the 2D Conformal Lie Algebra Action. *Acta Applicandae Mathematicae - An International Survey Journal on Applying Mathematics and Mathematical Applications* 109, 61–73 (2010).

- Marthe Måløy. Samarbeid for selvbevisste. Minerva 25.08.2016 *https://www.minervanett.no samarbeid-for-selvbevisste/*.

- Marthe Måløy. Immunterapi må tilbys til alle. Minerva 20.09.2016 *https://www.minervanett. no/immunterapi-ma-tilbys/*.

- Marthe Måløy. Hvordan vet kroppen vår når vi trenger mye blod, og når vi trenger lite? Forskning.no 12.04.2017 *https://forskning.no/menneskekroppen-blod/2017/04/hvordan-vet-kroppen-var-nar-vi-trenger-mye-blod-selvregulering*.

- Marthe Måløy. CRISPR kan revolusjonere behandlingen av kreft og genetiske sykdommer. Minerva 17.08.2017. *https://www.minervanett.no/crispr-revolusjonere-behandlingen-kreft-genetiske-sykdommer/*.

- Marthe Måløy. Helse-Norge bør ta i bruk nye og dyre medisiner. Minerva 04.11.2017 *https://www.minervanett.no/helse-norge-bor-ta-bruk-nye-dyre-medisiner/*.

- Marthe Måløy. Frp blør legge innvandrings—og integreringsministeren på bordet. Minerva 15.12.2017 *https://www.minervanett.no/frp-bor-legge-innvandrings-integreringsministeren-pa-bordet/*.

Contents

Part I

Introduction

The fundamental and defining principles of evolutionary theory are replication, selection and mutation [1]. Based on these three principles, it is possible to describe a biological system with mathematical equations. However, since living organisms reproduce, die and change strategy based on feedback from their environment, it is very hard to derive a mathematical model that is both general and gives an accurate description of a specific biological system.

An important question, both in this thesis as well as in mathematical modelling in general, is what level of complexity should be included in a model. For instance, flipping a coin could be represented by an extremely complex model that includes Newton's laws, the laws of thermodynamics and so on. However, a simple stochastic model, with equal probability of head and tail, is in general better.

For similar reasons, a very simplistic and stochastic model might be the best representation of blood formation, in general. That is, *haematopoiesis* is the generation of blood, and at the root of this process is a small group of slowly replicating cells, called the *haematopoietic stem cells* [2], [3], which have the capability to maintain themselves through *self-renewal* and produce mature blood cells through *differentiation*. Under normal conditions, the number of haematopoietic stem cells is approximately constant, and Lenaerts et al. [4] show that the haematopoietic stem cells dynamics can be described by the Moran process, which represents the simplest possible model of stochastic evolutionary dynamics in a finite population. However, the production of mature blood cells increases after blood loss, and research results by Gokoffski et al. [5] indicate that the number of haematopoietic stem cells increases when the number of mature blood cells decreases. A deterministic model of haematopoietic stem cell dynamics with only one regulation mechanisme, such as the model presented in paper 1, can give a better description of these specific data than the Moran process. Nevertheless, as I make a case for in Section 2, because the regulation of haematopoiesis is an extremely complex process, a model that treats stem cell behaviour as purely random events fits general data better than a deterministic model with a single regulation mechanism.

In a nutshell, if a process is "very complex", it is in general better to use a simple stochastic model, than a complex deterministic model. However, as discussed more thoroughly in Section 6, machine learning and neural networks challenge the concept of "very complex". That is, neural networks will almost certainly have a transformative impact on modelling high-dimensional complex systems in the years to come. Neural networks have challenged traditional mathematical models and outperformed competing methods without providing clear evidence why they are doing so [6].

1 Nonlinearity can generate stochasticity

A human child has 23 chromosomes from each parent and these chromosomes contain basic genetic rules. However, a child is more than half of each parent, and this illustrates why linear representation of genetic information is problematic because linear equations cannot produce something new. On the contrary, nonlinear rules can actually produce something new and unexpected. That is, even a very simple nonlinear relation can generate complex patterns that no mathematical tools can penetrate. A good example is the logistic map

$$Y_{T+1} = 4Y_T(1 - Y_T), \tag{1}$$

which is described in more details below.

In a nutshell, a linear mathematical equation is easy to understand and simple to solve, whereas a nonlinear mathematical equation might be hard or even impossible to understand and solve, and it can generate new knowledge. Since living organisms reproduce, die, and change strategy based on feedback from their environment, their behaviour is seldom captured by linear equations. Even some of the simplest biological systems require nonlinear representations. This is illustrated below, where it is shown how the behaviour of a system of replicating individuals changes dramatically when the description changes from linear to nonlinear, when time is represented by time steps instead of a continuous variable, and when the feedback from the environment is delated.

1.1 Population growth

When a population is in an isolated environment, only replication can change the population size. Suppose that the function $x(t)$ describes the size of an isolated population at time t, and that the derivative of $x(t)$ with respect to t, $\frac{dx}{dt}$, exists. Let b and d denote the birth rate and the death rate, respectively. Since all changes in $x(t)$ are caused by birth or death, we have that

$$\frac{dx}{dt} = bx(t) - dx(t) = rx(t), \tag{2}$$

where $r = b - d$. If r is a constant, then the differential equation given above is linear and have solutions on the form

$$x(t) = x(0)\exp(rt), \tag{3}$$

where $x(0)$ is the population size at time $t = 0$. Moreover, if r is positive, then $x(t)$ describes a growing population in an isolated environment where there is no competition for resources. Nevertheless, a population cannot expand infinitely. Eventually, resources become limited, and then competition occurs. Ecologists have introduced a variety of modifications to the linear differential equation given in (2) to take account of saturation effects. The main idea in all of them is to reduce the growth as $x(t)$ becomes large. One approach is the differential equation

$$\frac{dx}{dt} = rx(t)\left(1 - \frac{x(t)}{C}\right), \tag{4}$$

where the constant C is the carrying capacity of the population. As discussed in the beginning of the introduction, most nonlinear problems are very hard, and sometimes even impossible, to solve. However, the differential equation in (4) is one of relatively few nonlinear problems that have neat solutions, namely

$$x(t) = \frac{Cx(0)\exp(rt)}{C + x(0)(\exp(rt) - 1)}, \tag{5}$$

where $x(0)$ is the population size at time $t = 0$.

Some populations nurture their offspring before these reach reproductive age, and in this case it might be better to include different generations. The following nonlinear recurrence relation captures this competition dynamics:

$$X_{T+1} = RX_T\left(1 - \frac{X_T}{C}\right), \tag{6}$$

6

where T is a non-negative integer denoting time steps, X_T is the population size at time step T, and R is a positive constant less than or equal to 4. Like for the continuous differential equation given in (4), C is the carrying capacity. But in contrast to the continuous case, if the population described by the discrete representation reaches the the carrying capacity, the population will get extinct the next time step. The reason why the behaviour of the discrete description of the population is very different from the continuous description, is that the feedback from the environment is implemented differently in the two cases.

By substituting the function $Y_T = X_T/C$ into the recurrence relation given in (6), we obtain the logistic map

$$Y_T + 1 = RY_T(1 - Y_T).$$

The logistic map does not have any neat formula for the solution in terms of T. However, for $R < 4$ it is possible to predict the outcome of the logistic map by using advanced mathematical analysis [7]. On the contrary, for $R = 4$ we obtain the logistic map given in Equation (1). It has been proved that the statistical mechanics of this system has exactly the same statistical properties as a random system. Actually, according to Ian Stewart [7]: "Random means that no obvious structure exists, but that on average we can say various things, such as how often the values occur in a given range. Random has carried the connotation of *indeterministic*, that is, a system is deterministic if it follows exactly some regular law, random if not". Nevertheless, the logistic map given in (1) shows there is no clear distinction between deterministic and random behaviour, because this map is a deterministic system which behaves randomly.

2 Randomness and self-organisation

The logistic map given in Equation (1) is an example of a deterministic equation which behaves randomly. And in point of fact, for deterministic, nonlinear, dynamical systems, strange attractors, chaos and random behaviour are the rule rather than the exception, and as the number of variables increase, the phenomena become more peculiar [8].

The DNA of cells in a multicellular organism contains deterministic rules. However, complex and nonlinear interactions between cells can create patterns that not only seem totally random, but that actually are random. Furthermore, it is possible that the organisms make use of the randomness, because randomly organised systems can exhibit *self-organisation* which is a spontaneous order that arises from local interactions between parts of an initially disordered system. This possibility is explored in the next subsection.

2.1 Haematopoiesis and self-organisation

Even though the haematopoietic system in a healthy adult is stable and robust, the behaviour of each blood cell seems chaotic and random. This indicates that haematopoiesis is a self-organised process.

A healthy adult contains about five litres blood, which corresponds to about 37×10^{15} blood cells. Each day, the body produces around 10^{15} blood cells. Thus, most of the cells in the haematopoietic system is replaced each month [9].

At the root of the blood forming process are the haematopoietic stem cells, which are located within the bone marrow and segregated among different bones throughout the body. The haematopoietic stem cells differentiate into progenitor cells, which differentiate into red blood cells, white blood cells or platelets, through sequential division. Since the number of haematopoietic stem cells is much smaller than the number of more differentiated blood cells,

the haematopoietic stem cells must be tightly regulated and protected. The haematopoietic bone marrow niches may be crucial in both aspects [10], [11]. Since a niche cannot be reconstructed experimentally, it is difficult to study haematopoietic stem cells in vitro, because stem cell survival, self-renewal and differentiation are regulated by signals from the niche. Hence, relatively little is known about the exact behaviour of haematopoietic stem cells. For example, when a haematopoietic stem cell divides, exactly what determines whether a daughter cell becomes a stem cell or starts to differentiate, is still unclear. This is related to the symmetry of the stem cell division, which is discussed in the next subsection.

2.2   Symmetric and asymmetric stem cell division

An important concept related to stem cell self-renewal and differentiation, is the symmetry of the stem cell division. That is, an asymmetric stem cell division results in one daughter cell that has stem cell identity, and another daughter cell that starts to differentiate, whereas a symmetric stem cell division generates two daughter cells that are destined to the same fate [12], [13]. There are two types of symmetric stem cell division, namely *symmetric self-renewal*, which results in two stem cells, and *symmetric differentiation*, where both daughter cells start to differentiate. Under normal conditions, the number of cells in a given tissue is approximately constant, and the stem cells differentiate and self-renew at relatively constant rates to replace mature cells and to keep the stem cell number at a certain normal level [14], [15]. By dividing asymmetrically, the stem cells manage to both self-renew and produce differentiated cells in a single division. However, a disadvantage of asymmetric stem cell division is that it leaves stem cells unable to expand in number. It is, in general, believed that the stem cells can regenerate [12], [13]. For instance, haematopoietic stem cells can expand rapidly in response to injury to the bone marrow, such as stem cell transplantation [16]. Hence, asymmetric self-renewal cannot be the complete story, since it leaves stem cells unable to expand in number. The number of stem cells increases by one after symmetric self-renewal. Since the haematopoietic bone marrow can regenerate after injury, it is likely that the rate of symmetric self-renewal depends on the number of haematopoietic stem cells. On the contrary, the number of stem cells decreases by one after a symmetric commitment. Thus, the two types of symmetric divisions must occur at the same rate under normal conditions.

As discussed more thoroughly in paper 2, several experiments on Safari cats by Abkowitz et al. indicate that haematopoietic stem cells divide mostly asymmetrically under normal conditions, whereas when the haematopoietic bone marrow niche regenerates after injury, the haematopoietic stem cells start to divide symmetrically [16], [17], [18]. But does this mean that a stem cell somehow "knows" that it must divide asymmetrically under normal conditions and self-renew symmetrically when stem cells need to be replaced? This would also mean that the daughter cells inherit this "knowledge". However, as discussed in paper 2, the assumption that each cell "knows" how to behave in different situations is too rigorous and potentially misleading. Thus, it is more likely that each stem cell behaves completely random. Nevertheless, the stem cells divide mostly asymmetrically under normal conditions and symmetrically under regeneration, due to dynamic regulation and self-organisation in the haematopoietic bone marrow niche.

As discussed more thoroughly in paper 2, several experiments on *Drosophila* germline stem cells indicate that the stem cell niche can contain up to a certain number of cells, and that the niche is approximately full under normal conditions. When a stem cell divides, one of the daughters inherits the mother's place in the niche and retains stem cell identity. The fate of the second daughter depends on whether there is a vacant place in the niche or not. In the first case, the second daughter daughter remains in the niche and retains stem cell identity. If the niche is full, the second daughter is placed outside the niche, and loses its stem cell identity. Thus,

8

research on *Drosophila* germline stem cells indicates that the stem cells do not "know" whether they must divide asymmetrically or symmetrically. That is, the stem cells divide randomly, and the availability of the niche, and perhaps some other factors, determines whether the division is asymmetric or symmetric. This implies that an undifferentiated cell must be in the niche to function as a stem cell: Once a cell is placed outside the niche, it is no longer a stem cell.

3 Moran process and the invasion of mutants

As discussed in the previous subsection, under normal conditions, the rate of symmetric self-renewal must equal the rate at which the stem cells leave the niche, and in this case, the haematopoietic stem cell dynamics can be described by the Moran process, which assumes that the population size is constant and that at each time step, a random cell is selected to self-renew symmetrically and a random cell is selected to leave the growth environment. Dingli et al. present a version of the Moran process which includes all types of stem cell division [19]. However, for simplicity, we only consider symmetric stem cell division here.

Genetic changes called *mutations* can occur in any cell that divides [20]. Even though most mutations are harmless to the body, progressive accumulation of mutations can lead to cancer [21].

Results from theoretical work indicate that the tissue architecture of the haematopoietic system, where only a small number of stem cells have the ability to self-renew, has evolved to minimise the risk of malignant transformations [19]– [22]. That is, if a mutation occurs in a mature blood cell, it is likely to be washed out of the system before it becomes a cancer phenotype, because these cells do not self-renew. On the other hand, a mutation in a haematopoietic stem cell can generate a different type of stem cell, denoted *mutant stem cell*. This leads to an evolutionary process with competition between the mutant stem cells and the normal stem cells [1]. Lenaerts et al. [4] show that this competition dynamics might be captured by the Moran process. That is, the Moran process assumes that initially, all the stem cells are normal. When a normal stem cell self-renews, a mutation that creates a mutant stem cell occurs with probability u. The normal stem cell self-renews at rate 1 whereas the mutant stem cells self-renew at rate r. All stem cells are selected to leave the niche at the same rate. Hence, the mutant type is advantageous if $r > 1$, neutral if $r = 1$ and disadvantageous if $r < 1$. At each time step, the number of mutant stem cells can either increase by one, remain constant or decrease by one. The probability for these three events are

$$P(i+1|i) = \frac{u(N-i)+ri}{N-i+ir}\frac{N-i}{N}, \qquad (7)$$

$$P(i-1|i) = \frac{(1-u)(N-i)}{N-i+ir}\frac{i}{N}, \qquad (8)$$

$$P(i|i) = 1 - P(i+1|i) - P(i-1|i), \qquad (9)$$

respectively, where N is the number of stem cells in the niche and i is the number of mutants. If u is sufficiently small, the mutant type typically has time to take over the whole niche or get extinct before another mutant is created from the normal type. By using the approximation $u \approx 0$, Wodarz and Komarova [23] show that the probability that i mutant stem cells eventually invade the whole niche is

$$\rho_i = \frac{r^{N-i}\left(1-r^i\right)}{1-r^N} \qquad (10)$$

if $r \neq 1$ and

$$\rho_i = \frac{i}{N} \tag{11}$$

if it is a neutral Moran process, i.e. $r = 1$.

The reason why the mutant type can invade the whole niche, starting from a single mutant stem cell, is that the stem cells self-renew symmetrically. Note that if the stem cells only divided asymmetrically and no new mutation in a normal stem cell occurred, the number of mutants would remain constant. However, most type of cancers require more than one mutation, and it is illustrated in the paper by Shahriyari and Komarova [24] that symmetrically dividing cells can delay a second mutant production compared to an equivalent system with only asymmetrically dividing stem cells. More precisely, if stem cells only divide asymmetrically, then a mutation acquired in a stem cell will remain in the system indefinitely, and it is only a matter of time before the second mutation occurs. On the contrary, a mutant stem cell generated in a symmetric division has a less certain fate – half of the lineages will differentiate out after the very first division and only 1/K of all lineages will expand to size K. Thus, that the uncertainty of the fate of single mutant stem cells can be the reason for the statistically longer time it takes for the symmetrically dividing stem cell model to produce a double-hit mutant.

Unlike most types of cancers, the first phase of *chronic myeloid leukaemia* is caused by a single mutation in a haematopoietic stem cell that creates a *leukemic stem cell*. Since the mutation rate from normal cells to leukemic cells is nonzero, it follows from the Moran process that any person will eventually develop chronic myeloid leukaemia, given that he or she has a sufficiently long life. This might seem to contradict the phrase *the survival of the fittest*, that originated from Darwinian evolutionary theory as a way of describing the mechanism of natural selection. However, as discussed more thoroughly in the next sections, the cooperation among cells in a multicellular organism is based on an unsolvable dilemma, and, hence, it will sooner or later dissolve. So maybe we should rather say: *the survival of the one that keeps it together until after successful reproduction.* That is, the one thing that protects us from most types of cancer, is death.

4 The rise and fall of unconditional co-operators: the prisoner's dilemma and evolutionary stable games

There are approximately 5×10^{30} bacteria on Earth, and their biomass exceeds that of all plants and animals. Bacteria are present in most habitats: soil, water, radioactive waste, acidic hot springs, the deep portions of Earth's crust as well as in symbiotic and parasitic relationships with plants and animals. Furthermore, bacteria were among the first life forms and will most likely exist longer than multicellular organisms. It might seem like a mystery that multicellular organisms evolved from bacteria about 1.5 billion years ago, given that bacteria in so many ways are fitter than multicellular organisms.

The first life forms adopted the most basic strategy, which is to outcompete other individuals by dividing as fast as possible, when life started to evolve about four billion years ago [1]. Nevertheless, proliferation requires resources such as nutrient molecules and space, and different individuals can have access to different resources. Thus, cooperation can be beneficial in these situations [25], [26]. A simplified example of cooperation among single-celled organisms is that one cell has access to enough nutrient molecules for two cell divisions but no space, whereas another cell has access to enough space for two cell divisions but no nutrient molecules. Hence, if both cells share their resources, i.e. mutual cooperation, they will both reproduce. On the

contrary, if both cells do not share their resources, i.e. mutual defection, none of the cells reproduce. However, if only one cell shares its resources and the other does not share, then the co-operator does not reproduce and loses its resources whereas the defector reproduces twice. This simple example illuminates the dilemma of cooperation: even though mutual cooperation leads to higher payoff than mutual defection, a defector has higher payoff than a co-operator when they meet. Indeed, this example is a version of the well-known game called the prisoner's dilemma [1]. Moreover, it illustrates why unconditional cooperation is an unstable strategy: Consider a group of co-operators. If a mutation causes a cell to change strategy to defection, this cell increases its payoff. On the contrary, a strategy is a *Nash equilibrium* if no player, which in our example are cells, can deviate from this strategy and increase its payoff [27]. Defection is a Nash equilibrium both in our example with cells and in prisoner's dilemma in general, because if a defector mutates into a co-operator, it increases its payoff.

A Nash equilibrium is also an evolutionarily stable strategy if selection opposes the invasion of an alternative strategy [25]. That is, if a sufficiently large population adopts an evolutionarily stable strategy, it cannot be invaded by a alternative strategy that is initially rare. For prisoner's dilemma, defection is an evolutionarily stable strategy. Hence, co-operators cannot invade a large population of defectors. However, as illustrated by the Moran process, a relatively small group of defectors can be invaded by co-operators. Moreover, if the co-operators develop regulation mechanisms that control the cooperation, for instance by modifying the microenvironment such that the defectors lose their advantages, then the group can survive in the long term. Indeed, the evolution of multicellular organism was driven by increasingly advanced regulation mechanisms among cooperating cells [1].

5 Evolution of multicellular organisms: From randomness to strict regulation and back

The healthy life and development of an advanced multicellular organism, for instance a human being, depend upon the cooperation between millions of cells. Nevertheless, as discussed in Section 4, unconditional cooperation, such as the cooperation given in the prisoner's dilemma, is an unstable strategy. Hence, cooperation among cells in an advanced multicellular organism must be regulated by by a complex network of cellular checkpoints and signals.

Multicellular organisms consists of more than one cell. Similar to single-celled organisms that belong to a colony, the cells in a multicellular organism must cooperate. Be that as it may, even the simplest multicellular organisms have cells that depend on each other to survive, whereas the single-celled organisms that live in colonies, can survive on their own.

Multicellular organisms evolved from colonies of single-celled organisms. As discussed in the previous section, cooperating cells are vulnerable to mutants that change strategy to defection since these cells can invade the colony by exploiting the cooperation.

In paper 3, an extension of the Moran process with non-constant fitness is presented. This model captures the competition between co-operators and defectors, but also how the co-operators can change their environment such that the fitness of the defectors is reduced.

Advanced multicellular organisms, such as human beings, are maintained by very complex regulation networks. And, as illustrated in Section 1, as interactions get more complex and nonlinear, they can generate chaos and random behaviour. Indeed, as discussed in Section 2, it is possible that advanced multicellular organisms have evolved to make use of the randomness that is generated by the complexity of the multiple signals.

5.1 Multicellularity and cancer

The cooperation among cells in a multi-cellular organism is regulated by advanced control mechanisms that promote stability for a relatively long time. Nevertheless, multicellular organisms and all other forms of cooperation will eventually break down because natural selection favours defection.

Since children need care of for several years, humans must have a long life to reproduce successfully. As discussed in Section 3, mutations can occur in any cell that divides [20], and even though most mutations do not harm the body, progressive accumulation of mutations can lead to cancer. That is, mutations in the genetic code can make a cell ignore signals from other cells. For example, a mutant cell can divide when it is not needed and fail to undergo apoptosis. In general, mutant cells that stop cooperating are attacked by the defence system of the body, for instance killer T-cells. Thus, as discussed more thoroughly in paper 3, mutant cells are in most cases disadvantageous when they first appear in the body. However, if the mutant type manages to change the microenvironment such that at least some variants of the mutant cells become advantageous, the mutant type is likely to spread and cause cancer, which is the breakdown of cellular cooperation. That is, cancer is really a calculated risk of multicellularity: Cooperation is not a stable state, and hence, it will eventually break down. Control mechanisms increase the probability that the cooperation lasts long enough to lead to successful reproduction. However, there is no guarantee that mutations create defective cells that escape the control network and destroy the body.

Some cancer cells are programmed to adopt the strategy of primitive single-celled organisms: divide as fast as possible and outcompete all other cells. Healthy human cells cannot survive on their own. On the other hand, cancer cells might behave more like single-celled organisms and in some cases they can survive on their own. An example is the HeLa cell line, which are the cancer cells of Henrietta Lacks who died of cervical cancer in 1951. The HeLa cell is still used for scientific pursuits [28].

As discussed more thoroughly in Section 11, even though treatment for cancer and other genetic diseases are getting better every year, many types of cancer are so complex that we might never fully understand them. Nevertheless, machine learning and deep neural networks challenge the concept of "very complex", and will almost certainly have a transformative impact on modelling high-dimensional complex systems, such as cancer, in the years to come.

6 Machine learning

Machine learning enables computer systems to learn through progressively improving performance on a specific task. That is, the computer system is not explicitly programmed, but uses statistical techniques on big data. Deep neural networks are a type of machine learning that is inspired by biological neural networks. Deep neural networks have become the dominant mining tool for big data applications in the last decade, and it is expected that this type of machine learning will make their mark in the general area of high-dimensional, complex dynamical systems [6].

Neural networks are inspired by the work of Hubel and Wiesel on the primary visual cortex of cats [29], which they won the Nobel prize for. Their experiments demonstrated that neuronal networks were organised in hierarchical layers of cells that process visual stimulus. The first mathematical model of a neural network was presented in 1980 by Fukushima et. al [30], but up until the last decade, the neural networks have not been widely used. The recent success of deep neural networks has two major reasons [6], namely:

1. The continued growth of computional power.

2. Exceptionally large labelled data sets which take advantage of the power of multi-layers (deep) architecture.

Despite the success of deep neural networks, several basic questions remain wide open, for instance:

1. How many layers are necessary for a given data set?

2. How many nodes at each layer are needed?

3. How big must my data set be to properly train the network?

4. What guarantees exist that the mathematical architecture can produce a good predictor of the data?

5. What is the uncertainty and/or statistical confidence in deep neural network output?

6. Can I actually predict data well outside my training data?

7. Can I guarantee that I am not overfitting my data with such a large network.

The next decade will most likely witness significant progress in addressing these issues.

Part II

Results

In this part, the three papers which make up the thesis, are presented and discussed.

Paper 3, which was published in 2018, presents a model that capture the dynamics that occurs when a relatively stable population is invaded by an alternative strategy. Since this model is the most general, it is discussed first. Paper 2 was published in 2017 and presents a model of haematopoiesis that links self-organisation with symmetric and asymmetric cell division. This model can reproduce several experimental results. In paper 1 from 2015 we use a mathematical model to show that if symmetric cell division is regulated by differentiated cells, then changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells.

7 Paper 3: Extended Moran process that captures the struggle for fitness

Natural selection can cause evolution if there is enough variation in a population. When a mutant is generated in a stable population, the ability to create new variants is important for the mutant type if it is going to have any chance to invade the population. However, as discussed in Section 4, no individual has anything to gain from changing only its strategy in an evolutionary stable population [27], and this indicates that the mutant type must also change its environment to become advantageous.

That is, when a mutant is generated in a relatively stable population, it is most likely not advantageous. However, the number of mutants can grow due to stochasticity, and indeed, the mutants can invade a relatively small population, as illustrated by the Moran process presented in Section 3. Nevertheless, the mutants become advantageous only if they change their environment such that their fitness increases. This dynamics was present in the evolution of cooperation among bacteria and multicellularity [26], [31], [32], the invasion of cancer [33] and evolution of ideas that contradict social norms [34], [35]. In paper 3, we propose an extension of the Moran process with non-constant fitness that captures this dynamics. To be ore specific, individuals of the population can change the environment in such a way that the fitness landscape of the population is modified. That is, the model presented in paper 3 captures the struggle for fitness as well as the competition between different types of individuals.

Interestingly, the model can capture the invasion of defection as well as invasion of co-operators. That is, unconditional co-operators are expected to be exploited until they are extinct if they appear in a large group of defectors. The best possible scenario for this type of co-operators is that they manage to change their environment such that another type of co-operators that only cooperate under certain conditions, becomes advantageous. Similarly, when defectors appear in a regulated cooperation, the first generation of defectors typically dies while changing the environment such that coming generations become more advantageous.

7.1 Invasion of co-operators

As discussed in Section 4, it might seem like a mystery that multicellular organisms evolved from bacteria since natural selection favours defection over cooperation. The model presented in paper 3 illustrates that a small group of co-operators can invade a large population of defectors if they manage to change the environment such that defection becomes a disadvantageous strategy. That

is, the model presented in paper 3 assumes that initially the population dynamics is captured by a neutral Moran process, described in Subsection 3, and that all the individuals in the population are single-celled organisms that defect. A mutation can create a single-celled organism that co-operate. In this model, cooperation is captured by so-called detain entities that are activated by the co-operators. If a detain entity and a co-operator located at the same site, then the co-operator can be selected to die to give room for reproduction. On the other hand, the defectors ignore the detain entities, but reproduce if there is room for new daughter cells. Thus, the defectors are initially fitter than the co-operators. However, due to stochasticity, the co-operators can avoid extinction, and what is more, the model has an additional parameter called the *temperature*. In this example, the temperature represents regulation mechanisms. That is, initially, when there are only defectors in the population, the temperature is zero. However, the co-operators raise the temperature, and when the temperature reaches a certain limit, Υ, the regulation mechanisms start to kill the single-celled organisms that defect, and hence, defection becomes a disadvantageous strategy in the population.

7.2   Invasion of defectors and the Warburg effect

The model presented in paper 3 can also capture the dynamics of cancer invasion in solid tissues. As discussed more thoroughly in Subsection 5, mutant cells are in general not advantageous when they first appear in a human body because these cells are attacked by the defence system of the body. In this example, the detain entities represent the immune response. However, it is assumed that the immune cells are activated only if the mutant cells are harming healthy tissue. Moreover, the body can limit the blood flow to the microenvironment where the mutant cells are located. Consequently, the mutant cells break down the end product of glycolysis anaerobically, and this causes an acidic microenvironment. Hence, the temperature represents the acid level in the model, and it is assumed that the death rate of the healthy cells and the mutant cells that are not acid-resistant, increase when the acid level reaches the limit Υ.

To be more concrete, when the first mutant cell is generated, the acid level is zero. The mutant cells raise the acid level, but as long as the healthy tissue is not harmed, the competition dynamics between the healthy cells and the mutants cells is captured by the neutral Moran process. However, when the acid level reaches Υ, the healthy tissue is damaged, and consequently, the immune cells are activated. If none of the mutant cells are acid resistant, then the mutant type becomes disadvantageous, whereas if the mutant cells have generated a type of cells that are acid-resistant, then these cells are advantageous as long as there are less than N immune cells in the microenvironment. Hence, there is a race between the resistant mutants and the immune cells to reach population size N. If the immune cells respond quickly and reach population size N before the healthy cells in the microenvironment are extinct, the acid-resistant mutants are neutralised. In this case, the mutant cells are vulnerable to new immune attacks. On the other hand, the invasion of the resistant mutants represents the onset of a much more aggressive form of cancer, and in point of fact, in many cases, cancer cells exhibit glucose fermentation even when there is enough oxygen present. This is called the Warburg effect [33].

7.3   Further work

The model presented in paper 3 could be extended by including the interplay between evolution and learning, which is an important issue in evolutionary computation [36]. This plays a significant role in application areas that were used as examples in paper 3, such as biological modelling, multi-agent systems, economics and politics. All of these studies involve systems of interacting autonomous individuals in a population, and this raises several questions, like "Is

there any equilibrium?" and "How can cooperative behaviours evolve?". It is possible to apply methods from machine learning to seek an answer to these questions.

In the version of the prisoner's dilemma given in Subsection 4, single-celled organisms, that either cooperate or defect, compete in a finite population. Each cell has a fixed strategy, and after a normal cell division, both daughter cells inherit the strategy of the mother cell. However, when a cell divides, a mutation that changes the strategy of the daughter cells, can occur. Thus, if the population is relatively small, the co-operators can invade the population due to stochasticity. Nevertheless, if the population is sufficiently large, it will, eventually, be dominated by defectors, since defection is an evolutionarily stable strategy.

In contrast, a mix between cooperation and defection is often observed in several examples of prisoner's dilemma, for instance in human society [37]. One explanation is that in human life, a player often expects to meet the same opponent in the future, and he might remember a previous defection and take revenge [36]. On the other hand, if all the opponents know that a player always cooperate, they are likely to exploit him and defect. In a nutshell, the players use previous knowledge to decide whether to cooperate or defect.

In [36], the authors investigate the challenge of developing intelligent machine learning applications to address the problems of adaptation that arise in multi-agent systems, like expected long tem profit optimization. Moreover, the authors propose a learning algorithm for the emphiterated prisoner's dilemma problem and show that it performs strictly better than the tit-for-tat algorithm and many other adaptive and non-adaptive strategies. It would be interesting to study how these examples apply to the model presented in paper 3.

8   Paper 2: Dynamic self-organisation of haematopoiesis and (a)symmetric cell division

As discussed in Section 2, the blood system consists of approximately 37 trillion cells, and most of the differentiated blood cells are replaced each month. Hence, it is likely that haematopoiesis is regulated by self-organisation.

The model presented in paper 2 has a flexible and dynamically regulated self-organisation based on cell–cell and cell–environment interactions and extracellular regulations. What is more, the model links symmetric and asymmetric cell division with self-organisation, and as far as we know, our model is the first to make this connection.

The classical definition of a stem cell is an undifferentiated cell capable of self-renewal, production of a large number of differentiated cells, regenerating tissue after injury and a flexibility in the use of these options. This definition is fundamentally based on a functional perspective. As discussed by Loeffler and Roeder [38], the flexibility criterion attracted little attention when the definition of stem cells was first introduced. Yet considerable experimental results indicate that flexibility is a fundamental property of the stem cells [39], [40], [41]. For example, Zhang et al. [39] managed to bias the degree of lineage commitment by several maneuvers that altered the growth environment of the haematopoietic system.

Furthermore, many experiments show that haematopoietic stem cells can be manipulated such that they act as stem cells for another tissue such as neuronal and myogenic [40]. These experiments indicate that the growth environment is an important factor when tissue specification of stem cells are redirected.

The bone marrow niche contains both localised signalling cells and an extracellular matrix that support stem cell behaviour and control the fate of the undifferentiated cells [10], [11]. However, since it is not possible to reconstruct a bone marrow niche experimentally, the exact behaviour of haematopoietic stem cells is unknown. On the other hand, research on *Drosophila*

germline stem cells provides a clearcut example of how the niche maintains stem cell behaviour. That is, experiments on *Drosophila* germline stem cells support the following conjectures:

1. The stem cell niche promotes stem cell maintenance.

2. The stem cells self-renew at random.

3. When a stem cell self-renews, one of the daughter cells inherits the mother's place in the stem cell niche and retains stem cell identity, whereas the fate of the second daughter depends on the availability of space in the stem cell compartment – it either slips into a random vacant place in the stem cell compartment and remains a stem cell (symmetric self-renewal), or the second daughter leaves the stem cell compartment and loses its stem cell identity (asymmetric self-renewal).

4. Under normal conditions, the stem cell compartment is approximately full, and the stem cells typically self-renew asymmetrically.

5. When the stem cell compartment is not full, the rate of symmetric self-renewal generally increases, which leads to an expansion in the number of stem cells. The cells swift back to asymmetric self-renewal as the stem cell compartment reaches normal conditions.

The model presented in paper 2 assumes that Conjecture 1—6 also hold for the haematopoietic system. More specifically, the model assumes that all haematopoietic cell divisions occur randomly and that a haematopoietic stem cell is an undifferentiated cell located in a niche. That is, if a stem cell leaves the niche, it loses its stem cell identity. What is more, the daughter cells resulting from a stem cell division are phenotypically identical regardless of whether the division was asymmetric or symmetric. Due to self-organisation, the daughter cells remain in the niche and obtain stem cell identity or are placed outside the niche and commit to differentiation, depending on the need for self-renewal and differentiation. This is implemented by subdividing the niche into sites which represent signals and the environment as well as physical space.

8.1 Results from experiments on Safari cats can be explained by a self-organised model

As discussed above, relatively little is known about the exact behavior of the haematopoietic stem cells. On the contrary, haematopoietic progenitors have been studied both in vivo and in vitro. Loosely speaking, progenitors are cells on the first stage of the differentiation process.

Abkowitz et al. designed a set of experiments, using female Safari cats, to predict the contribution of haematopoietic stem cells to progenitor cells [17], [16], [42]. The Safari cat is a hybrid of the Geoffroy's cat (a South American wildcat) and a domestic cat (which is of Eurasian origin). These two species have evolved independently for twelve million years, and have distinct phenotypes of the X chromosome-linked enzyme glucose-6-phosphate dehydrogenase (G6PD). Female Safari cats have some cells that contain Geoffroy-type G6PD (G G6PD) and other cells that contain domestic-type G6PD (d G6PD). The G6PD phenotype is retained after replication and differentiation, and is functionally neutral. Therefore, it provides a binary marker of each cell and its offspring. In particular, this means that a progenitor cell that expresses G G6PD is the daughter of a stem cell that expresses G G6PD, and likewise, a progenitor cell that is d G6PD-positive is the daughter of a stem cell that is d G6PD-positive. Abkowitz et al. tracked the contributions of haematopoietic stem cells to the progenitor cells by observing the G6PD phenotype of haematopoietic progenitor cells.

Abkowitz et al. found that the percentage of progenitor cells expressing d G6PD remained relatively constant in normal female Safari cats. On the contrary, they observed that the percentage

of progenitor cells expressing d G6PD variated while the cells in the bone marrow regenerated, and, what is more, they found that the pattern of clonal contribution to haematopoiesis in each cat was unique. For instance, some of the cats that both had cells expressing d G6PD and cells expressing D G6PD when the regeneration started, had only cells expressing either d G6PD or D G6PD when the production of bone marrow cells stabilised after regeneration. Thus, one of the phenotypes had got extinct during the regeneration. On the contrary, in other cats, the percentage of cells expressing d G6PD and D G6PD remained on average relatively constant. Moreover, in some cats, significant variation in the percentage extended for years after the number of cells reached normal population levels, whereas in other cats, the percentage remained approximately constant.

Since the percentage of cells expressing d G6PD remained relatively constant when normal female Safari cats were observed, the experiments by Abkowitz et al. indicate that haematopoietic cells divide asymmetrically under normal conditions, because this type of division cannot change the number of stem cells expressing d G6PD. On the other hand, the number of stem cells expressing d GPD can increase or decrease by one after a symmetric stem cell division. Hence, since wide fluctuations in the percentage of cells expressing d G6PD were observed when the bone marrow regenerated, the experiments by Abkowitz et al. indicate that the rate of symmetric stem cell division increases during regeneration of the stem cell niche.

Other mathematical models of the haematopoietic system that include symmetric and asymmetric stem cell division, have been proposed, and they can reproduce several of the results obtained by Abkowitz et al. For instance, Wodarz and Komarova [23] propose a model where the haematopoietic stem cells divide asymmetrically under normal conditions and to symmetric division during regeneration. On the contrary, in the model presented by Abkowitz et al. [18], the haematopoietic stem cells only divide symmetrically. That is, under normal conditions, the haematopoietic stem cells undergo symmetric self-renewal and symmetric commitment at the same, constant rate, and under regeneration, the rate of the former type of division increases. Even though these models capture important aspects related to stem cell behaviour, it is a drawback that stem cell self-renewal and differentiation do not depend on local growth conditions because this implies that a stem cell somehow "knows" that it must self-renew symmetrically when stem cells need to be replaced. However, as discussed in Section 2, this assumption is potentially misleading and too rigorous. On the contrary, since the model presented in paper 2 links self-organisation with symmetric and asymmetric cell division, the rate of symmetric and asymmetric stem cell division is regulated by the needs of the haematopoietic system.

8.2 Differentiated cells

The model presented in paper 2 also includes the differentiated cells. It is assumed that these cells go through N stages of differentiation and that the cells that are at stage i in the differentiation process, are located in the i-th compartment. These compartments represent the sum of signals in the environment of the cells and not just physical locations. Moreover, it is assumed that the commited cells can only differentiate symmetrically. That is, if a cell in the i-th compartment divide, then both daughter cells migrate to the $i+1$-th compartment. Under normal conditions, there are approximately $2^i M$ cells in the i-th compartment, where M is the number of cells is stem cell niche when it is full, and the cells commit symmetrically to differentiation at the same, constant rate. However, the model assumes that there is a feedback from compartment i to compartment $i-1$, such that the system regenerates if there are less cells than under normal conditions.

8.3 Further work

The model presented in paper 2 is very simple with two parameters only, M and K, which are the number of sites in the stem cell niche and the number of compartments of differentiated cells, respectively.

In an extended version of the model, the committed haematopoietic cells should be divided into the erythroid lineage, the lymphoid lineage and the myeloid lineage. The first lineage is composed of red blood cells, the second of immune cells and the third includes granulocytes, megakaryocytes and macrophages [9]

It is still not clear exactly how differentiation of haematopoietic cells is regulated. In 1957, Waddington presented an epigenetic landscape that describes the differentiation of cells as the trajectories of balls rolling at random into branching valleys, where each branch represents a developmental state [43]. Based on Waddington's model, Furusawa and Kaneko propose a dynamical system model of cells with intracellular protein expression dynamics and interactions with each other [44]. The model predicts that cells with irregular, or chaotic, oscillations in gene expression dynamics have the potential to differentiate into other cell types. During development, such complex oscillations are lost successively, leading to loss of pluripotency. Their results are consistent with the view that pluripotency is a statistical property defined at the cellular population level, correlating with intra-sample heterogeneity, and driven by the degree of signalling promiscuity in cells.

To extend the model to include different lineages of the committed haematopoietic cells, it could be an advantage to use methods from big data analysis, such as machine learning, because these methods offer new ways to study the genome, transcriptome, proteome, and epigenome at the single-cell level. An increasing number of single-cell sequencing data makes it possible to carry out statistical inferences of pluripotency regulating genetic networks. In the work by Lin et. al, the authors develop a framework based on machine learning which explicitly account for the promoter architectures and gene state-switching dynamics. Their framework is useful for disentangling the various contributions that gene switching, external signaling, and network topology make to the global heterogeneity and dynamics of transcription factor populations. Their findings indicate that the pluripotent state of the network might be a steady state which is robust to global variations of gene-switching rates.

Differentiation modifies molecular properties of stem and progenitor cells, which leads to changing shape and movement characteristics. Buggenthin et al. present a method based on machine learning that predicts lineage choice in differentiating haematopoietic progenitors. Their method can detect lineage choice up to three generations before conventional molecular markers are observable. Thus, their approach manages to identify cells with differentially expressed lineage-specifying genes without molecular labelling.

9 Paper 1: Stem cell regulation: Implications when differentiated cells regulate symmetric stem cell division

Similar to the model discussed in the previous section, the model presented in paper 1 is used to study how stem cell division is regulated by other cells. However, the main focus of this paper is that changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells if symmetric stem cell division is regulated by differentiated cells, and this means that the relative fitness of the stem cells can be affected by modifying the death rate of the differentiated cells. This result is interesting because stem cells are in general less sensitive to medical therapy than differentiated cells, and our result implies that stem cells

can be manipulated indirectly by medical treatments that target the differentiated cells.

9.1 Symmetric stem cell division and cancer

As discussed more thoroughly in Subsection 2.2 and Section 8, the number of stem cells increases by one after a symmetric self-renewal, whereas after a symmetric differentiation, the number of stem cells decreases by one.

It is assumed that under normal condition the number of stem cells is approximately constant, and that these cells self-renew and differentiate at relatively constant rates to keep the number of stem cells at normal level and replace mature cells [45] Moreover, it has been shown that the haematopoietic stem cells can expand rapidly in response to stem cell transplantation and other injuries to the bone marrow. This indicates that the rate of symmetric self-renewal depends on the number of stem cells in the niche, since this is the only type of division that increases the number of stem cells.

As discussed more thoroughly in Section 2, the production of mature blood cells increases after blood loss. A symmetric differentiation produces two daughter cells that commit to differentiations, whereas an asymmetric stem cell division produces only one. However, since asymmetric stem cell division leaves the stem cell number unchanged, the stem cell niche is protected against fluctuations if only the differentiated regulate asymmetric stem cell division. Indeed, Wodarz propose a model where the rate of symmetric self-renewal depends only on the number of stem cells in the niche, whereas the differentiated cells regulate the rate of asymmetric stem cell division when there are only healthy cells in the system [46]. This means that the population dynamics of the stem cell niche is not influenced by the differentiated cells. However, as I make a case for in Section 8, it is likely that tissues, such as the haematopoietic system, is regulated by self-organisation and that all three types of stem cell divisions depend on both the number of stem cells and the number of differentiated cells. In particular, it is also possible that the rate of symmetric self-renewal increases when the number of differentiated cells is less than under normal conditions, which means that the number of stem cells increases and that more differentiated cells are produced than under normal conditions. Research by Gokoffski et al. (2011) on mice indicates that when there are less differentiated cells than under normal conditions, then the stem cell populations expand [5]. Indeed, this is the case for the model presented by Lander et. al [47]. Similarly, the model presented in paper 1 assumes that symmetric self-renewal is regulated by differentiated cells.

A mutant haematopoietic differentiated cell is likely to be washed out of the system before it becomes a cancer cell because haematopoietic differentiated cells do not in general self-renew. On the other hand, if a mutation occur in a haematopoietic stem cell, an evolutionary process with competition between the normal stem cells and the mutant stem cells might take place [1]. A critical aspect is whether the mutation affects how the mutant stem cells divide. That is, the population size of the mutants remains constant if they only divide asymmetrically. Since symmetric differentiation decreases the population size, the mutant stem cells have decreased fitness if the rate of this type of division increases. And finally, an increased rate of symmetric self-renewal increases the fitness of the mutant stem cells, because this type of division increases the population size.

9.2 Treatment of chronic myeloid with the tyrosine kinase inhibitors

Treatment of chronic myeloid with the *tyrosine kinase inhibitors* such as imatinib, represents a successful application of molecularly targeted anti-cancer therapy [48] (Druker et al., 1996, 2001; Kantarjian et al., 2002). These inhibitors reduce the fitness of Philadelphia-positive differentiated

cells. Nevertheless, the effect on Philadelphia-positive stem cells remain incompletely understood. For many patients, discontinuation of tyrosine kinase inhibitors results in a relapse of the disease within a few months [49]. Several explanations have been proposed to explain this phenomenon. For instance, tyrosine kinase inhibitors might not have any effect on the Philadelphia-positive stem cells [50], or the Philadelphia-positive stem cells can be susceptible to therapy when they are in an active state, but they are not be susceptible when they are in quiescent state [51]. Be that as it may, a small study involving 12 patients has shown that in some individuals the disease has remained undetected for two years after discontinuation of tyrosine kinase inhibitors. This raises the possibility that tyrosine kinase inhibitors have cured chronic myeloid leukaemia in these patients [52]. Furthermore, all studies indicate that the effect of tyrosine kinase inhibitors increases when treatment starts early. The model presented in paper 1 can explain these results: Tyrosine kinase inhibitors have most likely no direct effect on the Philadelphia-positive stem cells. Nevertheless, since differentiation regulates the proliferation of the stem cells, the tyrosine kinase inhibitors can change the population dynamics of the stem cells. More precisely, the following results observed in studies of chronic myeloid leukaemia patients treated with tyrosine kinase inhibitors, can be reproduced by the model:

1. The effect of tyrosine kinase inhibitors increases when treatment starts early in disease progression.

2. In some cases the treatment slows down the disease progression without erasing the Philadelphia-positive stem cells, which drive the disease.

3. In other cases the treatment reverses the disease progression and seems to erase the Philadelphia-positive stem cells stem cells.

If a model which assumes that stem cell activity is not regulated by the differentiated cells, is used, result 1–3 seem contradictory [50]. However, our model implies that these results can be explained by a negative feedback from the differentiated cells that regulate symmetric stem cell division [53].

9.3   Further work

The model presented in paper 1 is a simplification of the one presented in paper 2. Hence, the extensions discussed in Subsection 8.3 apply to both models.

Part III
Discussion

In 1953, James Watson and Francis Crick discovered that the DNA molecule exists in the form of a three-dimensional helix, and this brought new energy to the *paradigm of genetic determinism* [53], which claims that any characteristic of a living organism is directly proportional to the genes expressed in the DNA. This implies that the genetic rules that determine the behaviour of an organism, can be represented by linear equations, and hence, complex organisms, such as human beings, should have a much higher number of genes than a less complex organism, such plants.

Another significant milestone in molecular biology was the publication of the complete sequence of the human genome in 2003 [54]. The complete human genome is composed of over three billion bases and contains approximately 20,000 genes that code for proteins. This is much lower than earlier estimates of 80,000 to 140,000 and astonished the scientific community when revealed through human genome sequencing. Equally surprising was the finding that genomes of much simpler organisms contained a higher number of protein-coding genes than humans. For example, the mustard plant, *Arabidopsis thaliana*, which used as a model for studying plant genetics, has a genome size of 125 bases but a higher number of protein-coding genes than humans [55]. It is now clear that the size of a genome does not correspond with the number of protein-coding genes, and these do not determine the complexity of an organism.

As I make a case for in Subsection 2, the haematopoietic system is not regulated deterministically, but by self-organisation. That is, the body can regenerate blood cells to compensate for a loss of more than 15 percent of the circulating blood cells, and after a bone marrow transplantation, the haematopoietic stem cells, which are located in the bone marrow, can regenerate. Moreover, each day the body produces around a billion new blood cells. And since the human body contains about 37 billion blood cells, this means that most of the circulating blood cells are replaced each month [9].

Nevertheless, some biological traits are actually determined by a single gene. In contrast to the haematopoietic system, the fingers only grow out once. The Sonic Hedgehog gene is essential for normal limb development [56]. When a foetus, lying in the womb, develops fingers, the Sonic Hedgehog gene sends out a signal to shape the pattern of digits. Normally, five fingers are made. However, if a mutation occurs in the Sonic Hedgehog gene that turns down the effect of this gene, then fewer fingers are made, whereas if the mutation increases the effect of the Sonic Hedgehog gene, then each hand gets an extra finger.

Even though there are some examples where biological traits are determined by a given set of genes, the publication of the complete sequence of the human genome illuminates that that the paradigm of genetic determinism does not in general hold true, since complex organisms such as humans have a lower number of protein-coding genes than much simpler organisms such as the mustard plant, *Arabidopsis thaliana*. As illustrated in Subsection 1.1, nonlinear, high-dimensional and complex interaction between genes and regulation mechanisms can create new phenomena that cannot be explained by simply analysing the genetic code. Thus, the high complexity of humans compared to *Arabidopsis thaliana* might be explained by a higher complexity in the interaction between genes and regulation mechanisms.

Despite the fact that the paradigm of genetic determinism does not in general hold true, a new and generally accepted paradigm has not yet been established. There is almost no general information about nonlinear systems, except that they very often are chaotic and it is quite often impossible to find an exact solution. Nonlinear systems are often sensitive to starting conditions. For example, if you were given a list of numbers generated by the logistic map given in Equation (1) with starting condition $Y_0 = A$, you could not use this list to predict the outcome generated

by the logistic map with another starting condition, say $Y_0 = B \neq A$. Moreover, nonlinear systems quite often respond dramatically to changes in the feedback from the environment. For example, the differential equation in (4) is one of relatively few nonlinear problems that has a neat solution. Nevertheless, if the feedback from the environment is delayed, then the system can start to behave chaotically, as described by the logistic map given in (1). This illustrates that there is no set of mathematical tools that can be used on any nonlinear system. Thus, each nonlinear system must in general be analysed individually, and this requires more than basic mathematical knowledge. Moreover, small changes in a nonlinear system can lead to new behaviour.

10   New methods to analyse complex interactions

Genome editing, such as CRISPR/Cas9, machine learning and big data offer new ways to tackle the problems described above. Machine learning is explored in the previous two parts, whereas big data and CRISPR/Cas9 are briefly discussed in the following subsections.

10.1   Big data and the impact of the Human Genome Project

"The Human Genome Project led to a paradigm shift in the way science is conducted and data is shared," according to researcher in biotechnology, Rehma Chandaria [57]. The *Bermuda Principles* are rules for publication of DNA sequence data, and were proposed in 1996 by a group of international scientists who came together on Bermuda to discuss how sequence data from the Human Genome Project should be released. Challenging traditional practice in the sciences, which is to make experimental data available only after publication, the Bermuda Principles ensures that the data is immediately shared. The original Bermuda Principles were:

1. Automatic release of sequence assemblies larger than 1 kb (preferably within 24 hours).

2. Immediate publication of finished annotated sequences.

3. Aim to make the entire sequence freely available in the public domain for both research and development in order to maximise benefits to society.

The Bermuda Principles demonstrated how a global community of scientists could collectively produce and use data far more efficiently than a small group of scientists could.

The price of the Human Genome Project was 3 billion US dollars and it lasted for 13 years [55]. Today it is possible to sequence a human genome within days and it costs less than 1000 dollars. This big data requires that researchers from different specialities co-operate to process, analyse, store and utilise the vast quantities of data.

10.2   CRISPR/Cas9

CRISPR/Cas9 is one of the most effective gene-editing tools the word have seen, and originates from the immune system of bacteria [58]. With CRISPR/Cas9 the genome can be edited almost as easily as the text in a book [59]. CRISPR is an abbreviation of *Clustered regularly interspaced short palindromic repeats*, and is a family of DNA sequences in archaea and bacteria [60]. The sequences contain snippets of DNA from viruses that have attacked the prokaryote. Cas9, which is short for CRISPR-associated protein 9, is an RNA-guided DNA endonuclease enzyme associated with CRISPR. Cas9 uses the snippets to detect and destroy viruses with similar DNA [61], [62].

A study published on the 2th of August 2017, describe how a group of American and South Korean scientists for the first time successfully edited genes in human embryos to repair a common

and serious disease-causing mutation, producing apparently healthy embryos [63]. These results potentially open the door to preventing 10.000 disorders that are passed down the generations.

However, the main use of CRISPR/Cas9 is less spectacular. CRISPR/Cas9 is mostly used as a laboratory tool and is for instance used to study the expression of genes. Even though the complete sequence of the human genome was published 15 years ago, relatively little is known about how the genes are turned on and off and which traits they influence. Scientists can gain valuable knowledge by using CRISPR/Cas9 to switch on and off genes in a laboratory, for example they can study which genes must be turned off in a haematopoietic stem cell for it to become an immune cell.

Nevertheless, it is not possible to get the complete picture of how genes are regulated inside the body by just studying cells in a laboratory. For instance, it is still unclear why immune therapy can cure some cervical cancer patients but has no effect on others. To make hypotheses about why a medicine does not cure some patients, the researchers can make use of big data sets from patients by applying methods from machine learning. With solid hypotheses, the use of CRISPR/Cas9 becomes more efficient.

11   New medicines, new dilemmas

In Norway, lung cancer patients did not get the cost of immune therapy covered by the state until the end of 2016, whereas in Denmark the same patient group has been offered this treatment at public hospitals since September 2015 [64]. Similarly, patients with cervical cancer get immune therapy at public hospitals in Denmark, but not in Norway. In both cases, the state argued that the prize was too high because immune therapy does not cure all patients.

However, the price of each drug is only high in the beginning. When the effect of the medicine is more thoroughly documented and more patients can use it, the prize decreases. Large data sets from patients is of great value for scientific purposes, and there is no obvious reasons why Norway should contribute less to this research than other countries such as Denmark.

It is not only cancer treatment that creates new dilemmas. ®Spinraza is the first medicine that has any effect on *Spinal Muscular Atrophy (SMA)*, which is a rare neuromuscular disorder defined by progressive muscle wasting and loss of motor neurons. Approximately ten children are born with SMA each year in Norway. About 95 percent of the children with the most severe form of SMA die before they are two years old, whereas others can have a normal lifespan with a varying degree of disability. No previous drugs has any proven effect on SMA. Thus, Spinraza represents a big breakthrough. Moreover, the drug does not only slow down the progression of the disease – the patients improve motoric function and strengthen the muscles.

In 2017, Biogen, the company that developed Spinraza, let ten Norwegian children with SMA test the drug for free [65]. One of these children was a 11-month-old baby girl called Olivia. Olivia has the most severe form of SMA, and after starting the Spinraza treatment, she manages to sit, eat, hold a toy and turn around. This was impossible for Olivia before the treatment started. When the free trial ended, the government continued to pay for the Spinraza treatment for the ten children that had tried the drug for free. However, the government did not offer the remaining 40 children with SMA Spinraza treatment because the prize on the drug was too high. One of these children is Thea, a two year old girl, and like Olivia, Thea has the most severe form of SMA. Whereas other toddlers develop control of muscles, which enable walking, running, jumping and climbing, Thea managed to do less each day. Her parents lived in fear that she would die while the government negotiated with Biogen, until February 2017, when it was announced that the government would pay for Spinraza-treatment for all children with SMA. But the government does not cover this treatment for persons over 18 years of age. The reason

is that the effect of Spinraza on adults is not documented properly.

However, to evaluate the effect of Spinraza on adults, data from adult patients who actually use the drug is essential, and it seems reasonable that a rich country such as Norway should be one of the first to contribute to this research, whereas poorer countries might have no choice but to wait until the prize decreases and more is known about which patients the drug has any effect on.

One might argue that Biogen should give a discount on Spinraza-treatment for adult patients because the company can increase their final reckoning by gaining more information about the effect of the drug. However, this might increase the prize of the treatment for children.

At the end of the day, Biogen depends on the willingness of the owners to invest in the company. It is very expensive to develop new medicines. The company Bristol-Myers Squibb, which developed the immune therapy ®Opdivo, used 40 billion Norwegian Krones on research and development only in 2014. It is even more risky and potentially costly to try to develop a drug against disorders which no other drugs has any effect on, which was the case with SMA. Thus, it is a reasonable assumption that the owners of Biogen were willing to invest in the research that lead to the development of Spinraza, because there was a big payoff if they succeeded. And indeed, after the launch of Spinraza in 2016, the revenue of Biogen increased by four percent within a year. In the period from July to September 2017, the income from Spinraza is estimated to 2.1 billion Norwegian crones. And what is more, the Government Pension Fund of Norway owns one percentage of the company.

The fact that the Government Pension Fund of Norway is an owner of Biogen, does not mean that the Norwegian government should accept any price offer. However, to me it seems unethical if Norway waits until the prices drop to buy the drug, and at the same time earn money because other countries buy the drug when the price is high.

Moreover, the trust in the Norwegian government decreases when Norwegian patients die while the government is negotiating prices with the drug company, whereas the same patient group is offered the treatment at public hospitals in other countries.

Maybe the use of new medicines is an investment, both in research and in the welfare state.

References

[1] Martin A. Nowak and Karl Sigmund. Evolutionary dynamics of biological games. *Science*, 303(5659):793–799, 2004.

[2] C M Baum, I L Weissman, A S Tsukamoto, A M Buckle, and B Peault. Isolation of a candidate human hematopoietic stem-cell population. *Proceedings of the National Academy of Sciences*, 89(7):2804–2808, 1992.

[3] Sean J. Morrison and Irving L. Weissman. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. *Immunity*, 1(8):661 – 673, 1994.

[4] Tom Lenaerts, Jorge M. Pacheco, Arne Traulsen, and David Dingli. Tyrosine kinase inhibitor therapy can cure chronic myeloid leukemia without hitting leukemic stem cells. *Haematologica*, 2009.

[5] Kimberly K. Gokoffski, Hsiao-Huei Wu, Crestina L. Beites, Joon Kim, Euiseok J. Kim, Martin M. Matzuk, Jane E. Johnson, Arthur D. Lander, and Anne L. Calof. Activin and gdf11 collaborate in feedback control of neuroepithelial stem cell proliferation and fate. *Development*, 2011.

[6] J. Nathan Kutz. Deep learning in fluid dynamics. *Journal of Fluid Mechanics*, 814:1–4, 2017.

[7] I. Stewart. *Concepts of Modern Mathematics*. Dover Books on Mathematics. Dover Publications, 2012.

[8] M.W. Hirsch, S. Smale, and R.L. Devaney. *Differential Equations, Dynamical Systems, and an Introduction to Chaos*. Academic Press. Academic Press, 2013.

[9] C. Cooper and C.S. Cooper. *Blood: A Very Short Introduction*. Very Short Introductions Series. Oxford University Press, 2016.

[10] Tong Yin and Linheng Li. The stem cell niches in bone. *The Journal of Clinical Investigation*, 116(5):1195–1201, 5 2006.

[11] Chun-Chun Cheng, Yen-Hua Lee, Shau-Ping Lin, Wei-Chun HuangFu, and I-Hsuan Liu. Cell-autonomous heparanase modulates self-renewal and migration in bone marrow-derived mesenchymal stem cells. *Journal of Biomedical Science*, 21(1):21, Mar 2014.

[12] Sean J. Morrison and Judith Kimble. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature*, 441(15):1068–1074, 2006.

[13] Yukiko M Yamashita, D Leanne Jones, and Margaret T Fuller. Orientation of asymmetric stem cell division by the apc tumor suppressor and centrosome. *Science (New York, N.Y.)*, 301(5639):1547–50, Sep 2003.

[14] H E Wichmann, M Loeffler, and S Schmitz. A concept of hemopoietic regulation and its biomathematical realization. *Blood cells*, 14(2-3):411–29, 1988.

[15] K. Shortman and S.H. Naik. Steady-state and inflammatory dendritic-cell development. *Nature Reviews Immunology*, 7(1):19–30, 2006.

[16] J L Abkowitz, M L Linenberger, M A Newton, G H Shelton, R L Ott, and P Guttorp. Evidence for the maintenance of hematopoiesis in a large animal by the sequential activation of stem-cell clones. *Proceedings of the National Academy of Sciences of the United States of America*, 87(22):9062–6, Nov 1990.

[17] J L Abkowitz, R M Ott, R D Holly, and J W Adamson. Clonal evolution following chemotherapy-induced stem cell depletion in cats heterozygous for glucose-6-phosphate dehydrogenase. *Blood*, 71(6):1687–92, Jun 1988.

[18] J L Abkowitz, S N Catlin, and P Guttorp. Evidence that hematopoiesis may be a stochastic process in vivo. *Nature medicine*, 2(2):190–7, Feb 1996.

[19] David Dingli, Arne Traulsen, Franziska Michor, and Fransizka Michor. (a)symmetric stem cell replication and cancer. *PLoS computational biology*, 3(3):e53, Mar 2007.

[20] David J Araten, David W Golde, Rong H Zhang, Howard T Thaler, Lucia Gargiulo, Rosario Notaro, and Lucio Luzzatto. A quantitative measurement of the human somatic mutation rate. *Cancer research*, 65(18):8111–7, Sep 2005.

[21] B. Vogelstein and K.W. Kinzler. Cancer genes and the pathways they control. *Nature medicine*, 10(8):789–799, 2004.

[22] N. L. Komarova and P. Cheng. Epithelial tissue architecture protects against cancer. *Mathematical biosciences*, 200(1):90–117, Mar 2006.

[23] D. Wodarz and N. L. Komarova. *Computational Biology Of Cancer: Lecture Notes And Mathematical Modeling*. World Scientific Publishing Company, 2005.

[24] L. Shahriyari and N.L. Komarova. Symmetric vs. asymmetric stem cell divisions: an adaptation against cancer? *PloS one*, 8(10):e76195, 2013.

[25] J.M. Smith. The theory of games and the evolution of animal conflicts. *Journal of theoretical biology*, 47(1):209–221, 1974.

[26] Guillaume Lambert, Saurabh Vyawahare, and Robert H Austin. Bacteria and game theory: the rise and fall of cooperation in spatially heterogeneous environments. *Interface focus*, 4(4):20140029, Aug 2014.

[27] John Nash. Non-cooperative games. *Annals of mathematics*, pages 286–295, 1951.

[28] Raheleh Rahbari, Tom Sheahan, Vasileios Modes, Pam Collier, Catriona Macfarlane, and Richard M Badge. A novel l1 retrotransposon marker for hela cell line identification. *BioTechniques*, 46(4):277–84, Apr 2009.

[29] David H Hubel and Torsten N Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology*, 160(1):106, 1962.

[30] K Fukushima. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, 36(4):193–202, 1980.

[31] P E Turner and L Chao. Prisoner's dilemma in an rna virus. *Nature*, 398(6726):441–3, Apr 1999.

[32] M A Nowak and K Sigmund. Phage-lift for game theory. *Nature*, 398(6726):367–8, Apr 1999.

[33] Marco Archetti. Evolutionary dynamics of the warburg effect: glycolysis as a collective action problem among cancer cells. *Journal of theoretical biology*, 341:1–8, Jan 2014.

[34] Karl Marx and Friedrich Engels. *Manifesto of the communist party*. CH Kerr and Company, 1906.

[35] Marthe Måløy. Frp bør legge innvandrings- og integreringsministeren på bordet. *Minerva*, December 2017.

[36] Anurag Agrawal and Deepak Jaiswal. When machine learning meets ai and game theory. 2012.

[37] Steven Pinker, Martin A Nowak, and James J Lee. The logic of indirect speech. *Proceedings of the National Academy of Sciences*, 105(3):833–838, 2008.

[38] Markus Loeffler and Ingo Roeder. Tissue stem cells: definition, plasticity, heterogeneity, self-organization and models–a conceptual approach. *Cells, tissues, organs*, 171(1):8–26, 2002.

[39] P Zhang, G Behre, J Pan, A Iwama, N Wara-Aswapati, H S Radomska, P E Auron, D G Tenen, and Z Sun. Negative cross-talk between hematopoietic regulators: Gata proteins repress pu.1. *Proceedings of the National Academy of Sciences of the United States of America*, 96(15):8705–10, Jul 1999.

[40] Timothy R Brazelton, Fabio MV Rossi, Gilmor I Keshet, and Helen M Blau. From marrow to brain: expression of neuronal phenotypes in adult mice. *Science*, 290(5497):1775–1779, 2000.

[41] M Z Ratajczak, M Kucia, R Reca, M Majka, A Janowska-Wieczorek, and J Ratajczak. Stem cell plasticity revisited: Cxcr4-positive cells expressing mrna for early muscle, liver and neural cells 'hide out' in the bone marrow. *Leukemia*, 18(1):29–40, Jan 2004.

[42] J L Abkowitz, M L Linenberger, M Persik, M A Newton, and P Guttorp. Behavior of feline hematopoietic stem cells years after busulfan exposure. *Blood*, 82(7):2096–103, Oct 1993.

[43] C.H. Waddington. *The Strategy of the Genes*. Routledge, 1957.

[44] Chikara Furusawa and Kunihiko Kaneko. Chaotic expression dynamics implies pluripotency: when theory and experiment meet. *Biology direct*, 4:17, 2009.

[45] Ken Shortman and Shalin H Naik. Steady-state and inflammatory dendritic-cell development. *Nature reviews. Immunology*, 7(1):19–30, Jan 2007.

[46] Dominik Wodarz. Stem cell regulation and the development of blast crisis in chronic myeloid leukemia: Implications for the outcome of imatinib treatment and discontinuation. *Medical Hypotheses*, 70(1):128 – 136, 2008.

[47] Arthur D Lander, Kimberly K Gokoffski, Frederic YM Wan, Qing Nie, and Anne L Calof. Cell lineages and the logic of proliferative control. *PLoS biology*, 7(1):e1000015, 2009.

[48] B J Druker, S Tamura, E Buchdunger, S Ohno, G M Segal, S Fanning, J Zimmermann, and N B Lydon. Effects of a selective inhibitor of the abl tyrosine kinase on the growth of bcr-abl positive cells. *Nature medicine*, 2(5):561–6, May 1996.

[49] Jorge Cortes, Susan O'Brien, and Hagop Kantarjian. Discontinuation of imatinib therapy after achieving a molecular response. *Blood*, 104(7):2204–5, Oct 2004.

[50] Franziska Michor, Timothy P Hughes, Yoh Iwasa, Susan Branford, Neil P Shah, Charles L Sawyers, and Martin A Nowak. Dynamics of chronic myeloid leukaemia. *Nature*, 435(7046):1267–70, Jun 2005.

[51] Ingo Roeder, Matthias Horn, Ingmar Glauche, Andreas Hochhaus, Martin C Mueller, and Markus Loeffler. Dynamic modeling of imatinib-treated chronic myeloid leukemia: functional insights and clinical implications. *Nature medicine*, 12(10):1181–4, Oct 2006.

[52] Philippe Rousselot, Francoise Huguet, Delphine Rea, Laurence Legros, Jean Michel Cayuela, Odile Maarek, Odile Blanchet, Gerald Marit, Eliane Gluckman, Josy Reiffers, Martine Gardembas, and François-Xavier Mahon. Imatinib mesylate discontinuation in patients with chronic myelogenous leukemia in complete molecular remission for more than 2 years. *Blood*, 109(1):58–60, 2007.

[53] Richard C Strohman. The coming kuhnian revolution in biology. *Nature biotechnology*, 15(3):194–200, 1997.

28

[54] L Goodman. The human genome project aims for 2003. *Genome research*, 8(10):997–9, Oct 1998.

[55] Leroy Hood and Lee Rowen. The human genome project: big science transforms biology and medicine. *Genome medicine*, 5(9):79, 2013.

[56] P C Edwards, S Ruggiero, J Fantasia, R Burakoff, S M Moorji, E Paric, P Razzano, D A Grande, and J M Mason. Sonic hedgehog gene-enhanced tissue engineering for bone regeneration. *Gene therapy*, 12(1):75–86, Jan 2005.

[57] Julie Gould. Big data: The impact of the human genome project. *NatureJobs Blog*, October 2015.

[58] Lauren R Polstein and Charles A Gersbach. A light-inducible crispr-cas9 system for control of endogenous gene activation. *Nature chemical biology*, 11(3):198–200, 2015.

[59] Marthe Måløy. Crispr kan revolusjonere behandlingen av kreft og genetiske sykdommer. *Minerva*, August 2017.

[60] Isabel Lokody. Genetic therapies: Correcting genetic defects with crispr-cas9. *Nature Reviews Genetics*, 2013.

[61] Seung Woo Cho, Sojung Kim, Jong Min Kim, and Jin-Soo Kim. Targeted genome engineering in human cells with the cas9 rna-guided endonuclease. *Nature biotechnology*, 2013.

[62] Prashant Mali, Kevin M Esvelt, and George M Church. Cas9 as a versatile tool for engineering biology. *Nature methods*, 10(10):957–963, 2013.

[63] Hong Ma, Nuria Marti-Gutierrez, Sang-Wook Park, Jun Wu, Yeonmi Lee, Keiichiro Suzuki, Amy Koski, Dongmei Ji, Tomonari Hayama, Riffat Ahmed, Hayley Darby, Crystal Van Dyken, Ying Li, Eunju Kang, A.-Reum Park, Daesik Kim, Sang-Tae Kim, Jianhui Gong, Ying Gu, Xun Xu, David Battaglia, Sacha A. Krieg, David M. Lee, Diana H. Wu, Don P. Wolf, Stephen B. Heitner, Juan Carlos Izpisua Belmonte, Paula Amato, Jin-Soo Kim, Sanjiv Kaul, and Shoukhrat Mitalipov. Correction of a pathogenic gene mutation in human embryos. *Nature*, 548(7):413–419, 2017.

[64] Marthe Måløy. Immunterapi må tilbys til alle. *Minerva*, September 2016.

[65] Marthe Måløy. Helse-norge bør ta i bruk nye og dyre medisiner. *Minerva*, November 2017.

Stem cell regulation: Implications when differentiated cells regulate symmetric stem cell division

Marte Rørvik Høyem [a,*], Frode Måløy [b], Per Jakobsen [a], Bjørn Olav Brandsdal [c]

[a] Department of Mathematics and Statistics, University of Tromsø, Norway
[b] Department of Computer Science, University of Stavanger, Norway
[c] Department of Chemistry, University of Tromsø, Norway

HIGHLIGHTS

- Differentiated cells (DCs) might regulate symmetric stem cell (SC) division.
- This implies that changes in the dynamics of DCs can affect the fitness of SCs.
- Tyrosine kinase inhibitors (TKIs) are used to treat chronic myeloid leukaemia (CML).
- TKIs increase the death rate of DCs, but have most likely no direct effect on SCs.
- TKIs might have an indirect effect on SCs if DCs regulate symmetric SC division.

ARTICLE INFO

ABSTRACT

We use a mathematical model to show that if symmetric stem cell division is regulated by differentiated cells, then changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells. More precisely, the relative fitness of the stem cells can be affected by modifying the death rate of the differentiated cells. This result is interesting because stem cells are less sensitive than differentiated cells to environmental factors, such as medical therapy. Our result implies that stem cells can be manipulated indirectly by medical treatments that target the differentiated cells.

1. Introduction

Most tissues of the body go through continuous cell turnover du to apoptosis. This cell turnover can also give tissues the ability to self-repair after injury. In general, tissues are maintained by a small group of slowly replicating cells with the capacity to both self-renew and generate differentiated progeny required by a given tissue (Morrison et al., 1997; Reya et al., 2001). Cells that have these two capabilities are called *stem cells*. Differentiated cells perform their function and eventually die – they go through a number of divisions, obtaining various stages of differentiation, until the fully differentiated cells stop dividing (Donohue et al., 1958; Cronkite and Fliedner, 1964; Ogawa, 1993). Although it seems reasonable to propose that all tissues arise from tissue-specific stem cells, rigorous identification and isolation of these

stem cells have only been accomplished in a few instances. For example, *haematopoietic stem cells* have been isolated and shown to be responsible for the generation and regeneration of the blood-forming system and the immune system, called the *haematopoietic system* (Baum et al., 1992; Morrison and Weissman, 1994). The haematopoietic stem cells are located within the bone marrow and segregated among different bones throughout the body. Like several other models (Loeffler and Wichmann, 1980; Agur et al., 2002; Østby et al., 2003; Østby and Winther, 2004; Coiljn and Mackey, 2005; Adimy et al., 2006; Dingli and Michor, 2006; Dingli et al., 2007a,b; Wodarz, 2008; Marciniak-Czochra et al., 2009; Stiehl and Marciniak-Czochra, 2012; Lenaerts et al., 2010; Manesso et al., 2013), the model presented in this paper is inspired by the haematopoietic system. However, it applies to all other tissues that have similar architecture.

An important aspect, related to self-renewal and generation of differentiated cells, is the fate of the two daughter cells when a stem cell divides (Dingli et al., 2007b; Morrison and Kimble, 2006; Yamashita et al., 2003). *Symmetric division* is defined as generation

* Corresponding author.
  E-mail address: marte.rorvik.hoyem@gmail.com (M.R. Høyem).

of daughter cells destined to acquire the same fate. In this paper, symmetric stem cell division is defined as *symmetric self-renewal* if both daughter cells are stem cells and *symmetric differentiation* if both daughter cells are differentiated. In the former case the number of stem cells increases by one, whereas in the latter case the number of stem cells decreases by one. Stem cells can rely completely on symmetric division. On the other hand, if one daughter cell has stem cell identity and the other daughter cell starts to differentiate, it is called an *asymmetric stem cell division*. This type of division is particularly attractive because the stem cells manage to both self-renew and produce differentiated cells with a single division (Yamashita et al., 2003). However, a disadvantage of asymmetric stem cell division is that it leaves stem cells unable to expand in number. Serial haematopoietic transplantation supports the existence of all three types of divisions (McKenzie et al., 2006).

1.1. Stem cell niche

Since the number of stem cells is much smaller than the number of differentiated cells, the stem cells must be protected and tightly regulated. As discussed by Gentry and Jackson (2013), the *stem cell niche*, which is the restricted region in an organ that supports stem cell behaviour, may be crucial in both aspects (Fuchs et al., 2004; Nikolova et al., 2006; Yin and Li, 2006; Simons and Cleavers, 2011). The niche is composed of both localised signalling cells and an extracellular matrix that control stem cell fate. However, relatively little is known about the exact behaviour of most types of stem cells, and one of the reasons for this is that it is not possible to reconstruct niches scientifically, which makes it difficult to maintain stem cells *in vitro*, because signals from the niche affects stem cell survival, self-renewal, and differentiation.

Germline stem cells are unique stem cells in that they are solely dedicated to reproduction and transmission of genetic information from generation to generation. Through the use of genetic techniques in *Drosophila germline stem cells*, exciting progress has been made in understanding molecular mechanisms underlying interactions between stem cells and stem cell niches (Morrison and Kimble, 2006; Yamashita et al., 2003; Wong et al., 2005). The knowledge gained from studying *Drosophila* germline stem cells has provided an intellectual framework for defining the stem cell niche and molecular regulatory mechanisms for other adult stem cells, such as the haematopoietic stem cells.

The number of cells in a given tissue is approximately constant under normal conditions. It is generally believed that the number of stem cells is approximately constant under normal conditions, and that they differentiate and self-renew at relatively constant rates to replace mature cells and to keep the stem cell number at a certain normal level (Loeffler et al., 1988; Shortman and Naik, 2009). One strategy which stem cells can accomplish these two tasks is asymmetric stem cell division. A classical example of asymmetric division is provided by *Drosophila* germline stem cells. The outcome of a *Drosophila* germline stem cell division depends on the spindle orientation relative to the Hub cells in the stem cell niche, and results from the unequal distribution of intracellular regulators and extracellular (Hub-derived) signals between daughter cells during mitosis (Morrison and Kimble, 2006; Yamashita et al., 2003; Wong et al., 2005). The result is that when a *Drosophila* germline stem cell divides, one daughter remains in the stem cell niche and retains stem cell identity, and one daughter is left outside the stem cell niche and begins to differentiate. Research on *Drosophila* germline stem cells has provided a clear-cut example of how the stem cell niche promotes stem cell maintenance. Similarly, the haematopoietic microenvironment in the bone marrow also plays an important role in the regulation of haematopoietic

stem cell organisation (Lemischka, 1997; Bertolini et al., 1997; Aiuti et al., 1998; Thiemann et al., 1998). Self-renewal depends on local growth conditions, namely, on the direct contact between stem cells and stroma cells (Wineman et al., 1996; Verfaillie, 1998; Koller et al., 1999). However, there are no *in vivo* experiments that reveal exactly how proliferation of haematopoietic stem cells is regulated. Thus, it is not clear whether these cells divide asymmetrically or symmetrically under normal conditions. Serial haematopoietic transplantation indicates that both types of divisions occur under steady state (McKenzie et al., 2006). As discussed later in Section 1.3, theoretical work by Shahriyari and Komarova (2013) and McHale and Lander (2014) illustrate that the symmetric stem cell division can protect against cancer, and this indicates that stem cells divide symmetrically.

Although the number of haematopoietic stem cells remains nearly constant under normal conditions, they can expand rapidly in response to injury to the bone marrow, such as stem cell transplantation (McKenzie et al., 2006). This means that asymmetric stem cell division cannot be the complete story, because it leaves stem cells unable to expand in number. Since the number of stem cells increases with one after symmetric self-renewal, it is likely that the rate of such divisions depends on the number of stem cells, since the haematopoietic stem cells can regenerate after tissue damage. Indeed, *Drosophila* germline stem cells, which normally divide asymmetrically, can be induced to self-renew symmetrically to regenerate an additional stem cell after an experimental manipulation in which one stem cell is removed from the stem cell niche (Morrison and Kimble, 2006; Yamashita et al., 2003; Wong et al., 2005).

1.2. Extracellular regulation

Extracellular signalling molecules regulate the dynamics of cell proliferation and differentiation. However, the precise nature of these processes are in general not known (Layton et al., 1989; Aglietta et al., 1989; Metcalf, 2008; Fried, 2009). An example of extracellular signalling molecules is the *haematopoietic cytokines* that control the production of haematopoietic cells. Each of these cytokines has multiple actions mediated by receptors that can initiate various responses – survival, proliferation, differentiation, maturation, and functional activation. Individual haematopoietic cytokines can either regulate one specific lineage or multiple lineages (Metcalf, 2008). Moreover, for some haematopoietic cell types, such as stem cells or megakaryocyte progenitors, the simultaneous action of multiple cytokines is required for proliferative responses. Unlike other extracellular signalling molecules, like hormones, that have a limited, or single, organ source, the haematopoietic cytokines have many tissue sources, e.g. kidney, liver, lung, muscle and membrane-displayed factors on local stromal cells (Aglietta et al., 1989; Metcalf, 2008). This is one of the reasons why it is difficult to establish the precise source of a haematopoietic cytokine in any particular situation and to predict its ultimate fate. Results from theoretical work regarding the haematopoietic system (Wodarz, 2008) and crypt cells (Potten and Loeffler, 1990) indicate that changes in stem cell number and their cyclic activity are associated with changes in the demand of the mature cell stages. Marciniak-Czochra et al. (2009) designed a six-compartment model to test different hypotheses concerning regulation of self-renewal and differentiation by a feedback signalling factor. Since the precise nature of how extracellular signalling molecules such as cytokines control proliferation and differentiation is still unknown, Marciniak-Czochra et al. assume that the signal intensity is

$$s = \frac{1}{1 + kC_6},\qquad(1)$$

where k is a constant and C_6 is the number of mature cells. Marciniak-Czochra et al. compare three different cases:

1. Only proliferation rates are regulated by feedback signals.
2. Only differentiation rates are regulated by feedback.
3. Both proliferation and differentiation rates are regulated by feedback.

They show that the best results are obtained when both proliferation and differentiation rates are regulated by feedback.

Lander et al. (2009) investigate how secreted negative feedback factors may be used to control the output of multistage cell lineages, as exemplified by the actions of GDF11 and activin in a self-renewing neural tissue, the mammalian olfactory epithelium. Similar to Marciniak-Czochra et al. (2009), Lander et al. find that two feedback loops are in general better than one. That is, when feedback loops are added, good control (robustness, stability, low progenitor load, and fast regeneration from a variety of conditions) is found over an increasing fraction of the parameter space. Lander et al. discuss different strategies for how stem cell self-renewal and generation of differentiated progeny can be regulated by negative feedback from differentiated cells. The first scenario is that asymmetric stem cell division is regulated by differentiated cells. In this case, the rate of asymmetric stem cell division increases when the number of differentiated cells is less than under normal conditions, which means that more differentiated cells are produced while the number of stem cells remains constant. On the other hand, it is also possible that symmetric stem cell division is regulated by differentiated cells. In this case, the rate of symmetric self-renewal increases when the number of differentiated cells is less than under normal conditions, which means that the number of stem cells increases and that more differentiated cells are produced than under normal conditions. Since research by Gokoffski et al. (2011) on mice indicates that stem cell populations expand when there are less differentiated cells than under normal conditions, Lander et al. consider the latter case in their model. Similarly, in the model presented in this paper, symmetric stem cell division is regulated by differentiated cells.

Manesso et al. (2013) propose a model where mild perturbations of differentiated cells do not influence the stem cell dynamics – steady state is re-established by increasing the self-renewal rate of the differentiated cells. After a critical threshold level is reached in terms of cell numbers, a second response is activated by increasing the commitment rates from the directly upstream cell types. The second response can influence the stem cell dynamics. The model was able to recapitulate the fundamental steady-state features of haematopoiesis and simulate the re-establishment of steady-state conditions after haemorrhage and bone marrow transplantation in adult mice. However, as discussed in Section 1.3, increasing the self-renewal rate of the differentiated cells can increase the risk of cancer. This might be one of the reasons why several other models, like the ones proposed by Loeffler and Wichmann (1980), Østby and Winther (2004), Wodarz (2008), Gentry and Jackson (2013) and Rodriguez-Brenes et al. (2013), assume that stem cell self-renewal and differentiation are regulated by a negative feedback from more mature cells. In particular, the models proposed by Gentry et al. and Wodarz include both extrinsic and intrinsic chemical signalling and interaction with the niche to control self-renewal, and this novel feature is also investigated in this paper. However, unlike our model, Wodarz' model assumes that when there are only healthy cells in the system, the rate of symmetric stem cell division depends only on the number of stem cells and the rate of asymmetric stem cell division depends only on the number of differentiated cells. Thus, according to Wodarz' model, changes in the population dynamics of the differentiated cells do not influence the dynamics of the stem cell population when there are only healthy cells in the system. On the contrary, if symmetric stem cell division is regulated by the differentiated cells, then changes in population dynamics of the differentiated cells, such as increased death rate, can influence the dynamics of the stem cell population. In Section 2, we investigate the implications when the rate of symmetric self-renewal depends on both the number of stem cells and the number of differentiated cells.

1.3. Mutations and stem-cell-driven tumours

Genetic changes called *mutations* can occur in any cell that divides (Araten et al., 2005). Even though most mutations are harmless to the body, progressive accumulation of mutations can lead to cancer (Vogelstein and Kinzler, 2004). Indeed, results from theoretical work regarding stem cell self-renewal and differentiation indicate that the tissue architecture, where only a small number of stem cells have the ability to self-renew, has evolved to minimise the risk of malignant transformations (Dingli et al., 2007b; Wodarz and Komarova, 2005; Komarova and Cheng, 2006). That is, if a mutation occurs in a differentiated cell, it is likely to be washed out of the system before it becomes a cancer phenotype, because differentiated cells do not self-renew. On the other hand, mutation in a stem cell can generate a different type of stem cell, denoted *mutant stem cell*. This can lead to an evolutionary process with competition between the mutant stem cells and the normal stem cells (Nowak, 2006a; Dingli et al., 2010). A critical aspect is the fate of the daughter cells when the stem cells divide (Morrison and Kimble, 2006). The model proposed by Dingli et al. (2007b) shows that if the mutant stem cells divide only asymmetrically, their population size remains constant. A high probability of symmetric self-renewal increases the fitness of the stem cells, because this type of division increases the population size. Symmetric differentiation, on the other hand, decreases the population size. Thus, stem cells that differentiate symmetrically with a high probability have decreased fitness.

Shahriyari and Komarova (2013) and McHale and Lander (2014) illustrate that symmetrically dividing cells might delay double-hit mutant production compared to an equivalent system with asymmetrically dividing stem cells. More precisely, if stem cells only divide asymmetrically, then a mutation acquired in a stem cell will remain in the system indefinitely, and it is only a matter of time before the second mutation occurs. On the contrary, a mutant stem cell generated in a symmetric division has a less certain fate – half of the lineages will differentiate out after the very first division and only $1/K$ of all lineages will expand to size K. Thus, that the uncertainty of the fate of single mutant stem cells can be the reason for the statistically longer time it takes for the symmetrically dividing stem cell model to produce a double-hit mutant.

Rodriguez-Brenes et al. (2011) propose a model that illustrates that a key event in the development of cancer is the escape from feedback loops. In a genetically heterogeneous population, selection favours cells with advantageous traits (Wodarz and Komarova, 2005; Nowak, 2006b). Since cancer is a product of somatic evolution, it is important to investigate how mutants that originally appear in very small numbers are able to invade a cell population that is initially at dynamic equilibrium (Mangel and Bonsall, 2008). Rodriguez-Brenes et al. use computational models that are applied to experimental data, to study the evolutionary dynamics of feedback escape. Their model predicts different patterns of emerging tumour growth that fit previously published experimental data that describe tumour growth dynamics *in vitro* and *in vivo* (Rozenblum et al., 1997; Massagué, 2000, 2001; Derynck et al., 2001; Woodford-Richens et al., 2001; Wu et al., 2008). Of particular interest are non-standard growth patterns, both predicted by the model and found in published experimental

data, which indicates that feedback regulatory mechanisms are still partly at work in growing tumours (Rozenblum et al., 1997; Massagué, 2000, 2001; Derynck et al., 2001; Woodford-Richens et al., 2001; Wu et al., 2008). This gives rise to the notion that tumours not only retain some of the architectural aspects of the underlying healthy tissue, but also some of the regulatory mechanisms.

Stiehl and Marciniak-Czochra (2012) present a model of cancer cell dynamics where it is assumed that the leukemic cell population consists of an ordered sequence of cell statuses similar to the healthy haematopoietic cell lines. Moreover, it is assumed that leukemic stem cells are stimulated by the same cytokines as healthy stem cells. Similar to the models presented in this paper, a negative feedback function regulates self-renewal. However, unlike our model, the feedback function in the model proposed by Stiehl et al. only depends on the fully mature cells, namely,

$$s(t) = \frac{1}{1 + k^c c_m + k^l l_m},$$

where c_m and l_m are the number of fully mature healthy cells and fully mature leukemic cells, respectively, and k^c and k^l are constants. Moreover, the feedback that regulates healthy cells and cancer cells is the same – the difference between leukemic cells and healthy cells is captured by different constants associated with rates of self-renewal, differentiation and cell death. On the contrary, the model investigated in Section 2.1 assumes that the only difference between healthy cells and leukemic cells is the strength at which they regulate self-renewal.

2. Mathematical models

In Sections 2.1 and 2.2, we explore a simple model that only considers two types of cells, namely, stem cells and differentiated cells. This model provides analytic results and captures the basic idea of this paper, which is that changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells when symmetric stem cell division is regulated by differentiated cells. An extension of the model, which includes various stages of differentiation, is presented in Section 2.3. The extended model is explored numerically, since it is too complex to analyse analytically. The numerical analysis shows that the analytic results obtained from the simple model also apply to the extended model.

2.1. Model with two layers of differentiation

The basic model considers two layers of the differentiation hierarchy: Stem cells have the potential for indefinite self-renewal and to give rise to differentiated cells. The differentiated cells are the cells without stem cell characteristics. Let x_s denote the number of stem cells and x_d the number of differentiated cells. As discussed in the introduction, signalling molecules such as cytokines and interaction with the stem cell niche control stem cell behaviour, but the precise nature of this regulation is still unknown (Fuchs et al., 2004; Nikolova et al., 2006; Yin and Li, 2006; Simons and Cleavers, 2011; Layton et al., 1989; Aglietta et al., 1989; Metcalf, 2008; Fried, 2009). We assume that the signalling intensity is approximately

$$\Psi = \exp(-\theta x_s - \gamma x_d),$$

(2)

where θ and γ are positive constants. This function captures the fact that the secretion of cytokines is very fast in comparison to cell proliferation and differentiation (Metcalf, 2008). Moreover, the signal intensity reaches its maximum under complete absence of cells, and it decreases exponentially towards zero as the number of

cells increases. In the simple model presented in this subsection, only symmetric self-renewal is regulated by the feedback signals. It is assumed that the stem cells produce immature differentiated cells by asymmetric division and symmetric differentiation at constant rates, g and d_0, respectively, and die at constant rate, d_1. The differentiated cells go though a number of divisions, obtaining various stages of differentiation, until the fully mature cells stop dividing. This differentiation process is investigated in more details in Section 2.3. Here we simply assume that the process occurs at constant rate, f, which means that differentiated cells are generated at rate $P = (2d_0 + g)f$. The differentiated cells die at constant rate Q. Hence, the model is given by the following set of ordinary differential equations:

$$\frac{dx_s}{dt} = (r\Psi - d)x_s,$$

(3)

$$\frac{dx_d}{dt} = Px_s - Qx_d,$$

(4)

where $d = d_0 + d_1$ and r is a positive constant. The system has two equilibrium solutions, namely,

$$(x_s^{0*}, x_d^{0*}) = (0, 0),$$

(5)

$$(x_s^*, x_d^*) = \left(\frac{1}{\theta + \frac{P}{Q}\gamma} \ln\left(\frac{r}{d}\right), \frac{P}{Q} \frac{1}{\theta + \frac{P}{Q}\gamma} \ln\left(\frac{r}{d}\right) \right).$$

(6)

We only consider the case when $r > d$, which means that (x_s^*, x_d^*) is stable, whereas (x_s^{0*}, x_d^{0*}) is unstable (Appendix B). The former equilibrium solution describes the system under normal conditions. Note that the number of differentiated cells is much larger than the number of stem cells, and that the death rate of the differentiated cells, Q, is much higher than the rate at which the stem cells die and differentiate, d. The pseudo-steady state hypothesis is that the population dynamics of the differentiated cells occurs at a very high rate compared with the stem cell population dynamics. Hence, it is assumed that the differentiated cells are always in equilibrium. Mathematically, we use the approximation $\frac{dx_d}{dt}\frac{1}{Q} = \frac{P}{Q}x_s - x_d \approx 0$ to obtain $x_d \approx \frac{P}{Q}x_s$. Thus, the population dynamics of the stem cells is approximately described by the following differential equation:

$$\frac{dx_s}{dt} = \left(r\exp\left(-\left(\theta + \gamma\frac{P}{Q} \right)x_s \right) - d \right)x_s.$$

(7)

Starting with any population size (x_s^0, x_d^0), where $x_s^0 > 0$, the system given in Eqs. (3) and (4), converges towards (x_s^*, x_d^*) (Appendix B). Fig. 1 shows an example where the whole system is regenerated, starting with a single stem cell. For comparison, the figure also shows the regeneration in the absence of feedback from differentiated cells (dashed line). From Fig. 1, we can see that feedback from differentiated cells enables the system to regenerate faster.

Changes in the population dynamics of the differentiated cells lead to changes in the rate of symmetric stem cell division, since the function Ψ is dependent on the variable x_d. The factors that influence the population dynamics of the differentiated cells are included in the model by modifying the death rate to $\hat{Q} \neq Q$. If $\hat{Q} > Q$, then the number of differentiated cells starts decreasing, whereas if $\hat{Q} < Q$, then the number of differentiated cells starts increasing. This triggers changes in the function Ψ as follows: Ψ increases if the number of differentiated cells decreases, and Ψ decreases if the number of differentiated cells increases. The number of stem cells converges towards the following steady
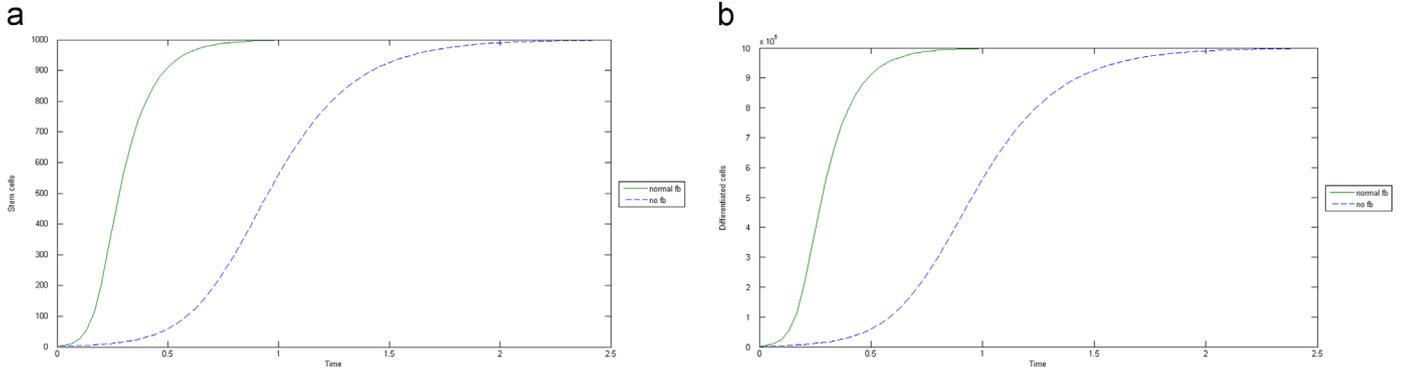
a

b

Fig. 1. *Regeneration.* The whole system is regenerated, starting with a single stem cell. The green line is an example where stem cell self-renewal is regulated by both stem cells and differentiated cells. The blue, dashed line is the regeneration with the same feedback from the stem cells, but no feedback from the differentiated cells. Both examples have the following parameter sizes: $\theta = 10^{-3}$, $d = 0.1353$, $P = 10^6$, $Q = 10^3$. In addition, the example with normal feedback has $\gamma = 10^{-6}$, $r = 1$, whereas the example without feedback from differentiated cells has $\gamma = 0$ and $r = \exp\left(-10^{-6} \times y_d^*\right) = \exp(-1)$. (a) and (b) display the stem cells and the differentiated cells, respectively.

a

b
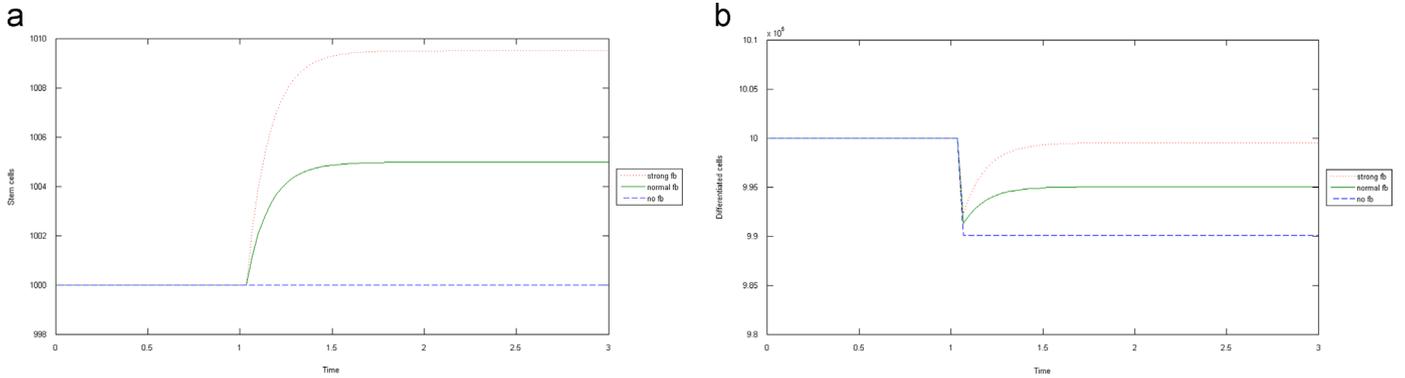
Fig. 2. *Increased death rate of the differentiated cells.* If the death rate of the differentiated cells increases, then the number of differentiated cells starts decreasing, leading to an increased self-renewal rate, resulting in an increased number of stem cells, and ultimately the number of differentiated cells increases. The red, dotted line shows an example where the feedback from the differentiated cells is much stronger than the feedback from the stem cells. In this case, the number of differentiated cells remains approximately the same as under normal conditions when the death rate of the differentiated cells changes. The blue, dashed line is an example where stem cell self-renewal is not regulated by feedback from differentiated cells, and the number of stem cells remains constant when the death rate of the differentiated cells changes. Consequently, the number of differentiated cells decreases sufficiently. The blue line shows an example where the feedback from the stem cells and the differentiated cells have approximately the same strength.
All examples have the following parameter sizes: $r = 1$, $d = 0.1353$, $P = 10^6$ and $Q = 10^3$. In addition, the example with strong feedback from the differentiated cells has $\theta = 10^{-4}$ and $\gamma = 1.9 \times 10^{-6}$, whereas the example where the feedback from stem cells and differentiated cells are the same has $\theta = 10^{-3}$ and $\gamma = 10^{-6}$, and, finally, the example with no feedback from the differentiated cells has $\theta = 2 \times 10^{-3}$ and $\gamma = 0$. (a) and (b) display the stem cells and the differentiated cells, respectively.

state:

$$x_s^{W*} = \frac{1}{\theta + \frac{P}{\hat{Q}}\gamma}\ln\left(\frac{r}{d}\right). \tag{8}$$

Note that if $\hat{Q} > Q$, then the number of stem cells increases, whereas if $\hat{Q} < Q$, then the number of stem cells decreases. Thus, for the former case the number of differentiated cells ultimately increases, and for the latter case the number of differentiated cells decreases to the steady state

$$x_d^{W*} = \frac{P}{\hat{Q}}x_s^{W*}.$$

Fig. 2 illustrates the cell dynamics when the death rate of the differentiated cells is increased. Note that it follows from Eq. (7), that if the pseudo-steady state hypothesis holds, then two different examples of the system given in Eqs. (3) and (4), with $(\theta_0, \gamma_0) \neq (\theta_1, \gamma_1)$, where

$$\theta_0 + \frac{P}{Q}\gamma_0 = \theta_1 + \frac{P}{Q}\gamma_1, \tag{9}$$

and all other parameters are the same, behave approximately identically. Indeed, this is the case in Fig. 2, which shows three different examples of the system given in Eqs. (3) and (4). Because the parameters satisfy the relations described in Eq. (9) when time is less than one, they behave approximately identically in this time interval. When time equals one, the death rate of the differentiated cells changes from Q to $\hat{Q}$, and the parameters do not satisfy the relations described in Eq. (9) anymore. The blue, dashed line is an example where stem cells are not regulated by feedback from differentiated cells. Hence, when the death rate of the differentiated cells changes to $\hat{Q}$, the number of stem cell remains constant, x_s^* given in Eq. (6), and the number of differentiated cells decreases to $\frac{P}{\hat{Q}}x_s^*$. The green line shows an example where stem cells are regulated by feedback from differentiated cells, and $\frac{P}{Q}\gamma_{no}$ has the same order as θ_{no}. When the death rate of the differentiated cells changes to $\hat{Q}$, the number of stem cells increases to $x_d^{W_{no}*}$, given in Eq. (8), and the number of differentiated cells converges towards $\frac{P}{\hat{Q}}x_d^{W_{no}*}$. The red, dotted line is an example where the stem cells are regulated by strong feedback from the differentiated cells. That is, $\frac{P}{Q}\gamma_{st}$ has a much higher order than θ_{st}. The number of stem cells increases to $x_d^{W_{st}*}$, given in Eq.
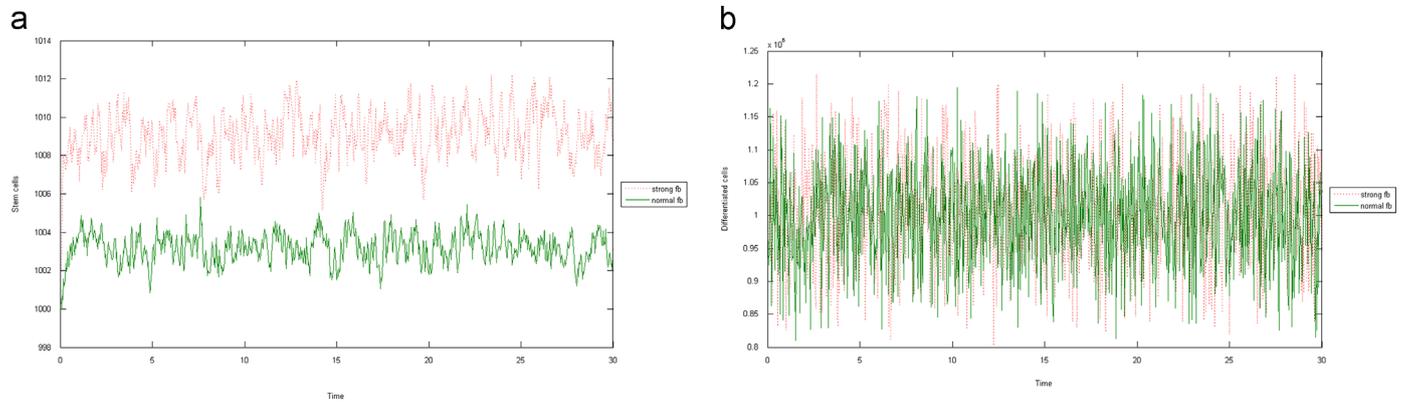
a

b

Fig. 3. *Parameter sensitivity.* In the examples displayed in this figure, all six parameters of the system are continuously varying with up to 20 per cent to test the robustness of the system. More precisely, every time interval $[T, T+1]$ is subdivided into 30 000 time steps. At each time step, every parameter is given a new random value within the interval $[P \times 0.9, P \times 1.1]$, where P is the mean value of the parameter. The green line shows an example where the feedback from the stem cells and the differentiated cells are of the same strength on average. The red, dotted line shows an example where the feedback from the differentiated cells is on average stronger than the feedback from the stem cells. Even though both examples are robust, the former example is less parameter sensitive than the latter.
Both examples have the following average parameter sizes: $r = 1$, $d = 0.1353$, $P = 10^6$ and $Q = 10^3$. In addition, the example with strong feedback from the differentiated cells has average parameter sizes $\theta = 10^{-4}$ and $\gamma = 1.9 \times 10^{-6}$, and the example where the feedback from stem cells and differentiated cells are the same has average parameter sizes $\theta = 10^{-3}$ and $\gamma = 10^{-6}$.
(a) and (b) display the stem cells and the differentiated cells, respectively.

(8), and the number of differentiated cells converges towards $\frac{P}{Q} x_d^{W_{st} 0*}$ when the death rate of the differentiated cells changes to Q. Note that

$$\frac{P}{Q} x_s^* < \frac{P}{Q} x_d^{W_{no}*} < \frac{P}{Q} x_d^{W_{st}*} \approx x_d^*,$$

where x_d^* is as given in Eq. (6). Thus, Fig. 2 illustrates that the worst outcome is obtained in the absence of feedback from differentiated cells. Moreover, it also shows that if the feedback from the differentiated cells is strong, i.e. $\frac{P}{Q}\gamma \approx \theta$, then the number of differentiated cells remains approximately the same as under normal conditions when the death rate of the differentiated cells changes. However, the system is less parameter sensitive when $\frac{P}{Q}\gamma \approx \theta$ than when $\frac{P}{Q}\gamma \gg \theta$, and as discussed by Lander et al. (2009), robustness, which is the ability to maintain performance when perturbations and uncertainties occur, is a key property of living systems (Stelling et al., 2004). How the system responds to perturbations and uncertainties when $\frac{P}{Q}\gamma \approx \theta$ and $\frac{P}{Q} \gg \theta$ is illustrated in Fig. 3. The green graph is an example where the former relation holds, and the red, dotted graph is an example where the latter relation holds. Even though both examples are robust, Fig. 3 shows that the variance and the mean number of stem cells increase more when $\frac{P}{Q} \gg \theta$ than when $\frac{P}{Q}\gamma \approx \theta$. This is the reason why we choose parameters that satisfy the latter relation in all examples in this paper, and denote it *normal feedback*. However, as illustrated in Fig. 2, a system with strong feedback from the differentiated cells performs much better than a system with normal feedback when the death rate of the differentiated cells decreases. One way to compensate for this is to assume that the differentiated cells also have the ability to self-renew. This is investigated by Manesso et al. (2013). However, as discussed in the introduction, increasing capacity of self-renewal among differentiated cells can lead to increasing probability of cancer (Dingli et al., 2007b; Wodarz and Komarova, 2005; Komarova and Cheng, 2006). In Section 2.3, where various stages of differentiation are included, we investigate another strategy to increase the production of differentiated cells.

2.2. Competition dynamics

When the stem cells divide, a mutation might occur (Araten et al., 2005; Vogelstein and Kinzler, 2004). The stem cells that

harbour a mutation are denoted *mutant stem cells*, whereas the other stem cells are denoted *wild-type stem cells*. When a mutant stem cell divides, both daughter cells also harbour the mutation. The differentiated cells that harbour the mutation are denoted *mutant differentiated cells*, and the other differentiated cells are denoted *wild-type differentiated cells*. Like Rodriguez-Brenes et al. (2011), we want to investigate the case when the mutant cells not only retain the architectural aspects of the wild-type cells, but also the regulatory mechanisms. Similar to Stiehl and Marciniak-Czochra (2012), we assume that the mutant stem cells are stimulated by the same cytokines as the wild-type stem cells, but the two cell types respond to these cytokines with different strength. More precisely, it is assumed that the only difference between the mutant cells and the wild-type cells is that the functions that regulate symmetric self-renewal of the wild-type stem cells and the mutant stem cells, denoted Ψ_x and Ψ_y, respectively, are different. Moreover, we neglect continuous production of mutant stem cells from wild-type stem cells. Let y_s denote the number of mutant stem cells and y_d denote the number of mutant differentiated cells. The basic model is given by the following set of ordinary differential equations:

$$\frac{dx_s}{dt} = (r\Psi_x - d)x_s, \tag{10}$$

$$\frac{dx_d}{dt} = Px_s - Qx_d, \tag{11}$$

$$\frac{dy_s}{dt} = (r\Psi_y - d)y_s, \tag{12}$$

$$\frac{dy_d}{dt} = Py_s - Qy_d, \tag{13}$$

where

$$\Psi_z = \exp\left(-\theta_x^z x_s - \theta_y^z y_s - \gamma_x^z x_d - \gamma_y^z y_d\right),$$

and $\theta_v^z, \gamma_v^z > 0$ for $z, v \in \{x, y\}$. Moreover, it is assumed that

$$\theta_z^x \neq \theta_z^y \quad \text{and} \quad \gamma_z^x \neq \gamma_z^y.$$

This means that wild-type cells can either inhibit growth of mutant stem cells *more* than they inhibit growth of wild type stem cells, or they inhibit growth of mutant stem cells *less* than they inhibit growth of wild type stem cells. Clearly, the fitness of

the wild-type cells is higher in former case than in the latter case. Similarly, the mutant cells have higher fitness if they inhibit growth of the wild-type stem cells more than they inhibit growth of the mutant stem cells. Thus, the terms Ψ_x and Ψ_y introduce competition between mutant stem cells and wild-type stem cells.

The system given in Eqs. (10)–(13) has three equilibrium solutions where at least one of the populations gets extinct, namely,

$$(x_s^{0*}, x_d^{0*}, y_s^{0*}, y_d^{0*}) = (0, 0, 0, 0),$$

$$(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*}) = \left(\frac{1}{\theta_x^x + \frac{P}{Q}\gamma_x^x} \ln\left(\frac{r}{d}\right), \frac{P}{Q}x_s^{1*}, 0, 0 \right),$$

$$(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*}) = \left(0, 0, \frac{1}{\theta_y^y + \frac{P}{Q}\gamma_y^y} \ln\left(\frac{r}{d}\right), \frac{P}{Q}y_s^{2*} \right),$$

and one equilibrium solution with coexistence, $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$, where

$$\begin{bmatrix} x_s^{3*} \\ y_s^{3*} \end{bmatrix} = \begin{bmatrix} \theta_x^x + \frac{P}{Q}\gamma_x^x & \theta_y^x + \frac{P}{Q}\gamma_y^x \\ \theta_x^y + \frac{P}{Q}\gamma_x^y & \theta_y^y + \frac{P}{Q}\gamma_y^y \end{bmatrix}^{-1} \begin{bmatrix} \ln\left(\frac{r}{d}\right) \\ \ln\left(\frac{r}{d}\right) \end{bmatrix},$$

$$x_d^{3*} = \frac{P}{Q}x_s^{3*}, \quad y_d^{3*} = \frac{P}{Q}y_s^{3*}. \tag{14}$$

It is assumed that the matrix is non-degenerate. As discussed in Section 2.1, the number of differentiated cells is much larger than the number of stem cells, and we expect the pseudo-steady state hypothesis

$$x_d \simeq \frac{P}{Q}x_s, \quad y_d \simeq \frac{P}{Q}y_s,$$

to hold when the system approaches the given equilibrium solution. Moreover, it is assumed that $r > d$. This means that the equilibrium solution where all types of cells get extinct is unstable. The stability of the remaining equilibrium solutions depends on the following four parameter regimes (Appendix C):

(I) $\theta_y^y + \frac{P}{Q}\gamma_y^y > \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y > \theta_x^x + \frac{P}{Q}\gamma_x^x$. For these parameter relations both the wild-type cells and the mutant cells inhibit growth of mutant stem cells more than growth of wild-type stem cells. The only stable equilibrium solution is extinction of the mutant cells and survival of the wild-type cells, $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$. Moreover, starting with any population size $(x_s^0, x_d^0, y_s^0, y_d^0)$, where $x_s^0, y_s^0 > 0$, the system converges towards $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$.

(II) $\theta_y^y + \frac{P}{Q}\gamma_y^y < \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y < \theta_x^x + \frac{P}{Q}\gamma_x^x$. For these parameter relations both the wild-type cells and the mutant cells inhibit growth of wild-type stem cells more than growth of mutant stem cells. The only stable equilibrium solution is extinction of the wild-type cells and survival of the mutant cells, $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$. Furthermore, starting with any population size $(x_s^0, x_d^0, y_s^0, y_d^0)$, where $x_s^0, y_s^0 > 0$, the system converges towards $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$.

(III) $\theta_y^y + \frac{P}{Q}\gamma_y^y > \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y < \theta_x^x + \frac{P}{Q}\gamma_x^x$. For these parameter relations the wild-type cells inhibit reproduction of wild-type stem cells more than reproduction of mutant stem cells, and likewise, the mutant cells inhibit reproduction of mutant stem cells more than reproduction of wild-type stem cells. In this case the only stable equilibrium solution is coexistence, $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$. Starting with any population size $(x_s^0, x_d^0, y_s^0, y_d^0)$, where $x_s^0, y_s^0 > 0$, the system converges towards $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$.

(IV) $\theta_y^y + \frac{P}{Q}\gamma_y^y < \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y > \theta_x^x + \frac{P}{Q}\gamma_x^x$. When the mutant cells inhibit reproduction of wild-type stem cells more than

reproduction of mutant stem cells, and likewise, the wild-type cells inhibit reproduction of mutant cells more than reproduction of wild-type stem cells more both the equilibrium solutions where only one type of cells survives, $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$ and $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$, are stable. Starting with any population size $(x_s^0, x_d^0, y_s^0, y_d^0)$ where $y_s^0, x_s^0 > 0$, then if $y_s^0 < x_s^0 \Upsilon$, the system converges towards $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$, whereas if $y_s^0 > x_s^0 \Upsilon$, the system converges towards $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$, and if $y_s^0 = x_s \Upsilon$, the system converges towards the equilibrium solution $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$, where

$$\Upsilon = \frac{\theta_x^y + \frac{P}{Q}\gamma_x^y - \left(\theta_x^x + \frac{P}{Q}\gamma_x^x\right)}{\theta_y^x + \frac{P}{Q}\gamma_y^x - \left(\theta_y^y + \frac{P}{Q}\gamma_y^y\right)}.$$

2.2.1. Changes in the competition dynamics

In this subsection we show how changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells. We include changes in the population dynamics of the differentiated cells by modifying the death rate of these cells. Clearly, changes in this death rate can effect the dynamics of the whole system, since the stability of all the equilibrium solutions depends on the inequality

$$\theta_z^v + \frac{P}{Q}\gamma_z^v < \theta_z^w + \frac{P}{Q}\gamma_z^w,$$

for $z, v, w \in \{x, y\}$, $v \neq w$. The stability of the system is changed when the death rate is modified from Q to $\hat{Q}$, such that the inequality is changed to

$$\theta_k^i + \frac{P}{\hat{Q}}\gamma_k^i > \theta_z^j + \frac{P}{\hat{Q}}\gamma_k^j,$$

for at least one triple $i, j, k \in \{z, v, w\}$. There are three different cases:

(I) $\theta_z^v < \theta_z^w$ and $\gamma_z^v < \gamma_z^w$. This inequality cannot be changed for any $\hat{Q} > 0$.

(II) $\theta_z^v < \theta_z^w$ and $\gamma_z^v > \gamma_z^w$. This inequality is changed for any $\hat{Q} < \frac{\gamma_z^v - \gamma_z^w}{\theta_z^w - \theta_z^v}P$.

(III) $\theta_z^v > \theta_z^w$ and $\gamma_z^v < \gamma_z^w$. This inequality is changed for any $\hat{Q} > \frac{\gamma_z^w - \gamma_z^v}{\theta_z^v - \theta_z^w}P$.

These mathematical results can be summarised as follows:

- The equilibrium solution where the mutant cells survive and the wild-type cells get extinct is stable when the mutant cells inhibit growth of wild-type stem cells more than growth of mutant stem cells. If the death rate of the differentiated cells is changed such that the mutant cells inhibit the mutant stem cells more than the wild-type stem cells, then this equilibrium solution becomes unstable.

- The equilibrium solution where the wild-type cells survive and the mutant cells get extinct is stable when the wild-type cells inhibit growth of mutant stem cells more than growth of wild-type stem cells. If the death rate of the differentiated cells is changed such that the wild-type cells inhibit the wild-type stem cells more than the mutant stem cells, then this equilibrium solution becomes unstable.

- The equilibrium solution with coexistence is stable when the mutant cells inhibit growth of mutant stem cells more than growth of wild-type stem cells, and likewise, the wild-type cells inhibit growth of wild-type stem cells more than growth of mutant stem cells. This equilibrium solution becomes

unstable if either the death rate of the differentiated cells is changed such that the mutant cells inhibit growth of wild-type stem cells more than growth of mutant stem cells and/or if the death rate of the differentiated cells is changed such that the wild-type cells inhibit growth of mutant stem cells more than growth of wild-type stem cells.

2.2.2. Numerical simulations

We have performed numerical simulations for different parameter regimes to illustrate how changes in the population dynamics of the differentiated cells can affect the competition dynamics of the stem cells. The goal of this paper is to point out that the relative fitness of stem cells can be affected by changes in the population of differentiated cells. Thus, the parameters are not scaled with respect to a specific tissue. Moreover, since the feedback mechanism within the stem cell area cannot be measured directly, it is not possible to give a precise estimate for all parameters. Just like the examples in Wodarz' (2008) paper, the time is given in an arbitrary unit.

Note that if

$$\theta_x^x = \theta_y^x, \quad \gamma_x^x = \gamma_y^x \quad \text{and} \quad \theta_y^y = \theta_x^y, \quad \gamma_y^y = \gamma_x^y,$$

then exactly one of the equilibrium solutions, where one type of cell gets extinct, is stable. An example of this is shown in Fig. 4. Initially we have that

$$\theta_z^y + \frac{P}{Q}\gamma_z^y < \theta_z^x + \frac{P}{Q}\gamma_z^x,$$

for $z \in \{x, y\}$. Thus, starting with only one mutant stem cell, the system converges towards the equilibrium solution where the mutant cells invade and the wild-type cells get extinct:

$$(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*}) = \left(0, 0, \frac{1}{\theta_y^y + \frac{P}{Q}\gamma_x^x}\ln\left(\frac{r}{d}\right), \frac{P}{Q}y_0^{2*}\right).$$

At time 350 the death rate of the differentiated cells is increased to $\hat{Q}$, such that the inequalities

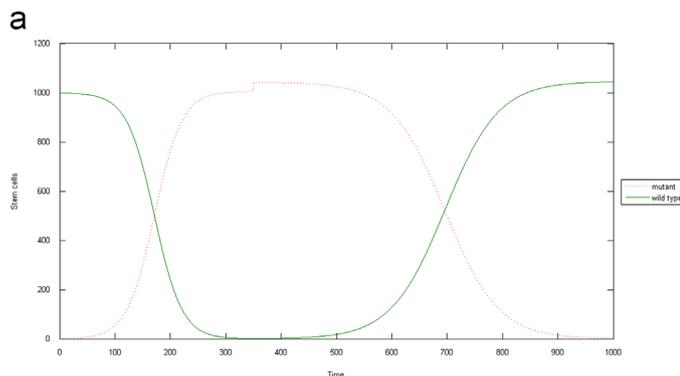$$\theta_z^y + \frac{P}{\hat{Q}}\gamma_z^y > \theta_z^x + \frac{P}{\hat{Q}}\gamma_z^x,$$

hold, and the system converges towards the equilibrium solution where the mutant cells get extinct and the wild-type cells survive:

$$(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*}) = \left(\frac{1}{\theta_x^x + \frac{P}{\hat{Q}}\gamma_x^x}\ln\left(\frac{r}{d}\right), \frac{P}{\hat{Q}}x_0^{1*}, 0, 0\right).$$

Fig. 5 shows an example where initially the inequalities

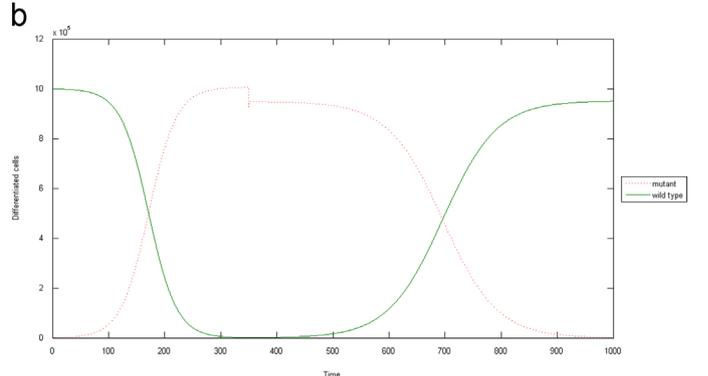$$\theta_z^y + \frac{P}{Q}\gamma_z^y < \theta_z^x + \frac{P}{Q}\gamma_z^x$$

hold for $z \in \{x, y\}$. Thus, only $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$ is stable, and the system converges towards this equilibrium solution. By changing the death rate of the differentiated cells to $\hat{Q}$, we obtain that

$$\theta_y^y + \frac{P}{\hat{Q}}\gamma_y^y < \theta_y^x + \frac{P}{\hat{Q}}\gamma_y^x \quad \text{and} \quad \theta_x^y + \frac{P}{\hat{Q}}\gamma_x^y > \theta_x^x + \frac{P}{\hat{Q}}\gamma_x^x.$$

This means that both $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$ and $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$ become stable. Thus, which of the equilibrium solutions the system converges towards, depends on the time that the death rate is modified.

2.3. Multi-compartment model

In this subsection, we present an extension of the simple model proposed in Section 2.1, which includes various stages of the differentiation process. As discussed in the introduction, the differentiated cells are produced by the stem cells through asymmetric division and symmetric differentiation, and they go though a number of divisions, obtaining various stages of differentiation, until the fully mature cells stop dividing (Donohue et al., 1958; Cronkite and Fliedner, 1964; Ogawa, 1993). However, as discussed by Dingli et al. (2007a), there is no unambiguous determination of the number of stages connecting stem cells and fully differentiated cells, let alone how fast cells go through different stages of maturation (Donohue et al., 1958; Cronkite and Fliedner, 1964). Similar to Dingli et al., we model differentiation as a multi-step process where cell replication and differentiation are coupled with cells moving through successive stages – compartments – of maturation in a series of steps from the stem cells all the way down to the fully differentiated cells. More precisely, when differentiated cells are produced by stem cells through asymmetric division and symmetric differentiation, they move to compartment 1. Furthermore, it is assumed that when a cell in compartment i
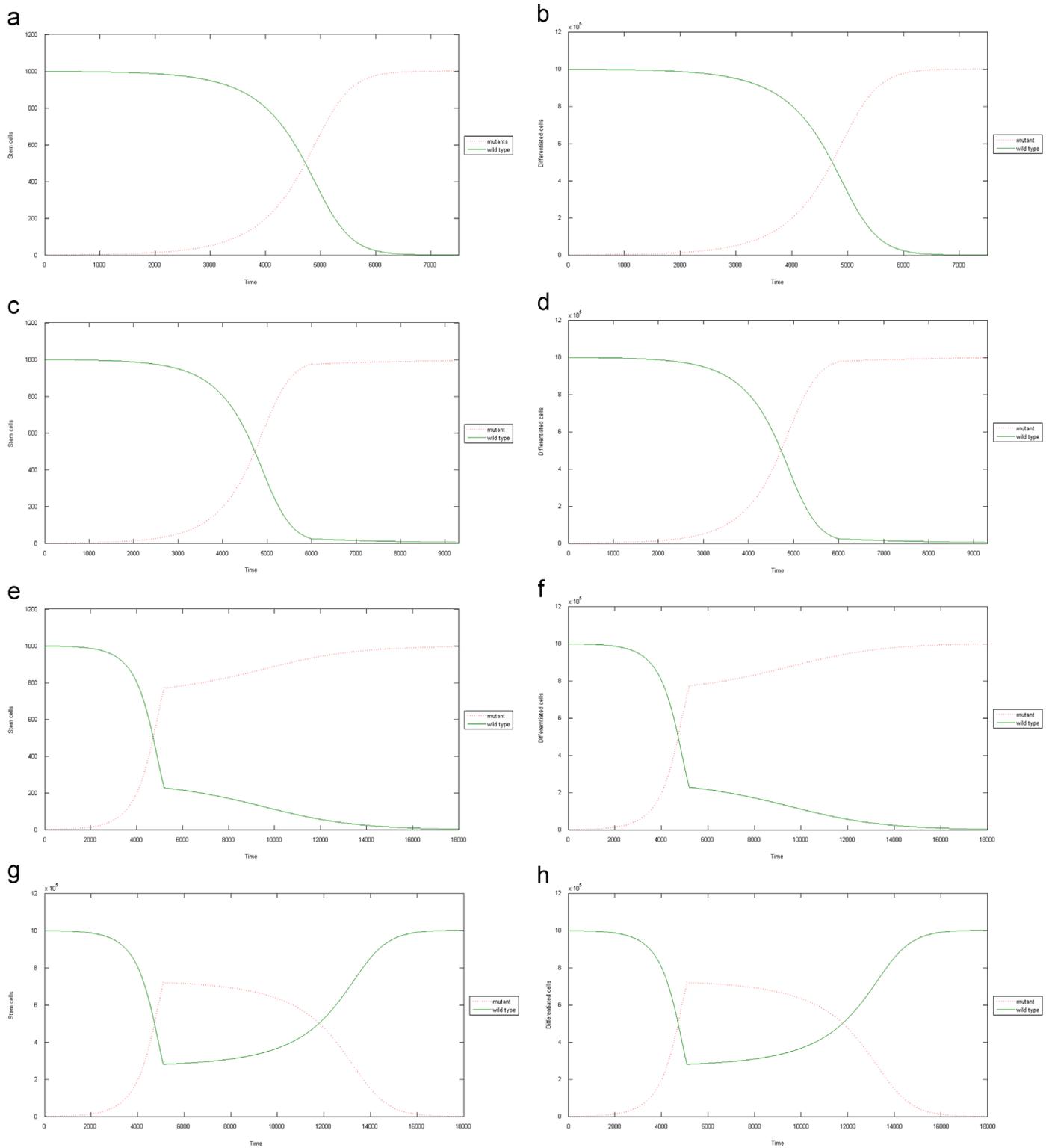

Fig. 4. *One stable equilibrium solution.* Initially, both the wild-type cells and the mutant cells inhibit growth of wild-type stem cells more than growth of mutant stem cells. Thus, if one mutant stem cell is generated at time zero, the system converges towards the only stable equilibrium solution, which is extinction of the wild-type cells and survival of the mutant cells. At time 350 the death rate of the differentiated cells is modified such that both the wild-type cells and the mutant cells inhibit growth of mutant stem cells more than growth of wild-type stem cells. Hence, extinction of the mutant cells and survival of the wild-type cells become the only stable equilibrium solution, and the system converges towards this solution.
The parameter sizes are: $\theta_x^x = \theta_y^x = 0.0012$, $\gamma_x^x = \gamma_y^x = 1.15 \times 10^{-6}$, $\theta_x^y = \theta_y^y = 1.18 \times \theta_x^x$, $\gamma_x^y = \gamma_y^y = 0.8 \times \gamma_x^x$, $r = 1$, $d = \exp\left(-\left(\theta_x^x + \frac{P}{Q}\gamma_x^x\right)10^{10}\right)$, $P = 10^6$ and $Q = 10^3$.
(a) and (b) display the stem cells and the differentiated cells, respectively.

Fig. 5. *Two stable equilibrium solutions.* Initially, both the wild-type cells and the mutant cells inhibit growth of wild-type stem cells more than growth of mutant stem cells. Thus, if one mutant stem cell is generated at time zero, the system converges towards the only stable equilibrium solution, which is extinction of the wild-type cells and survival of the mutant cells. By modifying the death rate of the differentiated cells, the equilibrium solution, where the wild-type cells survive and the mutant cells get extinct, also becomes stable. Which of the equilibrium solutions the system converges towards, depends on the time that the death rate is modified.

The parameter sizes are: $\theta_x^x = \theta_y^x = 0.0012$, $\gamma_x^x = \gamma_y^x = 1.15 \times 10^{-6}$, $\theta_x^y = \theta_y^y = 0.8 \times \theta_x^x$, $\gamma_x^y = 1.2083 \times \gamma_x^x$, $\gamma_y^y = 1.2077 \times \gamma_x^x$ $r = 1$, $d = \exp\left(-\left(\theta_x^x + \frac{P}{Q}\gamma_x^x\right)10^{10}\right)$, $P = 10^6$ and $Q = 10^3$.

(a) and (b) display the stem cells and the differentiated cells, respectively, when the death rate is not modified.

(c) and (d) display the stem cells and the differentiated cells, respectively, when the death rate is modified at time 6000.

(e) and (f) display the stem cells and the differentiated cells, respectively, when the death rate is modified at time 5200.

(g) and (h) display the stem cells and the differentiated cells, respectively, when the death rate is modified at time 5100.

divide, both daughter cells are placed in compartment $i+1$, for $1 \leq i < N$ where N is the total number of compartments of differentiated cells. When the cells reach compartment N, they stop dividing and eventually die. Let x_0 denote the number of stem cells and x_i denote the number of differentiated cells in compartment i. It is assumed that when the cells in all compartments are approximately in normal conditions, then the cells in compartments 1 to $N-1$ divide and die at the approximately same, constant rates, c and s, respectively, where $c > s$. For simplicity, it is assumed that the death rate of the cells in compartment N is $q = c+s$, and that $p = 2c = 2d_0 + g$, where d_0 and g are the rates at which the stem cells differentiate symmetrically and divide asymmetrically, respectively. Hence, if the number of stem cells is in equilibrium, x_0^*, then the number of differentiated cells in compartment i is expected to converge towards

$$x_i^* = \left(\frac{p}{q}\right)^i x_0^* \qquad (15)$$

(approximately). The approximation of the signalling intensity given in (2) considers the average feedback from all differentiated cells. Here, an approximation of the signalling intensity that includes different stages of differentiation is presented:

$$\Psi = \exp(-\theta x_0 - \sum \gamma_{x_i} x_i).$$

It is assumed that for any pair $1 \leq i, j \leq N$, $\gamma_{x_i} \left(\frac{2p}{p+q}\right)^i$ and $\gamma_{x_j} \left(\frac{2p}{p+q}\right)^j$ have the same order and that $\sum_{i=1}^{N} \gamma_{x_i} \left(\frac{2p}{p+q}\right)^i$ has the same order as θ, because our numerical results indicate that the systems with these parameter relations are most robust.

A second feedback mechanism is considered in this subsection, namely, that cells in compartment i inhibit cell division in compartment $i-1$ for $1 < i \leq N$, and that cells in compartment one inhibit asymmetric stem cell division. As discussed in the introduction and in Section 2.1, molecules such as cytokines regulate cell behaviour, and the secretion of cytokines is very fast compared with cell activity such as differentiation. However, the precise nature of this regulation is still unknown (Layton et al., 1989; Aglietta et al., 1989; Metcalf, 2008; Fried, 2009). We assume that the signalling intensity from compartment j is approximately

$$\Gamma_{x_j} = \exp(-\nu_j x_j),$$

for $1 \leq j \leq N$. Since the rates of differentiated cell division and asymmetric stem cell division are approximately constant under normal conditions, ν_j must be sufficiently large, such that

$$\exp(-\nu_j x_j^*) < \epsilon, \qquad (16)$$

for some small number $\epsilon \approx 0$, where x_j^* is given in Eq. (15). The extended model is given by the following set of ordinary differential equations:

$$\frac{dx_0}{dt} = (r\Psi - d)x_0, \qquad (17)$$

$$\frac{dx_i}{dt} = (p + 2W\Gamma_{x_i})x_{i-1} - (q + W\Gamma_{x_{i+1}})x_i, \qquad (18)$$

$$\frac{dx_N}{dt} = (p + 2W\Gamma_{x_N})x_{N-1} - qx_N, \qquad (19)$$

for $1 \leq i < N$, and where W is a positive constant. The system has two equilibrium solutions, namely,

$$(x_0^{0*}, x_1^{0*}, \ldots, x_N^{0*}) = (0, 0, \ldots, 0),$$
$$(x_0^{1*}, x_1^{1*}, \ldots, x_N^{1*}) \approx (x_0^*, x_1^*, \ldots, x_N^*),$$

where x_i^* is given in (15) for $1 < i \leq N$ and

$$x_0^* = \frac{1}{\theta + \sum_{j=1}^{N} \gamma_{x_j} \left(\frac{p}{q}\right)^j} \ln\left(\frac{r}{d}\right).$$

For $r > d$ the former equilibrium solution is unstable and the latter is stable. Moreover, the numerical analysis shows that starting with any population size $(x_0^0, x_1^0, \ldots, x_N^0)$ where $x_0^0 > 0$, the system converges towards the stable equilibrium solution.

The work by Komarova (2013) indicates that a well-regulated N-compartment model must have at least $N+1$ control loops, and that all the $N+1$ different cell populations must control at least one process. Moreover, the differentiation decision for stem cells must be controlled by another population, and the control of stem cell divisions must be negative. The multi-compartment model presented in this subsection satisfy all these conditions. Fig. 6 illustrates that this model performs better than a model that contains less control loops. That is, the figure shows an example where the whole system is regenerated, starting with a single stem cell. For comparison, the figure also shows the regeneration in the absence of feedback between the compartments. From Fig. 6, we can see that feedback between the compartments enables the system to regenerate faster.
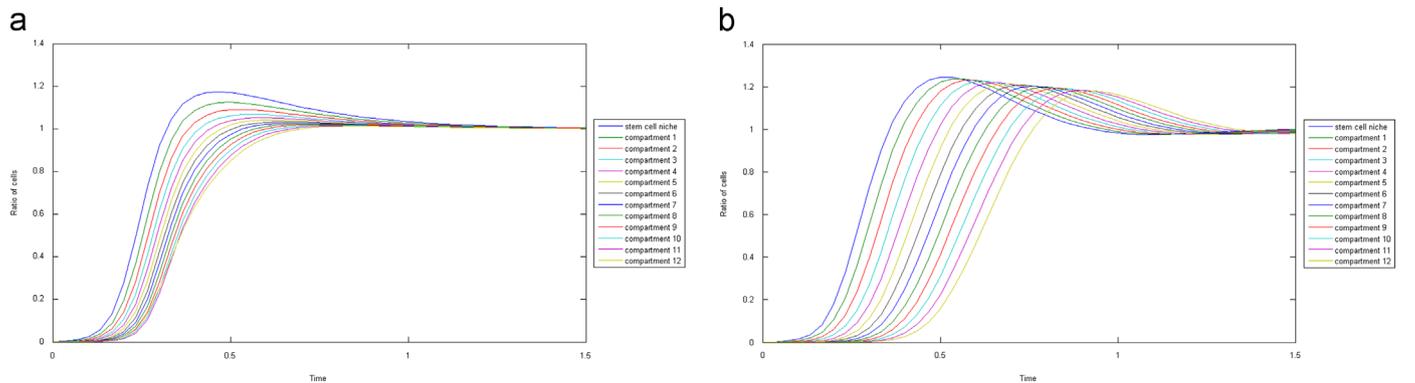

Fig. 6. *Regeneration of the multi-compartmental model.* The whole system is regenerated, starting with a single stem cell. The stem cells reach their normal population size first, and the differentiated cells in compartment i reach their normal population size before the differentiated cells in compartment $i+1$. The figure displays the following ratios:

$$\frac{\text{number of cells}}{\text{normal population size}}$$

for the stem cell niche and all compartments of differentiated cells. The regeneration-time depends on the strength of the feedback and on the number of feedback-loops. The parameter sizes are: $\theta = 0.0012$, $\gamma_{x_i} = 10^{-4} \times \left(\frac{q}{p}\right)^i r = 1$, $W = 2$, $d = 0.0907$, $p = 2$, $q = 1.1$, $\nu_{x_i} = \frac{\ln(10^3)}{x_i^*}$.
(a) displays regeneration with two feedback-loops. (b) displays regeneration with one feedback-loop.

We will now consider competition dynamics in the multi-compartment model. Like in Section 2.2, it is assumed that the wild-type cells and the mutant cells have the same differentiation hierarchy. However, the mutant cells and the wild-type cells inhibit symmetric stem cell self-renewal at different strength. Moreover, in this subsection it is also assumed that the mutant differentiated cells have a lower death rate than the wild-type cells. Let y_0 denote the number of mutant stem cells and y_i denote the number of differentiated cells in compartment i. The competition dynamics is given by the following set of ordinary differential equations:

$$\frac{dx_0}{dt} = (r\Psi_x - d)x_0,$$

$$\frac{dx_i}{dt} = (p + 2W\Gamma_{x_i+y_i})x_{i-1} - (q_x + W\Gamma_{x_{i+1}+y_{i+1}})x_i,$$

$$\frac{dx_N}{dt} = (p + 2W\Gamma_{x_N+y_N})x_{N-1} - q_x x_N,$$

$$\frac{dy_0}{dt} = (r\Psi_y - d)y_0,$$

$$\frac{dy_i}{dt} = (p + 2W\Gamma_{x_i+y_i})y_{i-1} - (q_y + W\Gamma_{x_{i+1}+y_{i+1}})y_i,$$

$$\frac{dy_N}{dt} = (p + 2W\Gamma_{x_N+y_N})y_{N-1} - q_y y_N,$$

for $1 \leq i < N$, where $q_y < q_x$,

$$\Gamma_{x_k+y_k} = \exp(-\nu_k(x_k + y_k))$$

where ν_k satisfies the inequality given in (16) for $1 \leq k \leq N$, and

$$\Psi_z = \exp\left(-\theta_x^z x_0 - \theta_y^z y_0 - \sum_{j=1}^N \left(\gamma_{x_j}^z x_j + \gamma_{y_j}^z y_j\right)\right)$$

for $z \in (x, y)$. Note that if the parameters $\gamma_{y_j}^z$ are of the same order as $\gamma_{x_j}^z$, respectively, then the total number of stem cells decreases if the mutant population starts to grow. Since we are interested in investigating the case where the number of mutant differentiated cells increases beyond the normal level, when the total number of stem cells remains approximately constant, it is assumed that $\gamma_{y_j}^z \left(\frac{q_y}{q_x}\right)^j$ is of the same order as $\gamma_{x_j}^z$.

The system has three equilibrium solutions where at least one type of cells gets extinct, namely,

$$(x_0^{0*}, \ldots, x_N^{0*}, y_0^{0*}, \ldots, y_N^{0*}) = (0, \ldots, 0), \qquad (20)$$

$$(x_0^{1*}, \ldots, x_N^{1*}, y_0^{1*}, \ldots, y_N^{1*}) = (x_0^*, \ldots, x_N^*, 0, \ldots, 0), \qquad (21)$$

$$(x_0^{2*}, \ldots, x_N^{2*}, y_0^{2*}, \ldots, y_N^{2*}) = (0, \ldots, 0, y_0^*, \ldots, y_N^*), \qquad (22)$$

where

$$z_0^* \approx \frac{1}{\theta_z^z + \sum_{j=1}^N \gamma_{z_j}^z \left(\frac{p}{q_z}\right)^j} \ln\left(\frac{r}{d}\right),$$

and

$$z_j^* \approx \left(\frac{p}{q_z}\right)^j z_0,$$

for $z \in \{x, y\}$ and $1 \leq j \leq N$. The system has also one equilibrium solution with coexistence, $(x_0^{3*}, \ldots, x_N^{3*}, y_0^{3*}, \ldots, y_N^{3*})$, where

$$\begin{bmatrix} x_0^{3*} \\ y_0^{3*} \end{bmatrix} \approx \begin{bmatrix} \theta_x^x + \sum_{j=1}^N \gamma_{x_j}^x \left(\frac{p}{q_x}\right)^j & \theta_y^x + \sum_{j=1}^N \gamma_{y_j}^x \left(\frac{p}{q_x}\right)^j \\ \theta_x^y + \sum_{j=1}^N \gamma_{x_j}^y \left(\frac{p}{q_y}\right)^j & \theta_y^y + \sum_{j=1}^N \gamma_{y_j}^y \left(\frac{p}{q_y}\right)^j \end{bmatrix}^{-1} \begin{bmatrix} \ln(\frac{r}{d}) \\ \ln(\frac{r}{d}) \end{bmatrix},$$

$$z_j^{3*} \approx \left(\frac{p}{q_z}\right)^j z_0.$$

The equilibrium solution where all cells get extinct is unstable for $r > d$. The numerical analysis shows that the analytic results

obtained in Section 2.2 also apply to the extended model. That is, the equilibrium solution with survival of the wild-type cells and extinction of the mutant cells given in (21) is stable if the wild-type cells inhibit reproduction of mutant cells more than reproduction of wild-type cells, i.e.

$$\theta_x^x + \sum_{j=1}^N \gamma_{x_j}^x \left(\frac{p}{q_x}\right)^j < \theta_y^x + \sum_{j=1}^N \gamma_{y_j}^x \left(\frac{p}{q_y}\right)^j.$$

On the other hand, if the wild-type cells inhibit reproduction of wild-type cells more than reproduction of mutant cells, then the equilibrium solution is unstable. Likewise, the equilibrium solution with survival of the mutant cells and extinction of the wild-type cells given in (22) is stable if the mutant cells inhibit reproduction of wild-type cells more than reproduction of mutant cells, i.e.

$$\theta_x^y + \sum_{j=1}^N \gamma_{x_j}^y \left(\frac{p}{q_y}\right)^j > \theta_y^y + \sum_{j=1}^N \gamma_{y_j}^y \left(\frac{p}{q_y}\right)^j.$$

On contrary, if the mutant cells inhibit reproduction of mutant cells more than reproduction of wild-type cells, then the equilibrium solution is unstable. If both the equilibrium solutions given in (21) and (21) are unstable, then the equilibrium solution with coexistence is stable. Moreover, if there is only one stable equilibrium solution and both types of stem cells are present, then the system converges towards this solution. On the other hand, if there are two stable equilibrium solutions and both types of stem cells are present, then the system converges towards one of the equilibrium solutions.

Fig. 7 shows an example where a mutant stem cell is generated when the wild-type cells are in normal condition. Since both the wild-type cells and the mutant cells inhibit growth of wild-type cells more than mutant cells, the only stable equilibrium is extinction of the wild-type cells and invasion of the mutant cells, and the system converges towards this solution. The death rate of the mutant differentiated cells is lower than the death rate of the wild-type cells. Hence, the number of differentiated cells increases beyond the normal level. Moreover, since the mutant differentiated cells have weak feedback to the stem cells, the number of stem cells remains approximately constant. When the death rate of the differentiated cells is reduced, the equilibrium solution where the wild-type cells survive and the mutant cells get extinct also becomes stable. Which of the two stable solutions the system converges, depends on the time that the death rate is reduced.

3. Discussion

In this paper we use a mathematical model to investigate implications when the rate of symmetric self-renewal is regulated by both differentiated cells and stem cells, and show that changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells. This result implies that a medical treatment that targets differentiated cells can change the competition dynamics of the stem cells, even if the treatment has no direct effect on the stem cells.

Research suggests that a subset of cancer cells within some tumours, the so-called *cancer stem cells*, may drive the growth and metastasis of these tumours (Reya et al., 2001; Clarke and Fuller, 2006). Understanding the pathways that regulate proliferation, self-renewal, survival and differentiation of malignant and normal stem cells may shed light on mechanisms that lead to cancer and suggest better modes of treatment (Rodriguez-Brenes et al., 2011). For most types of cancer, the target cell of transforming mutation is unknown. However, there is considerable evidence that certain types of leukaemia, such as *chronic myeloid leukaemia* (CML), arise
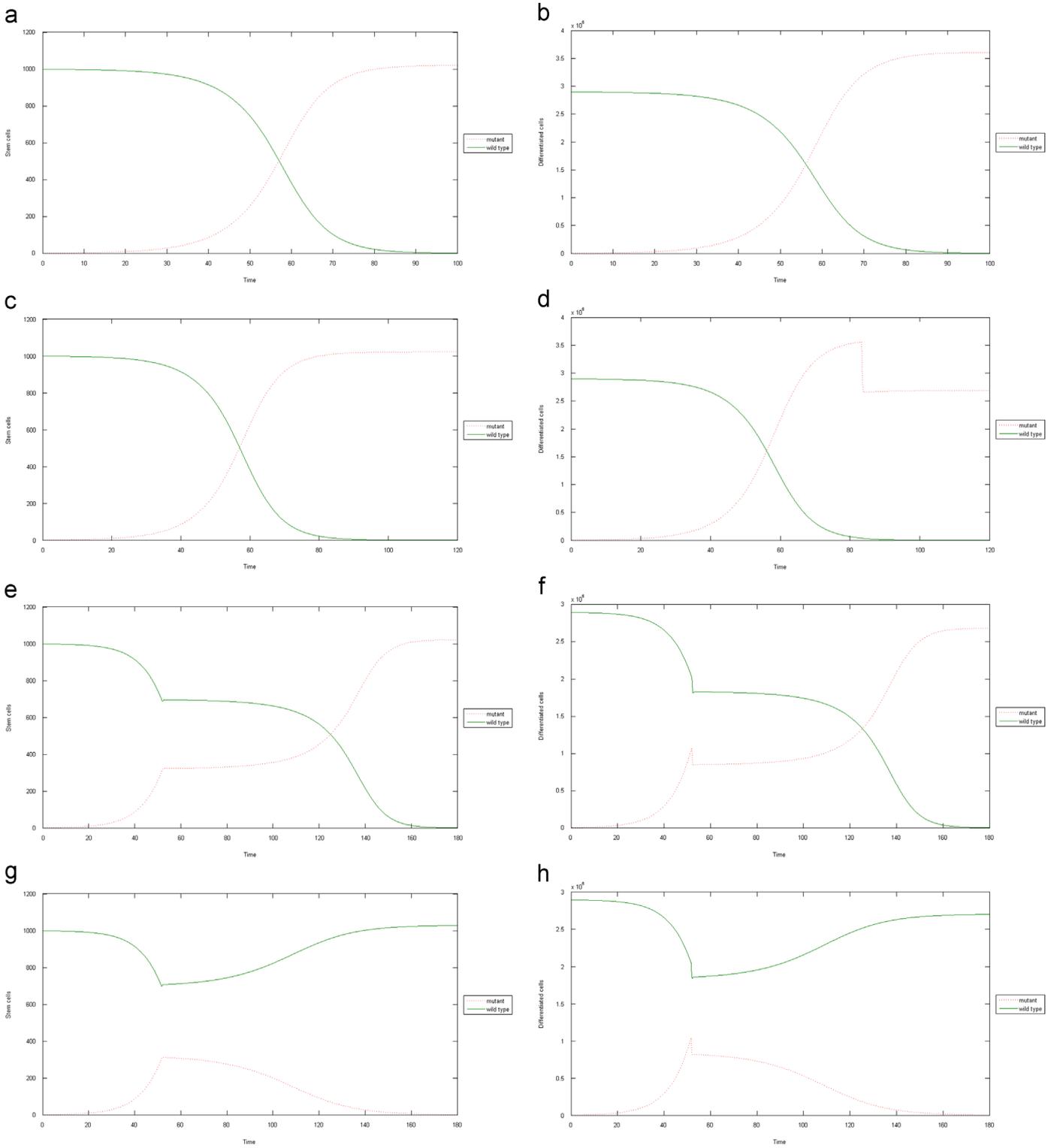
Fig. 7. *Two stable equilibrium solutions in the multi-compartmental model.* Initially, the mutant differentiated cells have a low death rate and weak feedback to the stem cells. Moreover, both the wild-type cells and the mutant cells inhibit growth of wild-type stem cells more than growth of mutant stem cells. Consequently, if one mutant stem cell is generated, the total number of differentiated cells increases and mutants invade the system, whereas the wild-type cells get extinct. If the death rate of the differentiated cells is decreased enough, the equilibrium solution, where the wild-type cells survive and the mutant cells get extinct, also becomes stable. Which of the equilibrium solutions the system converges to, depends on the time that the death rate is modified.

The parameters sizes are: $\theta_x^x = 0.0012$, $\theta_y^x = 0.0024$, $\theta_x^y = 0.0024$, $\theta_y^y = 0.0023$ $r = 1$, $d = 0.0907$, $p = 2$, $q_x = 1.1$, $q_y = 1.08$, $\hat{Q} = 1.11$ $\gamma_{x_i}^x = 10^{-4} \times (\frac{q_x}{p})^i$, $\gamma_{y_i}^x = 10^{-7} \times 8.3 \times (\frac{q_y}{p})^i$, $\gamma_{x_i}^y = 10^{-7} \times 8.3 \times (\frac{q_x}{p})^i$, $\gamma_{y_i}^y = 10^{-7} \times 8.3 \times (\frac{q_y}{p})^i$, $W = 1$, $\nu_{x_i+y_i} = \frac{\ln(10^3)}{x_i^4}$.

(a) and (b) display the stem cells and differentiated cells, respectively, when the death rate of the differentiated cells is not modified.
(c) and (d) display the stem cells and the sum of all differentiated cells, respectively, when the death rate of the differentiated cells is modified at time 83.
(e) and (f) display the stem cells and the sum of all differentiated cells, respectively, when the death rate of the differentiated cells is modified at time 52.
(g) and (h) display the stem cells and the sum of all differentiated cells, respectively, when the death rate of the differentiated cells is modified at time 51.5.

from mutation in haematopoietic stem cells (Reya et al., 2001; Wang and Dick, 2005; Hope et al., 2004).

Treatment of CML with the *tyrosine kinase inhibitors* (TKIs) imatinib and nilotinib represents a successful application of molecularly targeted anti-cancer therapy (Druker et al., 1996, 2001; Kantarjian et al., 2002). TKIs reduce the fitness of leukemic differentiated cells. However, the effect of TKIs on leukemic stem cells remains incompletely understood. Several mathematical models of CML and treatment with TKIs have been proposed (Dingli and Michor, 2006; Wodarz, 2008; Michor et al., 2005; Rodriguez-Brenes et al., 2011; Roeder et al., 2006). These models are discussed and compared by Michor (2008). Discontinuation of TKIs results in a relapse of the disease in many patients within a few months (Cortes et al., 2004). Explanations have been put forward for this phenomenon. For example, the drug might have no effect on the CML stem cells (Dingli and Michor, 2006; Michor et al., 2005), or the CML stem cells can be susceptible to drug therapy when they are in an active state, but are not be susceptible when they are in quiescent state (Roeder et al., 2006). In contrast to these arguments a small study involving 12 patients has shown that in some individuals the disease has remained undetected for two years after discontinuation of TKIs, raising the possibility that TKIs have eradicated the disease in these patients (Rousselot et al., 2006). Moreover, all studies indicate that the effect of TKIs increases when treatment starts early in disease progression (Rousselot et al., 2006; Gorre et al., 2001; Houchhause et al., 2002; Roche-Lestienne et al., 2002). These results can be explained by the mechanisms described in our model: Suppose that the treatment with TKIs has no direct effect on the leukemic stem cells. However, since the treatment changes the population dynamics of the differentiated cells, and the differentiated cells regulate the proliferation of the stem cells, treatment indirectly effects the stem cells and can lead to changes in the competition dynamics of the stem cells. More precisely, let us revisit the examples illustrated in Figs. 5 and 7. In both figures the wild-type cells represent the healthy cells, the mutant cells represent the leukemic cells, and treatment is represented by modifying the death rate of the differentiated cells. Subfigures (a) and (b) show the disease progression without any treatment – the number of leukemic cells expands and the healthy cells get extinct. In (c) and (d), treatment starts too late to have any significant effect on the disease progression. In (e) and (f), treatment starts early enough to slow down the disease progression and the healthy cells survive a bit longer. However, ultimately, the leukemic cells invade the population and the healthy cells get extinct. Finally, in (g) and (h), treatment starts early enough to reverse the competition dynamics – the healthy cells survive and the leukemic cells get extinct. Fig. 7 shows an example of the extended model, which captures the fact that the number of differentiated leukemic cells increases beyond the normal level, whereas the number of stem cells remains approximately constant (Wang and Dick, 2005; Hope et al., 2004). However, the competition dynamics in both examples is determined by the feedback functions that regulate self-renewal, and this is best captured by the example of the simple model illustrated in Fig. 5.

Lenaerts et al. (2010) illustrate that the results from studies of TKIs treatment (Cortes et al., 2004; Rousselot et al., 2006; Gorre et al., 2001; Houchhause et al., 2002; Roche-Lestienne et al., 2002) can also be explained by the stochastic nature of the haematopoietic stem cells. A deterministic model does not capture neither neutral drift nor that a disadvantageous phenotype can outcompete an advantageous phenotype in a finite population. Since stem cell populations in general are small, their population dynamics are highly sensitive to stochastic fluctuations. Under steady state, the number of stem cells is approximately constant, and Lenaerts et al. show that the stem cell population dynamics can be captured by the Moran process, which describes the probabilistic dynamics in a finite population of constant size N. The Moran process

predicts that if there are i mutant stem cells and $N-i$ wild-type stem cells in the population, while the mutants have relative fitness r and the wild-types relative fitness 1, then the probability that the mutant cells eventually invade the whole population is

$$p_i = \frac{i}{N} \tag{23}$$

if $r = 1$, and

$$p_i = \frac{1 - r^{-i}}{1 - r^{-N}}$$

if $r \neq 1$. The mutants are advantageous if $r > 1$, disadvantageous if $r < 1$, and neutral if $r = 1$. Moreover, the probability that the mutant population eventually gets extinct is $1 - x_i$. Hence, the Moran process predicts that coexistence is only temporary – ultimately the population consists of only one type of cells. Lenaerts et al. assume that the competition between the healthy stem cells and the CML stem cells is captured by a neutral Moran process and that TKIs treatment has no effect on stem cells. If CML is discovered early, then the number of CML stem cells, i, is in general much smaller than the total number of stem cells, N. It follows from Eq. (23) that the probability that the CML stem cells get extinct is $1 - \frac{i}{N} \approx 1$. Hence, there is a very good chance of full recovery, even though the TKIs treatment has no effect on the stem cells. On the other hand, if CML is discovered relatively late, then the number of CML stem cells is typically very high, such that $1 - \frac{i}{N} \approx 0$. This means that full recovery is very unlikely.

Lenaerts et al. (2010) illustrate the importance of stochastic fluctuations in stem cell populations, and the response dynamics predicted by the model closely matches data from clinical trials. Since stem cell regulation is an extremely complex process, a model that treats self-renewal and differentiation as purely random events fits general data better than a deterministic model with a single regulation mechanism. Thus, the model proposed by Lenarts et al. gives a general picture of how stem cells behave under steady state. However, Lander et al. (2009) show that linear models, e.g. the one proposed by Lenarts et. al., are very parameter sensitive. Since parameter sensitivities tend to be undesirable in well-regulated biological systems, stochastic behaviour cannot be the complete story. A deterministic model of stem cell dynamics with only one regulation mechanism can be designed to describe more specific data. For instance, research results by Gokoffski et al. (2011) indicate that the number of stem cells increases when the number of differentiated cells decreases. This can be explained by a model of stem cell self-renewal and differentiation, where symmetric stem cell division is regulated by differentiated cells, like the model proposed by Rodriguez-Brenes et al. (2013) and the models presented in this paper.

The model proposed by Lenaerts et al. (2010) and the model presented in this paper have different explanations for successful TKIs treatment. However, this does not mean that one of the conclusions must be false. It is possible that the CML stem cells are advantageous before TKIs treatment and, because the differentiated cells regulate symmetric stem cell division, the CML stem cells are disadvantageous during the TKIs treatment. In this case, the average behaviour of the CML stem cells can be approximately neutral, as assumed by Lenaerts et al.

The main purpose of the simple model proposed in Section 2.1 is to investigate implications when symmetric stem cell division is regulated by differentiated cells. Similar results can be obtained by replacing the signal intensity function given in (2) with another function that reaches its maximum under complete absence of cells and decreases towards zero as the number of cells decreases. For instance

$$S = 1/(\theta x_s + \gamma x_d + 1),$$

which is similar to the function proposed by Marciniak-Czochra et al. (2009) given in (1). Yet, the function in (2) is used in our model because it makes the stability analysis simple. It would also be interesting to investigate implications when differentiated cells also regulate symmetric differentiation. However, in this paper, the model is kept simple to obtain analytic results. In Sections 2.2 and 2.3 the competition dynamics is investigated. It is possible that the mutant cells have other properties than the ones investigated in this paper. For instance, the mutant differentiated cells in compartment i could have a much weaker feedback to compartment $i-1$ than the wild-type cell in compartment i. This could radically change the dynamics of the cells. However, investigation of these types of mutation is beyond the scope of this paper. Finally, the timescale in all examples are the same. This illustrates that regeneration of the system in general occurs much faster than the invasion of a mutation. Moreover, Figs. 4, 5 and 7 indicate that the timescale of the competition dynamics depends on the ratios

$$\frac{\theta_v^w + \frac{P}{Q}\gamma_v^w}{\theta_v^z + \frac{P}{Q}\gamma_v^z}$$

where $v, w, z \in \{x, y\}$, $w \neq z$. That is, the closer these ratios are to one, the slower the competition dynamics occurs.

4. Conclusion

In this paper, we use a mathematical models where symmetric stem cell division is regulated by negative feedback from the differentiated cells, to show that changes in the population dynamics of the differentiated cells can lead to changes in the population dynamics of the stem cells. This result is interesting because it can explain how medical treatments that have no direct effect on the stem cells can change the competition dynamics of these cells. For example, the model can reproduce some of the results from studies of TKIs treatment of CML patients (Cortes et al., 2004; Rousselot et al., 2006; Gorre et al., 2001; Houchhause et al., 2002; Roche-Lestienne et al., 2002):

- The effect of TKIs increases when treatment starts early in disease progression.
- In some cases the treatment slows down the disease progression without erasing the CML stem cells, which drive the disease.
- In other cases the treatment reverses the disease progression and seems to erase the CML stem cells.

The results from these studies seem contradictory if a classical deterministic model of stem cells and differentiation is used, where stem cell activity is not regulated by the differentiated cells (Dingli and Michor, 2006; Michor et al., 2005). Our model shows that the results from the different studies can be explained by negative feedback from differentiated cells that regulate symmetric stem cell division: TKIs treatment reduces the fitness of the CML differentiated cells, but has little or no direct effect on the CML stem cells. However, since the differentiated cells regulate the proliferation of the stem cells, the treatment indirectly affects the stem cells and can lead to changes in the competition dynamics of the stem cells, which in some cases results in the extinction of the CML stem cells.

Appendix A

Proposition 1. *Consider the systems of differential equations given in* (3) and (4) *and* (10)–(13). *If Q is sufficiently large, then the pseudo-state hypothesis* (Appendix C)

$$\frac{dz_d}{dt}\frac{1}{Q} = \left(\frac{P}{Q}z_s - z_d\right) \simeq 0 \tag{A.1}$$

holds for $z \in \{x, y\}$ and $t \geq \frac{1}{r}$, when it is given that the parameters r and d are of significantly lower order than the parameters P and Q, $r > d$ and $P > Q$, and initial values are non-negative.

Proof. By re-scaling the systems of given in (3) and (4) and (10)–(13) with respect to the constant r, we obtain

$$\frac{dv_s}{dT} = (\Psi_v - d_r)v_s, \tag{A.2}$$

$$\frac{dv_d}{dT} = P_r v_s - Q_r v_d, \tag{A.3}$$

where $d_r = \frac{d}{r}$, $P_r = \frac{P}{r}$, $Q_r = \frac{Q}{r}$, and $T = r \times t$. For the system with only one type of cells $v = x$, whereas $v \in \{x, y\}$ for the system with both wild-type cells and mutant cells. Note that if $v_s = 0$, then $\frac{dv_s}{dT} = 0$, and if $v_d = 0$ and $v_s \geq 0$, then $\frac{dv_d}{dT} \geq 0$. Consequently, a solution of the system with non-negative initial values will never obtain negative values.

We now use the perturbation methods presented by Fowler (1997) to analyse the system. Given that $0 < d_r < 1$, $Q_r \gg 1$ and $P_r > Q_r$, the pseudo-state hypothesis states that if Q_r is sufficiently large, then

$$\frac{dv_d}{dT}\frac{1}{Q_r} = \left(\frac{P_r}{Q_r}v_s - v_d\right) \simeq 0 \tag{A.4}$$

when $T \geq O(1)$. However, the approximation given in (A.4) does not generally hold for the initial values, i.e. when T is close to zero. This means that the neglect of $\frac{dv_d}{dT}\frac{1}{Q_r}$ is wrong in a region that contains $T = 0$. We will now show that if Q_r is sufficiently large, then this is only a thin region, termed the *boundary layer*. We bring back the term $\frac{dv_d}{dT}$ in the boundary layer by rescaling the time as

$$T = \frac{1}{Q_w}\tau,$$

where Q_w is $O(Q_r)$. To obtain variables that are $O(1)$, we rescale as follows:

$$v_s = v_s(0)V_s,$$
$$v_d = v_d(0)V_d,$$

where $v_s(0)$ and $v_d(0)$ are the initial values of v_s and v_d, respectively. By substituting this into the system of differential equations given in (A.2) and (A.3), we obtain

$$\frac{dV_s}{d\tau} = \frac{1}{Q_w}(\Psi_v - d_r)V_s,$$

$$\frac{dV_d}{d\tau} = \frac{P_r}{Q_w}\frac{v_s(0)}{v_d(0)}V_s - \frac{Q_r}{Q_w}V_d.$$

Note that V_s is $O(1)$ when τ is close to zero. Thus, since $0 < d_r, \Psi_v < 1$ and $P_r \gg 1$, the variable V_s is approximately constant, i.e. $\frac{dV_s}{d\tau} \simeq 0$ when τ is $O(1)$. On the other hand, $\frac{Q_r}{Q_w}$ is $O(1)$, while $\frac{P_r}{Q_w} \geq O(1)$. Thus, we obtain the approximate solution

$$V_d(\tau) = \frac{P_r}{Q_r}\frac{v_s(0)}{v_d(0)}V_s(\tau) + \left(1 - \frac{P_r}{Q_r}\frac{v_s(0)}{v_d(0)}\right)\exp\left(-\frac{Q_r}{Q_w}\tau\right).$$

By substituting the original variables, we obtain

$$v_d(T) = \frac{P_r}{Q_r}v_s(T) + \left(v_d(0) - \frac{P_r}{Q_r}v_s(0)\right)\exp\left(-\frac{Q_r}{Q_w^2}T\right).$$

Hence, outside the boundary layer, i.e. when $T \geq O(1)$, we obtain the approximation

$$v_d(T) = \frac{P_r}{Q_r}v_s(T),$$

which satisfies the pseudo-steady state hypothesis given in (A.4). Thus, we have proved that for sufficient large Q, any solution with non-negative initial values of either the system given in (3) and (4) or the system given in (10)–(13) satisfies the pseudo-state hypothesis given in (A.1) for $t \geq \frac{1}{r}$. □

Appendix B

Proposition 2. *The system of differential equations given in (3) and (4) has one stable equilibrium solution:*

$$(x_s^*, x_d^*) = \left(\frac{1}{\theta + \frac{P}{Q}\gamma} \ln\left(\frac{r}{d}\right), \frac{P}{Q} \frac{1}{\theta + \frac{P}{Q}\gamma} \ln\left(\frac{r}{d}\right) \right),$$

and one unstable equilibrium solution:

$$(x_s^{0*}, x_d^{0*}) = (0, 0),$$

for $r > d$. The following domain

$$\mathcal{V} = \{(x_s, x_d) \in \mathbb{R}^2 \mid x_s > 0, x_d \geq 0\}$$

is in the basin of attraction of (x_s^*, x_d^*).

Proof. We prove the proposition for the case when Q is sufficiently large so that Proposition 1 holds. Then the system given in (3) and (4) is reduced to the following differential equation:

$$\frac{dx_s}{dt} = (r\Psi - d)x_s,$$

where

$$\Psi = \exp\left(-\left(\theta + \gamma\frac{P}{Q}\right)x_s \right).$$

The Jacobian of the system is

$$\mathbf{J}(x_s) = (r\Psi - d) - x_s\left(\theta + \gamma\frac{P}{Q}\right)r\Psi.$$

We have that

$$\mathbf{J}(x_s^{0*}) = (r - d) > 0$$

for $r > d$. Hence, the equilibrium solution (x_s^{0*}, x_d^{0*}) is unstable. Since

$$\mathbf{J}(x_s^*) = -x_s^*\left(\theta + \gamma\frac{P}{Q}\right)r\Psi < 0,$$

the equilibrium solution (x_s^*, x_d^*) is stable. Note that $\frac{dx_s}{dt} > 0$ for $0 < x_s < x_s^*$. Hence, if the initial number of stem cells is less than x_s^*, then the solution converges towards the stable equilibrium solution. Likewise, $\frac{dx_s}{dt} < 0$ for $x_s > x_s^*$. Hence, if the initial number of stem cells is greater than x_s^*, then the solution converges towards the stable equilibrium solution. Consequently, $\mathcal{V}$ is in the basin of attraction of (x_s^*, x_d^*). □

Appendix C

Proposition 3. *The system given in Eqs. (10)–(13) has three equilibrium solutions where at least one of the populations gets extinct, namely,*

$$(x_s^{0*}, x_d^{0*}, y_s^{0*}, y_d^{0*}) = (0, 0, 0, 0),$$

$$(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*}) = \left(\frac{1}{\theta_x^x + \frac{P}{Q}\gamma_x^x} \ln\left(\frac{r}{d}\right), \frac{P}{Q}x_s^{1*}, 0, 0 \right),$$

$$(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*}) = \left(0, 0, \frac{d}{\theta_y^y + \frac{P}{Q}\gamma_y^y} \ln\left(\frac{1}{d}\right), \frac{P}{Q}y_s^{2*} \right),$$

and one equilibrium solution with coexistence, $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$, given in Eq. (14). For $r > d$, $(x_s^{0*}, x_d^{0*}, y_s^{0*}, y_d^{0*})$ is unstable. Moreover, given that Q is sufficiently large such that Proposition 1 holds, the behaviour of system depends on the following four parameter relations:

(a) *For $\theta_y^y + \frac{P}{Q}\gamma_y^y > \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y > \theta_x^x + \frac{P}{Q}\gamma_x^x$, the only stable equilibrium solution is $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$. Moreover, the domain*

$$\mathcal{U} = \{(x_s, x_d, y_s, y_d) \in \mathbb{R}^4 \mid z_s > 0, \ z_d \geq 0, z \in \{x, y\}\}, \quad (C.1)$$

is in the basin of attraction of $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$.

(b) *For $\theta_y^y + \frac{P}{Q}\gamma_y^y < \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y < \theta_x^x + \frac{P}{Q}\gamma_x^x$, the only stable equilibrium solution is $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$, and $\mathcal{U}$ is in the basin of attraction of $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$.*

(c) *For $\theta_y^y + \frac{P}{Q}\gamma_y^y > \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y < \theta_x^x + \frac{P}{Q}\gamma_x^x$, the only stable equilibrium solution is $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$. Moreover, $\mathcal{U}$ is in the basin of attraction of this equilibrium solution.*

(d) *For $\theta_y^y + \frac{P}{Q}\gamma_y^y < \theta_y^x + \frac{P}{Q}\gamma_y^x$ and $\theta_x^y + \frac{P}{Q}\gamma_x^y > \theta_x^x + \frac{P}{Q}\gamma_x^x$, both $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$ and $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$ are stable and $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$ is unstable. Moreover, the domain*

$$\mathcal{D}^1 = \left\{ (x_s, x_d, y_s, y_d) \in \mathcal{U} \mid y_s < x_s \frac{\Theta_x^y - \Theta_x^x}{\Theta_y^x - \Theta_y^y} \right\}$$

is in the basin of attraction of $(x_s^{1*}, x_d^{1*}, y_s^{1*}, y_d^{1*})$, and

$$\mathcal{D}^2 = \left\{ (x_s, x_d, y_s, y_d) \in \mathcal{U} \mid y_s > x_s \frac{\Theta_x^y - \Theta_x^x}{\Theta_y^x - \Theta_y^y} \right\}$$

is in the basin of attraction of $(x_s^{2*}, x_d^{2*}, y_s^{2*}, y_d^{2*})$. The basin of attraction of the equilibrium solution $(x_s^{3*}, x_d^{3*}, y_s^{3*}, y_d^{3*})$ is the line

$$\mathcal{L}^3 = \left\{ (x_s, x_d, y_s, y_d) \in \mathcal{U} \mid y_s = x_d \frac{\Theta_x^y - \Theta_x^x}{\Theta_y^x - \Theta_y^y} \right\},$$

where

$$\Theta_w^v = \theta_w^v + \frac{P}{Q_v}\gamma_w^v,$$

for $v, w \in \{x, y\}$.

Proof. Given that Q is sufficiently large, such that Proposition 1 holds, the system given in (10)–(13) is reduced to the following two differential equations:

$$\frac{dx_s}{dT} = (\Psi_x - D)x_s, \quad (C.2)$$

$$\frac{dy_s}{dT} = (\Psi_y - D)y_s, \quad (C.3)$$

where $D = \frac{d}{r}$, $T = r \times t$, and

$$\Psi_v = \exp\left(-\Theta_x^v x_s - \Theta_y^v y_s \right).$$

for $v \in \{x, y\}$. We have that $\frac{dv_s}{dT} = 0$ for $v_s = 0$ and $\Psi_v = D$. Note that $0 < D < 1$. Thus, there are three equilibrium solutions where at least one of the variables is zero, namely,

$$(x_s^{0*}, y_s^{0*}) = (0, 0),$$

$$(x_s^{1*}, y_s^{1*}) = \left(\frac{1}{\Theta_x^x} \ln\left(\frac{1}{D}\right), 0 \right),$$

$$(x_s^{2*}, y_s^{2*}) = \left(0, \frac{1}{\Theta_y^y} \ln\left(\frac{1}{D}\right) \right).$$

The equilibrium solution with coexistence, (x_s^{3*}, y_s^{3*}), must satisfy

$$\Omega \begin{bmatrix} x_s^{3*} \\ y_s^{3*} \end{bmatrix} = \ln\left(\frac{1}{D}\right) \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

where

$$\Omega = \begin{bmatrix} \Theta_x^x & \Theta_y^x \\ \Theta_x^y & \Theta_y^y \end{bmatrix}.$$

If $\det\Omega = 0$, we have a solution with coexistence if and only if $\Theta_x^x = \Theta_x^y$ and $\Theta_y^x = \Theta_y^y$, which means that there is no difference between the wild-type cells and mutant cells. Thus, we will only consider the case when $\det\Omega \neq 0$, and obtain the following solution with coexistence:

$$\begin{bmatrix} x_s^{3*} \\ y_s^{3*} \end{bmatrix} = \frac{\ln\left(\frac{1}{D}\right)}{\Theta_x^x\Theta_y^y - \Theta_y^x\Theta_x^y} \begin{bmatrix} \Theta_y^y & -\Theta_y^x \\ -\Theta_x^y & \Theta_x^x \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Note that $x_s^{3*}, y_s^{3*} > 0$ if and only if both $\Theta_y^y > \Theta_y^x$ and $\Theta_x^x > \Theta_x^y$ or both $\Theta_y^y < \Theta_y^x$ and $\Theta_x^x < \Theta_x^y$. The Jacobian of the system is

$$\mathbf{J}(x_s, y_s) = \begin{bmatrix} (\Psi_x - D) - \Theta_x^x\Psi_x x_s & -\Theta_y^x\Psi_x x_s \\ -\Theta_x^y\Psi_y y_s & (\Psi_y - D) - \Theta_y^y\Psi_y y_s \end{bmatrix}.$$

We have that

$$\mathbf{J}(x_s^{0*}, y_s^{0*}) = \begin{bmatrix} (1-D) & 0 \\ 0 & (1-D) \end{bmatrix}.$$

Since $D < 1$, both eigenvalues are positive. Hence, (x_s^{0*}, y_s^{0*}) is unstable. Moreover,

$$\mathbf{J}(x_s^{1*}, y_s^{1*}) = \begin{bmatrix} -\Theta_x^x\Psi_x x_s^{1*} & -\Theta_y^x\Psi_x x_s^{1*} \\ 0 & (\Psi_y - \Psi_x) \end{bmatrix}.$$

Thus, if $\Psi_y > \Psi_x$, i.e. $\Theta_x^y < \Theta_x^x$, then one eigenvalue is positive and (x_s^{1*}, y_s^{1*}) is unstable, whereas if $\Psi_y < \Psi_x$, i.e. $\Theta_x^y > \Theta_x^x$, then both eigenvalues are negative and (x_s^{1*}, y_s^{1*}) is stable. Likewise,

$$\mathbf{J}(x_s^{2*}, y_s^{2*}) = \begin{bmatrix} (\Psi_x - \Psi_y) & 0 \\ -\Theta_x^y\Psi_y y_s^{2*} & -\Theta_y^y\Psi_y y_s^{2*} \end{bmatrix}.$$

Hence, if $\Psi_y < \Psi_x$, i.e. $\Theta_y^y > \Theta_y^x$, then one eigenvalue is positive and (x_s^{2*}, y_s^{2*}) is unstable, whereas if $\Psi_y > \Psi_x$, i.e. $\Theta_y^y < \Theta_y^x$, then both eigenvalues are negative and (x_s^{2*}, y_s^{2*}) is stable. Finally,

$$\mathbf{J}(x_s^{3*}, y_s^{3*}) = \begin{bmatrix} -\Theta_x^x\Psi_x x_s^{3*} & -\Theta_y^x\Psi_x x_s^{3*} \\ -\Theta_x^y\Psi_y y_s^{3*} & -\Theta_y^y\Psi_y y_s^{3*} \end{bmatrix}.$$

The characteristic equation is

$$(\Theta_x^x\Psi_x x_s^{3*} + \lambda)(\Theta_y^y\Psi_y y_s^{3*} + \lambda) - \Theta_y^x\Theta_x^y\Psi_x\Psi_y x_s^{3*}y_s^{3*} = 0.$$

Hence, (x_s^{3*}, y_s^{3*}) is stable if $\Theta_x^x\Theta_y^y > \Theta_y^x\Theta_x^y$. Thus, the equilibrium solution with coexistence is both stable and positive if $\Theta_x^x > \Theta_x^y$ and $\Theta_y^y > \Theta_y^x$.

By analysing the nullclines of the system given in (C.1) and (C.2), namely,

$$\Theta_x^x x_0 + \Theta_y^x y_0 = -\ln D,$$
$$\Theta_x^y x_0 + \Theta_y^y y_0 = -\ln D$$

we can predict the global behaviour. Since these nullclines are the same as the nullclines of the two-species Lotka–Volterra competition model (Smith, 1978), the global stability analysis is identical. Hence, the basins of attraction of the equilibrium solutions are as described in (a)–(d) (Smith, 1978).□

References

Adimy, M., Crauste, F., Ruan, S., 2006. A mathematical study of hematopoiesis process with application to chronic myeloid leukemia. SIAM J. Appl. Math. 65, 1328–1352.

Aglietta, M., Piacibello, W., Sanavio, F., Stacchini, A., Apra, F., Schena, M., Mossetti, C., Carnino, F., Caligaris-Cappio, F., Gavosto, F., 1989. Kinetics of human hemopoietic cells after in vivo administration of granulocyte-macrophage colony-stimulating factor. J. Clin. Investig. 83, 551–557.

Agur, Z., Daniel, Y., Ginosar, Y., 2002. The universal properties of stem cells as pinpointed by a simple discrete model. J. Math. Biol. 44, 79–86.

Aiuti, A., Friedrich, C., Sie, C.A., Gutierrez-Ramos, J.C., 1998. Identification of distinct elements of the stromal microenvironment that control human hematopoietic stem/progenitor cell growth and differentiation. Exp. Hematol. 26, 143157.

Araten, D.J., Golde, D.W., Zhang, R.H., et al., 2005. A quantitative measurement of human somatic mutation rate. Cancer Res. 65, 8111–8117.

Baum, C.M., Weissman, I.L., Tsukamoto, A.S., Buckle, A.M., Peault, B., 1992. Isolation of a candidate human hematopoietic stem-cell population. Proc. Natl. Acad. Sci. U.S.A. 89, 2804–2808.

Bertolini, F., Battaglia, M., Soligo, D., et al., 1997. "Stem cell candidates" purified by liquid culture in the presence of Steel factor, IL-3, and 5FU are strictly stroma-dependent and have myeloid, lymphoid, and mega karyocytic potential. Exp. Hematol. 26, 350–356.

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

Coiljn, C., Mackey, M.C., 2005. A mathematical model of hematopoiesis—I. Periodic chronic myeloid leukemia. J. Theor. Biol. 237, 117–132.

Cortes, J., O'Brien, S., Kantarjian, H., 2004. Discontinuation of imatinib therapy after achieving a molecular response. Blood 104, 2204–2205.

Cronkite, E.P., Fliedner, T.M., 1964. Granulocytopoiesis. N. Engl. J. Med. 270, 1347–1352.

Derynck, R., Akhurst, R.J., Balmain, A., 2001. Tgf-beta signaling in tumor suppression and cancer progression. Nat. Genet. 29, 117–129.

Dingli, D., Michor, F., 2006. Successful therapy must eradicate cancer stem cells. Stem Cells 24, 2603–2610.

Dingli, D., Traulsen, A., Pacheco, J.M., 2007a. Compartmental architecture and dynamics of hematopoiesis. PLoS One 2, e345.

Dingli, D., Traulsen, A., Michor, F., 2007b. (A) Symmetric stem cell replication and cancer. PLoS Comput. Biol. 3, 1.

Dingli, D., Traulsen, A., Lenaerts, T., Pacheco, J.M., 2010. Evolutionary dynamics of chronic myeloid leukemia. Genes Cancer 1, 309–315.

Donohue, D.M., Reiff, R.H., Hanson, M.L., et al., 1958. Quantitative measurement of the erythrocytic and granulocytic cells of the marrow and blood. J. Clin. Investig. 37, 1571–1576.

Druker, B.J., Tamura, T., Buchdunger, E., et al., 1996. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcr-Abl positive cells. Nat. Med. 2, 561–566.

Druker, B.J., Talpaz, M., Resta, D.W., et al., 2001. Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia. N. Engl. J. Med. 344, 1031–1037.

Fowler, A.C., 1997. Mathematical Models in the Applied Sciences. Cambridge University Press, Cambridge, pp. 45–55.

Fried, W., 2009. Erythropoietin and erythropoiesis. Exp. Hematol. 37, 1007–1015.

Fuchs, E., Tumbar, T., Guasch, G., 2004. Socializing with the neighbors: stem cells and their niche. Cell 116, 769–778.

Gentry, S., Jackson, T., 2013. A mathematical model of cancer stem cell driven tumor initiation: implications of niche size and loss of homeostatic regulatory mechanisms. PLoS One 8, e71128.

Gokoffski, K.K., Wu, H.H., Beites, C.L., et al., 2011. Activin and GDF11 collaborate in feedback control of neuroepithelial stem cell proliferation and fate. Development 138, 4131–4142.

Gorre, M.E., Mohammed, M., Eliwood, L., et al., 2001. Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science 5531, 876–880.

Hope, K.J., Jin, L., Dick, J.E., 2004. Acute myeloid leukemia originates from a hierarchy of leukemic stem cell classes that differ in self-renewal capacity. Nat. Immunol. 5, 738–743.

Houchhause, A., Kreail, S., Corbin, A.S., et al., 2002. Molecular and chromosomal mechanisms to resistance to imatinib (STI571) therapy. Leukemia 16, 2190–2196.

Kantarjian, H., Sawyers, C., Hochhaus, A., et al., 2002. Hematologic and cytogenic responses to imatinib mesylate in chronic myelogenous leukemia. N. Engl. J. Med. 346, 645–652.

Koller, M.R., Oxender, M., Jensen, T.C., Goltry, K.L., Smith, A.K., 1999. Direct contact between CD34+lin− cells and stroma induces a soluble activity that specifically increases primitive hematopoietic cell production. Exp. Hematol. 27, 734–742.

Komarova, N.L., 2013. Principles of regulation of self-renewing cell lineages. PLoS One 8 (9), e72847.

Komarova, N.L., Cheng, P., 2006. Epithelial tissue architecture protects against cancer. Math. Biosci. 200, 90–117.

Lander, A.D., Gokoffski, K.K., Wan, F.Y., et al., 2009. Cell lineages and the logic of proliferative control. PLoS Biol. 7, e15.

Layton, J.E., Hockman, H., Sheridan, W.P., Morstyn, G., 1989. Evidence for a novel in vivo control mechanism of granulopoiesis: mature cell-related control of a regulatory growth factor. Blood 74, 1303–1307.

Lemischka, I.R., 1997. Microenvironmental regulation of hematopoietic stem cells. Stem Cells 15, 63–68.

Lenaerts, T., Pacheco, J., Traulsen, A., Dingli, D., 2010. Tyrosine kinase inhibitor therapy can cure chronic myeloid leukemia without hitting leukemic stem cells. Haematologica 95, 900–907.

Loeffler, M., Wichmann, H.E., 1980. A comprehensive mathematical model of stem cell proliferation which reproduces most of the published experimental results. Cell Tissue Kinet. 13, 543–561.

Loeffler, M., Wichmann, H.E., Schmitz, S., 1988. A concept of hemopoietic regulation and its biomathematical realization. Blood Cells 14, 411–429.

Manesso, E., Bryder, D., Peterson, C., 2013. Dynamical modelling of haematopoiesis: an integrated view over the system in homeostasis and under perturbation. J. R. Soc. Interface 10, 20120817.

Mangel, M., Bonsall, M.B., 2008. Phenotypic evolutionary models in stem cell biology: replacement, quiescence, and variability. PLoS One 3, e1591.

Marciniak-Czochra, A., Stiehl, T., Ho, A., Jaeger, W., Wagner, W., 2009. Modeling of asymmetric cell division in hematopoietic stem cells-regulation of self-renewal is essential for efficient repopulation. Stem Cells Dev. 18, 377–386.

Massagué, J., 2000. Tgf in cancer. Cell 103, 295–309.

Massagué, J., 2001. G1 cell cycle control and cancer. Nature 432, 298–306.

McHale, P.T., Lander, A.D., 2014. The protective role of symmetric stem cell division on the accumulation of heritable damage. PLoS Comput. Biol. 10, e1003802.

McKenzie, J.L., Gan, O.I., Doedens, M., Wang, J.C., Dick, J.E., 2006. Individual stem cells with highly variable proliferation and self-renewal properties comprise the human hematopoietic stem cell compartment. Nat. Immunol. 7, 1225–1233.

Metcalf, D., 2008. Hematopoietic cytokines. Blood 111, 485–491.

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

Michor, F., Hughes, T.P., Iwasa, Y., et al., 2005. Dynamics of chronic myeloid leukaemia. Nature 435, 1267–1270.

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

Morrison, S.J., Weissman, I.L., 1994. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity 1, 661–673.

Morrison, S.J., Shah, N.M., Anderson, D.J., 1997. Regulatory mechanisms in stem cell biology. Cell 88, 287–298.

Nikolova, G., Strilic, B., Lammert, E., 2006. The vascular niche and its basement membrane. Trends Cell Biol. 17, 19–25.

Nowak, M., 2006a. Evolutionary Dynamics. Harvard University Press. Cambridge, Massachusetts.

Nowak, M.A., 2006b. Evolutionary Dynamics: Exploring the Equations of Life. Belknap Press of Harvard University Press, Cambridge, MA.

Ogawa, M., 1993. Differentiation and proliferation of hematopoietic stem cells. Blood 81, 2844–2853.

Østby, I., Winther, R., 2004. Stability of a model of human granulopoiesis using continuous maturation. J. Math. Biol. 49, 501–536.

Østby, I., Rusten, L.S., Kvalheim, G., Grtturn, P., 2003. A mathematical model for reconstituting of granulopoiesis after high dose chemotherapy with autologous stem cell transplantation. J. Math. Biol. 47, 101–136.

Potten, C.S., Loeffler, M., 1990. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the crypt. Development (Cambridge, England) 110, 1001–1020.

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

Roche-Lestienne, C., Soenen-Cornu, V., Grardel-Dulflos, N., et al., 2002. Several types of mutations of the Abl gene can be found in chronic myeloid leukemia patients resistant to STI571, and they can pre-exist to the onset of treatment. Blood 100, 1014–1018.

Rodriguez-Brenes, I.A., Komarova, N.L., Wodarz, D., 2011. Evolutionary dynamics of feedback escape and the development of stem-cell-driven cancers. Proc. Natl. Acad. Sci. 108, 18983–18988.

Rodriguez-Brenes, I., Wodarz, D., Komarova, N., 2013. Stem cell control, oscillations, and tissue regeneration in spatial and non-spatial models. Front. Oncol. 3, 82.

Roeder, I., Horn, M., Glauche, I., et al., 2006. Dynamic modeling of imatinib treated chronic myeloid leukemia: functional insights and clinical implication. Nat. Med. 12, 1181–1184.

Rousselot, P., Huguet, F., Rea, D., et al., 2006. Imatinib mesylate discontinuation in patients with chronic myelogenous leukemia in complete molecular remission for more than two years. Blood 109, 58–60.

Rozenblum, E., et al., 1997. Tumor-suppressive pathways in pancreatic carcinoma. Cancer Res. 57, 1731–1734.

Shahriyari, L., Komarova, N.L., 2013. Symmetric vs. asymmetric stem cell divisions: an adaptation against cancer?. PloS One 8, e76195.

Shortman, K., Naik, S.H., 2009. Steady-state and inflammatory dendritic-cell development. Nat. Rev. Immunol. 7, 19–30.

Simons, B., Cleavers, H., 2011. Strategies for homeostatic stem cell self-renewal in adult tissues. Cell 1454, 851–862.

Smith, J.M., 1978. Models in ecology. CUP Arch., 59–62.

Stelling, J., Sauer, U., Szallasi, Z., et al., 2004. Robustness of cellular functions. Cell 118, 675–685.

Stiehl, T., Marciniak-Czochra, A., 2012. Mathematical modeling of leukemogenesis and cancer stem cell dynamics. Math. Model. Nat. Phenom. 7, 166–202.

Thiemann, F.T., Moore, K.A., Smogorzewska, E.M., Lemiscka, I.R., Crooks, G.M., 1998. The murine stromal cell line AFT024 acts specifically on human CD34+CD38- progenitorsto maintain primitive function and immuniphenotype in vitro. Exp. Hematol. 26, 612–619.

Verfaillie, C.M., 1998. Adhesion receptors as regulators of the hematopoietic process. Blood 92, 2609–2612.

Vogelstein, B., Kinzler, K.W., 2004. Cancer genes and the pathways they control. Nat. Med. 10, 789–799.

Wang, J.C., Dick, J.E., 2005. Cancer stem cells: lessons from leukemia. Trends Cell Biol. 15, 494–501.

Wineman, J., Moore, K., Lemischka, I., Muller-Sieburg, C., 1996. Functional hetero-geneity of the hematopoietic microenvironment: rare stromal elements maintain long-term repopulating stem cells. Blood 87, 4082–4090.

Wodarz, D., 2008. Stem cell regulation and the development of blast crisis in chronic myeloid leukemia: implications for the outcome of imatinib treatment and discontinuation. Med. Hypotheses 70, 128–136.

Wodarz, D., Komarova, N.L., 2005. Computational Biology of Cancer: Lecture Notes and Mathematical Modeling. World Scientific, London.

Wong, M., Jin, Z., Xiw, T., 2005. Molecular mechanisms of germline stem cell regulation. Ann. Rev. Genet. 39, 172–195.

Woodford-Richens, K.L., et al., 2001. Smad4 mutations in colorectal cancer probably occur before chromosomal instability, but after divergence of the microsatellite instability pathway. Proc. Natl. Acad. Sci. U.S.A. 98, 9719–9723.

Wu, A., et al., 2008. Persistence of cd133+ cells in human and mouse glioma cell lines: detailed characterization of gl261 glioma cells with cancer stem cell-like properties. Stem Cells Dev. 17, 173–184.

Yamashita, Y.M., Jones, D.L., Fuller, M.T., 2003. Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. Science 12, 1547–1550.

Yin, T., Li, L., 2006. The stem cell niches in bone. J. Clin. Investig. 116, 1195–1201.

CrossMark

Dynamic self-organisation of haematopoiesis and (a)symmetric cell division

Marthe Måløy[a,*], Frode Måløy[a], Per Jakobsen[a], Bjørn Olav Brandsdal[b]

[a] Department of Mathematics and Statistics, University of Tromsø, Norway
[b] Department of Chemistry, University of Tromsø, Norway

ABSTRACT

A model of haematopoiesis that links self-organisation with symmetric and asymmetric cell division is presented in this paper. It is assumed that all cell divisions are completely random events, and that the daughter cells resulting from symmetric and asymmetric stem cell divisions are, in general, phenotypically identical, and still, the haematopoietic system has the flexibility to self-renew, produce mature cells by differentiation, and regenerate undifferentiated and differentiated cells when necessary, due to self-organisation. As far as we know, no previous model implements symmetric and asymmetric division as the result of self-organisation. The model presented in this paper is inspired by experiments on the *Drosophila* germline stem cell, which imply that under normal conditions, the stem cells typically divide asymmetrically, whereas during regeneration, the rate of symmetric division increases. Moreover, the model can reproduce several of the results from experiments on female Safari cats. In particular, the model can explain why significant fluctuation in the phenotypes of haematopoietic cells was observed in some cats, when the haematopoietic system had reached normal population level after regeneration. To our knowledge, no previous model of haematopoiesis in Safari cats has captured this phenomenon.

1. Introduction

Haematopoiesis is the generation of the blood-forming system. At the root of this process is a small group of slowly replicating cells, the *haematopoietic stem cells*, which are undifferentiated cells with the capacity to both *self-renew* and generate all types of blood cells (Baum et al., 1992; Morrison and Weissman, 1994). The haematopoietic stem cells are located within the bone marrow and segregated among different bones throughout the body. Through sequential division, the haematopoietic stem cells differentiate into *progenitor cells*, which in turn differentiate into red blood cells, white blood cells or platelets. Since the number of haematopoietic stem cells is much smaller than the number of more differentiated blood cells, the haematopoietic stem cells must be protected and tightly regulated. *Haematopoietic bone marrow niches*, which are restricted regions in the bone marrow that contain undifferentiated cells and support stem cell behaviour, may be crucial in both aspects (Wineman et al., 1996; Lemischka, 1997; Bertolini et al., 1997; Aiuti et al., 1998; Thiemann et al., 1998; Verfaillie, 1998; Koller et al., 1999; Yin and Li, 2006; Zhang and Li, 2008; Cheng et al., 2014). Since it is not possible to reconstruct a niche experimentally, it is difficult to maintain haematopoietic stem cells in vitro, because signals from the niche affect stem cell survival, self-renewal, and differentiation. This is one of the reasons why relatively

little is known about the exact behaviour of haematopoietic stem cells. On the other hand, haematopoietic progenitors have been studied both in vivo and in vitro (Abkowitz et al., 1988, 1990, 1993; Gehling et al., 2000; Akita et al., 2013; Herrmann et al., 2014). A set of experiments was designed by Abkowitz et al., using female *Safari cats*, in order to get an idea of the contribution of haematopoietic stem cells to progenitor cells (Abkowitz et al., 1988, 1990, 1993). The Safari cat is a hybrid of the Geoffroy cat (a South American wildcat) and a domestic cat (which is of Eurasian origin). These two species have evolved independently for twelve million years, and have distinct phenotypes of the X chromosome-linked enzyme glucose-6-phosphate dehydrogenase (G6PD) (Molecular genetics in the domestic cat and its relatives, 1986). Female Safari cats have some cells that contain Geoffroy-type G6PD (G G6PD) and other cells that contain domestic-type G6PD (d G6PD). The G6PD phenotype is retained after replication and differentiation, and is functionally neutral. Therefore, it provides a binary marker of each cell and its offspring. In particular, this means that a progenitor cell that expresses G G6PD is the daughter of a stem cell that expresses G G6PD, and likewise, a progenitor cell that is d G6PD-positive is the daughter of a stem cell that is d G6PD-positive. Abkowitz et al. (1988), Abkowitz et al. (1990), Abkowitz et al. (1993) tracked the contributions of haematopoietic stem cells to the progenitor cells by observing the G6PD phenotype of haematopoietic progenitor cells. In the first trials,

the percentage of committed progenitor cells expressing d G6PD was observed over a period of almost six years in normal female Safari cats, and Abkowitz et al. found that the percentage remained relatively constant (Abkowitz et al., 1988, 1990). On the contrary, the G6PD phenotype of haematopoietic progenitors varied extensively when six Safari cats were lethally irradiated, in order to kill the cells in their bone marrow, and a small number of bone marrow cells, collected prior to the radiation, were transplanted back (Abkowitz et al., 1990, 1993). Abkowitz et al. observed the percentage of progenitor cells expressing d G6PD while the cells in the bone marrow regenerated, and they found that the pattern of clonal contribution to haematopoiesis in each cat was unique. For instance, some of the cats that both had cells expressing d G6PD and cells expressing D G6PD when the regeneration started, had only cells expressing either d G6PD or D G6PD when the production of bone marrow cells stabilised after regeneration. Thus, one of the phenotypes had got extinct during the regeneration. On the contrary, in other cats, the percentage of cells expressing d G6PD and D G6PD remained on average relatively constant. Moreover, in some cats, significant variation in the percentage extended for years after the number of cells reached normal population levels, whereas in other cats, the percentage remained approximately constant. Several mathematical models (Guttorp et al., 1990; Newton et al., 1995; Abkowitz et al., 1996; Golinelli et al., 2006; Fong et al., 2009) have been proposed to explain the results from the experiments on female Safari cats (Abkowitz et al., 1988, 1990, 1993). These models are discussed in Section 1.4.

1.1. Symmetric and asymmetric stem cell division

Stem cells are, in general, undifferentiated cells that can both self-renew and generate differentiated progeny required by a given tissue (Morrison et al., 1997; Reya et al., 2001). An important aspect is the fate of the two daughter cells when a stem cell divides (Yamashita et al., 2003; Morrison and Kimble, 2006; McKenzie et al., 2006; Dingli et al., 2007). If one daughter cell has stem cell identity and the other daughter cell commits to differentiation and loses the stem cell identity, it is called as an *asymmetric stem cell division* or *asymmetric self-renewal*. Under normal conditions, the number of cells in a given tissue is approximately constant. It is generally believed that the number of stem cells is also approximately constant under normal conditions, and that they differentiate and self-renew at relatively constant rates to replace mature cells and to keep the stem cell number at a certain normal level (Wichmann et al., 1988; Shortman and Naik, 2007). By dividing asymmetrically, the stem cells manage to both self-renew and produce differentiated cells in a single division. The experiments by Abkowitz et al. indicate that haematopoietic cells divide asymmetrically under normal conditions, because the percentage of cells expressing d G6PD remained relatively constant when normal female Safari cats were observed over a period of almost six years (Abkowitz et al., 1988, 1990). However, a disadvantage of asymmetric stem cell division is that it leaves stem cells unable to expand in number. It is, in general, believed that the stem cells can regenerate (Morrison et al., 1997; Reya et al., 2001; Yamashita et al., 2003; Morrison and Kimble, 2006; McKenzie et al., 2006; Dingli et al., 2007). In particular, haematopoietic stem cells can expand rapidly in response to injury to the bone marrow, such as stem cell transplantation (Abkowitz et al., 1990, 1993; McKenzie et al., 2006). Hence, asymmetric self-renewal cannot be the complete story, since it leaves stem cells unable to expand in number.

Symmetric division is defined as generation of daughter cells destined to acquire the same fate. In this paper, symmetric stem cell division is defined as *symmetric self-renewal* if both daughter cells are stem cells and *symmetric commitment* if none of the daughters are stem cells. The number of stem cells increases by one after symmetric self-renewal. Hence, since the haematopoietic bone marrow can regenerate after injury (Abkowitz et al., 1990, 1993; McKenzie et al., 2006), it is likely that the rate of symmetric self-renewal depends on

the number of haematopoietic stem cells. On the contrary, the number of stem cells decreases by one after a symmetric commitment. Thus, this type of division can cause the extinction of a stem cell phenotype. The experiments on female Safari cats indicate that both types of symmetric stem cell division occur when the haematopoietic bone marrow niche regenerates after injury (Abkowitz et al., 1990, 1993). Wide fluctuation in the percentage of progenitors with d G6PD was observed for one to four years, before the percentage stabilised and became relatively constant. This indicates that when there are significantly less haematopoietic stem cells in the niche than under normal conditions, the rate of symmetric self-renewal increases such that the number of haematopoietic stem cells also increases. When the number of haematopoietic stem cells reaches its normal population level, the rate of symmetric self-renewal decreases, and proliferation in the haematopoietic niche stabilises. Moreover, some of the cats that both had cells expressing d G6PD and D G6PD when the regeneration started, only had cells expressing either d G6PD or D G6PD when the production of bone marrow cell stabilised after regeneration. This indicates that the haematopoietic stem cells commit symmetrically to differentiation under regeneration, since this type of division can cause the extinction of a phenotype. Clearly, the rate of symmetric self-renewal must, on average, be higher than the rate of symmetric commitment when the haematopoietic niche regenerates, such that the number of stem cells increases. On the other hand, under normal conditions, the number of stem cells remains constant, and hence, the two types of symmetric division must occur at the same rate. Thus, the experiments by Abkowitz et al. indicate that haematopoietic stem cells divide mostly asymmetrically under normal conditions, whereas when the haematopoietic bone marrow niche regenerates after injury, the haematopoietic stem cells start to divide symmetrically (Abkowitz et al., 1988, 1990, 1993; McKenzie et al., 2006). Does this mean that a stem cell "knows" that it must divide asymmetrically under normal conditions and self-renew symmetrically when stem cells need to be replaced? This would also imply that the daughter cells inherit this "knowledge". As discussed by Loeffler and Roeder (2002), the assumption that each cell "knows" how to behave in different situations is too rigorous and potentially misleading. In the next subsection, it is argued that each stem cell behaves completely random. However, the stem cells divide mostly asymmetrically under normal conditions and symmetrically under regeneration due to dynamic regulation and self-organisation in the haematopoietic bone marrow niche.

Several mathematical models that include symmetric and asymmetric stem cell division have been proposed (Abkowitz et al., 1988, 1990, 1993; Dingli et al., 2007; Wodarz and Komarova, 2005). Wodarz and Komarova (2005) present a model where the haematopoietic stem cells only divide asymmetrically under normal conditions, whereas during regeneration, the stem cells switch to symmetric division. On the contrary, in the model proposed by Abkowitz et al. (1996), the haematopoietic stem cells can only divide symmetrically: Under normal condition, the stem cells undergo symmetric self-renewal and symmetric commitment at the same, constant rate, and under regeneration, the rate of the former type of division increases. Even though all the models presented in Abkowitz et al. (1988), Abkowitz et al. (1990), Abkowitz et al. (1993), Dingli et al. (2007), Wodarz and Komarova (2005) capture important aspects related to stem cell behaviour, it is a drawback that stem cell self-renewal and differentiation do not depend on local growth conditions. The model proposed by Roeder and Loffler in Loeffler and Roeder (2002) and Roeder and Loeffler (2002) considers the dependence of proliferation control on the local growth conditions. However, no implications about symmetric or asymmetric stem cell division are included in this model.

1.2. Haematopoietic bone marrow niche

The haematopoietic bone marrow niche is composed of both localised signalling cells and an extracellular matrix that control the

Fig. 1. The population dynamics in the compartments of undifferentiated cells The bone marrow niche is represented as the compartment of stem cells and the compartment of undifferentiated cells committed to differentiation. In this figure, the former compartment is green, whereas the latter compartment is blue. Both the compartments have M sites. In this figure, $M = 100$, and each site is represented by a square. Each site can either be full, i.e. contain one cell, or be vacant, i.e. contain no cell. In this figure, the full sites are the squares that contain a circle, and the vacant sites are the squares that do not contain a circle. At each elementary event, a random site and a random stem cell are selected. In this figure, the selected stem cell is in the red box and the selected site is in the yellow box, in the two compartments to the left. (a) Asymmetric stem cell division: a site in the compartment of stem cells is selected, and hence, the selected stem cell, in the red box, divides. One of the daughter cells inherits the mother's site. Since the selected site, in the yellow box, is full, the second daughter cell migrates to the compartment of committed undifferentiated cells and is placed in a random vacant site. (b) Symmetric self-renewal: a site in the compartment of stem cells is selected, and hence, the selected stem cell, in the red box, divides. One of the daughter cells inherits the mother's site. Since the selected site, in the yellow box, is vacant, the second daughter cell is placed in this site. (c) Symmetric commitment: a vacant site in the compartment of undifferentiated cells committed to differentiation, in the yellow box, is selected, and hence, the selected stem cell, in the red box, divides. Both daughter cell migrate to the compartment of undifferentiated cells committed to differentiation. One of the daughter cells is placed in the selected site, and the other daughter cell is placed in a random vacant site. (d) Symmetric differentiation a full site in the compartment of undifferentiated cells committed to differentiation, in the yellow box, is selected. The cell in the selected site divides, and both daughter cells leave the compartment of undifferentiated cells committed to differentiation, and begin to differentiate.

fate of the undifferentiated cells. Not all undifferentiated cells can self-renew. Research indicates that stem cells are located in a restricted region of the bone marrow niche (Wineman et al., 1996; Lemischka, 1997; Bertolini et al., 1997; Aiuti et al., 1998; Thiemann et al., 1998; Verfaillie, 1998; Koller et al., 1999; Yin and Li, 2006; Zhang and Li, 2008; Cheng et al., 2014). In this paper, this region is referred to as the compartment of stem cells, whereas the compartment of committed undifferentiated cells refers to the region of the bone marrow niche which contains undifferentiated cells that can not self-renew. However, it is still unknown whether this representation gives an accurate description of the bone marrow niche in vivo: As discussed in the introduction, it is not possible to reconstruct a niche experimentally, and hence, relatively little is known about the exact behaviour of most types of undifferentiated cells, including the haematopoietic stem cells (Wineman et al., 1996; Lemischka, 1997; Bertolini et al., 1997; Aiuti et al., 1998; Thiemann et al., 1998; Verfaillie, 1998; Koller et al., 1999; Yin and Li, 2006; Zhang and Li, 2008; Cheng et al., 2014; Fuchs et al., 2004; Nikolova et al., 2007; Simons and Clevers, 2011). On the other hand, research on *Drosophila germline stem cells* provides a clear-cut example of how the stem cell compartment promotes stem cell maintenance (Yamashita et al., 2003; Morrison and Kimble, 2006; Wong et al., 2005). Germline stem cells are unique because they are solely dedicated to reproduction and transmission of genetic information. Exciting progress has been made in understanding molecular mechanisms underlying interactions between stem cells and stem cell compartments through the use of genetic techniques in *Drosophila* germline stem cells. The knowledge gained from studying the *Drosophila* germline stem cells has provided an intellectual framework for defining the niche and molecular regulatory mechanisms for other adult stem cells. The results on *Drosophila* germline stem cells have previously been used to describe systems and construct models of other types of stem cells, including the haematopoietic stem cells (Lemischka, 1997; Cinquin, 2009; He et al., 2009; Xia et al., 2012; Sada and Tumbar, 2013). The outcome of a *Drosophila* germline stem cell division depends on the spindle orientation relative to the Hub cells in the stem cell compartment, and the results from the unequal distribution of intracellular regulators and extracellular (Hub-derived) signals between daughter cells during mitosis. The result is that when a *Drosophila* germline stem cell divides under normal conditions, one daughter remains in the stem cell compartment and retains stem cell identity, and the other daughter is left outside the stem cell compartment and commits to differentiation. Yamashita et al. (2003), Morrison and Kimble (2006), Wong et al. (2005). This is a classical example of asymmetric stem cell division. Even though *Drosophila* germline stem cells normally divide asymmetrically, they can be induced to self-renew symmetrically to regenerate an additional stem cell after an experimental manipulation in which one stem cell is removed from the stem cell compartment. Thus, the experiments on *Drosophila* germline stem cells indicate that the stem cell compartment can contain up to a certain number of cells, and that the stem cell compartment is full under normal conditions. When a stem cell divides, one of the daughters inherits the mother's place in the stem cell compartment and retains stem cell identity. The fate of the other daughter depends on whether there is a vacant place in the stem cell compartment or not. If there is a vacant place in the stem cell compartment, the latter daughter remains in the stem cell compartment and retains stem cell identity. If the stem cell compartment is full, it is placed outside, and loses its stem cell identity. Hence, research on *Drosophila* germline stem cells implies that the stem cells do not "know" that they must divide asymmetrically or symmetrically, as discussed in Section 1.1. On the contrary, the stem cells divide at random, and the availability of the stem cell compartment, and perhaps other regulatory factors, determines whether the division is symmetric or asymmetric. This indicates that there are, in general, no phenotypic differences between daughter cells resulting from a symmetric and asymmetric stem cell division, which means that a cell must be in the stem cell compartment to

function as a stem cell: Once a cell is placed outside, it is no longer a stem cell.

Similar to the *Drosophila* germline stem cell compartment, the stem cell compartment in the haematopoietic bone marrow niche plays an important role in the regulation of haematopoietic stem cell organisation (Wineman et al., 1996; Lemischka, 1997; Bertolini et al., 1997; Aiuti et al., 1998; Thiemann et al., 1998; Verfaillie, 1998; Koller et al., 1999; Yin and Li, 2006; Zhang and Li, 2008; Cheng et al., 2014). Even though there are no in vivo experiments that reveal exactly how proliferation of the haematopoietic stem cells is regulated, it is known that self-renewal depends on local growth conditions, namely, on the direct contact between stem cells and stroma cells (Wineman et al., 1996; Verfaillie, 1998; Koller et al., 1999). The model presented in this paper assumes that the results obtained from the experiments on *Drosophila* germline stem cell compartment and the implications that follow from these results, also hold true for the bone marrow niche. The main idea is illustrated in Fig. 1 and explained more thoroughly in Section 2.

1.3. Haematopoietic cytokines and extracellular regulation

It is commonly accepted that all types of blood cells are generated by haematopoietic stem cells (Baum et al., 1992; Morrison and Weissman, 1994), and that these cells go through a number of divisions, obtaining various stages of differentiation, until the fully mature haematopoietic cells stop dividing. However, as discussed by Dingli et al. (2007) and Furusawa and Kaneko (2009), Furusawa and Kaneko (2012), there is no unambiguous determination of the number of stages connecting haematopoietic stem cells and fully mature cells, let alone how fast cells go through different stages of maturation and exactly how these processes are regulated (Donohue et al., 1958; Cronkite and Fliedner, 1964; Ogawa, 1993). *Haematopoietic cytokines* are extracellular signalling molecules that regulate the generation of haematopoietic cells (Aglietta et al., 1989; Layton et al., 1989; Metcalf, 2008; Fried, 2009). Each of these cytokines can regulate one specific lineage or multiple lineages. Individual haematopoietic cytokines have multiple actions mediated by receptors that can initiate various responses – differentiation, maturation, functional activation, survival and proliferation (Metcalf, 2008). Furthermore, for some cell types, such as haematopoietic stem cells and megakaryocyte progenitors, the simultaneous action of multiple cytokines are required for proliferative responses. One of the reasons why it is very challenging to establish the precise source of cytokines and predict their ultimate fate, is that the haematopoietic cytokines have many tissue sources, for instance lung, kidney, muscle, liver and membrane-displayed factors on local stromal cells (Aglietta et al., 1989; Metcalf, 2008). Several models have been proposed to investigate different feedback mechanisms (Roeder and Loeffler, 2002; Fuchs et al., 2004; Nikolova et al., 2007; Simons and Clevers, 2011; Wong et al., 2005; Cinquin, 2009; He et al., 2009; Xia et al., 2012; Sada and Tumbar, 2013; Furusawa and Kaneko, 2009, 2012; Donohue et al., 1958; Cronkite and Fliedner, 1964; Ogawa, 1993; Aglietta et al., 1989; Layton et al., 1989; Metcalf, 2008; Fried, 2009; Potten and Loeffler, 1990; Wodarz, 2008; Lander et al., 2009; Høyem et al., 2015; Larsen, 2016; Mangel et al., 2016; Rompolas et al., 2016). Results from theoretical work modelling the haematopoietic system (Wodarz, 2008) and crypt cells (Potten and Loeffler, 1990) imply that changes in stem cell number and their cyclic activity are associated with changes in the demand of the mature cell stages. Lander et al. (2009) explore how secreted negative feedback factors may be used to control the output of multistage cell lineages, as exemplified by the actions of GDF11 and activin in a self-renewing neural tissue, the mammalian olfactory epithelium. The results by Lander et al. indicate that two feedback loops are in general better than one. That is, good control (robustness, stability, low progenitor load, and fast regeneration from a variety of conditions) is found over an increasing fraction of the parameter space when feedback loops are

added. These results might also apply to the haematopoietic system. Similar to the models presented in Dingli et al. (2007) and Høyem et al. (2015), we model differentiation as a multi-step process where cell replication and differentiation are coupled with cells moving through successive stages – compartments – of maturation in a series of steps from the haematopoietic stem cells all the way down to the fully differentiated haematopoietic cells.

1.4. Models for haematopoiesis in female Safari cats

The experiments on female Safari cats (Abkowitz et al., 1988, 1990, 1993) have inspired several mathematical models (Guttorp et al., 1990; Newton et al., 1995; Abkowitz et al., 1996; Golinelli et al., 2006; Fong et al., 2009). In 1990, Guttorp et al. proposed a state-space Markov model for haematopoiesis in Safari cats (Guttorp et al., 1990). It is assumed that in each cat there is a large pool of haematopoietic stem cells, and that a proportion p of these stem cells express d G6PD. The proportion p may vary between cats, but remains constant within each cat. The authors suppose that most haematopoietic stem cells are not involved in the production of mature blood cells, but are members of a primary pool of slowly self-replicating cells. A relatively small number of haematopoietic stem cells produce mature blood cells through asymmetric division and differentiation, and are referred to as *active stem cells*. It is assumed that the number of active stem cells is constant, N, and that the active stem cells do not have the ability to self-renew symmetrically. Consequently, when an active stem cell dies, a member of the primary pool of slowly self-replicating stem cells must become an active stem cell, in order to keep the number of active stem cells constant. Since N is much smaller than the total number of haematopoietic stem cells, the number of active stem cells that express d G6PD can be between 0 and N, even though the proportion of haematopoietic stem cell expressing d G6PD is constant. Indeed, the probability that i of the active stem cells express d G6PD is given by the probability mass function of the binomial distribution:

$$P_i = \binom{N}{i} p^i (1 - p)^{N-i}.$$

Moreover, suppose that there are i active stem cells expressing d G6PD. When an active stem cell dies, the number of active stem cells expressing d G6PD can either increase by one, decrease by one or remain constant. The conditional probabilities for these three events are

$$P(i + 1, i) = \left(1 - \frac{i}{N}\right)p, \, P(i - 1, i) = \frac{i}{N}(1 - p), \, P$$
$$(i, i) = 1 - \left(1 - \frac{i}{N}\right)p - \frac{i}{N}(1 - p),$$

respectively. Although the model proposed by Guttorp et al. can explain some of the results from the experiments on female Safari cats (Abkowitz et al., 1988, 1990, 1993), for instance that the proportion of cells expressing d G6PD remained relatively constant under normal conditions, the model cannot explain the results that indicate that the proportion of cells expressing d G6PD can change during regeneration. The reason for this is that Guttorp et. al. assume that stem cell self-replication is a deterministic process such that the proportion of d G6PD remains constant. The models presented in Abkowitz et al. (1996), Golinelli et al. (2006), Fong et al. (2009) and the model presented in this paper assume that self-replication is a stochastic process.

In 1995, Newton et al. (1995) used a simple stochastic model, similar to the model presented by Guttorp et al. (1990), to quantify the relationship between observed proportions of progenitors expressing d G6PD and unobserved haematopoietic stem cell populations.

Abkowitz et al. stimulated haematopoiesis by assuming that all stem cell decisions, that is, replication, apoptosis and initiation of

differentiation, are determined by chance (Abkowitz et al., 1996). The paper was published in 1996. They show that stochastic stem cell behaviour can result in a wide spectrum of discrete outcomes observed in vivo (Abkowitz et al., 1988, 1990, 1993), and that clonal dominance can occur by chance. More precisely, each haematopoietic stem cell is randomly selected for replication, apoptosis (cell death) and differentiation at constant rates λ, α and μ, respectively. Furthermore, the probability that a stem cell is selected for replication is much higher than the probability that a stem cell is selected for apoptosis or differentiation, i.e.

$$\lambda > \alpha + \mu.$$

This means that the number of stem cells increases when the haematopoietic system regenerates after injury. When the number of stem cells reaches a certain limit, the stem cells ignore the signals that tell them to reproduce. This means that each stem cell must keep track of the total number of stem cells. In our paper, an alternative strategy is investigated, where the rates of replication and differentiation depend on the number of stem cells and undifferentiated committed cells. That is, when cells need to be replaced, the rate of symmetric stem cell division increases, whereas under normal conditions, the stem cells divide mostly asymmetrically.

In 2006, Golinelli et al. published a paper (Golinelli et al., 2006) that describe a stochastic process used to model early haematopoiesis in continuous time. The haematopoietic stem cells follow a simple linear birth-death process where each stem cell can either self-renew symmetrically or differentiate into a progenitor cell at constant rates λ and ν, respectively. Similar to the model presented in Abkowitz et al. (1988), the rates satisfy

$$\lambda > \nu,$$

so the stem cells can regenerate after injury. Moreover, if the stem cell compartment is full and a stem cell self-renews symmetrically, then a random stem cell dies.

Fong et al. (2009) performed Bayesian statistical inference on extensions of the model proposed by Golinelli et al. (2006), in order to determine if haematopoietic stem cell decisions are linked to cell divisions or occur independently. This paper was published in 2009. Their results show that haematopoietic stem cells must divide symmetrically in order to maintain haematopoiesis. They also demonstrate that a model that adds asymmetric division events provides a better fit to the competitive transplantation data. The conclusions drawn by Fong et al. correspond well with the results of this paper. However, unlike the model investigated by Fong et al., stemness is not treated as an explicit cellular property in this paper, but as the result of a dynamic process of regulation and self-organisation similar to the models presented by Loeffler and Roeder (2002), Roeder and Loeffler (2002).

2. Model of haematopoiesis with self-organisation

In this section, we present a compartmental model of the haematopoietic system with self-organisation. The model can reproduce several of the results from the experiments with female Safari cats (Abkowitz et al., 1988, 1990, 1993). At the root of the model are the stem cells, located in the SC-compartment. It is assumed that the committed cells go through K stages of differentiation. A committed cell at stage i is denoted DC^i and is located in the DC^i-compartment for $0 \leqslant i \leqslant K$. The dynamics of the compartments of undifferentiated cells are described in Section 2.1, whereas in Section 2.4, the differentiated cells are also included.

The results from the experiments on *Drosophila* germline stem cells (Yamashita et al., 2003; Morrison and Kimble, 2006; Wong et al., 2005) and female Safari cats (Abkowitz et al., 1988, 1990, 1993) which can be reproduced by our model, are discussed in Section 3, and the biological processes that the model are based on are examined in Section 4, whereas in this section we mainly focus on describing the model.

2.1. Compartments of undifferentiated cells

As discussed in Section 1, undifferentiated haematopoietic cells are, in general, located in the bone marrow. The model presented in this paper subdivides these cells into two groups: the undifferentiated cells located in the SC-compartment and the undifferentiated cells located in the DC^0-compartment. It is assumed that these two groups of cells are phenotypically identical. However, the cells located in the former compartment are stem cells because they self-renew and produce differentiated cells, whereas the cells in the DC^0-compartment are committed to differentiation and cannot self-renew, and hence, they are not stem cells. The compartments of undifferentiated cells regulate symmetric and asymmetric stem cell division. The main idea is that under steady-state the stem cells divide mostly asymmetrically, whereas when cells need to be replaced due to tissue damage, the stem cells start to divide symmetrically. Both compartments contain M sites. Each of the $2M$ sites can either contain exactly one cell or no cell, denoted *full* sites and *vacant* sites, respectively. Thus, $2M$ represents the carrying capacity of the bone marrow niche. Under steady-state there are approximately M cells in both compartments, and the stem cells typically divide asymmetrically – one daughter cell inherits the mother's site and the other daughter is placed in a vacant site in the DC^0-compartment. The DC^0s migrate to the DC^1-compartment when they divide and obtain the first stage of differentiation.

It is known that the number of undifferentiated cells can increase markedly when they are regenerated after injury to the bone marrow (Abkowitz et al., 1990, 1993; Morrison et al., 1997; Reya et al., 2001; Yamashita et al., 2003; Morrison and Kimble, 2006; McKenzie et al., 2006; Dingli et al., 2007). This type of injury is modelled by decreasing the number of cells in the SC-compartment and DC^0-compartment well below M. The stem cells start to divide symmetrically after injury to the compartments of undifferentiated cells. It is symmetric self-renewal if one daughter cell inherits the mother's site while the other daughter is placed in a vacant site in the SC-compartment, and symmetric commitment if both daughter cells are placed in vacant sites in the DC^0-compartment.

2.2. Markov process

The population dynamics in the compartments of undifferentiated cells, described in Section 2.1, are implemented by the following Markov process: At each elementary event, a random site and a random stem cell are selected. If a site in the SC-compartment is selected, the selected stem cell divides. One of the daughter cells inherits the mother's site. If the selected site is full, then the second daughter cell migrates to the DC^0-compartment, and is placed in a random vacant site, i.e. the division is asymmetric (see Fig. 1 (a)). If the selected site is vacant, the second daughter is placed in this site, resulting in symmetric self-renewal (see Fig. 1 (b)). On the other hand, suppose that a random site in the DC^0-compartment is selected. If the selected site is vacant, the selected stem cell commits symmetrically to differentiation, and both daughter cells are placed in random vacant sites in the DC^0-compartment (see Fig. 1 (c)). If the selected site is full, this cell leaves the DC^0-compartment (see Fig. 1 (d)). For boundary conditions, it is assumed that when all the sites in the SC-compartment are vacant, a cell from another SC-compartment migrates to the empty SC-compartment, so that symmetric division is possible. Moreover, it is assumed that when all the sites in the DC^0-compartment are full, then any cell that enters the DC^0-compartment undergoes apoptosis, i.e. programmed cell death. Thus, given that there are I stem cells and J DC^0s, we obtain the following transition probabilities:

$$P_{I,J}(I, J - 1) = \frac{1}{2}\frac{J}{M},$$
(1)

$$P_{I,J}(I, J + 1) = \frac{1}{2}\frac{I}{M},$$
(2)

$$P_{I,J}(I + 1, J) = \frac{1}{2}\left(1 - \frac{I}{M}\right),$$
(3)

$$P_{I,J}(I - 1, J + 2) = \frac{1}{2}\left(1 - \frac{J}{M}\right).$$
(4)

That is, the conditional probability that a cell leaves the DC^0-compartment is given in (1), a stem cell divides asymmetrically is given in (2), a stem cell self-renews symmetrically is given in (3), and a stem cell commits symmetrically to differentiation in given in (4). Let $X(\Gamma)$ and $Y(\Gamma)$ be the expected number of cells in the SC-compartment and DC^0-compartment, respectively, at elementary event Γ. It follows from Eqs. (1)–(4) that

$$X(\Gamma + 1) = X(\Gamma) + \frac{1}{2M}((M - X(\Gamma)) - (M - Y(\Gamma)))$$
$$= X(\Gamma) + \frac{1}{2M}(Y(\Gamma) - X(\Gamma)),$$
(5)

$$Y(\Gamma + 1) = Y(\Gamma) + \frac{1}{2M}(X(\Gamma) + 2(M - Y(\Gamma)) - Y(\Gamma))$$
$$= X(\Gamma) + \frac{1}{2M}(X(\Gamma) + 2M - 3Y(\Gamma)),$$
(6)

for $0 < X(\Gamma)$ and $Y(\Gamma) \leqslant M - 2$. When the SC-compartment is empty, the number of stem cells increases by two after symmetric self-renewal. Moreover, when there is only one vacant site in the DC^0-compartment, one of the daughters undergoes apoptosis when a stem cell commits symmetrically to differentiation, whereas if there are no vacant sites, both daughters undergo apoptosis. For simplicity, these boundary conditions are neglected in the following approximation of the mean functions: First, the system of linear difference equations given in (5)–(6) has exactly one equilibrium solution, namely

$$(X^*, Y^*) = (M, M),$$

which means that all sites in both compartments are full. The corresponding transition matrix is

$$\frac{1}{2M}\begin{bmatrix} -1 & 1 \\ 1 & -3 \end{bmatrix}$$

and the eigenvalues are

$$\lambda_1 = -2 + \sqrt{2}, \quad \lambda_2 = -2 - \sqrt{2}.$$

An eigenvector corresponding to λ_i is

$$\mathbf{v}_i = \begin{bmatrix} 1 \\ 1 - \lambda_i \end{bmatrix},$$

for $i \in \{1, 2\}$. One time step is defined as $2M$ elementary events. It follows that the expected number of stem cells and DC^0s at time step t are approximately

$$X(t) = M + c_1 \exp\left(\frac{\lambda_1}{2M}t\right) + c_2 \exp\left(\frac{\lambda_2}{2M}t\right),$$
(7)

$$Y(t) = M + c_1(1 + \lambda_1)\exp\left(\frac{\lambda_1}{2M}t\right) + c_2(1 + \lambda_2)\exp\left(\frac{\lambda_2}{2M}t\right),$$
(8)

respectively, where

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{(\lambda_2 - \lambda_1)M}\begin{bmatrix} 1 + \lambda_2 & -1 \\ -(1 + \lambda_1) & 1 \end{bmatrix}\left(\begin{bmatrix} M \\ M \end{bmatrix} - \begin{bmatrix} X(0) \\ Y(0) \end{bmatrix}\right)$$

and $X(0)$ and $Y(0)$ are the initial number of stem cells and DC^0s, respectively. It follows from (7) and (8) that it is expected that the system converges towards the steady state where both compartments are (approximately) full. However, given that the process runs long enough, stochastic realisation will lead to extinction of one of the phenotypes with probability one. As illustrated in the next subsection, for small populations, one phenotype gets extinct after a relative short time period, whereas for sufficiently large populations, both compart-
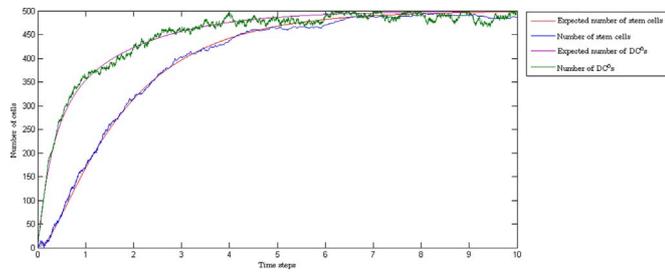
Fig. 2. Regeneration of the undifferentiated cells The compartments of undifferentiated cells are regenerated, starting with a single stem cell, with compartment size $M = 500$. The red and the purple smooth curves show the expected numbers of stem cells and DC^{0}s, respectively, and the jagged blue and green curves are simulations of stem cells and DC^{0}s, respectively.

ments of undifferentiated cells remain approximately full under normal conditions for any time interval corresponding to the lifetime of a mammal.

2.3. Numerical simulations

Fig. 2 shows the regeneration of the population of undifferentiated cells, starting with a single stem cell. The red and purple smooth curves illustrate the approximation of the expected number of stem cells and DC^{0}s, given in Eqs. (7) and (8), respectively, whereas the jagged curves are simulations of the population dynamics described in Section 2.2. The figure illustrates that when the compartment size is sufficiently large, the simulations fit the expected numbers of undifferentiated cells well: The number of cells in both compartments grow steadily until the compartments are approximately full. The DC^{0}s grow significantly faster than the stem cells. Under stable, normal conditions, the number of cells in both compartments remain close to M.

The approximations of the expected number of undifferentiated cells, given in Eqs. (7) and (8), indicate that under steady-state, both compartments remain approximately full. In general, the simulations become more similar to the expected functions as the number of sites increases. In Fig. 3, different compartment sizes are tested. Figs. 3 (a)–(b) and (c)–(d), with compartment size M=10 and M=20, respectively, illustrate that the model works poorly with relatively small compartment sizes. In Figs. 3 (a)–(b), the number of stem cells is zero 24 times during 10^4 time steps. When the compartment size is M=20, as illustrated in Figs. 3 (c)–(d), extinction of stem cells has not been observed during simulations when both compartments were initially full. However, the number of cells in both compartments vary too much to be a realistic representation of the bone marrow niche. Figs. 3 (e)–(f) show that when $M = 50$, the number of cells in each compartment remains relatively close to 50. Moreover, several results from experiments by Abkowitz et al. (1988, 1990, 1993) can be reproduced by the model when $M = 50$. This corresponds well with the model by Abkowitz et al. (1996) – 50 is the minimum size of the stem cell compartment in their model. However, in the remaining examples, the compartment sizes are larger than M=50. Figs. 3 (g)–(h), (i)–(j) and (k)–(l), where the compartment sizes are 100, 500 and 1000, respectively, illustrate that the number of undifferentiated cells varies less as the compartment size increases.

In all remaining examples, we use compartment size $M = 500$, which makes it easy to compare the results obtained in the different examples. Moreover, we want to compare our results with the results obtained from the previous models (Guttorp et al., 1990; Newton et al., 1995; Abkowitz et al., 1996; Golinelli et al., 2006; Fong et al., 2009) based on the experiments on Safari cats, and in particular, with the results obtained by Abkowitz et al. (1996). In the model by Abkowitz et al., all undifferentiated cells are stem cells, and in their numerical simulations, the stem cell compartment can contain up to 750 undifferentiated cells, whereas there can be up to 500 stem cells and

1000 undifferentiated cells in the continuing examples of this paper.

In the remaining examples, the value $\mu - s$ will be referred to as the lower limit for *normal population level*, where μ is the estimated mean number of cells in a given compartment and s is the estimated standard deviation. When the mean numbers of cells in all compartments are approximately the same as the estimated mean, and, at the same time, the standard deviations are approximately equal to the estimated standard deviations, the system is said be in *stable, normal state*.

When the cells are subdivided into two neutral phenotypes, such as cells expressing G G6PD and d G6PD for the Safari cats, the percentage of cells that express one type is expected to remain constant. Indeed, Figs. 4 (a)–(b) show a numerical example where the percentage of d G6PD-positive cells varies relatively little during stable, normal conditions. Fig. 4 (b) displays the percentage of self-renewal divisions that are symmetric and illustrates that under stable, normal conditions, the stem cells generally divide asymmetrically. Indeed, on average, 2.34% of the self-renewals are symmetric. On the other hand, Figs. 4 (c)–(d) show a numerical example where the percentage of cells expressing d G6PD varies extensively during regeneration. Initially, 5% of the sites in both compartments are full. After 4 time steps, the DC^{0}s reach the normal population level, whereas the stem cells reach the normal population level at time step $t = 5.7$. The percentage of self-renewals that are symmetric during regeneration is shown in Fig. 4 (d), and illustrates that when a large proportion of the sites are vacant, the stem cells divide symmetrically at a high rate, and as the number of cells in both compartment gradually increase, the rate of symmetric division steadily decreases. Fig. 5 shows twelve numerical examples of regeneration, where the initial conditions are the same as in the example illustrated in Figs. 4 (c)–(d). The curves in each of these examples are unique, which corresponds well with the experimental and theoretical work by Abkowitz et al. (1988), Abkowitz et al. (1990), Abkowitz et al. (1993), Abkowitz et al. (1996). Moreover, the time the cell population uses to reach normal population levels also varies – in Fig. 5 (d), the stem cells reach normal population level at time step $t = 7.5$, whereas in Fig. 5 (f), normal population level is reached after 5 time steps.

As illustrated in Figs. 5 and 6, the system has not, in general, gained stable, normal state when it reaches normal population level after regeneration – the DC^{0}s reach normal level before the stem cells, and this causes an intermediate time interval with relatively high variance in the cell number. For instance, consider Figs. 6 (a)–(b): In the time interval 6.5–50, where both compartments have reached normal level, the mean percentage of cells expressing the d G6PD phenotype is 62% and the standard deviation is 1.9%. On the contrary, the standard deviation is 0.8% in the numerical example plotted in Figs. 6 (c)–(d), where the system is in stable, normal state with mean percentage of cells expressing the d G6PD phenotype equal to 62%. Since symmetric stem cell division causes variation in the cell number, it is reasonable to expect that the stem cells self-renew symmetrically at a higher rate in the intermediate time interval with high variance than under stable, normal conditions, and, indeed, it follows from Fig. 6 (b) that at time step $t = 6$ and $t = 7$, the percentage of symmetric self-renewal is above 9% and 5%, respectively, which is rarely observed under stable, normal state. Moreover, the mean percentage of symmetric self-renewal is 2.7% during the time interval 6.5–50 in Fig. 6 (b), whereas the estimated mean during stable, normal state is 2.3%. The intermediate time interval with high variance has more apparent effect on the population dynamics when compartments of differentiated cells are included, and is investigated more thoroughly in Section 2.6.

In our simulations of regeneration, illustrated in Figs. 5 and 6, the average time the population of cells uses to reach normal population level is 6.2 time steps. There are no in vivo data for the undifferentiated cells in the bone marrow niche. However, experiments on Safari cats showed that bone marrow BFU-E and CFU-GM, as well as progenitor cell-cycle kinetics, returned to baseline values a hundred weeks after transplantation, on average (Abkowitz et al., 1988, 1990, 1993). Moreover, the pattern of clonal contribution to haematopoiesis in each
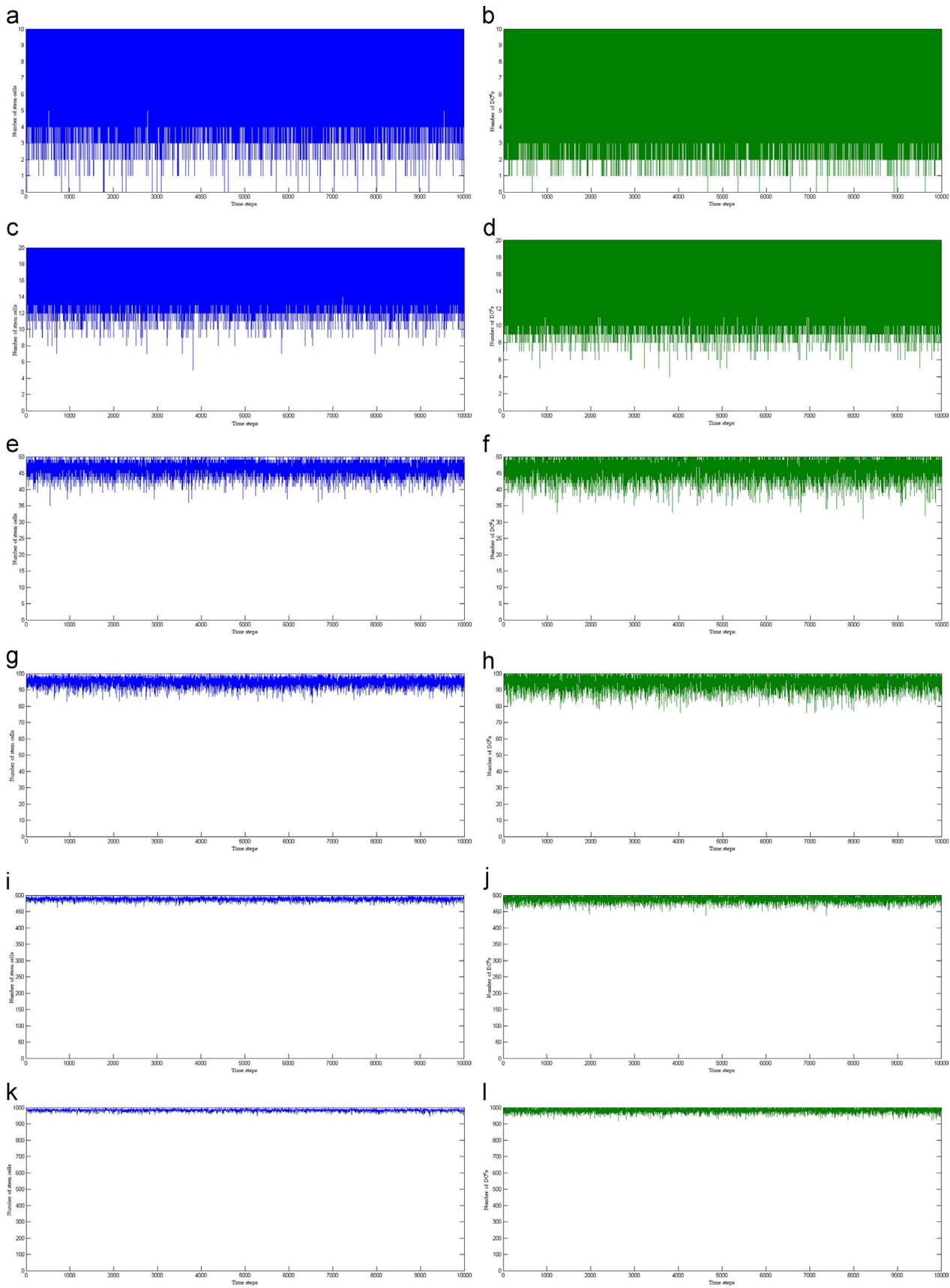
Fig. 3. Different compartment sizes Initially, all sites contain one cell. Different compartment sizes are tested, and it is verified that the simulations become increasingly more similar to the expected functions as the number of sites increases. (a) and (b) display the stem cells and DC^0, respectively, for $M = 10$. The system is highly unstable. (c) and (d) display the stem cells and DC^0, respectively, for $M = 20$. The system is unstable. (e) and (f) display the stem cells and DC^0, respectively, for $M = 50$. The system is quite stable. (g) and (h) display the stem cells and DC^0, respectively, for $M = 100$. The system is stable. (i) and (j) display the stem cells and DC^0, respectively, for $M = 500$. The system is stable. (k) and (l) display the stem cells and DC^0, respectively, for $M = 1000$. The system is stable.
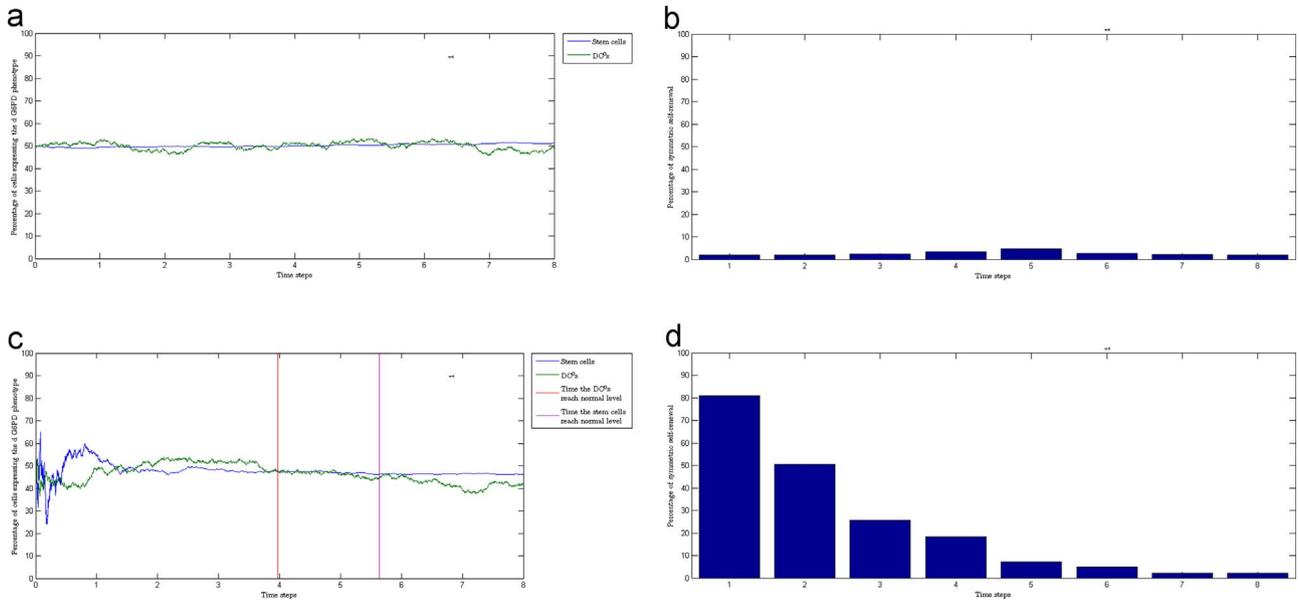
Fig. 4. Stable, normal conditions versus regeneration This figure illustrates that when the system is in stable, normal conditions, the percentage of cells expressing d G6PD is approximately constant and the stem cells typically divide asymmetrically, whereas when the system regenerates, the percentage of cells expressing d G6PD varies extensively, and the rate of symmetric division is relatively high. In both simulations, the compartment size is $M = 500$. (a) displays the percentage of stem cells and DC^0s expressing d G6PD when the system is in stable, normal state. (b) displays the percentage of self-renewals that is symmetric when the system is in stable, normal state. (c) displays the percentage of stem cells and DC^0s expressing d G6PD when the system regenerates and, initially, 20 (d) displays the percentage of self-renewals that are symmetric when the system regenerates.

cat was unique, and, in some cats, significant variation in the percentage of cells expressing d G6PD and D G6PD was observed for years after the number of cells reached normal population levels, whereas in other cats, the percentage remained approximately constant. The uniqueness and variation observed in vivo are, to some extension, captured by our model: In our simulations, the minimum number of time steps until normal population is reached, is five, and the maximum number of time steps is fifty percent greater, and, as discussed above, the system has not, in general, gained normal state when it reaches normal population level after regeneration – the system enters an intermediate time interval with high variance. On the contrary, for the model of haematopoiesis in Safari cats by Abkowitz et al. (1996), the time the system uses to regenerate varies little – less than five percent, and once the system reaches normal population size, it behaves exactly as under normal conditions.

2.4. Multi-compartmental model

In this subsection, the differentiated cells are also included in the model. That is, it is assumed that the committed cells go through K stages of differentiation, and that a cell at stage i in the differentiation process, denoted DC^i, is located in the DC^i-compartment for $0 \leqslant i \leqslant K$. All the cells in these compartments are *committed* to differentiation. However, the DC^0s are still undifferentiated whereas the DC^js, for $0 < j$, are actual differentiated cells. Moreover, when a cell in the DC^j-compartment divides, for $0 \leqslant j < K$, both daughter cells migrate to the DC^{j+1}-compartment. The cells in the DC^K-compartment are fully differentiated and stop dividing. The DC^i-compartment contains $2^i M$ sites. The sites in the compartments of differentiated cells are not just concrete, physical locations, but more abstract, representing the sum of signals in the environment of the cells. Similar to the compartments of undifferentiated cells, the sites in the compartments of differentiated cells are called vacant when they contain no cell, and unlike the compartments of undifferentiated cells, the full sites in a compartment of differentiated cells can contain more than one cell if all the other sites in this compartment are full. Under stable, normal conditions, there are approximately $2^i M$ cells in the DC^i-compartment for $0 \leqslant i \leqslant K$, and the cells commit symmetrically to differentiation at the same, constant rate. On the other hand, when the number of cells in

the DC^{i+1}-compartment is significantly less than under normal conditions, the rate of symmetric commitment in the DC^i-compartment increases.

2.5. Extended markov process

The population dynamics of the multi-compartmental model are implemented by the following Markov process: At each elementary event, a random site is selected. Each site in the $K + 2$ compartments has the same probability of being selected. If a site in a compartment of undifferentiated cells is selected, the elementary event is as described in Section 2.2, whereas if a site in the DC^i-compartment is selected, for $1 \leqslant i \leqslant K$, and the site is full, then, for $i < K$, a DC^i divides symmetrically and both daughter cells migrate to the DC^{i+1}-compartment, i.e. symmetric commitment, while for $i = K$, a cell in this compartment dies. On the other hand, if the selected site is vacant, then a random cell from the DC^{i-1}-compartment commits symmetrically to differentiation. For boundary conditions, it is assumed that if a vacant site in the DC^{i+1}-compartment is selected and the DC^i-compartment is empty, then the process finds the highest integer j, where $0 \leqslant j < i$, such that the DC^j-compartment is not empty, and a random DC^j commits symmetrically to differentiation. If all DC^j-compartments are empty for $j < i$, then a random stem cell commits symmetrically to differentiation.

Given that there are I and J^i full sites in the SC-compartment and DC^i-compartment, respectively, for $0 \leqslant i \leqslant K$ and $0 < J^i \leqslant 2^i M$ for $0 < i$, we obtain the following transition probabilities:

$$P_{J^K}(J^K - 1) = \frac{J^K}{2^{K+1}M}, \tag{9}$$

$$P_{J^i, J^{i+1}}(J^i - 1, J^{i+1} + 2) = \frac{2^{i+1}M + J^i - J^{i+1}}{2^{K+1}M}, \tag{10}$$

$$P_{I, J^0}(I, J^0 + 1) = \frac{I}{2^{K+1}M}, \tag{11}$$

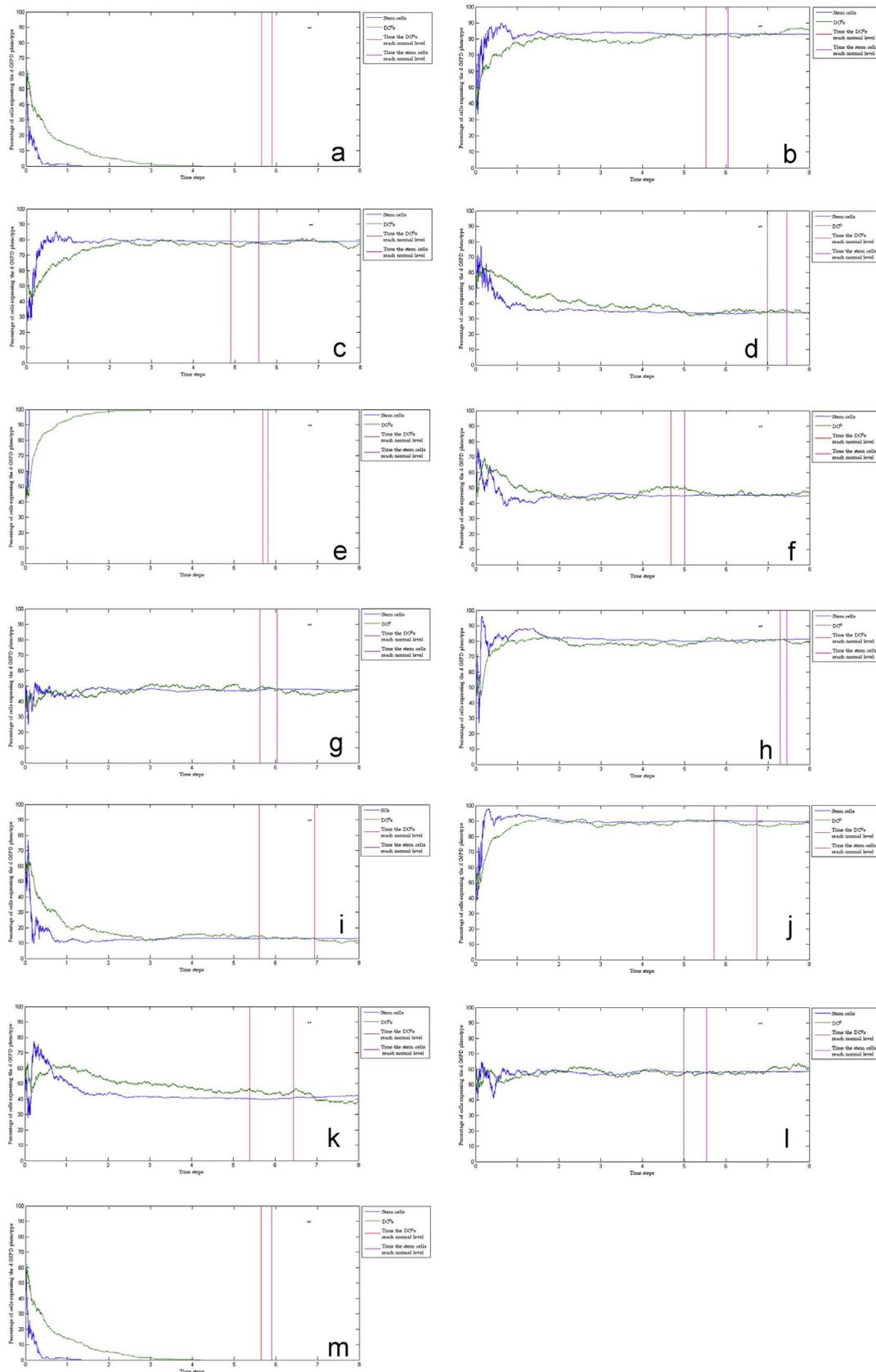$$P_{I, J^0}(I + 1, J^0) = \frac{M - I}{2^{K+1}M}, \tag{12}$$

Fig. 5. Unique traits of regeneration This figure displays twelve simulations of regeneration where, initially, 20% of the sites are full, and illustrates that every regeneration is unique. In all simulations, the compartment size is $M = 500$.
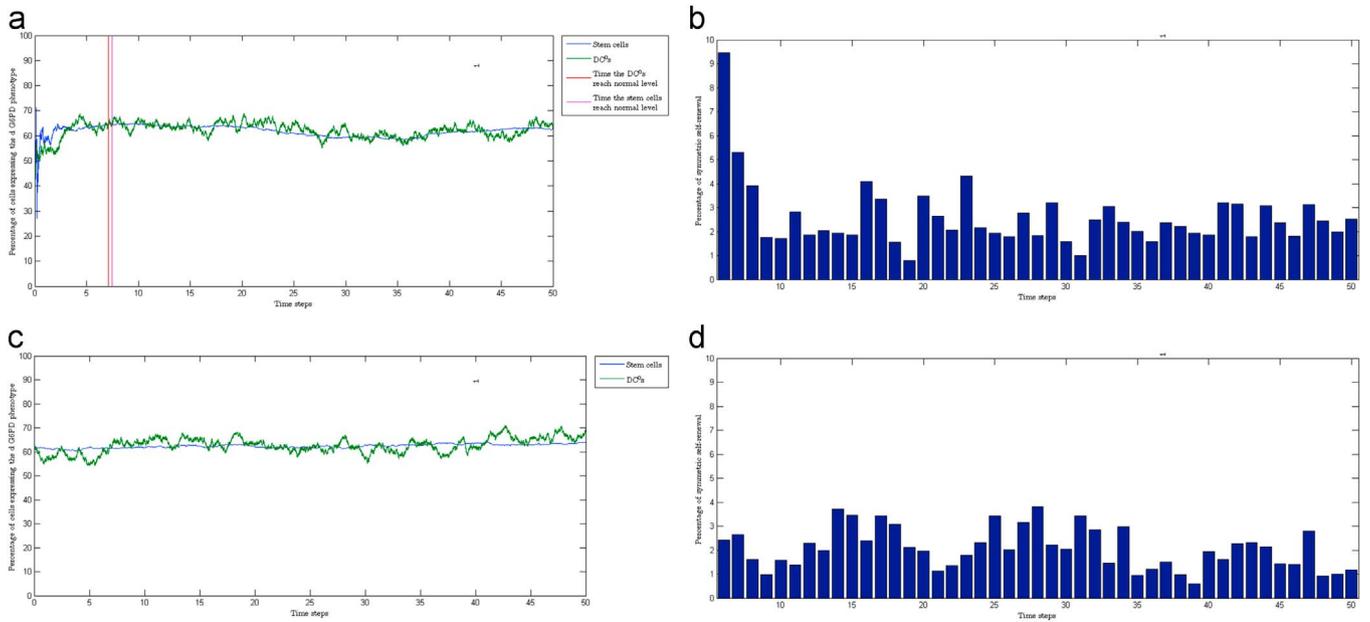
Fig. 6. Stable, normal conditions versus intermediate time interval with high variance This figure illustrates that when the system reaches normal population levels, the stem cells continue to divide symmetrically at a slightly higher rate than under stable, normal conditions, and consequently, the percentage of cells expressing d G6PD might vary more in the intermediate time interval with high variance than under stable, normal conditions. The compartment size is $M = 500$. (a) The percentage of stem cells and DC^0s expressing d G6PD under regeneration and the intermediate time interval with high variance. (b) The percentage of self-renewals that are symmetric in the intermediate time interval with high variance. (c) The percentage of stem cells and DC^0s expressing d G6PD under stable, normal conditions. (d) The percentage of self-renewals that are symmetric under stable, normal conditions.

$$P_{I,J^0}(I-1, J^0+2) = \frac{M-J^0}{2^{K+1}M}. \tag{13}$$

That is, the conditional probability that a DC^K is selected to die is given in (9), a DC^i commits symmetrically to differentiation is given in (10), a stem cell divides asymmetrically is given in (11), a stem cell self-renews symmetrically is given in (12), and a stem cell commits symmetrically to differentiation is given in (13). Let $X(\Gamma)$ and $Y^i(\Gamma)$ be the expected number of cells in the SC-compartment and DC^i-compartment, respectively, at elementary event Γ. It follows from Eqs. (9)–(13), given that $0 < X(\Gamma), Y^0(\Gamma) \leqslant M-2$ and $Y^j \leqslant 2^j M$, for $0 < j$, we have that

$$X(\Gamma + 1) = X(\Gamma) + \frac{1}{2^{K+1}M}(Y^0(\Gamma) - X(\Gamma)), \tag{14}$$

$$Y^0(\Gamma + 1) = Y^0(\Gamma) + \frac{1}{2^{K+1}M}(X(\Gamma) + Y^1(\Gamma) - 3Y^0(\Gamma)), \tag{15}$$

$$Y^j(\Gamma + 1) = Y^j(\Gamma) + \frac{1}{2^{K+1}M}(2Y^{j-1}(\Gamma) + Y^{j+1}(\Gamma) - 3Y^j(\Gamma)), \tag{16}$$

$$Y^K(\Gamma + 1) = Y^K(\Gamma) + \frac{1}{2^{K+1}M}(2Y^{2-1}(\Gamma) + 2^K M - 3Y^j(\Gamma)), \tag{17}$$

where $0 < j < K$. Because of the boundary conditions when the compartments of differentiated cells are empty, it is not possible to derive a simple approximation of the mean function, as it was for the model of undifferentiated cells illustrated in Fig. 2. Hence, we simply inspect the stability of the system of linear difference equations given in (14)–(17). The system has exactly one equilibrium solution, namely

$$(X^*, Y^{0*},...,Y^{i*}, ...Y^{K*}) = (M, M,...,2^j M,...,2^K M).$$

The corresponding transition matrix is:

$$\begin{bmatrix} -1 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -3 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 2 & -3 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 2 & -3 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 2 & -3 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 2 & -3 \end{bmatrix}, \tag{18}$$

It follows from the work by Kulkarni et al. (1999) that the correspond-

ing eigenvalues are negative (see Appendix A). Hence, if all sites are initially vacant or contain exactly one cell, it is expected that the number of cells increases until approximately all sites are full.

2.6. Numerical simulations

Since there are M sites in the SC-compartment and $2^i M$ sites in the DC^i-compartment, for $0 \leqslant i \leqslant K$, it follows that the total number of sites in the multi-compartmental model is

$$M\left(1 + \sum_{i=0}^{K} 2^i\right) = M\left(1 + \frac{1 - 2^{K+1}}{1 - 2}\right) = 2^{K+1}M.$$

In the numerical examples in this subsection, one time step consists of $2^{K+1}M$ elementary events. Since each site has the same probability of being selected at any elementary event, it follows that, on average, each site is selected once during a time step.

Fig. 7 shows the multi-compartmental model in stable, normal state. In Figs. 7 (a)–(b), the ratio

$$\frac{\text{number of cells in compartment}}{\text{number of sites in the compartment}}$$

is plotted for cells of all stages in the multi-compartmental model. The figures verify that under stable, normal state, all sites contain approximately one cell. Since the number of cells in the compartments of undifferentiated cells cannot exceed M, the corresponding ratios remain under one. On the other hand, the sites in the compartments of differentiated cells may contain more than one cell. Consequently, the corresponding ratios fluctuate over one. Fig. 7 (c) shows the percentage of self-renewal divisions that are symmetric. The estimated mean is 2.46%. This verifies that during normal conditions the stem cells divide mainly asymmetrically. Consequently, the number of stem cells fluctuate less than the number of DC^0s and the number of differentiated cells with compartments sizes that are relatively small. Indeed, Fig. 7 (d) displays the intervals $(\mu - s, \mu + s)$ for all compartments, where μ is the estimated mean in a given compartment and s is the estimated standard deviation. It can be verified that the estimated coefficient of variation, s/μ, is significantly larger for the DC^i-compart-
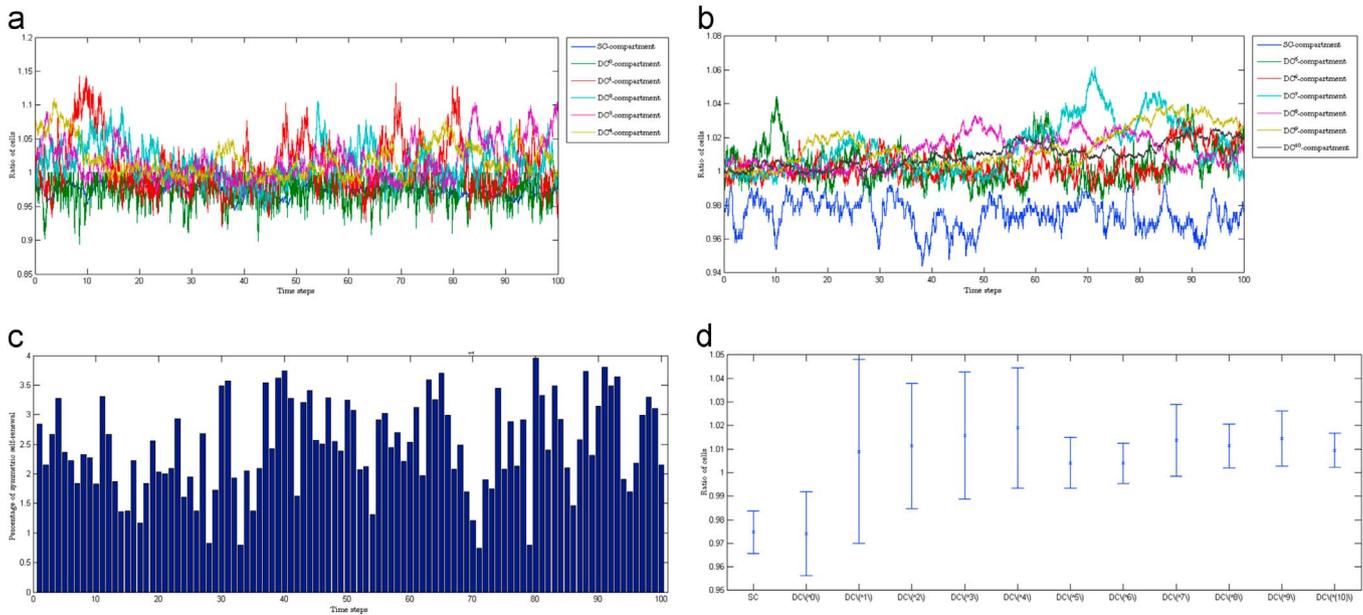
Fig. 7. The multi-compartmental model under stable, normal conditions This figure illustrates the multi-compartmental model under stable, normal conditions. The compartment size is $M = 500$. The ratio (number of cells in compartment)/(number of sites in compartment) is plotted for all compartments. (a) The ratio of cells in the SC-compartment and DC^i-compartment, for $0 \leqslant i \leqslant 4$. (b) The ratio of cells in the SC-compartment and DC^i-compartment, for $4 \leqslant i \leqslant 10$. (c) The percentage of self-renewal divisions that are symmetric. (d) The intervals $(\mu - s, \mu + s)$ for all compartments, where μ is the estimated mean in a given compartment and s is the estimated standard deviation.

ment than for the SC-compartment for $i \leqslant 4$, whereas for $4 < i$, the DC^i-compartment has smaller or approximately the same estimated coefficient of variation as the SC-compartment. We will use the values $\mu - s$, given in Fig. 7 (d), as lower limits for normal population levels in the examples where the multi-compartmental model is regenerated.

Fig. 8 shows the regeneration of the whole system, starting with a single stem cell, and verifies that the number of cells converges towards the steady-state where all compartments are approximately full. It follows from Fig. 8 (a) that the DC^{i+1} s grow, in general, faster towards the normal population level than the DC^is, for $0 \leqslant i < 10$, and that the stem cells typically grow slowest. Fig. 8 (b), which displays the percentage of self-renewal divisions that are symmetric, verifies that during regeneration, the rate of symmetric self-renewal increases. I.e. in the beginning of the regeneration, the percentage of symmetric self-renewal is close to 100%, and it decreases steadily down to approximately 2.5%. All the differentiated cells have reached normal population levels at time step $t = 19$. However, the stem cells continue to self-renew symmetrically at a higher rate than what is observed under stable, normal state. This illustrates the phenomenon, denoted intermediate time interval with high variance, which occurs in all of our numerical trials: When the cells reach normal population level, the stem cells continue to self-renew symmetrically at a relative high rate for some period of time, before the rate stabilises at normal level, and the whole system enters stable, normal state. The time-laps from the moment the cells reach normal population level to the system reaches

stable, normal state, varies both in length and in how much it affects the population dynamics of the multi-compartmental model. In particular, when the cells are subdivided into two neutral phenotypes, such as G G6PD-positive and d G6PD-positive cells for the Safari cat, the percentage of cells that expresses each type might change radically during the intermediate time interval with high variance. When the system is in stable, normal state, the percentage of each phenotype remains approximately constant. This is illustrated in Figs. 9 and 10. The blue and green curves plotted in Fig. 9 are, respectively, the ratio of full sites in the SC-compartment and the ratio of stem cells expressing d G6PD when the multi-compartmental model regenerates. The initial conditions are that 70% of the sites in all compartments are vacant and that 40% of the cells in the SC-compartment express d G6PD. It follows from Fig. 9 (a) that the ratio of d G6PD-positive stem cells fluctuates most intensely during the first ten time steps. At time step $t = 22$, when all the compartments of committed cells have reached their normal population level, 26.12% of the stem cells express d G6PD. The stem cells reach their normal population level at time step $t = 26$, followed by a relatively long period with high fluctuation in the population size. When the system stabilises at stable, normal state at time step $t = 60$, the percentage of stem cells that express d G6PD is on average 14.69%. However, in other numerical trials the percentage of d G6PD-positive cells does not change significantly after the committed cells reach normal population level. For instance, in the example displayed in Fig. 10, where the multi-compartmental model is regenerated, starting
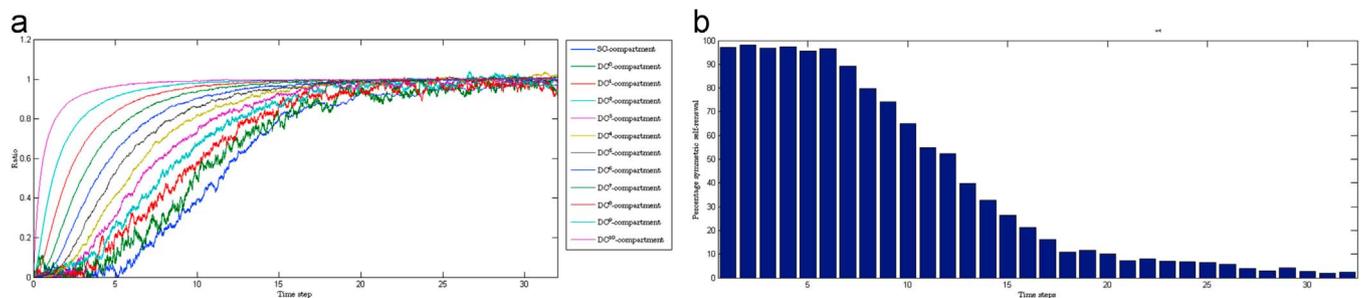

Fig. 8. Regeneration of the multi-compartmental model The whole system regenerates, starting with a single stem cell. The compartment size is $M = 500$. (a) displays the ratio (number of cells in compartment)/(number of sites in compartment) for all compartments. (b) displays the percentage of self-renewals that are symmetric.
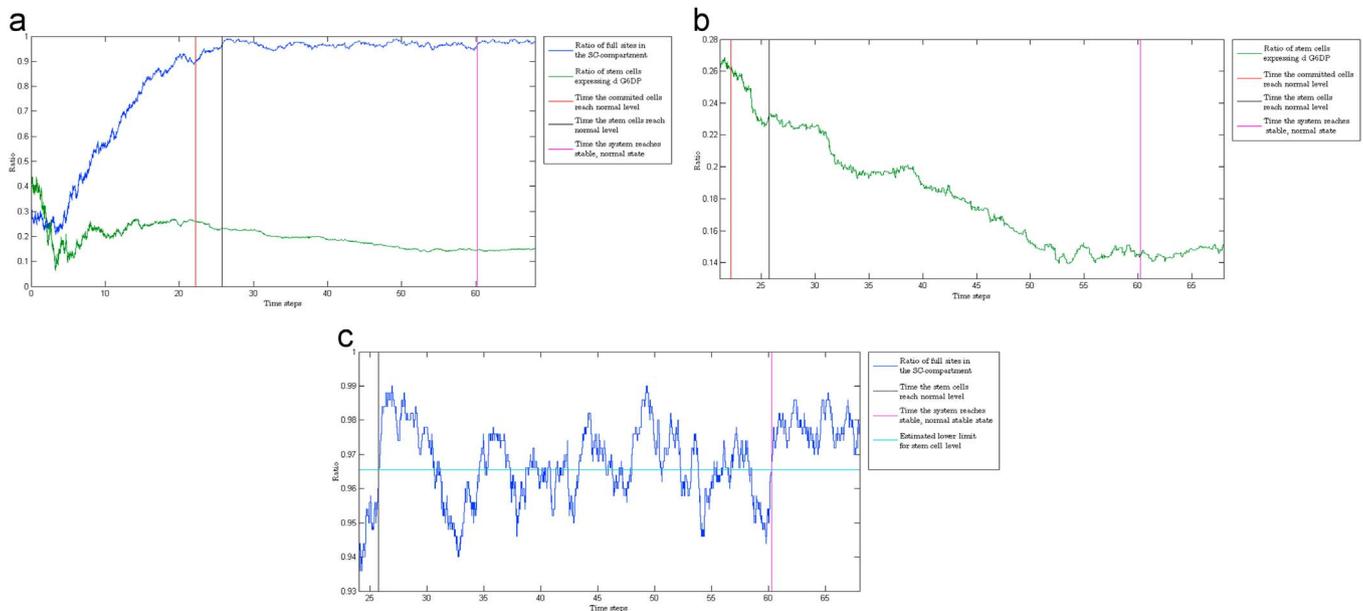
Fig. 9. The intermediate time interval with high variance This figure illustrates that the intermediate time interval with high variance can cause a significant fluctuation in the number of stem cells and the percentage of stem cells expressing d G6PD. Initially, 70% of the sites in all compartments are vacant and 40% of the stem cells express d G6PD. (a) displays the ratios (number of stem cell)/(number of sites) and (number of stem cell expressing d G6PD)/(number of stem cells) as blue and green curves, respectively, during regeneration, the intermediate time interval with high variance, and when the system reaches stable, normal conditions. (b) displays the ratio of stem cells expressing d G6PD during the intermediate time interval with high variance and when the system reaches stable, normal state. (c) displays the ratio of full sites in the SC-compartment during the intermediate time interval with high variance and when the system reaches stable, normal state. The horizontal, turquoise line is the estimated lower limit for stem cell level. During the intermediate time interval, the ratio of full sites in the SC-compartment is frequently below this limit, whereas under stable, normal conditions, the ratio is in general above this limit.

with 80% vacant sites in all compartments and 50% of the cells expressing d G6PD, the stem cells expressing d G6PD get extinct at time step $t = 3$, and eventually all the cells express G G6PD. As illustrated by Fig. 10 (b), the percentage of mature cells expressing d G6PD does not follow the fluctuation of d G6PD-positive stem cells during regeneration. In particular, the d G6PD-positive stem cells get extinct at time step $t = 3$ in the example illustrated in Fig. 10, though there are still mature cells expressing d G6PD at time step 13. However, at time step $t = 14$, all cells in the system are G G6PD-positive. Under stable, normal state, all the compartments have approximately the same percentage of d G6PD-positive cells as the SC-compartment. Hence, it is possible to estimate the percentage of d G6PD-positive stem cells by measuring the percentage of mature cells expressing d G6PD, under stable normal conditions.

In all examples each site is, on average, selected once during a time step. However, in 2.3, it is only the two compartments of undifferentiated cells that regenerate, whereas in this subsection, both the compartments of undifferentiated cells and the compartments of differentiated cells regenerate, and this is the reason why the average time the population of cells uses to reach normal population level is 6.2 time steps in the former subsection, whereas in this subsection, the average number of time steps is 27.6. As illustrated by Figs. 5, 6, 9 and

10, the intermediate time interval with high variance has more apparent effect on the population dynamics when compartments of differentiated cells are included. However, as discussed in Sections 4 and 5, none of our simulations could reproduce all the results from the experiments on Safari cats (Abkowitz et al., 1988, 1990, 1993).

3. Results

The compartmental model of haematopoiesis presented in this paper is inspired by the results from the experiments on the *Drosophila* germline stem cell compartment (Yamashita et al., 2003; Morrison and Kimble, 2006; Wong et al., 2005). As discussed in Section 1.2, these results support the following conjectures:

Conjectures.

I. The stem cell compartment promotes stem cell maintenance.
II. The stem cell compartment can contain up to a certain number of cells.
III. The stem cells self-renew at random.
IV. When a stem cell self-renews, one of the daughter cells inherits the mother's place in the stem cell compartment and retains stem cell identity, whereas the fate of the second daughter depends on the
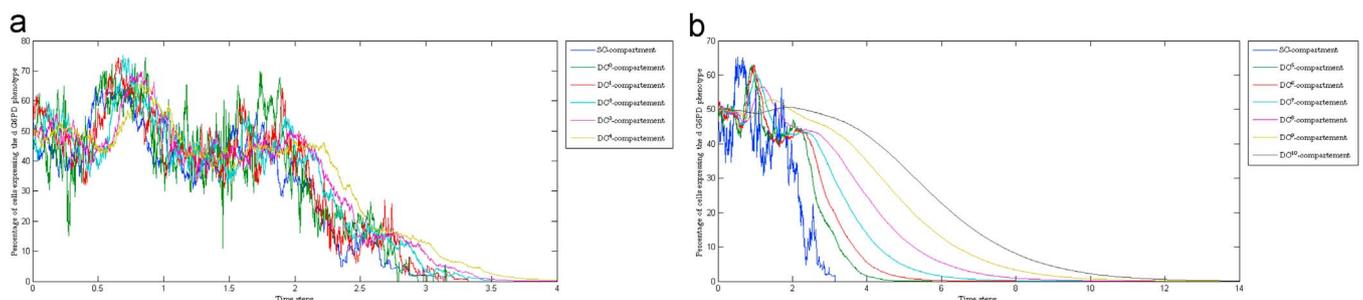

Fig. 10. The cells expressing d G6PD get extinct Initially, 80% of the sites in all compartments are vacant and 50% of the cells express d G6PD. At time step $t = 3$, the stem cells expressing d G6PD get extinct, and at time step $t = 14$, all the d G6PD-positive cells are extinct. The compartment size is $M = 500$. (a) displays the percentage of d G6PD-positive cells in the SC-compartment and DC^i- compartment, for $0 \leqslant i \leqslant 4$. (b) displays the percentage of d G6PD-positive cells in the SC-compartment and DC^i-compartment, for $5 \leqslant i \leqslant 10$.

availability of space in the stem cell compartment – it either slips into a random vacant place in the stem cell compartment and remains a stem cell (symmetric self-renewal), or the second daughter leaves the stem cell compartment and loses its stem cell identity (asymmetric self-renewal).

V. Under normal conditions, the stem cell compartment is approximately full, and the stem cells typically self-renew asymmetrically.

VI. When the stem cell compartment is not full, the rate of symmetric self-renewal generally increases, which leads to an expansion in the number of stem cells. The cells swift back to asymmetric self-renewal as the stem cell compartment reaches normal conditions.

As illustrated in Figs. 3 (a)–(d), if the number of sites in the SC-compartment is rather small, Conjectures IV−IV do not hold: The stem cells divide symmetrically at a relatively high rate when the SC-compartment is approximately full, causing a high variation in the number of cells, and more than not, the stem cell population continues to decrease when the number of cells in the SC-compartment is significantly less than the number of sites. In particular, when there are ten sites in the SC-compartment, the stem cell population frequently goes extinct. These results indicate that if the bone marrow niche can contain only a few active cells, then self-renewal is not a random process, but regulated deterministic. On the other hand, when there are fifty or more sites in the SC-compartment, all the conjectures hold (see Figs. 2, 3 (e)–(l), 4−10). This demonstrates that dynamic self-organisation of self-renewal and differentiation requires that the number of stem cells is sufficiently large.

The model can reproduce the following results from experiments with female Safari cats (Abkowitz et al., 1988, 1990, 1993):

Results from experiments.

I. The percentage of cells expressing d G6PD is approximately constant in healthy cats.

II. The pattern of clonal contribution to haematopoiesis is unique when the bone marrow regenerates. For instance, one of the G6PD phenotypes might get extinct during regeneration, but it is also possible that the percentage of each phenotype remains constant.

III. Significant variation in the percentage of cells expressing d G6PD might occur in a period after the cells have reached normal population level.

As shown in Figs. 4 (a)–(b), the percentage of cells expressing d G6PD varies relatively little under stable, normal conditions: The stem cells typically divide asymmetrically, and the daughter that inherits the mother's site has the same G6PD-phenotype as the mother. Figs. 5, 8–10 illustrate that each pattern during regeneration is unique: The rate of symmetric division increases, causing great fluctuation in the percentage of cells expressing d G6PD, and in some of the simulations, one of the G6PD phenotypes gets extinct. As demonstrated in Figs. 5, 6, 8–10, the system does not, in general, gain stable, normal condition when it reaches normal population levels after regeneration. Typically, the DC^{i+1} s grow faster towards normal population level than the DC^is, for $0 \leqslant i < K$, whereas the stem cells grow slowest. This causes an intermediate time interval where the number of cells varies more than under stable, normal condition. The time-laps from the moment the cells reach normal population level to the system reach stable, normal condition varies both in length and in how much it affects the population dynamics. In Fig. 9, the percentage of cells expressing d G6PD changes considerably during the intermediate interval with high variance, whereas in Figure 10, there is no significant change in the percentage after the cells reach normal population level. Hence, the model can reproduce Results I–III. As discussed in Section 1.4, several other models recreate Results I-II. However, to our knowledge, none of the previous models describing haematopoiesis in female Safari cats can explain Result III.

Results from simulations.

I. For sufficiently large population sizes, the percentage of each phenotype remained approximately constant under stable, normal conditions.

II. Each regeneration was unique, both with respect to the number of time steps until normal population was reached and with respect to the percentage of each phenotype.

III. The system did not, in general, gain stable, normal condition when it reached normal population levels after regeneration, and in some simulations, variation in the percentage of each phenotype occurred in a period after the system reached normal population level.

4. Discussion

The model of haematopoiesis presented in this paper, includes flexible and dynamically regulated self-organisation based on extra-cellular regulations and cell–cell and cell–environment interactions. The classical definition of stem cells – an undifferentiated cell capable of self-renewal, production of a large number of differentiated cells, regenerating tissue after injury and a flexibility in the use of these options – is fundamentally based on a functional perspective. As discussed by Loeffler and Roeder (2002), when the definition of stem cells was first introduced, the flexibility criterion attracted little attention. However, several experimental results indicate that flexibility is a fundamental property of the stem cells. For instance, a level of flexibility was found for lineage specifications within the haematopoietic system (Zhang et al., 1999): Zhang et al. managed to bias the degree of lineage commitment by several maneuvers that altered the growth environment. The present explanation of the fluctuations observed in lineage specification is based on a dynamic network of interacting transcription factors involving the PU-1 and GATA molecules. Cross and Enver introduced the concept of fluctuating levels of transcription factors within the haematopoietic system with threshold-dependent commitment (Cross and Enver, 1997). Moreover, several experiments indicate that stem cells specified for one type of tissue (e.g. haematopoiesis) can be manipulated in such a way that they can act as stem cells for another tissue (e.g. neuronal, myogenic) (Bjornson et al., 1999; Brazelton et al., 2000; Seale and Rudnicki, 2000; Goodell et al., 2001). The growth environment seems to be an important factor when tissue specification of stem cells are redirected. These results might support to our assumption that self-renewal is a property of undifferentiated cells located in the stem cell compartment, and that once a cell leaves the stem cell compartment, it loses the ability to self-renew. This implies that the cells located in the stem cell compartment and the compartment of undifferentiated cells committed to differentiation are phenotypically identical and cannot be distinguished in a laboratory.

Theoretical work also implies that flexibility is one of the most fundamental properties of the stem cells, because models without self-organisation must, in general, require that the cells somehow know how to behave under different circumstances (Loeffler and Roeder, 2002). For instance, as discussed in Section 1.4, the model presented by Abkowitz et al. in Abkowitz et al. (1996) has no self-organisation, and assumes that the stem cells ignore the signals that tell them to self-renew symmetrically when the number of stem cells reaches a certain limit. This means that each stem cell must keep track of the total number of stem cells, in order to make the right decision. Moreover, Loeffler and Roeder argue that a number of models include assumptions about symmetric and asymmetric stem cell division that in one way or another requires that the cells somehow explicitly "know" how to behave (Vogel et al., 1968; A comprehensive mathematical model, 1980; Loeffler and Grossmann, 1991; Loeffler et al., 1993, 1997). Loeffler and Roeder conclude that such concepts are too rigorous and potentially misleading, and hence, no implications about symmetric or asymmetric stem cell division are included in the definition of tissue

stem cells in Loeffler and Roeder (2002). The model presented in this paper includes symmetric and asymmetric stem cell division. Moreover, even though the stem cells have the flexibility to undergo self-renewal, produce mature cells by differentiation, and regenerate undifferentiated cells and differentiated cells when necessary, each cell in the system behaves completely random. This is implemented by subdividing the compartments into sites which represent physical space as well as signals and the environment (Wineman et al., 1996; Lemischka, 1997; Bertolini et al., 1997; Aiuti et al., 1998; Thiemann et al., 1998; Verfaillie, 1998; Koller et al., 1999; Yin and Li, 2006; Zhang and Li, 2008; Cheng et al., 2014; Fuchs et al., 2004; Nikolova et al., 2007; Simons and Clevers, 2011; Aglietta et al., 1989; Layton et al., 1989; Metcalf, 2008; Fried, 2009). Moreover, it is assumed that when a cell in the SC-compartment divides, both daughter cells remain undifferentiated, whereas when a cell in any of the DC-compartments divides, both daughter cells are more differentiated than the mother cell. This means that when a cell migrates from the SC-compartment to the DC^0-compartment, no change occurs in the phenotype. However, the cell is no longer a stem cell, because when the cell divides in the DC^0-compartment, none of the daughter cells remain undifferentiated, which means that the cell can no longer self-renew (Yamashita et al., 2003; Morrison and Kimble, 2006; Wong et al., 2005). Furthermore, it is assumed that the cells in the DC^0-compartment release signals that inhibit migration from the SC-compartment to the DC^0-compartment. If a cell in the SC-compartment does not receive these signals when it divides, then both daughter cells migrate to the DC^0-compartment, which means that the cell commits symmetrically to differentiation. This is implemented by selecting a random site in the compartments of undifferentiated cells and a random stem cell. The absence of signals that inhibit migration from the SC-compartment to the DC^0-compartment is represented by selecting a vacant site in the DC^0-compartment, in which case the selected stem cell commits symmetrically to differentiation. Hence, symmetric commitment is a random event, and the probability that this type of division occurs increases when the number of cells in the DC^0-compartment decreases. On the other hand, if a cell in the SC-compartment receives the signals that inhibit migration when it divides, then one of the daughter cells inherits the site of the mother, whereas the other daughter cell is placed by a random site in the SC-compartment. If this site is vacant, then the second daughter cell inhabits the site, which means that the division is symmetric self-renewal. On the contrary, if the site is occupied by another cell, then the second daughter cell migrates to the DC^0-compartment. That is, the division is asymmetric. Thus, both symmetric self-renewal and asymmetric division are random events, and the probability that each of these divisions occurs increases and decreases, respectively, when the number of cells in the SC-compartment decreases. The model presented in this paper also assumes that the cells in the DC^i-compartment are regulated by negative feedback from the cells in the DC^{i+1}-compartment, for $0 \leqslant i < K$ (Aglietta et al., 1989; Layton et al., 1989; Metcalf, 2008; Fried, 2009). More precisely, the cells in the DC^{i+1}-compartment release signals that inhibit symmetric commitment in the DC^i-compartment, such that under normal conditions, the cells in the latter compartment differentiate symmetrically at approximately constant rate. However, if the concentration of the signals that inhibit symmetric commitment to differentiation decreases, the rate of this type of division increases. This is implemented by selecting a random site in the compartments of differentiated cells. The absence of signals that inhibit symmetric commitment in the DC^i-compartment is represented by selecting a vacant site in the DC^{i+1}-compartment, in which case a random DC^i commits symmetrically to differentiation. On the other hand, if the selected site is full, a DC^{i+1} commits symmetrically to differentiation if $i < K - 1$ or dies if $i = K - 1$. Consequently, the K feedback loops from the DC^{i+1}-compartment to the DC^i-compartment, for $0 \leqslant i < K$, ensure that the system of cells regenerates the differentiated cells after injury, even though each cell in the system behaves completely random.

The model presented in this paper is very simplistic and has only two parameters, M and K – the number of sites in the SC-compartment and the number of compartments of differentiated cells, respectively. It is possible to add more parameters to the model, for instance letting the SC-compartment and DC^0-compartment have different number of sites, or selecting random cells to undergo apoptosis. However, the scope of this model is to link self-organisation with symmetric and asymmetric cell division, and these parameters do not lead to the revealing of new structures or any other relevant information. Hence, we choose to keep the model simple and comprehensible with two parameters only. If we want to extend the model such that it becomes more realistic and sophisticated, several aspects should be addressed. For one thing, the extended model should divide the committed haematopoietic cells into the erythroid lineage, the lymphoid lineage and the myeloid lineage. The first lineage is composed of red blood cells, the second of immune cells and the third includes granulocytes, megakaryocytes and macrophages (Morrison and Weissman, 1994; Verfaillie, 1998; Gehling et al., 2000). As discussed in Section 1.3, it is still unclear exactly how differentiation of haematopoietic cells is regulated. More than half a century ago, Waddington (1957) presented an epigenetic landscape to describe the differentiation of cells as the trajectories of balls rolling at random into branching valleys, each of which represents a developmental state. Based on Waddington's model, Furusawa and Kaneko (2009) propose a dynamical system model of cells with intracellular protein expression dynamics and interactions with each other. The model predicts that cells with irregular, or chaotic, oscillations in gene expression dynamics have the potential to differentiate into other cell types. During development, such complex oscillations are lost successively, leading to loss of pluripotency. Their results are consistent with the view that pluripotency is a statistical property defined at the cellular population level, correlating with intra-sample heterogeneity, and driven by the degree of signalling promiscuity in cells.

Another aspect that should be addressed in a more realistic and sophisticated version of the model, is that the cells in the SC-compartment are homogeneous with respect to functionality in the model presented in this paper. Nevertheless, phenotypic heterogeneity has been observed in haematopoietic stem cells with regards to various markers (e.g. CD34, CD38, c-kit, Sca 1) (Uchida et al., 1993; Lord, 1997). Moreover, experiments by Sato et al. indicate that both CD34-positive and CD34-negative cells can be effective stem cells and that the cells can even alter the CD34 property (Sato et al., 1999). As discussed by Huss (2000), CD34-negative stem cells are considered to be predominantly part of the quiescent stem cell pool of the haematopoietic system, and it is possible that haematopoietic stem cells alter the CD34 property from positive to negative when they go from active to quiescent state, and vice versa. Roeder and Loeffler propose a single-based stochastic model of haematopoietic stem cells that includes quiescence (Roeder and Loeffler, 2002). This model does not incorporate regulation of asymmetric and symmetric stem cell division. However, similar to the model presented in this paper, the model by Roeder and Loeffler introduces a perspective on stem cell organisation where stemness is not treated as an explicit cellular property but as the result of a dynamic process of self-organisation. That is, the model makes the novel concept of within-tissue plasticity operational – within a range of potential options, individual cells may reversibly change their actual set of properties, like going from active to quiescent state and vice versa, depending on the influence of the local growth environment. Stochastic switching between the growth environments introduces fluctuations that eventually generate heterogeneity.

5. Conclusion

In this paper, a simplistic model of haematopoiesis that links self-organisation with symmetric and asymmetric cell division is proposed. Each cell in the system behaves randomly and the daughter cells

resulting from symmetric and asymmetric stem cell divisions are, in general, phenotypically identical, and still, the haematopoietic system has the flexibility to self-renew, produce mature cells by differentiation, and regenerate undifferentiated cells and differentiated cells when necessary, due to self-organisation. Moreover, the compartments of committed cells are regulated by feedback loops, so that the system of cells regenerates the differentiated cells after injury. To our best knowledge, no previous model implements symmetric and asymmetric division as the result of self-organisation. Different models of self-organisation are discussed and compared by Osborne et al. in Osborn et al. (2016). The model of self-organisation proposed by Loeffler and Roeder (2002), Roeder and Loeffler (2002) and the potential impact of the stem cell niche on asymmetry of stem cell fate are discussed by Roeder and Lorenz in Roeder and Lorenz (2006). The authors state that since no implications about symmetric or asymmetric stem cell division are included in the definition of tissue stem cells in Loeffler and Roeder (2002) and Roeder and Loeffler (2002), this perspective of stem cell organisation does explicitly preclude asymmetric cell division. However, Roeder and Lorenz suggest that symmetric cell fates might be indirectly linked to self-organisation. On the contrary, our model implements symmetric and asymmetric division as the direct result of self-organisation.

The model can reproduce several of the results from experiments with female Safari cats (Abkowitz et al., 1988, 1990, 1993). Similar to previous models of haematopoiesis in female Safari cats (Guttorp et al., 1990; Newton et al., 1995; Abkowitz et al., 1996; Golinelli et al., 2006; Fong et al., 2009), the model presented in this paper can explain why the percentage of d G6PD-positive cells is approximately constant in healthy cats, whereas the pattern of clonal contributions to haematopoiesis is unique when the bone marrow regenerates. In addition, the model indicates that self-organisation of haematopoiesis might cause significant variation in the percentage of d G6PD-positive cells after the number of cells has reached normal population level. In general, the DC^{i+1} s reach normal population level before the DC^is, for $0 \leqslant i \leqslant K$, whereas the stem cells grow slowest, and this generates an intermediate time interval with relative high rate of symmetric stem cell division and corresponding high variance in the cell number. Eventually, the system self-regulates such that the rate of symmetric stem cell division decreases and the system enters stable, normal state.

Several of the results from experiments with female Safari cats (Abkowitz et al., 1988, 1990, 1993) cannot be reproduced by the model. For instance, for the first 10–12 weeks after transplantation, the percent of progenitors with d-G6PD was unchanged from that observed prior to transplantation in each cat (Abkowitz et al., 1990). This might indicate that when the number of stem cells is very small, self-renewal is strictly regulated. It may also be the case that a relative large number of quiescent stem cells are activated. However, after 10–12 weeks, the percents of progenitors with d-G6PD fluctuated widely, which indicates that self-renewal occurs more randomly. The model presented in this paper does not capture the difference between before and after 10–12 weeks because the model is very simplistic. For instance, it assumes that self-renewal always occurs at random, and quiescent stem cells are not included in the model. Moreover, Abkowitz et al. (1988) found that when the peripheral blood counts and the number of marrow progenitors detected in culture had reached normal level, the percentages of erythroid burst-forming cells and granulocyte/macrophage colony-forming cells in DNA synthesis increased. The main reason that the model presented in this paper does not capture this result, is that the committed haematopoietic cells are not divided into different lineages.

It is possible to extend the model such that it becomes more realistic and sophisticated by including quiescence for the stem cells and subdividing the committed cells into different lineages, similar to the models presented by Roeder and Loeffler (2002) and Furusawa and Kaneko (2009), respectively. As discussed in Section 4, these two models are based on similar assumptions as the model presented in this paper, namely that everything is totally random at single cell level. However, self-organisation of the system of cells ensures that self-renewal, production of mature cells, and regeneration of undifferentiated cells and differentiated cells are well orchestrated. The intermediate time interval with high variance has an apparent effect on the population dynamics when the differentiated cells are included. However, none of our simulations could reproduce all the results obtained by Abkowitz et al. (1988), Abkowitz et al. (1990), Abkowitz et al. (1993). By including quiescence for the stem cells and subdividing the committed cells into different lineages, the model would become more complex and richer, and it might capture a broader spectre of the results from experiment on female Safari cats.

Appendix A. Appendix

The transition matrix given in (18) might be the most the most natural representation of the system given in (14)–(17). However, the number of cells at elementary event Γ may be given by the vector

$$(Y^K(\Gamma),\dots,Y^i(\Gamma),\dots,Y^0(\Gamma), X(\Gamma)),$$

and in this case the corresponding transition matrix is

$$\begin{bmatrix} -3 & 2 & 0 & 0 & 0 & \cdots & 0 \\ 1 & -3 & 2 & 0 & 0 & \cdots & 0 \\ 0 & 1 & -3 & 2 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -3 & 2 & 0 \\ 0 & \cdots & 0 & 0 & 1 & -3 & 1 \\ 0 & \cdots & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \tag{A.1}$$

Clearly, the stability of the system given in Eqs. (14)–(17) does not depend on the representation of the transition matrix.

Kulkarni et al. (1999) present the general $n \times n$ tridiagonal Toeplitz matrix, denoted $T_n(a, b, c)$, in Section 2. By letting $a = -3$, $b = 1$ and $c = 2$, we obtain the following tridiagonal Toeplitz matrix:

$$T_n(-3, 2, 1) = \begin{bmatrix} -3 & 2 & 0 & \cdots & 0 \\ 1 & -3 & 2 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & -3 & 2 \\ 0 & \cdots & 0 & 1 & -3 \end{bmatrix}.$$

It from Theorem 2.2 given in Kulkarni et al. (1999), that the eigenvalues of $T_n(-3, 1, 2)$ are

$$\lambda_k = -3 - i2\sqrt{6}\cos(k\pi/(n+1))$$

for $k \in \{1, 2, \ldots, n\}$.

Kulkarni et al. study the eigenvalues of those tridiagonal matrices with upper left blocks which are Toeplitz matrices. If we let $a = -3$, $b = 1$, $c = 2$, $a_1 = -1$ and $b_1 = c_1 = 1$ in the matrix presented in the first example, where k=1, given in Section 4 of Kulkarni et al. (1999), the pseudo-Toeplitz matrix, denoted $T_n^1(-3, 1, 2)$, is the transition matrix given in (A.1). By examining the roots of the characteristic polynomial of $(1/\sqrt{6})T_n^1(-3, 1, 2)$ and using the substitution $-3/\sqrt{6} - \lambda = 2x$, Kulkarni et al. show that the roots must satisfy the equation

$$4(1/\sqrt{2} + x)U_n(x) - U_{n-1}(x) = 0,$$

where $U_n(x)$ denotes the nth degree Chebyshev polynomial of the second kind. By studying intersections of graphs in the xy-plane, Kulkarni et al. show that if $b_1c_1 > 0$ and $bc > 0$, then the $(n + 1) \times (n + 1)$-matrix $T_n^1(a, b, c)$ has $n + 1$ real distinct eigenvalues. It can be verified by these graphs that for $a = -3$, $b = 1$, $c = 2$, $a_1 = -1$ and $b_1 = c_1 = 1$, all eigenvalues are negative.

References

Abkowitz, J., Linenberger, M., Persik, M., et al., 1993. Behavior of feline hematopoietic stem cells years after busulfan exposure. Blood 82, 2096–2103.

Abkowitz, J.L., Catlin, S.N., Guttorp, P., 1996. Evidence that hematopoiesis may be a stochastic process in vivo. Nat. Med. 2, 190–197.

Abkowitz, K., Linenberger, M., Persik, M., et al., 1988. Clonal evolution following chemotherapy-induced stem cell depletion in cats heterzygous for glucose6-phosphate dehydrogenase. Blood 71, 1687–1692.

Abkowitz, J., Linenberger, M., Newton, M. et al., 1990. Evidence for the maintenance of hematopoiesis in larger animals by the sequential activation of stem-cel,l clones. Proceedings of the National Academy of Science USA, vol. 87, pp. 9062–9066.

Aglietta, M., Piacibello, W., Sanavio, F., Stacchini, A., Apra, F., Schena, M., Gavosto, F., 1989. Kinetics of human hemopoietic cells after in vivo administration of granulocyte-macrophage colony-stimulating factor. J. Clin. Investig. 83, 551.

Aiuti, A., Friedrich, C., Sie, C.A., Gutierrez-Ramos, J.C., 1998. Identification of distinct elements of the stromal microenvironment that control human hematopoietic stem/progenitor cell growth and differentiation. Exp. Hematol. 26, 143–157.

Akita, M., Tanaka, K., Matsumoto, S., et al., 2013. Detection of the hematopoietic stem and progenitor cell marker CD133 during angiogenesis in three-dimensional collagen gel culture. Stem Cells Int..

Baum, C.M., Weissman, I.L., Tsukamoto, A.S., et al., 1992. Isolation of a candidate human hematopoietic stem-cell population. In: Proceedings of the National Academy of Sciences of the United States of America. vol 89, pp. 2804–2808.

Bertolini, F., Battaglia, M., Soligo, D., et al., 1997. "Stem cell candidates" puried by liquid culture in the presence of Steel factor, IL-3, and 5FU are strictly stroma-dependent and have myeloid, lymphoid, and mega karyocytic potential. Exp. Hematol. 26, 350–356.

Bjornson, C.R., Rietze, R.L., B.A. Reynolds, B.A., et al., 1999. Turning brain into blood: a hematopoietic fate adopted by adult neural stem cells in vivo. Science 283, 534–537.

Brazelton, T.R., Rossi, F.M., Keshet, G.I., Blau, H.M., 2000. From marrow to brain: expression of neuronal phenotypes in adult mice. Science 290, 1775–1779.

Cheng, C.C., Lee, Y.H., Lin, S.P., HuangFu, W.C., Liu, I.H., 2014. Cell-autonomous heparanase modulates self-renewal and migration in bone marrow-derived mesenchymal stem cells. J. Biomed. Sci. 21, 1–12.

Cinquin, O., 2009. Purpose and regulation of stem cells: a systems-biology view from the Caenorhabditis elegans germ line. J. Pathol. 217, 186–198.

Cronkite, E.P., Fliedner, T.M., 1964. Granulocytopoiesis. New Engl. J. Med. 270, 1347–1352.

Cross, M., Enver, T., 1997. The lineage commitment of haemopoietic progenitor cells. Curr. Opin. Genet. Dev. 7, 609–613.

Dingli, D., Traulsen, A., Michor, F., 2007. (A)symmetric stem cell replication and cancer. PLoS Comput. Biol. 3.

Donohue, D., Reiff, R., Hanson, M., Betson, Y., Finch, C., 1958. Quantitative measurement of the erythrocytic and granulocytic cells of the marrow and blood. J. Clin. Investig. 37, 1571–1576.

Fong, Y., Guttorp, P., Abkowitz, J., 2009. Bayesian inference and model choice in a hidden stochastic two-compartment model of hematopoietic stem cell fate decisions. Ann. Appl. Stat. 3, 1696.

Fried, W., 2009. Erythropoietin and erythropoiesis. Exp. Hematol. 37, 1007–1015.

Fuchs, E., Tumbar, T., Guasch, G., 2004. Socializing with the neighbors: stem cells and their niche. Cell 116, 769–778.

Furusawa, C., Kaneko, K., 2009. Chaotic expression dynamics implies pluripotency: when theory and experiment meet. Biol. Direct 4, 17.

Furusawa, C., Kaneko, K., 2012. A dynamical-systems view of stem cell biology. Science 338, 215–217.

Gehling, U.M., Ergün, S., Schumacher, U., et al., 2000. In vitro differentiation of endothelial cells from AC133-positive progenitor cells. Blood 95, 3106–3112.

Golinelli, D., Guttorp, P., Abkowitz, J.A., 2006. Bayesian inference in a hidden stochastic two-compartment model for feline hematopoiesis. Math. Med. Biol. 23, 153–172.

Goodell, M.A., Jackson, K.A., Majka, S.M., et al., 2001. Stem cell plasticity in muscle and bone marrow. Ann. N. Y. Acad. Sci. i 938, 208–220.

Guttorp, P., Newton, M.A., Abkowitz, J.L., 1990. A stochastic model for haematopoiesis in cats. Math. Med. Biol. 7, 125–143.

He, S., Nakada, D., Morrison, S.J., 2009. Mechanisms of stem cell self-renewal. Annu. Rev. Cell Dev. Biol. 25, 377–406.

Herrmann, M., Binder, A., Menzel, U., et al., 2014. CD34/CD133 enriched bone marrow progenitor cells promote neovascularization of tissue engineered constructs in vivo. Stem Cell Res. 13, 465–477.

Høyem, M.R., Måløy, F., Jakobsen, P., Brandsdal, B.O., 2015. Stem cell regulation: implications when differentiated cells regulate symmetric stem cell division. J. Theor. Biol. 380, 203–219.

Huss, R., 2000. Isolation of primary and immortalized CD34- hematopoietic and mesenchymal. Stem Cells Var. Sources Stem Cells 18, 1–9.

Koller, M.R., Oxender, M., Jensen, T.C., et al., 1999. Direct contact between CD34+lin-cells and stroma induces a soluble activity that specically increases primitive hematopoietic cell production. Exp. Hematol. 27, 734–742.

Kulkarni, D., Schmidt, D, Tsui, S., 1999. Eigenvalues of tridiagonal pseudo-Toeplitz matrices. Linear Algebra Appl. 297, 63–80.

Lander, A.D., Gokoffski, K.K., Wan, F.Y., et al., 2009. Cell lineages and the logic of proliferative control. PLOS Biol. 7, e15.

Larsen, J.C., 2016. Models of cancer growth. J. Appl. Math. Comput., 1–33.

Layton, J.E., Hockman, H., Sheridan, W.H., Morstyn, G., 1989. Evidence for a novel in vivo control mechanism of granulopoiesis: mature cell-related control of a regulatory growth factor. Blood 74, 1303–1307.

Lemischka, I.R., 1997. Microenvironmental regulation of hematopoietic stem cells. Stem Cells 15, 63–68.

Loeffler, M., Wichmann, H.E., 1980. A comprehensive mathematical model of stem cell proliferation which re-1055 produces most of the published experimental results. Cell Tissue Kinet. 13, 543–561.

Loeffler, M., Grossmann, B., 1991. A stochastic branching model with formation of subunits applied to the growth of intestinal crypts. J. Theor. Biol. 150, 175–191.

Loeffler, M., Roeder, I., 2002. Tissue stem cells: definition, plasticity, heterogeneity, self-organization and models-a conceptual approach. Cells Tissues Organs 171, 8–26.

Loeffler, M., Birke, A., Winton, D., Potten, C., 1993. Somatic mutation, monoclonality and stochastic models of stem cell organization in the intestinal crypt. J. Theor. Biol. 160, 471–491.

Loeffler, M., Bratke, T., Paulus, U., et al., 1997. Clonality and life cycles of intestinal crypts explained by a state dependent stochastic model of epithelial stem cell organization. J. Theor. Biol. 184, 42–54.

Lord, B.I., 1997. Biology of the Haemopoietic Stem Cell. In: Potten, C.S. (Ed.), Stem Cells. Academic Press, Cambridge, 401–422.

Mangel, M., Bonsall, M.B., Aboobaker, A., 2016. Feedback control in planarian stem cell systems. BMC Syst. Biol. 10, 17.

McKenzie, J.L., Gan, O.I., Doedens, M., Wang, J.C., Dick, J.E., 2006. Individual stem cells with highly variable proliferation and self-renewal properties comprise the human hematopoietic stem cell compartment. Nat. Immunol. 7, 1225–1233.

Metcalf, D., 2008. Hematopoietic cytokines. Blood, 111, pp. 485–491.

Morrison, S.J., Weissman, I.L., 1994. The long-term repopulating subset of hematopoietic stem cells is deterministic and isolatable by phenotype. Immunity 1, 661–673.

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

Morrison, S.J., Shah, N.M., Anderson, D.J., 1997. Regulatory mechanisms in stem cell biology. Cell 88, 287–298.

Newton, M.A., Guttorp, P., Catlin, S., Assuno, R., Abkowitz, J.L., 1995. Stochastic modeling of early hematopoiesis. J. Am. Stat. Assoc. 90, 1146–1155.

Nikolova, G., Strilic, B., Lammert, E., 2007. The vascular niche and its basement membrane. Trends Cell Biol. 17, 19–25.

Ogawa, M., 1993. Differentiation and proliferation of hematopoietic stem cells. Blood 81, 2844–2853.

Osborn, J., Fletcher, A., Pitt-Fancis, J., et al., 2016. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. Cold Spring Harbor Labs Journals, ⟨http://biorxiv.org/content/early/2016/09/09/074351⟩.

Potten, C.S., Loeffler, M., 1990. Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons crypt. Dev. 110, 1001–1020.

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

Roeder, I., Loeffler, M., 2002. A novel dynamic model of hematopoietic stem cell organization based on the concept of within-tissue plasticity. Exp. Hematol. 30, 853–861.

Roeder, I., Lorenz, R., 2006. Asymmetry of stem cell fate and the potential impact of the niche. Stem Cell Rev., 171–180.

Rompolas, P., Mesa, K.R., Kawaguchi, K., et al., 2016. Spatiotemporal coordination of stemcell commitment during epidermal homeostasis. American Association for the Advancement of Science, vol. 352, pp. 1471–1474.

Sada, A., Tumbar, T., 2013. New insight into mechanisms of stem cell daughter fate determination in regenerative tissues. Int. Rev. Cell Mol. Biol. 300.

Sato, T., Laver, J.H., Ogawa, M., 1999. Reversible expression of CD34 by murine hematopoietic stem cells I. Blood 94, 2548–2554.

Seale, P., Rudnicki, M.A., 2000. A new look at the origin, function, and stem-cell status of muscle satellite cells. Dev. Biol. 218, 115–124.

Shortman, K., Naik, S.H., 2007. Steady-state and inflammatory dendritic-cell development. Nat. Rev. Immunol. 7, 19–30.

Simons, B.D., Clevers, H., 2011. Strategies for homeostatic stem cell self-renewal in adult tissues. Cell 145, 851–862.

Stephen J. O'Brien, 1986. Molecular genetics in the domestic cat and its relatives. Trends Genetics, vol. 2, pp. 137–142.

Thiemann, F.T., Moore, K.A., Smogorzewska, E.M., et al., 1998. The murine stromal cell line AFT024 acts specically on human CD34+CD38- progenitorsto maintain primitive function and immuniphenotype in vitro. Exp. Hematol. 26, 612–619.

Uchida, N., Fleming, W.H., Alpern, E.J., Weissman, I.L., 1993. Heterogeneity of hematopoietic stem cells. Curr. Opin. Immunol. 5, 177–184.

Verfaillie, C.M., 1998. Adhesion receptors as regulators of the hematopoietic process. Blood 92, 2609–2612.

Vogel, H., Niewisch, G., Matioli, G., 1968. The self renewal probability of hemopoietic stem cells. J. Cell. Physiol. 72, 221–228.

Waddington, C.H., 1957. The Strategy of the Genes. George Allen and Unwin, London.

Wichmann, H.E., Loeffler, M., Schmitz, S., 1988. A concept of hemopoietic regulation and its biomathematical realization. Blood Cells 14, 411–429.

Wineman, J., Moore, K., Lemischka, I., Muller-Sieburg, C., 1996. Functional heterogeneity of the hematopoietic microenvironment: rare stromal elements maintain long-term repopulating stem cells. Blood 87, 4082–4090.

Wodarz, D., 2008. Stem cell regulation and the development of blast crisis in chronic myeloid leukemia: implications for the outcome of Imatinib treatment and discontinuation. Med. Hypotheses 70, 128–136.

Wodarz, D., Komarova, N.L., 2005. Computational Biology of Cancer: lecture Notes and Mathematical Modeling. World Scientific, London.

Wong, M.D., Jin, Z., Xie, T., 2005. Molecular mechanisms of germline stem cell regulation. Annu. Rev. Genet. 39, 173–195.

Xia, L., Zheng, X., Zheng, W., et al., 2012. The niche-dependent feedback loop generates a bmp activity gradient to determine the germline stem cell fate. Curr. Biol. 22, 515–521.

Yamashita, Y.M., Jones, D.L., Fuller, M.T., 2003. Orientation of asymmetric stem cell division by the APC tumor suppressor and centrosome. Science 301, 1547–1550.

Yin, T., Li, L., 2006. The stem cell niches in bone. J. Clin. Investig. 116, 1195–1201.

Zhang, J., Li, L., 2008. Stem cell niche-Microenvironment and beyond. J. Biol. Chem..

Zhang, P., Behre, G., Pan, J., et al., 1999. Negative cross-talk between hematopoietic regulators: GATA proteins repress PU.1. Proc. Natl. Acad. Sci. USA 96, 8705–8710.

An extended Moran process that captures the struggle for fitness

Marthe Måløy[a,*], Frode Måløy, Rafael Lahoz-Beltrá[b], Juan Carlos Nuño[c], Antonio Bru[d]

[a]*Department of Mathematics and Statistics, The Arctic University of Norway.*
[b]*Department of Biodiversity, Ecology and Evolution, Complutense University of Madrid, Spain.*
[c]*Departamento de Matemática Aplicada a los Recursos Naturales, Universidad Politécnica de Madrid, Spain*
[d]*Department of Applied Mathematics, Complutense University of Madrid, Spain*

Abstract

When a new type of individual appears in a stable population, the newcomer is typically not advantageous. Due to stochasticity, the new type can grow in numbers, but the newcomers can only become advantageous if they manage to change the environment in such a way that they increase their fitness. This dynamics is observed in several situations in which a relatively stable population is invaded by an alternative strategy, for instance the evolution of cooperation among bacteria, the invasion of cancer in a multicellular organism and the evolution of ideas that contradict social norms. These examples also show that, by generating different versions of itself, the new type increases the probability of winning the struggle for fitness. Our model captures the imposed cooperation whereby the first generation of newcomers dies while changing the environment such that the next generations become more advantageous.

Keywords: Evolutionary dynamics, Nonlinear dynamics, Mathematical modelling, Game theory, Cooperation

1. Introduction

When unconditional cooperators appear in a large group of defectors, they are exploited until they become extinct. The best possible scenario for this type of cooperators is to change the environment such that another type of cooperators that are regulated and only cooperate under certain conditions becomes
5 advantageous. Furthermore, when defectors appear in a regulated cooperation, the first generation of defectors typically dies while changing the environment such that the next generations become more advantageous; hence, cooperation

---
*Corresponding author
Email address: marthe.maloy@uit.no (Marthe Måløy)

is imposed on the defectors. In this paper, we propose a model that captures
this dynamics. More specifically, we introduce an extension of the Moran process whereby individuals can change the fitness landscape of the population by modifying the environment.

1.1. The Moran process

The Moran process represents the simplest possible stochastic model that captures the three basic building blocks of evolution – replication, mutation and selection – in a finite population [1],[2]. The process assumes that the population size is constant and that each type of individual has constant fitness. In each time step, a random individual is selected to reproduce and a random individual is selected to die. In one implementation of the Moran process, all individuals are initially of the same type, denoted the *wild type*. When a wild-type individual reproduces, a mutation that creates a new type of individuals, denoted the *mutant type*, occurs with probability u. It is assumed that no other mutation can occur. The wild type has reproductive rate 1, whereas the mutant type has reproductive rate r, where r is a non-negative constant. All individuals are selected to die at the same rate. Hence, the mutant type is advantageous if $r > 1$, neutral if $r = 1$ and disadvantageous if $r < 1$. In each time step, the number of mutants can increase by one, decrease by one or remain constant. The probabilities for these three events are

$$\text{P}(i+1|i) = \frac{u(N-i)+ri}{N-i+ir}\frac{N-i}{N}, \tag{1}$$

$$\text{P}(i-1|i) = \frac{(1-u)(N-i)}{N-i+ir}\frac{i}{N}, \tag{2}$$

$$\text{P}(i|i) = 1 - P(i+1|i) - P(i-1|i), \tag{3}$$

respectively, where N is the population size and i is the number of mutants. The model is discussed more thoroughly in Appendix A.

If the timescale of the mutants' fixation is much shorter than the timescale of mutation, then a lineage of mutants is likely to take over the whole population or become extinct before another lineage of mutants is created from the wild type. In this case, the probability that i mutants will eventually invade the whole population is

$$\rho_i = \frac{r^{N-i}\left(1-r^i\right)}{1-r^N} \tag{4}$$

if $r \neq 1$ and

$$\rho_i = \frac{i}{N} \tag{5}$$

if it is a neutral Moran process, that is, $r = 1$ [3].

The Moran process can also capture the competition dynamics between three types of individuals [3]–[6]. As discussed more thoroughly in Subsections 1.2

and 1.3, a mutant created in a stable population has in general low fitness,
because it is attacked by defence mechanisms that protect the stability of the
population. However, the first type of mutants, denoted *intermediate mutants*,
typically has a higher mutation rate than the wild type and can produce a
new type of mutants that avoids most of these attacks. This type of mutants is
denoted *resistant mutants*. The reproductive rates of the wild type, intermediate
mutants and resistant mutants are 1, r and r_1, respectively, where $r \leq 1 < r_1$.

As discussed more thoroughly by Wodarz and Komarova [3], Nowak et al. [7]
and Breivik [8], the ability to create new variants is important when a mutant
type invades a population. However, as discussed more thoroughly in the next
subsection, no individual has anything to gain from changing only its strategy
in an *evolutionarily stable population*, and this indicates that the mutants must
also change their environment to become advantageous.

In the extended Moran model presented in this paper, the fitness is not
constant. Similar to the model presented by Wodarz and Komarova [3], the
model presented in this paper considers three types of individuals, namely the
wild type, intermediate mutants and resistant mutants, but, in contrast to the
previous model, the resistant mutants become advantageous only if the mutants
manage to change the environment. However, changing the environment reduces
the fitness of the intermediate mutants; thus, there is a cost for the mutants.
In particular, there is a chance that the mutants will not produce a resistant
type; in this case, the mutants actually reduce their own fitness. To analyse this
dynamics, we use the results from evolutionary game theory, which is presented
in the next subsection.

1.2. Evolutionary game theory

Evolutionary game theory is the generic approach to evolutionary dynamics
[9],[10]. In these games, the fitness depends on the frequencies of the different
types in the population [2]. In contrast to traditional game theory, evolutionary
game theory does not rely on rationality [11]. Instead it considers a population of
individuals with fixed strategies that interact randomly. When two individuals
interact, each receives a payoff that depends on the strategy of both individuals.
The payoff is interpreted as fitness [12].

Table 1 shows the payoffs in a well-known game called the *prisoner's dilemma*
[2]. This game has two strategies: cooperation and defection. A group of
cooperators has higher fitness than a group of defectors. However, if a defector
and a cooperator meet, the defector receives a higher payoff than the cooperator,
and, what is more, the defectors are fitter than the cooperators in a mixed group.

In an evolutionary game, a mutation can change the strategy of an individ-
ual. In some cases, the mutation increases the fitness of the individual. For
instance, consider a group of cooperators with interactions that are captured by
the prisoner's dilemma. If a mutation causes an individual to change strategy
to defection, the individual increases its payoff. This means that cooperation
is an unstable strategy. On the contrary, a strategy is a *Nash equilibrium* if no
individual can deviate from this strategy and increase its payoff [13]. Defection

3

in the prisoner's dilemma is a Nash equilibrium because, if a defector mutates into a cooperator, it decreases its payoff.

A Nash equilibrium is also an *evolutionarily stable strategy* if selection opposes the invasion of an alternative strategy [9]. That is, if a sufficiently large population adopts an evolutionarily stable strategy, it cannot be invaded by any alternative strategy that is initially rare. For the prisoner's dilemma, defection is an evolutionarily stable strategy. Hence, cooperators cannot invade a large population of defectors of which interaction is captured by the prisoner's dilemma. However, as discussed more thoroughly in Section 4, a relatively small group of defectors can be invaded by cooperators.

The prisoner's dilemma illustrates why a well-functioning cooperation, such as a multicellular organism or society, must have control mechanisms that stabilise the cooperation and protect against defective individuals. Even though cooperations are not stable in general [2], the control mechanisms make them behave similarly to an evolutionarily stable population within relatively short timescales.

In the next subsection, we discuss some of the mechanisms that regulate cooperation in a multicellular organism, whereas the regulation of human interaction is examined in Subsection 4.3.

1.3. Regulation of cooperation in a multicellular organism

In a large multicellular organism, such as a human being, millions of cells must cooperate [3],[14]. This cooperation is maintained by a very complex network of signals and cellular checkpoints, and the immune system is an important component of this network. The immune system must detect mutant cells that have stopped cooperating as well as foreign agents, from viruses to parasitic worms, and distinguish them from the organism's own healthy tissue [15].

Mutated cells can be detected and killed by T cells, which are a type of white blood cells [16]. The exact details of how the T cells are regulated and activated are still uncertain [17]. In a nutshell, a type of T cells, called antigen-presenting cells (APCs), circulates with the blood. If an APC recognises a foreign protein, called an antigen, on a cell, then it makes a copy of the antigen and transports it to the lymph nodes. When the lymph nodes receive the antigen, the production of a type of T cells called cytoxic T lymphocytes (CTLs) is activated. A CTL is programmed to find and kill the cells that display the type of antigen brought to the lymph nodes by the APCs [18].

The body can also prevent the growth of mutated cells by limiting the blood supply. As discussed more thoroughly in Section 4, this can lead to acidification of the microenvironment, which increases the death rate of both mutant cells and healthy cells. However, a new type of mutant that is resistant to the acidic environment, might be created [19]–[24]. This competition dynamics is captured by the extended Moran process, presented in the next section.

4

2. Extended Moran process with non-constant fitness

In this section, we present an extension of the Moran process with non-constant fitness. The model assumes that the population has constant size, N, and that it consists of three types of individuals, namely the wild type, intermediate mutants and resistant mutants. During reproduction, a wild-type individual can mutate into an intermediate mutant with probability u and an intermediate mutant can mutate into a resistant mutant with probability u_1. It is assumed that no other mutation can occur.

The environment is described by a parameter called the *fitness parameter*. As long as the fitness parameter is below the *fitness threshold*, Υ, all individuals have the same fitness. The mutants increase the fitness parameter, and, when the fitness parameter reaches Υ, the fitness of the non-resistant individuals decreases, whereas the resistant mutants become advantageous.

In each time step, the following four events occur:

1. A random individual is selected to reproduce and a random individual is selected to die.

2. If the fitness parameter is higher than Υ, a random individual is selected. If the selected individual is not resistant, it dies, and a random individual reproduces, whereas, if the selected individual is resistant, nothing occurs in this event.

3. A random individual is selected. If it is a mutant, then $1/N$ is added to the fitness parameter.

4. The fitness parameter is reduced by $F \times 100$ per cent, where $0 \leq F \leq 1$.

Similar to the original Moran process, it is assumed that all individuals are selected simultaneously and randomly in events 1–3. Hence, if there are i mutants at the beginning of the time step, the probability of selecting a mutant in event 3 is i/N. In events 1 and 2, the probability of selecting a mutant for reproduction is also constant. However, the same individual cannot die twice; hence, if the same individual is selected to die in events 1 and event 2, a new random individual must be selected to die. Nevertheless, as shown in Subsection 2.2, for sufficiently large population sizes, the probability of selecting a mutant is approximately i/N in both events 1 and 2.

Initially, all the cells are the wild type and the fitness parameter equals zero. Eventually, a mutant is created, and it is assumed that the timescale of the mutants' fixation is much shorter than the timescale of mutation. Hence, a lineage of mutants is likely to take over the whole population or become extinct before another lineage of mutants is created from the wild type.

2.1. Event 1

Let i and j denote the numbers of intermediate mutants and resistant mutants, respectively, at the beginning of a given time step. Since the population size is constant, N, the number of wild-type individuals is $N - i - j$.

All individuals are selected to die and reproduce at the same constant rate. It is assumed that the same individual can be selected to reproduce and to die and that a new individual cannot be selected to die in the same time step in which it is produced. Thus, by ignoring further mutations, the probabilities that an intermediate mutant, a resistant mutant and a wild-type individual is selected to reproduce or to die are

$$P_{im} = \frac{i}{N}, \tag{6}$$

$$P_{rm} = \frac{j}{N}, \tag{7}$$

$$P_w = \frac{N - i - j}{N}, \tag{8}$$

respectively. Hence, we obtain the following transition probabilities for event 1:

$$P^1(i+1, j|i, j) = \frac{i}{N} \frac{N - i - j}{N},$$

$$P^1(i+1, j-1|i, j) = \frac{i}{N} \frac{j}{N},$$

$$P^1(i, j+1|i, j) = \frac{j}{N} \frac{N - i - j}{N},$$

$$P^1(i-1, j+1|i, j) = \frac{i}{N} \frac{j}{N},$$

$$P^1(i-1, j|i, j) = \frac{i}{N} \frac{N - i - j}{N},$$

$$P^1(i, j-1|i, j) = \frac{j}{N} \frac{N - i - j}{N},$$

$$P^1(i, j|i, j) = 1 - 2\frac{i}{N} \frac{N - i - j}{N} - 2\frac{i}{N} \frac{j}{N} - 2\frac{j}{N} \frac{N - i - j}{N}.$$

145 *2.2. Event 2*

If the fitness parameter is below Υ, then nothing occurs in event 2. On the other hand, if the fitness parameter is higher than Υ, then a non-resistant individual can be selected to die.

To obtain a simplistic model, we want the probability of selecting a given type of individual to be constant throughout the time step. As discussed in Appendix A, this is the case for the standard Moran process.

By assuming that a new individual cannot be selected to reproduce or die in the same time step in which it was produced, and that the same individual can be selected to reproduce several times and to die in the same time step, the probabilities that an individual selected to reproduce in event 2 is an intermediate mutant, a resistant mutant and a wild type are given in (6), (7) and (8), respectively.

On the other hand, the same individual cannot die several times. That is, the probability that an intermediate mutant will be selected to die in event 1

is i/N, and, in this case, the probability of selecting an intermediate mutant to die in event 2 is $(i-1)/(N-1)$. The probability that the individual selected to die in event 1 is not an intermediate mutant, is $1 - i/N$, and, in this case, the probability of selecting an intermediate mutant to die in event 2 is $i/(N-1)$. Thus, it follows by the rule of total probability that the probability that an intermediate mutant is selected to die in event 2, given that the fitness parameter is higher than Υ, is

$$P^2_{-im} = \frac{i}{N}\frac{i-1}{N-1} + \left(1 - \frac{i}{N}\right)\frac{i}{N-1} = \frac{i}{N-1} - \frac{1}{N(N-1)}.$$

For large population sizes, $1/(N-1) \approx 1/N$ and $1/N >> 1/N^2$. Hence, P^2_{-im} tends to

$$P^2_{-im} = \frac{i}{N}.$$

For similar reasons, if the fitness parameter is higher than Υ, then the probability that a wild-type individual is selected to die in event 2 tends to

$$P^2_{-w} = \frac{N-i-j}{N}$$

for large population sizes, and consequently the transition probabilities for event 2 are

$$P^2(i-1, j|i, j) = \frac{i}{N}\frac{N-i-j}{N}, \tag{9}$$

$$P^2(i, j-1|i, j) = \frac{j}{N}\frac{N-i-j}{N}, \tag{10}$$

$$P^2(i-1, j+1|i, j) = \frac{i}{N}\frac{j}{N}, \tag{11}$$

$$P^2(i+1, j-1|i, j) = \frac{i}{N}\frac{j}{N}, \tag{12}$$

$$P^2(i, j|i, j) = 1 - \frac{i}{N}\frac{N-i-j}{N} - 2\frac{i}{N}\frac{j}{N} - \frac{j}{N}\frac{N-i-j}{N}. \tag{13}$$

2.3. Events 3 and 4

Event 3 captures the assumption that the mutants raise the fitness parameter. The main reason why the fitness parameter is raised by $1/N$ in this event is that the growth environment is subdivided into N sites in Section 3.

Event 4 captures the diffusion of the fitness parameter. If $F = 0$, then the population is in an isolated growth environment, whereas, if $F = 1$, the fitness parameter decreases to zero at the end of every time step.

2.4. Expected functions

When the fitness parameter is lower than the fitness threshold Υ, the competition dynamics between the mutants and the wild type is identical to an

ordinary neutral Moran process. Thus, we are interested in how long it takes for the mutants to change the competition dynamics by increasing the fitness parameter to a level higher than Υ. In this subsection, we derive the expected time until the fitness parameter reaches this limit.

We expect that the number of mutants must reach a certain limit, ν, before the fitness parameter approaches Υ. Since the population dynamics is identical to a neutral Moran process when the fitness parameter is below Υ, we can use the following theorem to find the probability that the number of mutants will reach ν.

Theorem 2.1. *The probability that the neutral Moran process will reach the state in which there are ν mutants, given that the present number of mutants is i, is*

$$P(\text{reach } \nu \mid i \,) = \frac{i}{\nu},$$

where $0 \leq i \leq \nu$.

Theorem 2.1 is a standard result in Markov chain analysis [25]; hence, the proof is left to Appendix B.

It follows from Theorem 2.1 that most lineages of mutant cells become extinct before they reach the state ν if $\nu > 2$. We are interested in investigating the lineages that survive long enough for the fitness parameter to reach the threshold Υ.

Theorem 2.2. *Conditioning on the fact that the neutral Moran process eventually reaches the state in which there are ν mutants, the transition probabilities for $0 < i < \nu$ are*

$$P_\nu(i+1|i) = \frac{i+1}{N}\left(1 - \frac{i}{N}\right) \tag{14}$$

$$P_\nu(i-1|i) = \frac{i-1}{N}\left(1 - \frac{i}{N}\right) \tag{15}$$

$$P_\nu(i|i) = 1 - 2\frac{i}{N}\left(1 - \frac{i}{N}\right) \tag{16}$$

where i is the present number of mutants.

Proof. We have four events:

- A_1: the next time step moves to state $i+1$.

- A_2: the next time step moves to state $i-1$.

- B: the process is currently in state i.

- C: the process will reach state ν.

8

For $k \in 1, 2$, we want to determine the conditional probability

$$\mathrm{P}(A_k|B \cap C) = \frac{A_k \cap B \cap C}{B \cap C}.$$

It follows from Theorem 2.1 that

$$\mathrm{P}(\text{reach } \nu|i) = \mathrm{P}(C|B) = \frac{\mathrm{P}(B \cap C)}{\mathrm{P}(B)} = \frac{i}{\nu},$$

$$\mathrm{P}(\text{reach } \nu|i + 1) = \mathrm{P}(C|A_1 \cap B) = \frac{\mathrm{P}(A_1 \cap B \cap C)}{\mathrm{P}(A_1 \cap B)} = \frac{i+1}{\nu},$$

$$\mathrm{P}(\text{reach } \nu|i - 1) = \mathrm{P}(C|A_2 \cap B) = \frac{\mathrm{P}(A_2 \cap B \cap C)}{\mathrm{P}(A_2 \cap B)} = \frac{i-1}{\nu},$$

and it follows from the transition probabilities given in (1)–(3), with $u = 0$ and $r = 1$, that

$$\mathrm{P}(i + 1|i) = \mathrm{P}(A_1|B) = \frac{\mathrm{P}(A_1 \cap B)}{\mathrm{P}(B)} = \frac{i}{N}\left(1 - \frac{i}{N}\right),$$

$$\mathrm{P}(i - 1|i) = \mathrm{P}(A_2|B) = \frac{\mathrm{P}(A_2 \cap B)}{\mathrm{P}(B)} = \frac{i}{N}\left(1 - \frac{i}{N}\right).$$

Thus, we obtain the following equality:

$$\begin{aligned}
\mathrm{P}(A_k|B \cap C) &= \frac{\mathrm{P}(A_k \cap B \cap C)}{\mathrm{P}(B \cap C)} \\
&= \frac{\mathrm{P}(A_k \cap B \cap C)}{\mathrm{P}(B \cap C)}\left(\frac{\mathrm{P}(A_k \cap B)}{\mathrm{P}(A_k \cap B)}\right)\left(\frac{\mathrm{P}(B)}{\mathrm{P}(B)}\right) \\
&= \left(\frac{\mathrm{P}(A_k \cap B \cap C)}{\mathrm{P}(A_k \cap B)}\right)\left(\frac{\mathrm{P}(A_k \cap B)}{\mathrm{P}(B)}\right)\left(\frac{\mathrm{P}(B \cap C)}{\mathrm{P}(B)}\right)^{-1} \\
&= \frac{\mathrm{P}(C|A_k \cap B)\mathrm{P}(A_k|B)}{\mathrm{P}(C|B)}.
\end{aligned}$$

Hence,

$$\mathrm{P}(A_1|B \cap C) = \frac{\mathrm{P}(C|A_1 \cap B)\mathrm{P}(A_1|B)}{\mathrm{P}(C|B)} = \frac{\frac{i+1}{\nu}\frac{i}{N}\left(1 - \frac{i}{N}\right)}{\frac{i}{\nu}} = \frac{i+1}{N}\left(1 - \frac{i}{N}\right),$$

$$\mathrm{P}(A_2|B \cap C) = \frac{\mathrm{P}(C|A_2 \cap B)\mathrm{P}(A_2|B)}{\mathrm{P}(C|B)} = \frac{\frac{i-1}{\nu}\frac{i}{N}\left(1 - \frac{i}{N}\right)}{\frac{i}{\nu}} = \frac{i-1}{N}\left(1 - \frac{i}{N}\right).$$

190 $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

Proposition 2.3. *Conditioning on the fact that the neutral Moran process eventually reaches the state in which there are ν mutants, the expected number of mutants before the process reaches ν is approximately*

$$\mu(t) = N - (N - 1)\exp(-2t/N) \qquad (17)$$

in generation t, where one generation is N time steps and the first mutant is generated at $t = 0$.

Proof. It follows from the transition probabilities given in Equations (14)–(16) that the expected number of mutant cells, $\mu(t)$, must satisfy

$$\mu(t + 1/N) = \mu(t) + \frac{\mu(t) + 1}{N}\left(1 - \frac{\mu(t)}{N}\right) - \frac{\mu(t) - 1}{N}\left(1 - \frac{\mu(t)}{N}\right)$$

$$= \mu(t) + \frac{2}{N}\left(1 - \frac{\mu(t)}{N}\right).$$

We use the following approximation:

$$\frac{\mathrm{d}\mu}{\mathrm{d}t}(t) \approx \frac{\mu(t + 1/N) - \mu(t)}{1/N} = 2\left(1 - \frac{\mu(t)}{N}\right).$$

The differential equation has general solutions of the following form:

$$\mu(t) = N + \alpha \exp(-2t/N),$$

where α is a constant. Since the first mutant was generated at $t = 0$, that is, $\mu(0) = 1$, we obtain the solution

$$\mu(t) = N - (N - 1)\exp(-2t/N).$$

$\square$

We finally arrive at an expression for the expected fitness parameter given that the mutants survive long enough to change the competition dynamics.

Proposition 2.4. *Conditioning on the fact that the extended Moran process with non-constant fitness eventually reaches the state in which there are ν mutants, given that the fitness parameter is below Υ, the expected fitness parameter in generation t is approximately*

$$\Gamma(t) = \frac{1 - F}{FN} - \exp(-2t/N)\left(\frac{(1 - F)(N - 1)}{N^2 F - 2}\right)$$

$$+ \exp(-NFt)\left(\Gamma(0) + \frac{1 - F}{FN} - \frac{(1 - F)(N - 1)}{N^2 F - 2}\right)$$

for $F \neq 0$ and

$$\Gamma(t) = t + \frac{N - 1}{2}(\exp(-2t/N) - 1) + \Gamma(0)$$

for $F \approx 0$, where the first mutant is generated at $t = 0$ and $\Gamma(0)$ is the fitness parameter when the first mutant in the lineage is generated.

Proof. It follows from events 3 and 4, given at the beginning of Section 2, that the fitness parameter, $\Gamma(t)$, must satisfy

$$\Gamma(t + 1/N) = (\Gamma(t) + \mu(t)/N^2)(1 - F),$$

10

where $\mu(t)$ is the expected number of mutant cells given in Equation (17). By using the approximation

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t}(t) \approx \frac{\Gamma(t+1/N) - \Gamma(t)}{1/N},$$

we obtain the differential equation

$$\frac{\mathrm{d}\Gamma}{\mathrm{d}t} + NF\Gamma = \frac{1-F}{N}\mu.$$

For $F = 0$, we have general solutions of the form

$$\Gamma(t) = 1/N \int \mu(t)dt$$

$$= 1/N \int N - (N-1)\exp(-2t/N)dt$$

$$= t + \frac{N-1}{2}\exp(-2t/N) + \alpha,$$

where α is a constant. Thus, we obtain

$$\Gamma(t) = t + \frac{N-1}{2}(\exp(-2t/N) - 1) + \Gamma(0),$$

where $\Gamma(0)$ is the fitness parameter when the first mutant is generated. For $F \neq 0$, we have

$$\Gamma(t) = \exp(-NFt)\left(\Gamma(0) + \frac{1-F}{N}\int_0^t \exp(NFy)\mu(y)\mathrm{d}y\right).$$

Since

$$\int_0^t \exp(NFy)(N - (N-1)\exp(-2y/N))\mathrm{d}y$$

$$= \frac{\exp\left(-\frac{2t}{N}\right)\left(\left(\left(FN^2 - 2\right)\exp\left(\frac{2t}{N}\right) - FN^2 + FN\right)\exp\left(FNt\right) + (2 - FN)\exp\left(\frac{2t}{N}\right)\right)}{F\left(FN^2 - 2\right)},$$

we obtain

$$\Gamma(t) = \frac{1-F}{FN} - \exp(-2t/N)\left(\frac{(1-F)(N-1)}{N^2F - 2}\right)$$

$$+ \exp(-NFt)\left(\Gamma(0) + \frac{1-F}{FN} - \frac{(1-F)(N-1)}{N^2F - 2}\right).$$

$\square$

2.4.1. Expected functions and numerical simulations

Figure 1 displays the expected functions and numerical simulations of the extended Moran process. In all cases, the fitness parameter remains below the fitness threshold; hence, the growth of the mutant population is only driven by stochasticity. Consequently, the population dynamics is characterised by great variation. Figures 1(a)–(f) display the expected functions and simulations of mutant populations that reach population size $\nu = 10^3$, starting with a single mutant. It follows from Theorem 2.1 that the probability that a single mutant will generate a lineage of mutants that reaches population size $\nu = 10^3$ is $\rho = 10^{-3}$, regardless of the total population size. Indeed, for all three population sizes, $N = 10^3$, $N = 10^4$ and $N = 10^6$, we performed on average a thousand simulations to obtain one simulation in which the mutant population size reached $\nu = 10^3$, starting with a single mutant.

Note that the transition probabilities given in (14)–(16) and the expected number of mutants given in (17) do not contain the term ν. It is shown in the respective proofs that the terms with ν cancel out. However, a more intuitive explanation is as follows. The expected functions plotted in Figure 1 condition on the fact that the mutant populations reach the size $\nu = 10^3$. However, suppose that we stopped the simulations when the mutant populations reached the size $\nu_0 = 10^2$. Should this change the expected function? Clearly not. This is also compatible with the fact that neither the transition probabilities given in (14)–(16) nor the expected number of mutants depend on the size of ν.

In Figures 1(a)–(d), the diffusion rate of the fitness parameter, F, equals the inverse of the total population size, $1/N$. On these terms, it is expected that the fitness parameter is approximately F times the number of mutants. In point of fact, the simulations of the fitness parameter are close to F times the simulations of the number of mutants. In Figures 1(e)–(h), the diffusion rate of the fitness parameter, F, equals zero. In this case, the fitness parameter cannot decrease but is expected to increase as long as there are mutant individuals in the population. Figures 1(e) and 1(f) display the expected function and simulation of a mutant population that reaches the population size $\nu = 1500$, starting with a single mutant, and the corresponding fitness parameter, respectively. As illustrated in Figure 1(e), the mutant population size decreases in some time intervals for the simulation. However, as displayed in Figure 1(f), the fitness parameter does not decrease. In the simulation displayed in Figure 1(g), the mutant population size fluctuates before the mutant type becomes extinct around generation $t = 1600$. Even though the number of mutants remains below $\nu = 750$, the fitness parameter reaches 54. On the other hand, in the simulation displayed in Figures 1(e) and 1(f), the population size is close to $\nu = 1500$ when the fitness parameter is approximately 54. Thus, the simulation displayed in Figures 1(g) and 1(h) illustrates that, when F is equal to or relatively close to zero, then the mutant population can raise the fitness parameter to relatively high levels by delaying extinction.

12

2.5. The fitness parameter reaches the fitness threshold

In this subsection, we consider the case in which the fitness parameter reaches the fitness threshold, Υ, which means that the death rate of both the intermediate mutants and the wild-type individuals decrease whereas the resistant mutants become advantageous.

Let ν be the number of mutants. If no resistant mutant has been generated, the competition between the wild-type individuals and the intermediate mutants can be captured by a neutral Moran process; hence, it follows from Equation (5) that the probability that the intermediate mutants will invade the whole population is ν/N, given that no resistant mutant is generated before the intermediate mutants reach fixation.

On the other hand, if at least one resistant mutant has been generated, this lineage has a great advantage, because these cells survive when the fitness parameter is high. Thus, when the fitness parameter is higher than Υ, the resistant mutants are expected to invade the whole population.

If the timescale of fixation of the resistant mutants is much shorter than the timescale of mutation from the intermediate to the resistant type, then a lineage of resistant mutants is likely to take over the whole population or become extinct before another resistant mutant is created from the intermediate type. In this case, the expected number of resistant mutants in generation t, denoted $\chi(t)$, can be approximated as follows. In event 1 of the time step described at the beginning of Section 2, all the cells are expected to reproduce and die at the same rate; thus, $\chi(t)$ remains constant. On the other hand, if a cell that is not resistant is selected in event 2, then the selected cell dies, and a random cell is selected to reproduce. As derived in Subsection 2.2, the number of resistant mutants can either increase by one with probability $\pi(j) = \frac{j}{N}\left(1 - \frac{j}{N}\right)$ or remain constant with probability $1 - \pi(j)$, where j is the number of resistant mutants. Consequently, the expected number of resistant mutants in generation t must satisfy the equality

$$\chi(t + 1/N) = \chi(t) + \pi(\chi(t)).$$

We use the approximation

$$\frac{\mathrm{d}\chi}{\mathrm{d}t}(t) \approx \frac{\chi(t + 1/N) - \chi(t)}{1/N}$$

and obtain the differential equation

$$\frac{\mathrm{d}\chi}{\mathrm{d}t} = \chi\left(1 - \frac{\chi}{N}\right)$$

which has the solution

$$\chi(t) = \frac{N\chi(0)\exp(t)}{\chi(0)(\exp(t) - 1) + N}, \tag{18}$$

where $\chi(0)$ is the number of resistant mutants when the fitness parameter reaches the threshold Υ. Clearly $\chi(t)$ converges to N, which means that the

13

resistant mutants are expected to invade the whole population. The expected number of individuals that are not resistant is $N - \chi(t)$. Since intermediate mutants and wild-type individuals are neutral variants, the relation

$$\frac{\text{expected number of intermediate mutants}}{\text{expected number of wild-type individuals}}$$

remains constant. Thus, the expected number of intermediate mutants is approximately

$$\phi(t) = \frac{\nu - \chi(0)}{N - \chi(0)}(N - \chi(t))$$

and the expected number of wild-type individuals is approximately

$$\zeta(t) = \frac{N - \nu}{N - \chi(0)}(N - \chi(t)),$$

where ν is the total number of mutants when the fitness parameter reaches the threshold, Υ. Clearly, both the wild-type individuals and the intermediate mutants are expected to become extinct.

2.5.1. Expected functions and numerical simulations

Figures 2–4 display the expected functions and numerical simulations of the extended Moran process. In all cases, a mutant is generated in generation $t = 0$, and, since the mutants and the wild type are neutral variants as long as the fitness parameter is below Υ, the mutant population grows due to stochasticity and the population dynamics is characterised by great variation.

For the simulation illustrated in Figures 2 and 4, the fitness parameter reaches the fitness threshold, Υ. This means that the death rate of both the intermediate mutants and the wild-type individuals increases, whereas the resistant mutants become advantageous. If there is no resistant mutant in the population when the fitness parameter is above Υ, the probability that the intermediate mutants will invade the whole population is i/N, where i is the number of mutants and N is the total population size. For the simulation illustrated in Figure 3, the number of intermediate mutants is approximately $i = 10^3$ when the fitness parameter reaches the fitness threshold $\Upsilon = 0.1$. The mutation rate, μ_1, is relatively low; hence, no resistant mutant has been generated. Since the total population size is $N = 10^5$, the probability that the intermediate mutants will invade the whole population, given that no resistant mutant is generated, is $P_{inv} = 10^{-2}$. Due to stochasticity, the mutant population size nearly doubles before it starts decreasing. Since the diffusion rate of the fitness parameter, F, is relatively high, the fitness parameter decreases to a level below Υ soon after the number of mutants decreases to $i = 10^3$, and ultimately the mutant population becomes extinct.

The simulation illustrated in Figure 2 has the same low mutation rate as the simulation illustrated in Figure 3, and therefore there is no resistant mutant in the population when the fitness parameter reaches Υ. However, due to stochasticity, the fitness parameter remains above the fitness threshold Υ, and,

14

after approximately $t = 200$ generations, a resistant mutant is generated. This type is expected to invade the population, because it survives when the fitness parameter is high, and this makes it a very advantageous type. In point of fact, the growth of resistant mutants lies close to the expected function given in (18), as illustrated in Figure 2d.

The simulation displayed in Figure 4 has a relatively high mutation rate, and thus there are resistant mutants present in the population when the fitness parameter reaches the fitness threshold, Υ. The resistant type invades the population, but, as illustrated in Figures 2d and 4d, for the simulations with high mutation rates, the growth of the mutants does not lie as close to the expected function as the simulations with a low mutation rate. The reason for this is that the expected function given in (18) assumes that $\mu_1 \approx 0$, and this assumption does not hold when μ_1 is high.

3. Extended Moran process with cooperation entities

In this section, *cooperation entities* that can kill mutants, are included in the extended Moran process. Cooperation entities can represent regulation mechanisms that defend a cooperation, for instance T-cells in a multicellular organism. This is discussed in greater detail in Subsection 4.2. However, cooperation entities can also represent the cost of cooperation, for instance when cooperators invade a group of defectors, as discussed in Subsection 4.1.

The population still consists of N individuals, which are subdivided into three types, namely the wild type, intermediate mutants and resistant mutants. However, in events 5–8, the intermediate mutants and the resistant mutants behave identically; consequently, we simply refer to them as mutants.

In addition, there are up to N cooperation entities. The growth environment in which the population is located is subdivided into N sites. Each site contains exactly one individual; furthermore, each site can contain exactly one cooperation entity or no cooperation entity. At the beginning of each time step, the process passes through events 1–4, which are described in Section 2. Afterwards, the following events occur:

5. A random site is selected. If the site contains both a mutant and a cooperation entity, the mutant dies and a random individual is selected to reproduce.

6. A random site is selected, and, if this site contains a cooperation entity, it reproduces. The new cooperation entity is placed in a random site that does not already contain a cooperation entity at the end of the time step.

7. A random site is selected, and, if the site contains a cooperation entity, it dies.

8. A random site is selected. If the site contains a mutant and no cooperation entity, then, with probability P_d, a cooperation entity is activated and placed in the selected site.

15

At the end of each time step, all the individuals of the population are mixed and placed in random sites. As discussed in Section 2, it is assumed that, in all the events for each time step, the individuals are selected simultaneously. This assumption also holds for the cooperation entities. That is, if there are k cooperation entities at the beginning of a time step, then the probability that the selected site will contain a cooperation entity is k/N in both event 6 and event 7.

3.1. Event 5

If the selected site contains a wild-type individual, then nothing occurs at event 5. On the other hand, if the selected site contains both a cooperation entity and a mutant, then the mutant dies.

Since the population is mixed at the end of each time step, the probability that the selected site will contain both a mutant and a cooperation entity is

$$P_{de\ m} = \frac{k}{N} \frac{i}{N},$$

where k is the number of cooperation entities and i is the number of mutants.

As discussed in Subsections 2.1 and 2.2 and in Appendix A, we want the probability of selecting a given type of individual to be constant throughout the time step to keep the model as simplistic as possible. In Subsection 2.2, we show that the probability of selecting an individual to reproduce or die is approximately constant in events 1 and 2 for large population sizes. Since the same argument holds for event 5, the probabilities that the number of mutants decrease by one and remain constant in event 5 are

$$\mathrm{P}^5(i-1|i) = \frac{k}{N} \frac{i}{N} \left(1 - \frac{i}{N}\right), \tag{19}$$

$$\mathrm{P}^5(i|i) = 1 - \frac{k}{N} \frac{i}{N} \left(1 - \frac{i}{N}\right), \tag{20}$$

respectively, where k and i are the number of cooperation entities and the number of mutants at the beginning of the time step, respectively.

3.2. Events 6–8

The cooperation entities are activated by the mutants. In addition, the cooperation entities can reproduce and die. We assume that the same cooperation entity can be selected to reproduce and die in the same time step and that a new cooperation entity cannot be selected to die in the time step in which it is produced. These are similar to the assumptions made in the Moran process, discussed in Appendix A.

Let k and i denote the numbers of cooperation entities and mutants at the beginning of the time step, respectively. The probability that a cooperation entity will be selected to reproduce and to die in events 6 and 7, respectively, is

$$P_{ce} = \frac{k}{N},$$

16

whereas the probability that a cooperation entity will be activated by a mutant in event 8 is

$$P_{ac\ ce} = P_d \frac{i}{N} \left(1 - \frac{k}{N}\right). \tag{21}$$

3.3. Implications of cooperation entities

As discussed more thoroughly in Subsection 4.1, cooperation entities can represent cooperation. Moreover, they can represent T cells. As discussed in Subsection 1.3, if an APC recognises an antigen on a mutated cell, the production of CTLs is activated. This activation is captured by event 8. The CTLs are programmed to find and kill mutated cells, which is captured by event 5. The exact details of how the T cells are regulated and activated are uncertain. For instance, it is still unknown why APCs do not always recognise antigens on mutated cells. One hypothesis is that APCs only activate CTLs if healthy tissue is being injured [26]. In our model, healthy tissue, which is represented by the wild type, is not injured as long as the fitness parameter is below Υ. Hence, in some examples, the cooperation entities are not activated until the fitness parameter reaches this limit, whereas, in other examples, the cooperation entities are activated earlier.

In Section 2, the intermediate mutants and the wild type are neutral variants. However, when the cooperation entities are included, there is much more at stake for the mutants. If the activation of cooperation entities only depends on the presence of mutated cells, the mutants are disadvantageous when the fitness parameter is below Υ. Hence, the survival of the mutants depends on how fast they raise the fitness parameter, because resistant mutants become advantageous when the fitness parameter is higher than Υ. On the other hand, if the cooperation entities are not activated until the fitness parameter reaches Υ, the mutants and the wild type are neutral variants when the fitness parameter is below Υ, whereas the fitness of the mutants depends on whether a resistant mutant has been generated when the fitness parameter reaches Υ. That is, if all the mutants are of the intermediate type, then these cells are disadvantageous, whereas, if resistant mutants have been generated, these cells become advantageous. Hence, the probability of mutant invasion increases if the fitness parameter remains below Υ until a resistant mutant has been generated.

3.4. The cooperation entities are activated before the fitness parameter reaches the fitness threshold

When the activation of cooperation entities only depends on the presence of mutant cells, the mutants are disadvantageous when the fitness parameter is below Υ. That is, all individuals have the same probability of being selected to die and reproduce in event 1. However, in event 5, mutants can be selected to die if they are located in a site with a cooperation entity, whereas wild-type individuals can only be selected to reproduce in this event.

In event 5, the number of mutants either decreases by one or remains constant with probabilities given in (19) and (20), respectively. Hence, the expected number of mutants in generation t, $\phi(t)$, must satisfy

$$\phi(t + 1/N) = \phi(t) - \frac{\kappa(t)}{N} \frac{\phi(t)}{N} \left(1 - \frac{\phi(t)}{N}\right),$$

where $\kappa(t)$ is the expected number of cooperation entities in generation t. By using the approximation

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}(t) \approx \frac{\phi(t + 1/N) - \phi(t)}{1/N},$$

we obtain the following differential equation:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\kappa \frac{\phi}{N} \left(1 - \frac{\phi}{N}\right). \tag{22}$$

As discussed in Subsection 3.2, the probability that a cooperation entity will reproduce equals the probability that a cooperation entity will die in events 6 and 7, respectively; hence, the number of cooperation entities is expected to remain constant after these two events. In event 8, a cooperation entity is activated by a mutant with probability $P_{ac\ ce}$, given in (21). Otherwise, the number of cooperation entities remains constant. Hence, the expected number of cooperation entities must satisfy the difference equation

$$\kappa(t + 1/N) = \kappa(t) + P_d \frac{\phi(t)}{N} \left(1 - \frac{\kappa(t)}{N}\right).$$

By using the approximation

$$\frac{\mathrm{d}\kappa}{\mathrm{d}t}(t) \approx \frac{\kappa(t + 1/N), -\kappa(t)}{1/N}$$

we obtain the following differential equation:

$$\frac{\mathrm{d}\kappa}{\mathrm{d}t} = P_d \phi \left(1 - \frac{\kappa}{N}\right). \tag{23}$$

3.4.1. Numerical simulations

Figures 5–7 display numerical simulations of the extended Moran process with cooperation entities. In all the cases, the activation of cooperation entities depends only on the presence of the mutant cells. Hence, the mutants are disadvantageous when the fitness parameter is below Υ. On the other hand, the resistant mutants become advantageous if the fitness parameter reaches the fitness threshold Υ. Thus, the survival of the mutants depends on how fast they raise the fitness parameter.

When the fitness parameter is below Υ, it follows from the differential equation given in (22) that, if there is at least one cooperation entity in the system,

18

it is expected that the number of mutants will decrease until the mutants are extinct. However, the extinction can be delayed due to stochasticity, and, given that the diffusion rates of both the fitness parameter, F, and the fitness threshold, Υ, are relatively low, it is possible that the mutants will survive long enough to raise the fitness parameter above Υ. Due to stochasticity, the mutants become extinct before the fitness parameter reaches the fitness threshold, $\Upsilon = 2.5$, in the simulation illustrated in Figure 5, whereas, in the simulation displayed in Figure 6, the mutants survive long enough for the fitness parameter to reach the fitness threshold. Furthermore, if the mutant population produces at least one resistant mutant, this type of cells becomes advantageous when the fitness parameter is above Υ and is expected to invade the whole population.

It follows from the differential equation given in (23) that, if there is at least one mutant in the system, the number of cooperation entities is expected to grow until it reaches N. However, if the activation rate, P_d, is relatively low, then the activation of the cooperation entities can be delayed due to stochasticity. In this case, the mutants and the wild type are initially neutral variants, and the mutants can grow in number due to stochasticity. On these terms, the fitness parameter can reach the fitness threshold, Υ, even when it is relatively high. Moreover, the probability that the mutant population will produce a resistant type increases as the number of mutants increases. This scenario is illustrated in Figure 7.

3.5. The fitness parameter reaches the fitness threshold

When the fitness parameter reaches the fitness threshold, Υ, the intermediate mutants and the wild-type individuals have the same probability of being selected to die and reproduce in events 1 and 2. However, in event 5, the intermediate mutants can be selected to die if they are located in a site with a cooperation entity, whereas the wild-type individuals can only be selected to reproduce in this event. Hence, the wild-type individuals are more advantageous than the intermediate mutants.

If the mutants produce a resistant lineage, these mutants will be more advantageous than the wild-type individuals when the fitness parameter is higher than Υ and there are relatively few cooperation entities. That is, in event 1, the resistant mutants and the wild-type individuals have the same probability of being selected to die and to reproduce, whereas each wild-type individual has a probability $1/N$ of being selected to die in event 2, and each resistant mutant has a probability k/N^2 of being selected to die in event 5, where k is the number of cooperation entities. Thus, if each site contains a cooperation entity, the competition dynamics between the resistant mutants and the wild-type individuals is neutral, and the resistant mutants are increasingly advantageous with a decreasing number of cooperation entities.

Let i, j and k denote the number of intermediate mutants, resistant mutants and cooperation entities, respectively. Since the total number of individuals in the population is constant, N, the number of wild-type individuals is $N - i - j$.

It follows from the transition probabilities given in (9)–(13) that the probabilities that the number of intermediate mutants will decrease by one, remain

19

constant and increase by one in event 2 are

$$q_{-1}^2(j, i, k) = \pi(i), \tag{24}$$

$$q_0^2(j, i, k) = 1 - 2\pi(i) - \frac{ij}{N^2}, \tag{25}$$

$$q_1^2(j, i, k) = \pi(i) - \frac{ij}{N^2}, \tag{26}$$

respectively, where $\pi(i) = i/N(1 - i/N)$. Moreover, since the intermediate mutants and the resistant mutants are neutral variants in event 5, it follows from Subsection 3.1 that the probabilities that the number of intermediate mutants will decrease by one, remain constant and increase by one in event 5 are

$$q_{-1}^5(j, i, k) = \frac{k}{N}\pi(i),$$

$$q_0^5(j, i, k) = 1 - \frac{k}{N}\left(\pi(i) + \frac{ij}{N^2}\right),$$

$$q_1^6(j, i, k) = \frac{k}{N}\frac{ij}{N^2},$$

respectively. Thus, the probabilities that the number of intermediate mutants will decrease by two, decrease by one, increase by one and increase by two after events 2 and 5 are

$$Q_{-2}^{2,5}(j, i, k) = \frac{k}{N}\pi(i)^2,$$

$$Q_{-1}^{2,5}(j, i, k) = \pi(i)\left(1 + \frac{k}{N}(1 - 3\pi(i))\right),$$

$$Q_1^{2,5}(j, i, k) = \pi(i)\left(1 - \frac{k}{N}\left(\pi(i) + \frac{ij}{N^2}\right)\right) + \frac{ij}{N^2}\left(-1 + \frac{k}{N}\left(1 - \pi(i) + 2\frac{ij}{N^2}\right)\right),$$

$$Q_2^{2,5}(j, i, k) = \frac{k}{N}\frac{ij}{N^2}\left(\pi(i) - \frac{ij}{N^2}\right),$$

respectively; hence, the expected number of intermediate mutants in generation t, $\phi(t)$, must satisfy

$$\phi(t + 1/N) = \phi(t) - 2Q_{-2}^{2,5}(\chi(t), \phi(t), \kappa(t)) - Q_{-1}^{2,5}(\chi(t), \phi(t), \kappa(t))$$
$$+ Q_1^{2,5}(\chi(t), \phi(t), \kappa(t)) + 2Q_2^{2,5}(\chi(t), \phi(t), \kappa(t)),$$

where $\chi(t)$ and $\kappa(t)$ are the expected numbers of resistant mutants and cooperation entities, respectively, in generation t. By using the approximation

$$\frac{\mathrm{d}\phi}{\mathrm{d}t}(t) \approx \frac{\phi(t + 1/N) - \phi(t)}{1/N},$$

we obtain the following differential equation:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\phi\left(\left(1 - \frac{\phi}{N}\right)\frac{\kappa}{N} + \frac{\chi}{N}\left(1 - \frac{\kappa}{N}\right)\right). \tag{27}$$

Since $\phi, \kappa \leq N$, it follows that the expected number of intermediate mutants
decreases towards zero.

We will now derive an approximation for the expected number of resistant
mutants. It follows from the transition probabilities given in (9)–(13) that the
probabilities that the number of resistant mutants will remain constant and
increase by one in event 2 are

$$p_0^2(j, i, k) = 1 - \pi(j),$$
$$p_1^2(j, i, k) = \pi(j),$$

respectively, where $\pi(j) = j/N(1-j/N)$. Since the mutants are neutral variants
in event 5, it follows from Subsection 3.1 that the probabilities that the number
of resistant mutants will decrease by one, remain constant and increase by one
in event 5 are

$$p_{-1}^5(j, i, k) = \frac{k}{N}\pi(j),$$
$$p_0^5(j, i, k) = 1 - \frac{k}{N}\left(\pi(j) + \frac{ij}{N^2}\right),$$
$$p_1^5(j, i, k) = \frac{k}{N}\frac{ij}{N^2},$$

respectively. Thus, the probabilities that the number of resistant mutants will
decrease by one, increase by one and increase by two after events 2 and 5 are

$$P_{-1}^{2,5}(j, i, k) = \frac{k}{N}\pi(j)(1 - \pi(j)),$$
$$P_1^{2,5}(j, i, k) = \pi(j)\left(1 - \frac{k}{N}\left(\pi(j) + \frac{ij}{N^2}\right)\right) + \frac{k}{N}\frac{ij}{N^2}(1 - \pi(j)),$$
$$P_2^{2,5}(j, i, k) = \frac{k}{N}\frac{ij}{N^2}\pi(j),$$

respectively; hence, the expected number of resistant mutants must satisfy

$$\chi(t + 1/N) = \chi(t) - P_{-1}^{2,5}(\chi(t), \phi(t), \kappa(t))$$
$$+ P_1^{2,5}(\chi(t), \phi(t), \kappa(t)) + 2P_2^{2,5}(\chi(t), \phi(t), \kappa(t)).$$

By using the approximation

$$\frac{d\chi}{dt}(t) \approx \frac{\chi(t + 1/N) - \chi(t)}{1/N},$$

we obtain the differential equation

$$\frac{d\chi}{dt} = \chi\left(\left(1 - \frac{\chi}{N}\right)\left(1 - \frac{\kappa}{N}\right) + \frac{\phi}{N}\frac{\kappa}{N}\right).$$

21

Since
$$N = \chi(t) + \phi(t) + \zeta(t),$$
where $\zeta(t)$ is the expected number of healthy cells, it follows that
$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = -\frac{\mathrm{d}\chi}{\mathrm{d}t} - \frac{\mathrm{d}\phi}{\mathrm{d}t} = \zeta\left(\left(1 - \frac{\zeta}{N}\right)\frac{\kappa}{N} - \frac{\chi}{N}\right).$$

The expected number of cooperation entities is described by the differential equation given in (23). Hence, the expected numbers of resistant mutants, wild-type individuals and cooperation entities are described by the following system of differential equations:

$$\frac{\mathrm{d}\chi}{\mathrm{d}t} = \chi\left(1 - \frac{\chi}{N} - \frac{\zeta}{N}\frac{\kappa}{N}\right), \tag{28}$$

$$\frac{\mathrm{d}\zeta}{\mathrm{d}t} = \zeta\left(\left(1 - \frac{\zeta}{N}\right)\frac{\kappa}{N} - \frac{\chi}{N}\right), \tag{29}$$

$$\frac{\mathrm{d}\kappa}{\mathrm{d}t} = P_d N\left(1 - \frac{\zeta}{N}\right)\left(1 - \frac{\kappa}{N}\right), \tag{30}$$

respectively. The system is in equilibrium on the line
$$\mathcal{L}^* = \{\kappa = N, \zeta + \chi = N\},$$
and the point
$$(\chi^*, \zeta^*, \kappa^*) = (0, 0, N).$$
The domain
$$\mathcal{D} = \{0 \leq \kappa, \zeta + \chi \leq N | 0 \leq \zeta, \chi\}$$
is bounded by the following five planes:

$$\begin{aligned}
\mathcal{P}^1 &= \{\kappa = N | 0 \leq \chi, \zeta \;;\; \zeta + \chi \leq N\}, \\
\mathcal{P}^2 &= \{\chi + \zeta = N | 0 \leq \chi, \zeta \;;\; 0 \leq \kappa \leq N\}, \\
\mathcal{P}^3 &= \{\chi = 0 | 0 \leq \kappa, \zeta \leq N\}, \\
\mathcal{P}^4 &= \{\zeta = 0 | 0 \leq \kappa, \chi \leq N\}, \\
\mathcal{P}^5 &= \{\kappa = 0 | 0 < \chi, \zeta \;;\; \zeta + \chi < N\}.
\end{aligned}$$

Clearly, a solution of the system of differential equations given in (28)–(30) cannot leave the domain $\mathcal{D}$. Moreover, it follows from the differential equations given in (28) and (30) that both χ and κ grow in the interior of $\mathcal{D}$, denoted $\mathcal{D}^*$. Hence, any solution with initial values in $\mathcal{D}^*$ will grow towards the equilibrium line, $\mathcal{L}^*$.

22

3.6. $\mathcal{P}^1$: N cooperation entities

When all the sites contain a cooperation entity, the resistant mutants and the wild-type individuals are neutral variants, whereas the intermediate mutants are disadvantageous. Substituting $\kappa = N$ into the differential equation given in (27), we obtain

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = -\phi\left(1 - \frac{\phi}{N}\right).$$

By scaling the generations such that $t = 0$ is the generation when the number of cooperation entities reaches N, we obtain

$$\phi(t) = \frac{N\phi(0)\exp(-t)}{\phi(0)(\exp(-t) - 1) + N}.$$

Clearly, $\phi(t)$ converges to zero. Since the resistant mutants and the wild-type individuals are neutral variants, the expected number of resistant mutants is

$$\chi(t) = \frac{\chi(0)}{N - \phi(0)}\left(N - \phi(t)\right)$$

and the expected number of wild-type individuals is

$$\zeta(t) = \frac{\zeta(0)}{N - \phi(0)}\left(N - \phi(t)\right).$$

Thus, $\chi(t)$ converges to $\frac{\chi(0)N}{N-\phi(0)}$, whereas $\zeta(t)$ converges to $\frac{\zeta(0)N}{N-\phi(0)} = N - \frac{\chi(0)N}{N-\phi(0)}$.

Figure 9 displays a numerical simulation of the extended Moran process with N cooperation entities. In this case, the resistant mutants and the wild type are neutral variants, whereas the intermediate mutants are disadvantageous. Hence, the ratio of resistant mutants and wild-type individuals is expected to remain constant, whereas the intermediate mutants are expected to become extinct. Since the resistant mutants and the wild type are neutral variants, their competition dynamics is characterised by great variation, whereas the simulation of the intermediate mutants lies close to the expected function, as illustrated by Figure 9.

3.7. $\mathcal{P}^2$: The intermediate mutants become extinct

In this subsection, we consider the case in which all the intermediate mutants become extinct. By substituting $\zeta = N - \chi$ into the differential equations given in (28) and (30), we obtain

$$\frac{\mathrm{d}\chi}{\mathrm{d}t} = \chi\left(1 - \frac{\chi}{N}\right)\left(1 - \frac{\kappa}{N}\right) \tag{31}$$

$$\frac{\mathrm{d}\kappa}{\mathrm{d}t} = P_d\chi\left(1 - \frac{\kappa}{N}\right). \tag{32}$$

23

It follows from the differential equations above that κ grows until $\kappa = N$, whereas χ grows as longs as $\chi < N$ and $\kappa < N$. Since $\frac{d\zeta}{dt} = -\frac{d\chi}{dt}$, it follows that ζ decreases as long as $\chi < N$ and $\kappa < N$. The case in which χ reaches N before κ corresponds to the invasion of resistant mutants and the extinction of wild-type individuals, whereas the case in which κ reaches N before κ corresponds to the survival of both the wild type and the resistant type, as described in Subsection 3.6

It follows from the differential equations given in (31) and (32) that

$$\frac{\frac{d\kappa}{dt}}{\frac{d\chi}{dt}} = \frac{d\kappa}{d\chi} = \frac{P_{APC}}{1 - \frac{\chi}{N}}$$

for $0 < \chi < N$ and $\kappa < N$. Thus,

$$\int d\kappa = NP_d \int \frac{d\chi}{(N - \chi)}.$$

Hence, a solution of the system given in 31 and 32, $\Omega(t) = (\chi(t), N - \chi(t), \kappa(t)) \in \mathcal{P}^1$, with initial value $\Omega(0) = (\chi(0), N - \chi(0), \kappa(0)) \in \mathcal{P}^1$, where $0 < \chi(0) < N$ and $\kappa(0) < N$, must satisfy

$$\kappa(t) = \kappa(0) + NP_d \ln \left(\frac{N - \chi(0)}{N - \chi(t)} \right)$$

$$\zeta(t) = N - \chi(t),$$

as long as $\chi(t) < N$ and $\kappa(t) < N$. We are interested in investigating whether the cooperation entities or the resistant mutants reach the population size N first. To achieve this, we make use of the fact that the process is discrete and investigate which population is expected to reach the population size $N - 1$ first. For sufficiently large population sizes, this is equivalent to reaching N. We have

$$\kappa(t) = \kappa(0) + NP_d \ln \left(\frac{N - \chi(0)}{N - \chi(t)} \right) = N - 1$$

for

$$\chi(t) = N - \frac{N - \chi(0)}{\exp \left(\frac{N - (\kappa(0) + 1)}{NP_d} \right)}.$$

Hence, $\kappa(t)$ reaches $N - 1$ before $\chi(t)$ if

$$1 < \frac{N - \chi(0)}{\exp \left(\frac{N - (\kappa(0) + 1)}{NP_d} \right)}.$$

The above inequality can be expressed as

$$\frac{N - (\kappa(0) + 1)}{N \ln(N - \chi(0))} < P_d. \tag{33}$$

24

3.7.1. Numerical simulations

Figures 10–13 display simulations in which the fitness parameter reaches the fitness threshold, Υ, and there is a race between the resistant mutants and the cooperation entities to reach population size N. The inequality given in (33) is derived under the assumption that the intermediate mutants are extinct, and this is not the case when the first resistant mutant appears in the simulations illustrated in Figures 10–13. However, since the intermediate mutants become very disadvantageous, their population size decreases rapidly, and therefore the inequality in (33) gives a good indication of whether the cooperation entities win the race to reach population size N.

The activation rate of the cooperation entities is very high in the example illustrated in Figure 11. Consequently, the left side, which equals 0.1, is less than the right side, which equals 1, in the inequality given in (33), and this indicates that the cooperation entities will win the race. In point of fact, the number of cooperation entities reaches N when the number of resistant mutants is approximately 4×10^4. The example illustrated in Figure 10 has a moderate activation rate of the cooperation entities, and both sides of the inequality given in (33) are approximately 0.1. Indeed, the number of cooperation entities grows more slowly towards N than the number of resistant mutants.

If there are no resistant mutants in the population when the fitness parameter reaches Υ, it is possible that the cooperation entities will win the race towards N even though the cooperation entities are activated at a moderate rate. That is, the examples given in Figures 10 and 13 have the same activation rate. However, in the example illustrated in Figure 13, there are approximately 7×10^3 cooperation entities in the system when the first resistant mutant is produced. Thus, the left side, which equals 0.08, is less than the right side, which equals 0.1, in the inequality given in (33). In fact, the number of cooperation entities reaches N when the number of resistant mutants is approximately 8×10^4.

However, if the activation rate of the cooperation entities is sufficiently low, the resistant mutants can invade the system even though the production of the first resistant mutant is delayed. In the example given in Figure 12, there are approximately 7×10^3 cooperation entities in the system when the first resistant mutant is produced. However, the left side, which equals 0.08, is higher than the right side, which equals 0.01, in the inequality given in (33). Indeed, the resistant mutants reach the population size N first.

4. Discussion

Several other models and texts describe situations in which relatively stable populations are invaded by an alternative strategy. Examples are the evolution of cooperation among bacteria and multicellularity [14],[27]–[29], the invasion of cancer [19]–[24] and the evolution of ideas that contradict social norms [30],[31]. These models and texts are more detailed and sophisticated than the model described in this paper. However, by keeping our model simplistic, it applies to different situations, as illustrated by the examples below. Hence, our model gives

a more general description of the dynamics that occur when a stable population is invaded by an alternative strategy.

4.1. Evolution of cooperation among bacteria and multicellularity

When life started to evolve about four billion years ago, the first life forms adopted the most basic strategy, which is to outcompete other individuals by dividing as fast as possible [2]. However, proliferation requires resources, such as space and nutrient molecules, and different individuals can have access to some resources and no access to other resources. In these situations, cooperation can be beneficial [10],[14],[27].

A simplified example of cooperation among single-celled organisms is that one cell has access to enough nutrient molecules for two cell divisions but no space, whereas another cell has access to enough space for two cell divisions but no nutrient molecules. Thus, if the two cells share their resources, that is, mutual cooperation, they will both reproduce. On the contrary, if the two cells do not share their resources, that is, mutual defection, neither of the cells will reproduce. However, if only one cell shares its resources and the other does not share, then the cooperator does not reproduce and loses its resources whereas the defector reproduces twice.

This simple example illustrates the dilemma of cooperation: even though mutual cooperation leads to a higher payoff than mutual defection, a defector has a higher payoff than a cooperator when they meet. Indeed, it is a version of the prisoner's dilemma, which is discussed in Subsection 1.2.

Moreover, a group of cooperating cells is vulnerable to intruders and mutants that stop cooperating, because these cells can invade the colony by exploiting the cooperating cells [32]. Hence, a group of cooperators can only survive in the long term if it develops regulation mechanisms that control the cooperation, for instance by modifying the microenvironment such that the defectors lose their advantages. Indeed, the evolution of multicellular organisms was driven by increasingly advanced regulation mechanisms among cooperating cells [14].

A small group of cooperators can invade a large population of defectors if they manage to change the environment such that defection becomes a disadvantageous strategy. This is illustrated by Figure 6. In the context of the evolution of cooperation among cells, the wild-type individuals represent defectors and the intermediate mutants represent unconditional cooperators. Cooperation is captured by the cooperation entities.

The fitness parameter represents the evolution of regulation mechanisms, whereas the resistant mutants represents conditional cooperators. In a nutshell, the conditional cooperators cooperate with cells that are of the same type and create a microenvironment that kills cells that are of a different type. As illustrated by Figure 5 and 6, the cooperators are disadvantageous when the fitness parameter is below Υ. Indeed, in the example given in Figure 5, the cooperators become extinct. On the other hand, due to stochasticity, the cooperators survive long enough to raise the fitness parameter above Υ in the example illustrated by Figure 5, and then the conditional cooperators invade the whole population.

26

4.2. The invasion of cancer

As discussed in Subsections 1.3 and 4.1, the human body has an advanced defence system that attacks mutant cells that have stopped cooperating. Hence, mutant cells are in general disadvantageous when they first appear in the body, and, what is more, mutant cells that progress into cancer typically change their microenvironment and create new variants that are advantageous in the new microenvironment [19]–[24]. This dynamics is captured by the model presented in this paper. In this context, the cooperation entities represent the immune response, such as T cells, whereas the wild type and the mutant type represent the healthy cells and the mutant cells, respectively.

Figures 5–7 illustrate the case in which T cells detect and kill mutant cells before they cause any harm, whereas Figures 8 and 10–13 illustrate the case in which T cells are only activated if they harm healthy tissue. In the first case, the healthy cells are initially advantageous, whereas, in the latter case, the competition dynamics is neutral. Given that mutants can evolve into cancer cells, it might seem that the best strategy is to kill them once they appear in the body. However, too aggressive an immune system poses a greater risk to the body than mutant cells with minor genetic errors [26],[33].

The body can limit the blood flow to mutant cells. Hence, these cells must break down the end product of glycolysis anaerobically, and this leads to an acidic microenvironment [20]. In the model, the acid level is represented by the fitness parameter. When the acid level reaches the limit Υ, the death rate of the cells that are not acid resistant increases. Moreover, since the mutants are harming the healthy cells, the T cells are activated. Thus, the non-resistant mutants become less advantageous than the healthy cells, and, as illustrated in Figure 8, they are expected to become extinct if they do not produce an acid-resistant variant. On the other hand, if the mutant cells survive long enough such that they produce a variant that is acid resistant, this cell type has a great advantage, because they can kill other cells by increasing the acid level, as illustrated in Figures 10–13.

As illustrated by the examples given in Figures 10–13, there is a race between the acid-resistant mutants and the T cells. If the T cells respond quickly, such that there is a T cell at every site in the microenvironment before the normal cells become extinct, the acid-resistant mutants are neutralised. In this case, the mutant cells are vulnerable to new attacks from the body's defence mechanisms.

On the other hand, if the normal cells become extinct before there is a T cell at every site, the resistant mutants are expected to take over the microenvironment. This represents the onset of a more aggressive form of cancer. Indeed, many observations reveal that cancer cells exhibit glucose fermentation even when there is enough oxygen present. This is called the Warburg effect and has been described in several other papers.

The model by Robertson-Tessi et al. includes several other mechanisms of immune evasion that tumours use, including immunosuppressive surface markers such as PD-L1, the down-regulation of antigen presentation machinery, the recruitment of immunosuppressive immune cells and the secretion of immuno-

suppressive factors such as TGF-beta [22]. Moreover, several other models indicate that tissue architecture and signals between different microenvironments play major roles in population dynamics and the progression of cancer [34]–[36].

It is possible to include these mechanisms in our model. However, the main scope of this paper is to give a general characterisation of the dynamics that occur when a stable population is invaded; therefore, we keep the model as simple as possible.

4.3. Evolution of ideas that contradict social norms

Game and evolutionary theory can also be used to study human behaviour and society [37]. For instance, the founders of Marxism, Karl Marx and Friedrich Engels, were inspired by Charles Darwin [38].

The model can also capture the dynamics of political changes that are less dramatic than revolutions and dictatorships, for instance when politicians use populist rhetoric or, depending on who has the power of definition, speak freely. Sylvi Listhaug is a Norwegian politician for the Progress Party who was Minister of Migration from December 2015 to March 2018. Listhaug has been called the Trump of Norway, both as a compliment and as a criticism [39],[40].

The consensus of the Norwegian political elite is to address problems related to immigration and integration in a polite and indirect way. Hence, Listhaug's direct and confrontational style has created waves of reactions. Her critics claim that her aggressive style creates conflicts with people who could become allies and that she should rather focus on building a broad and inclusive alliance. A paper by Pinker et al. [41], in which the authors apply ideas from evolutionary biology and game theory to illuminate possible advantages of indirect speech, lend some support to Listhaug's critics. Pinker et al. argue that most human communication involves a mixture of cooperation and conflict and that indirect speech is used to negotiate the type of relationship holding between the speaker and the hearer. Moreover, indirectness in speech appears to be nearly universal [42].

However, when it comes to integration, indirect speech might promote parallel societies, because it can create misunderstandings about what is socially acceptable and make the majority society seem very complex and unmanageable.

Regardless of whether indirect speech is an advantage, breaking an unwritten law is associated with social stigmatisation. Thus, Listhaug must pay a cost for bringing up unpleasant issues related to migration and integration in a direct and, perhaps, populistic way [31]. Listhaug's statements are almost automatically considered to be controversial, as philosopher Lars Kolbeinstveit, from the liberal think tank Civita, writes in a text about Listhaug and the media [43].

In point of fact, after Listhaug claimed that the Labour Party puts the rights of terrorists above the security of the nation in a Facebook post, the reactions were so strong that Listhaug announced her resignation from the Government to avoid a vote of confidence [44].

Even though Listhaug reduced her political influence, at least on the short term, after she resigned as Minister of Immigration, her political party, the

Progress Party, gained support after her fall [45]. Moreover, there has been an increased use of words regarding anti-elitism as well as a more heated immigration debate on the Internet [46].

5. Conclusion

In this paper, an extension of the Moran process with non-constant fitness is presented. The model captures not only the competition between different types of individuals but also the struggle for fitness. That is, the type of individuals that manage to change the environment such that they become advantageous is expected to outcompete other types of individuals.

The model captures the dynamics that occurs when a relatively stable population is invaded by a new type of individuals and can reproduce the following events:

1. When a new type of individual appears in a relatively stable population, the newcomer is not advantageous.

2. Due to stochasticity, the new type grows in number and generates different versions of itself.

3. The new type becomes advantageous if it manages to change the environment such that at least one of its variants increases its fitness.

These events occur in different examples in which a relatively stable population is invaded by a new type of individuals, for instance the evolution of cooperation among bacteria and multicellularity, the invasion of cancer and the evolution of ideas that contradict existing social norms. Several models have already been proposed to describe these situations; however, none of them generalise the phenomena. Indeed, to our knowledge, the model presented in this paper is the first general model of competition dynamics in relatively stable populations that captures events 1–3.

Appendix A.

In this subsection, we summarise the way in which Wodarz and Komarova [3] obtain the transition probabilities for the Moran process given in (1)–(3).

The Moran process assumes that the population has A constant size, N, and consists of two types of individuals, denoted the wild type and the mutant type. The individuals can reproduce, mutate and die.

When a wild-type individual reproduces, the probability that it will produce a wild-type individual is $1-u$, and the probability that it will produce a mutant is u, where $0 \leq u \leq 1$. It is assumed that when a mutant individual reproduces, it always produces a new mutant. Moreover, the wild-type individuals have reproductive rate 1 and the mutants have reproductive rate r, where $r > 0$. Both types are selected to die at the same rate. In each time step, one individual

29

reproduces and one individual dies. It is assumed that the same individual can be selected both to reproduce and to die.

Let i denote the number of mutants at the beginning of a given time step. Thus, the number of wild-type individuals at the beginning of the time step is $N-i$. The probability that a wild-type individual will reproduce is proportional to its frequency and the reproductive rate and is given by $(N-i)/(N-i+ri)$. Similarly, the probability that a mutant will reproduce is $ri/(N-i+ri)$. Thus, the probabilities that the new individual will be a wild type and a mutant type are

$$P_{+w} = (1-u)\frac{N-i}{N-i+ri},$$
$$P_{+m} = u\frac{N-i}{N-i+ri} + \frac{ri}{N-i+ri},$$

respectively.

The Moran process assumes that the new individual cannot be selected to die in the time step in which it was produced. Hence, the probability that a type of individual will be selected to die depends on its abundance at the beginning of the time step. That is, the probabilities that the individual selected to die is a wild-type individual and a mutant are

$$P_{-w} = \frac{N-i}{N}, \tag{A.1}$$

$$P_{-m} = \frac{i}{N}, \tag{A.2}$$

respectively.

Note that, if the new individual could be selected to die in the same time step in which it was produced, then the probability that an individual of a certain type will be selected to die would depend on which type the new individual is. Moreover, the population size would be $N+1$ before the selected individual dies. Thus, the conditional probabilities that the individual selected to die is of a certain type would be

$$\mathrm{P}(A|B) = \frac{N+1-i}{N+1}, \tag{A.3}$$

$$\mathrm{P}(A|C) = \frac{N-i}{N+1}, \tag{A.4}$$

$$\mathrm{P}(D|B) = \frac{i}{N+1}, \tag{A.5}$$

$$\mathrm{P}(D|C) = \frac{i+1}{N+1}, \tag{A.6}$$

where the events A–D are as follows:

- A: a wild-type individual is selected to die.

- B: the new individual is wild type.

30

- C: the new individual is a mutant.

- D: a mutant is selected to die.

Hence, if the new individual could be selected to die in the same time step in which it was produced, the probabilities associated with each time step would be more complex; consequently, it would become more complicated to compute the absorption time and the probabilities of being absorbed. Furthermore, for a sufficiently large population size, N, the probability given in (A.1) is a reasonable approximation of the probabilities given in (A.3) and (A.4); likewise, (A.2) is a tolerable approximation of (A.5) and (A.6). These approximations are very good when i is close to $N/2$. However, for $i = 1$

$$\mathrm{P}(D|C) = 2\mathrm{P}(D|B),$$

and for $i = N - 1$

$$\mathrm{P}(A|B) = 2\mathrm{P}(A|C).$$

690 Even though the approximations are not very precise when i is either very small or close to N, the request for simplicity weighs more in the Moran model.

In each time step of the Moran process, the number of mutants can increase by one, decrease by one or remain constant. By assuming that the new individual cannot be selected to die in the time step in which it was produced, the probabilities of these three events are given by

$$\mathrm{P}(i+1|i) = P_{+m}P_{-w} = \frac{u(N-i) + ri}{N - i + ir}\frac{N-i}{N},$$
$$\mathrm{P}(i-1|i) = P_{+w}P_{-m} = \frac{(1-u)(N-i)}{N - i + ir}\frac{i}{N},$$
$$\mathrm{P}(i|i) = 1 - \mathrm{P}(i+1|i) - \mathrm{P}(i-1|i),$$

respectively.

Appendix B.

Given that the present number of mutants is i, there are $i - 1$, i or $i + 1$ mutants after the next time step. The probabilities for these events are given in Equations (1)–(3), respectively, with $r = 1$ and $u = 0$. Thus, the conditional probability of reaching state ν, $\mathrm{P}(\text{reach } \nu \mid i)$, must satisfy

$$\mathrm{P}(\text{reach } \nu \mid i) = \mathrm{P}(\text{reach } \nu \mid i-1)\frac{i}{N}\left(1 - \frac{i}{N}\right) + \mathrm{P}(\text{reach } \nu \mid i)\left(1 - 2\frac{i}{N}\right)$$
$$+ \mathrm{P}(\text{reach } \nu \mid i+1)\frac{i}{N}\left(1 - \frac{i}{N}\right).$$

This equation can be reduced to the following second-order difference equation with constant coefficients:

$$\mathrm{P}(\text{reach } \nu \mid i-1) = 2\mathrm{P}(\text{reach } \nu \mid i) - \mathrm{P}(\text{reach } \nu \mid i+1).$$

31

Since the corresponding quadratic equation

$$r^2 - 2r + 1 = 0$$

has only one root, namely $r = 1$, the solutions of the difference equation have the form

$$P(\text{reach } \nu \mid i) = \alpha i + \beta$$

where α and β are constants. Note that the system has exactly two absorbing states, namely $i = 0$ and $i = \nu$, with corresponding transition probabilities $P(\text{reach } \nu \mid 0) = 0$ and $P(\text{reach } \nu \mid \nu) = 1$. Thus, we have the following boundary conditions:

$$0 = 2P(\text{absorbed in } \nu \mid 1) - P(\text{absorbed in } \nu \mid 2)$$

$$P(\text{absorbed in } \nu \mid \nu - 2) = 2P(\text{absorbed in } \nu \mid \nu - 1) - 1.$$

We obtain $\alpha = 1/\nu$ and $\beta = 0$. Hence, the conditional probability for reaching ν is

$$P(\text{reach } \nu \mid i) = \frac{i}{\nu}.$$

Appendix C.

	Co-operator	Defector
Co-operator	C,C	R,T
Defector	T,R	D,D

Table 1 Payoff matrix for prisoner's dilemma
This Table display the payoff matrix for a 2×2 game with two strategies, namely cooperation and defection. If the following inequalities $T > C > D > R$ hold, then the game is a version of the prisoner's dilemma.
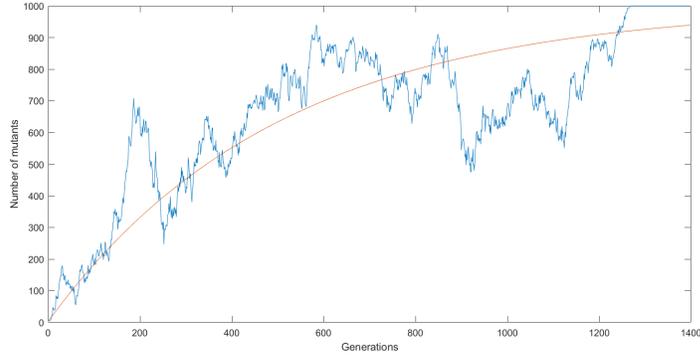
Figure 1a displays the number of mutants for $N = 10^3$ and $F = 10^{-3}$

Figure 1 Population dynamics when the fitness parameter is below Υ
When the fitness parameter is below Υ, the population dynamics is identical to a neutral Moran process. Moreover, starting with one mutant at generation $t = 0$, the probability that this lineage reaches population size $i = 10^3$ is 10^{-3}. The growth of the fitness parameter, depends on the diffusion rate, F. In Figure 1(a)(d), F equals the inverse of the total population size. On these terms, it is expected that the fitness parameter is approximately F times the number of mutants. And in point of fact, the fitness parameter is close to F times the number of mutants in the simulations displayed in (a)(d). On the other hand, in Figure 1(e)(h), F equals zero. In this case, the fitness parameter cannot decrease, but is expected to increase as long as there are mutants in the population. And indeed, the simulation displayed in Figure 1(e) and (f) illustrates that given that the number of mutants reaches $i = 1.5 \times 10^3$, it follows that the fitness parameter grows exponentially, whereas the simulation displayed in Figure 1(g) and (h) illuminates that the fitness parameter grows until the mutants are extinct.
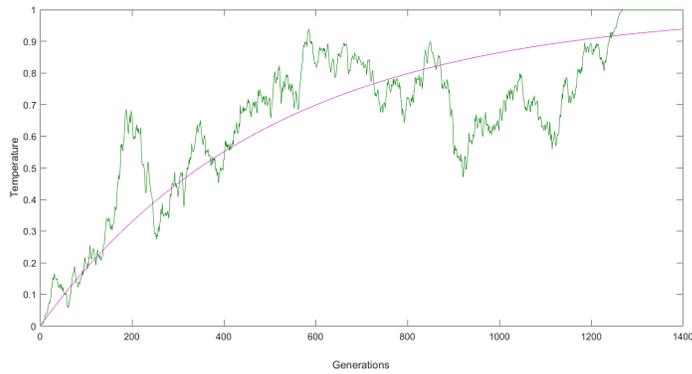
Figure 1b displays the fitness parameter for $N = 10^3$ and $F = 10^{-3}$.
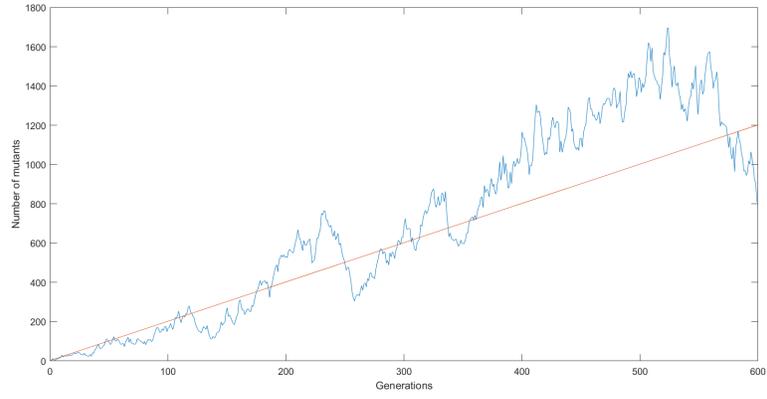
Figure 1c displays the number of mutants for $N = 10^6$ and $F = 10^{-6}$
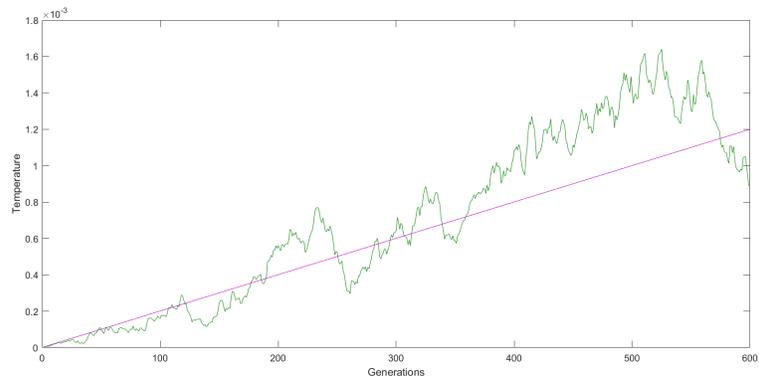
Figure 1d displays the fitness parameter for $N = 10^6$ and $F = 10^{-6}$
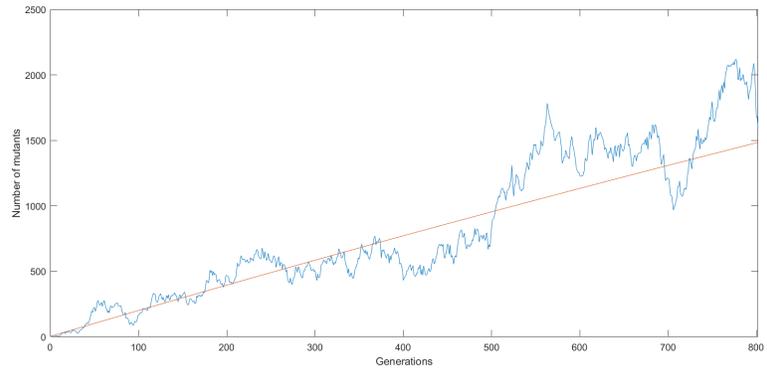
Figure 1e displays the number of mutants for $N = 10^4$ and $F = 0$

35
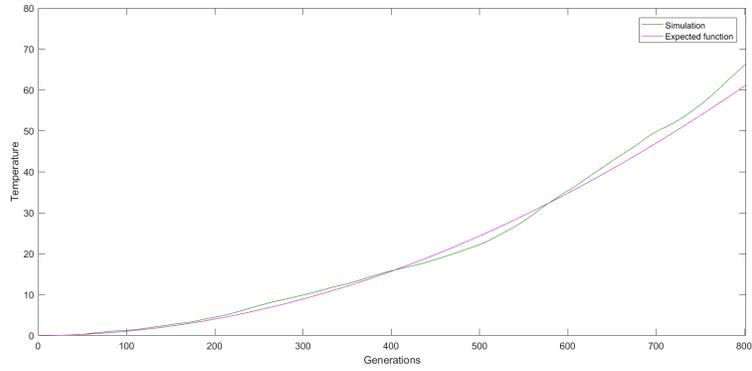
Figure 1f displays the fitness parameter for $N = 10^4$ and $F = 0$
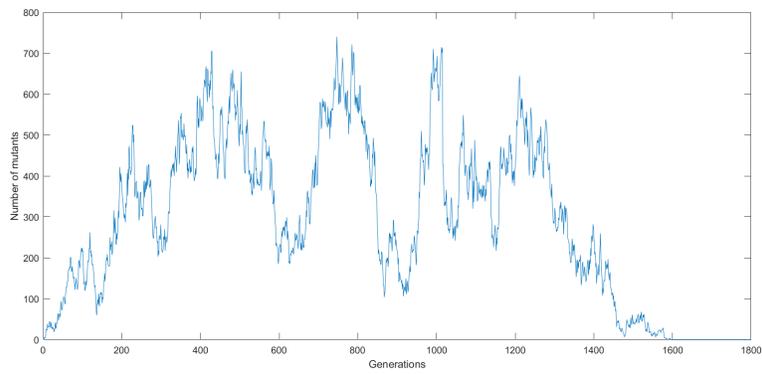
Figure 1g displays the number of mutants for $N = 10^4$ and $F = 0$
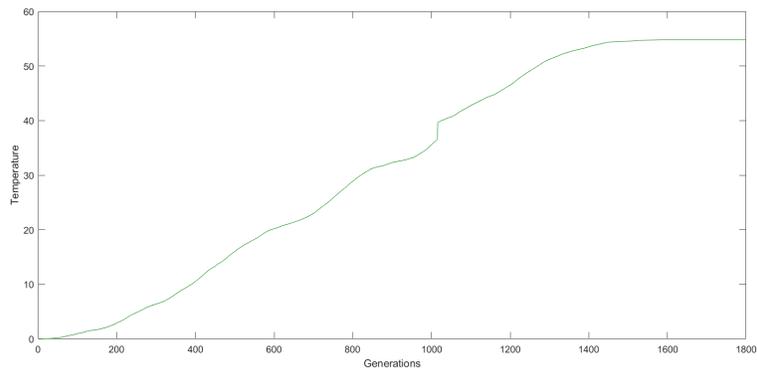
Figure 1h displays the fitness parameter for $N = 10^4$ and $F = 0$

36
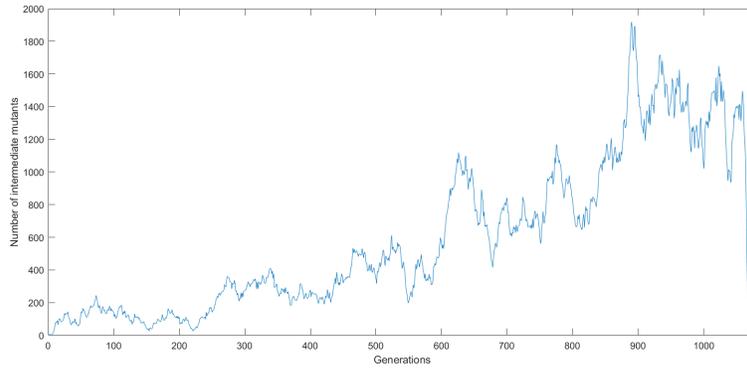
Figure 2a displays the number of intermediate mutants

Figure 2 Low mutation rate and high diffusion rate: Invasion of mutants
The first mutant is generated at generation $t = 0$, and due to stochasticity, the mutant population grows in number and the fitness parameter reaches the limit Υ at generation $t = 854$. Since both the mutant population size and the mutation rate, μ_1, are relatively small, no resistant individual is present in the population when the fitness parameter reaches the limit. However, at generation $t = 1058$, a resistant mutant is generated, and since this type of individual is very advantageous, it is expected to invade the whole population. And indeed, as illustrated in Figure 2(d), the growth of resistant mutants lies close to the expected function.
The parameter sizes are: $N = 10^5$, $\Upsilon = 0.1$, $\mu = 10^{-5}$ and $F = 10^{-6}$.
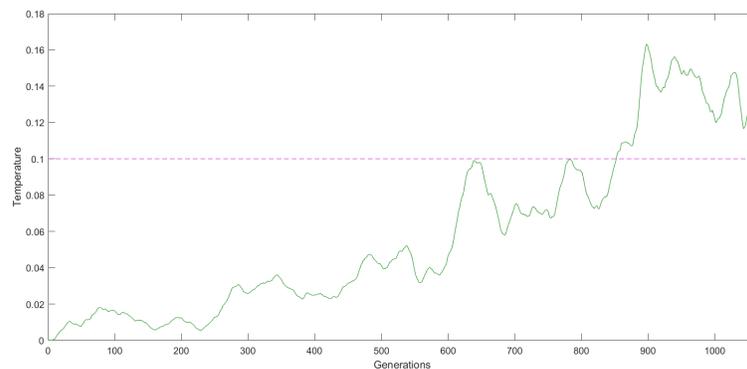
Figure 2b displays the fitness parameter and the pink dashed line marks the limit Υ

37
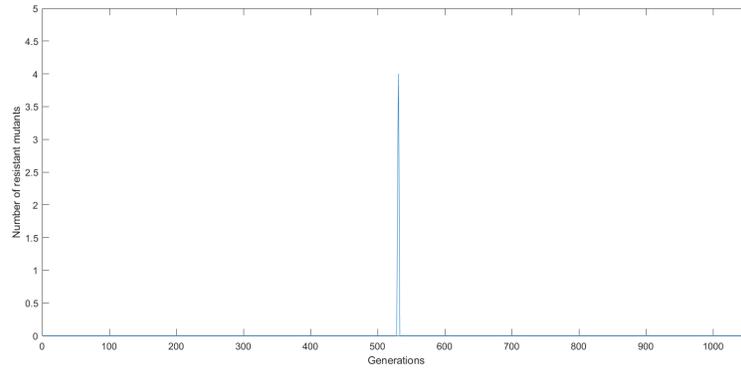
Figure 2c displays the number of resistant mutants before the invasion
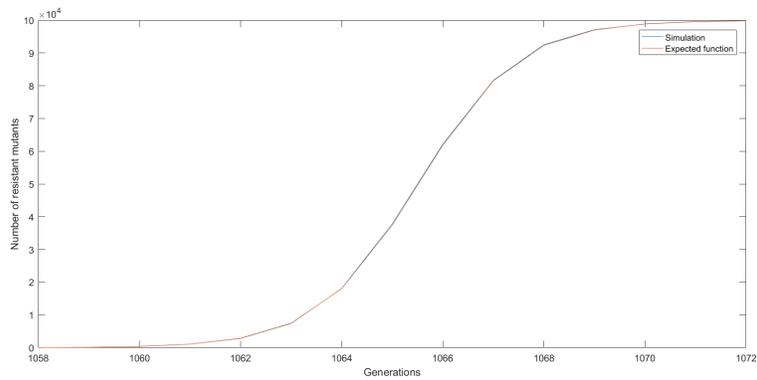
Figure 2d displays the invasion of resistant mutants, both the simulation and the expected function

38
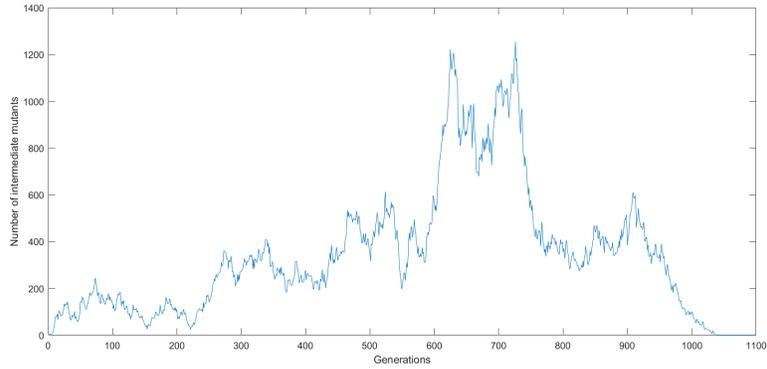
Figure 3a displays the number of intermediate mutants

Figure 3 Low mutation rate and high diffusion rate: Extinction of mutants

The first mutant is generated at generation $t = 0$, and due to stochasticity, the mutant population grows in number and the fitness parameter reaches the limit Υ at generation $t = 630$. Since both the mutant population size and the mutation rate, μ_1, are relatively small, no resistant individual is present in the population when the fitness parameter reaches the limit. Moreover, since the diffusion rate of the fitness parameter, F, is relatively high, the fitness parameter starts to decrease when the number of mutants decreases. Hence, the mutant population goes extinct.
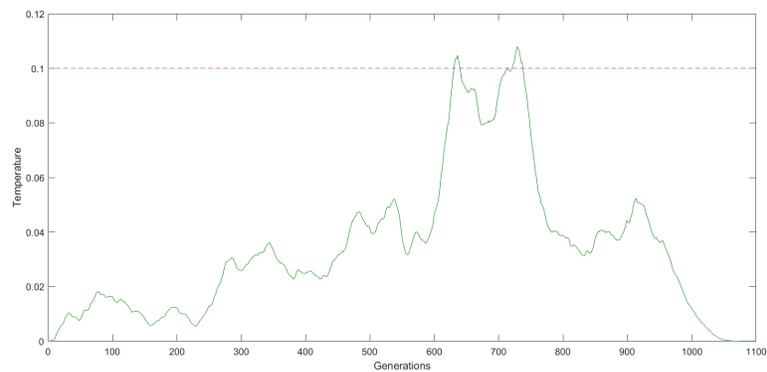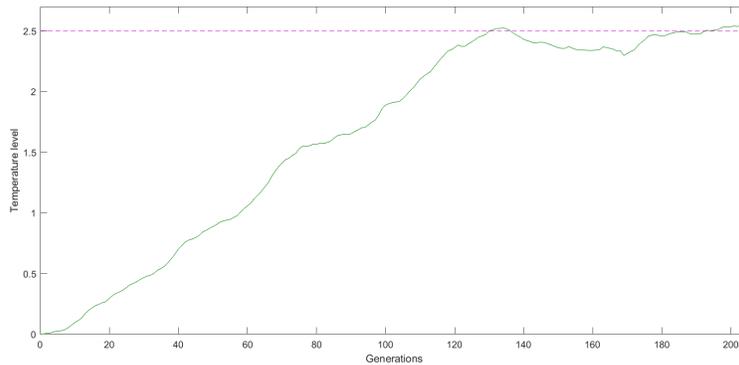

Figure 3b displays the fitness parameter and the pink dashed line marks the limit Υ

39
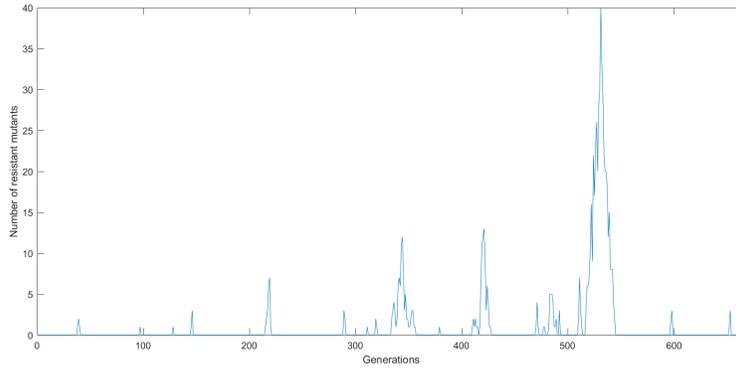
Figure 3c displays the number of resistant mutants
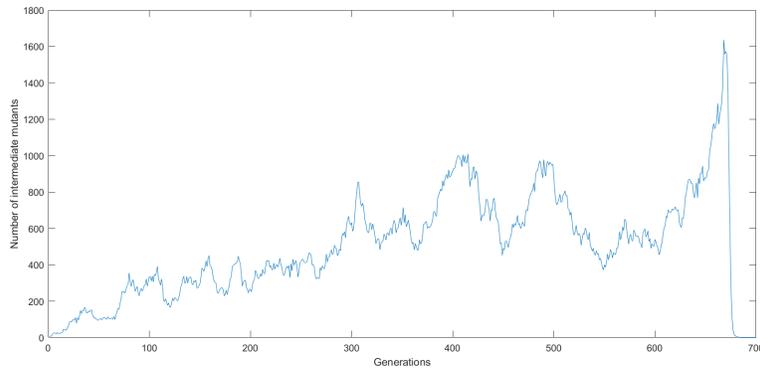
Figure 4a displays the number of intermediate mutants

High mutation rate: Invasion of mutants

The first mutant is generated at generation $t = 0$, and due to stochasticity, the mutant population grows in number and the fitness parameter reaches the limit Υ at generation $t = 662$. Since the mutation rate, μ_1, is relatively large, resistant individuals are present in the population when the fitness parameter reaches Υ. Moreover, since this type of individual is very advantageous, it is expected to invade the whole population. And indeed, as illustrated in Figure 3(d), the resistant mutants invade the population.

The parameter sizes are: $N = 10^5$, $\Upsilon = 0.1$, $\mu = 10^{-4}$ and $F = 10^{-6}$
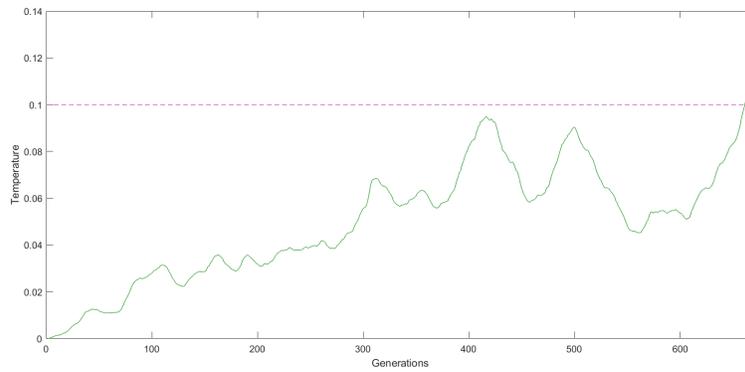
Figure 4b displays the fitness parameter and the pink dashed line marks the limit Υ
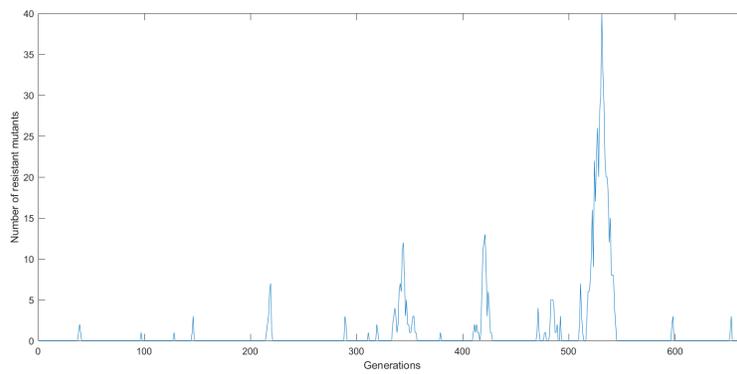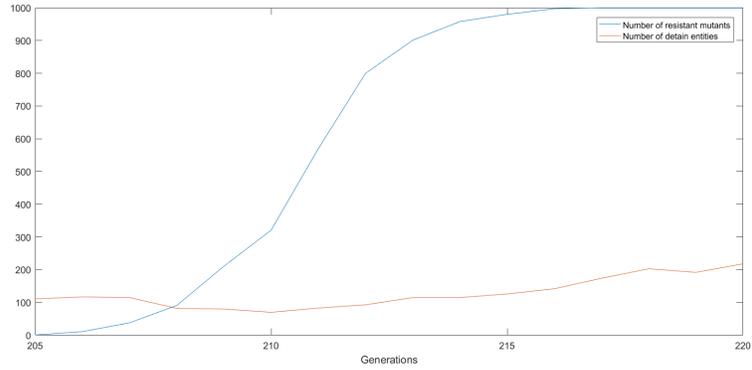

Figure 4b displays the number of resistant mutants before the invasion
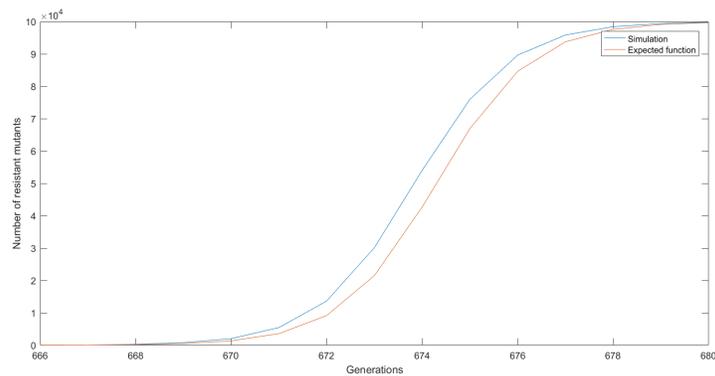
Figure 4c displays the invasion of resistant mutants, both the simulation and the expected function

41
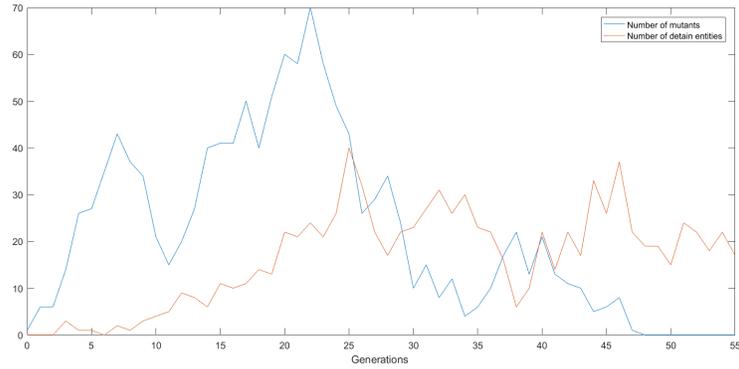
Figure 5a displays the number of mutants and the number of cooperation entities

The cooperation entities make the mutants disadvantageous
The first mutant is generated at generation $t = 0$, and almost immediately after, a cooperation entity is activated. The population of cooperation entities grows in number, whereas the mutant population gets extinct before the fitness parameter reaches the limit Υ.
The parameter sizes are: $N = 10^3$, $\Upsilon = 2.5$, $\mu = 10^{-3}$, $P = 0.01$, $F = 10^{-5}$.
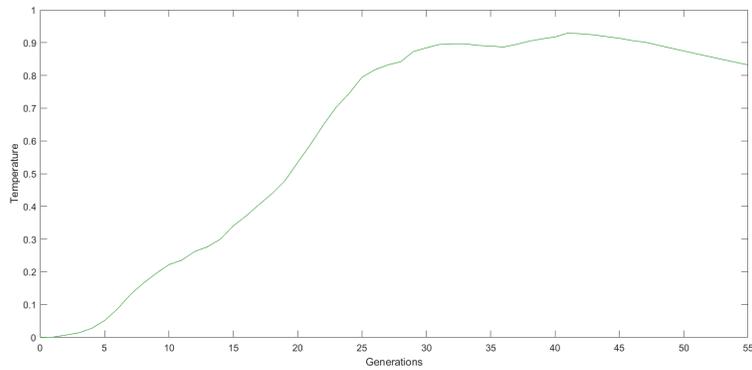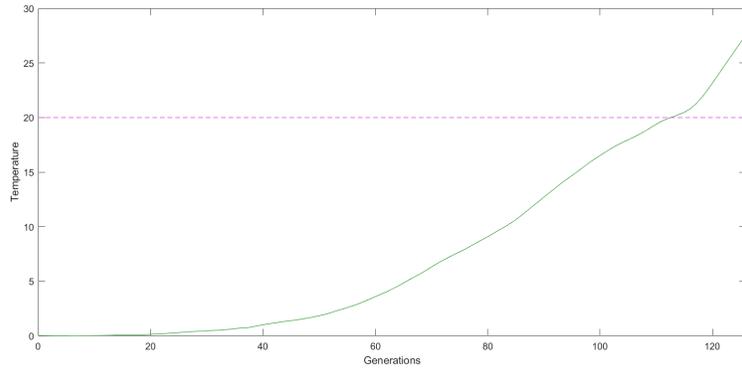

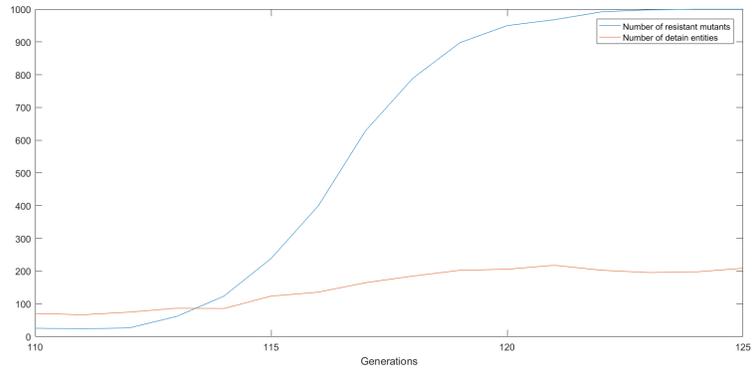Figure 5b displays the fitness parameter and the pink dashed line marks the limit Υ
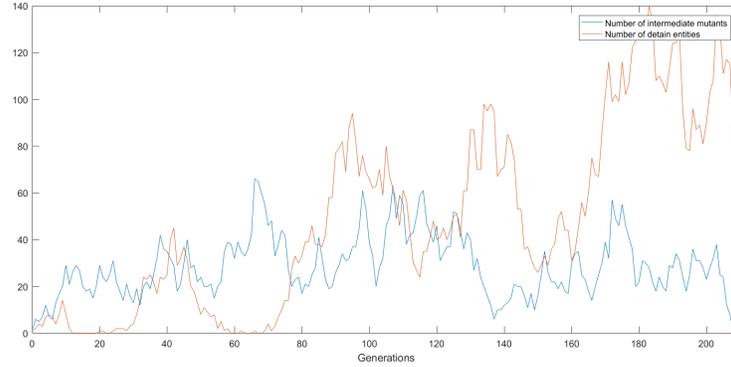
Figure 6a displays the number of intermediate mutants and the number of cooperation entities before the resistant mutants invade the whole population

Figure 6 The mutants lay low before they invade
The first mutant is generated at generation $t = 0$, and immediately after, a cooperaion entity is activated. The population of cooperation entities grows in number and prevent the mutant population from expansion. However, the mutants avoid extinction, and survive long enough to raise the fitness parameter above Υ. Since both the mutation rate, μ, and the mutant population size are relatively small, the mutant population contains no resistant when the fitness parameter reaches Υ. However, after 75 generations, a resistant mutant is generated, and since the resistant mutants are advantageous when the number of cooperation entities is less N and the fitness parameter is above Υ, the resistant mutants invade the whole population.
The parameter sizes are: $N = 10^3$, $\Upsilon = 2.5$, $\mu = 10^{-3}$, $P = 0.01$, $F = 10-5$.

Figure 6b displays the fitness parameter and the pink dashed line marks the limit Υ

43

Figure 6c displays the number of resistant mutants and the number of cooperation entities as the resistant mutants invade the whole population
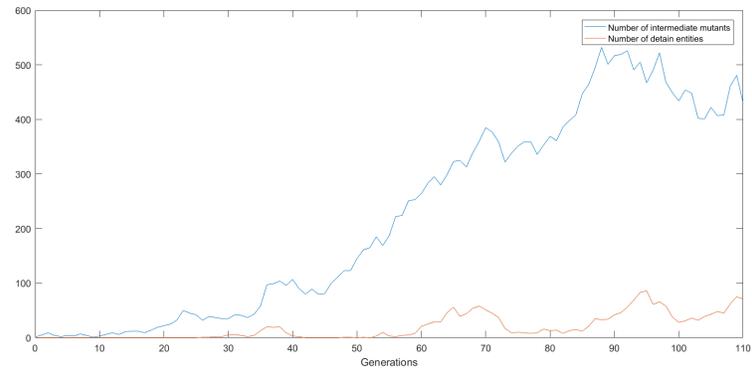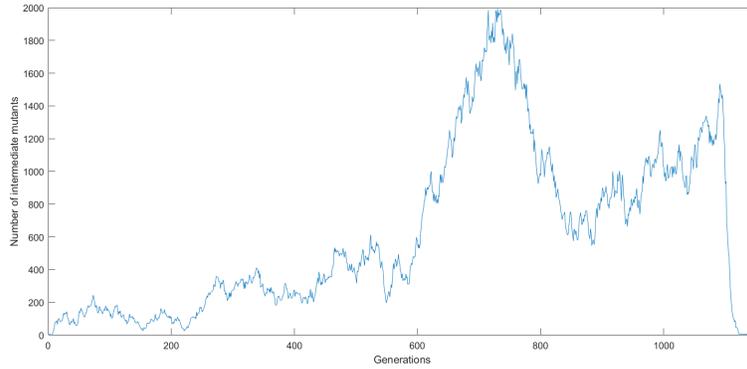

Figure 7a displays the number of intermediate mutants and the number of cooperation entities before the resistant mutants invade the whole population

Figure 7 The mutants grow faster than the cooperation entities and invade the whole population

The first mutant is generated at generation $t = 0$, and due to stochasticity, the mutant population grows fast whereas the growth of the cooperation entities is delayed. Consequently, the fitness parameter reaches the limit Υ. Since the population size is relatively large, there have already been generated resistant mutants, and since the resistant mutants are advantageous when the number of cooperation entities is less N and the fitness parameter is above Υ, the resistant mutants invade the whole population.

The parameter sizes are: $N = 10^3$, $\Upsilon = 20$, $\mu = 10^{-3}$, $P = 0.01$ and $F = 10^{-5}$.

Figure 7b displays the fitness parameter and the pink dashed line marks the limit Υ.

Figure 7c displays the number of resistant mutants and the number of cooperation entities as the resistant mutants invade the whole population

Figure 8a display the number of intermediate mutants

Figure 8 High activation rate of the cooperation entities: The mutants get extinct

A mutant appears in the population at generation $t = 0$ and generates a lineage of mutants that survives long enough such that the fitness parameter reaches the limit Υ at generation $t = 1092$. The mutation rate, μ_1 is relatively low and when the fitness parameter reaches Υ, there are no resistant mutants in the population. Moreover, the activation rate of the cooperation entities is very high, and hence, the number of cooperation entities grows rapidly, whereas the mutants become increasingly disadvantageous and decrease fast towards zero.

The parameter sizes are: $N = 10^5$, $\Upsilon = 5$, $\mu = 10^{(-5)}$, $P = 1$ and $F = 10^{(-8)}$.
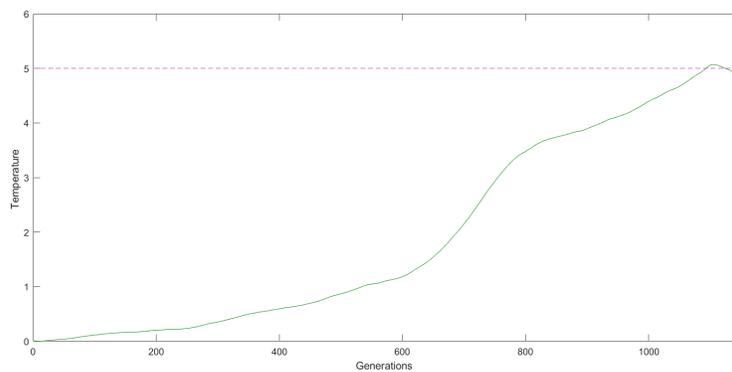
Figure 8b displays the fitness and the pink dotted line marks the limit Υ
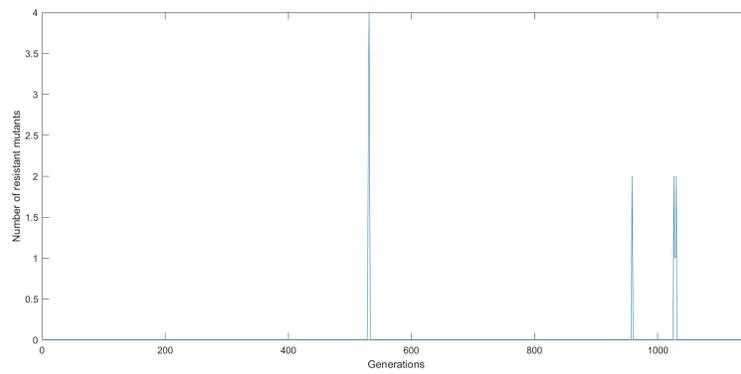
Figure 8c displays the number of resistant mutants
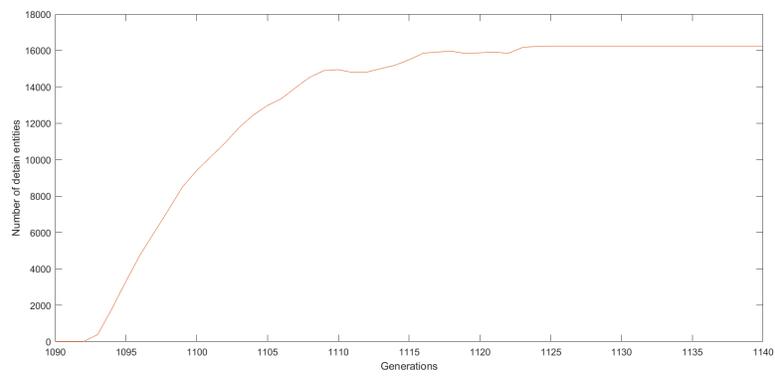
Figure 8d displays the number of cooperation entities
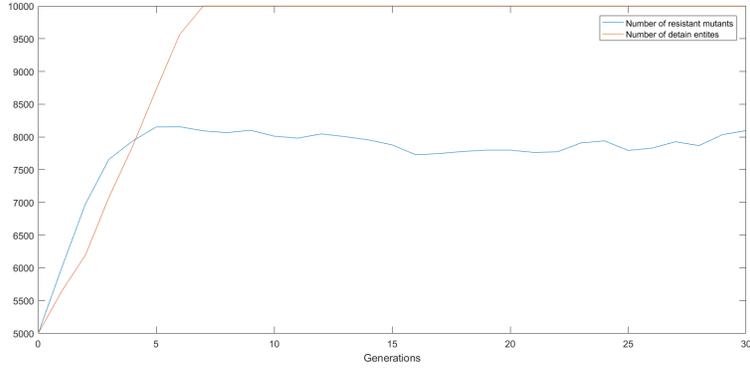
Figure 9 N **cooperation entities**
This Figure shows the case when the initial number of cooperation entities is $N = 10^4$, whereas the initial number of wild-type individuals, resistant mutants and intermediate mutants are 10, 10 and $N - 20$, respectively. Since the intermediate mutants are disadvantageous whereas the wild-type individuals and the resistant mutants are neutral variants, it is expected that the intermediate mutants get extinct while the wild-type individuals and the resistant mutants both grow towards $N/2$. The competition dynamics between the wild type and the resistant mutants is characterised by great variance whereas the number of intermediate mutants follows the expected function closely.
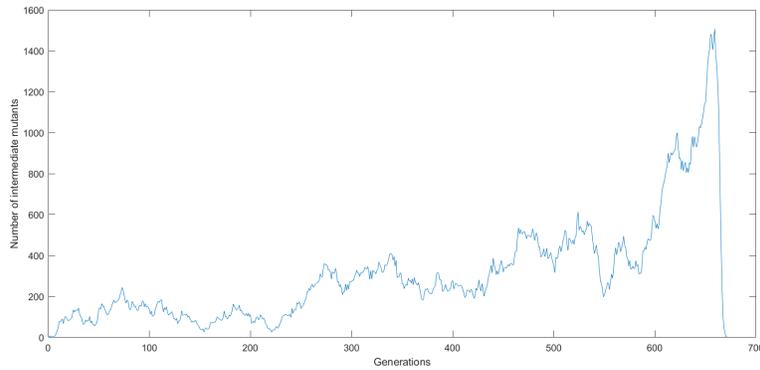
Figure 10a displays the number of intermediate mutants

Figure 10 Moderate activation rate of cooperation entities and invasion of resistant mutants
A mutant appears in the population at generation $t = 0$ and generates a lineage of mutants that survive long enough such that the fitness parameter reaches the limit Υ. When the fitness parameter reaches this limit, the population of mutants has already generated four resistant individuals, which are advantageous as long as the number of cooperation entities are lower than the population size, N. Even though the number of cooperation entities grows quit quickly, the number of resistant mutants reaches N first.
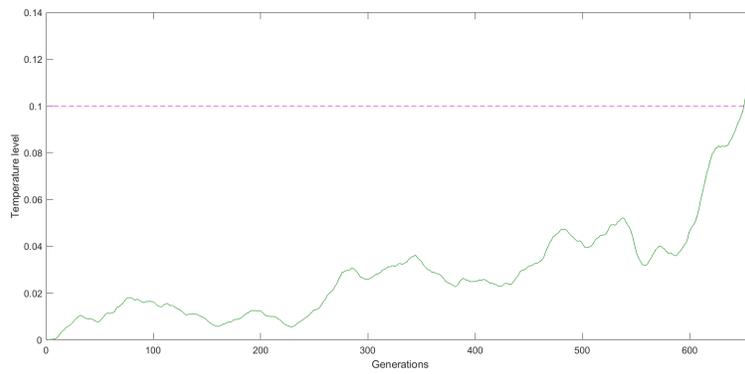The parameter sizes are: $N = 10^5$, $\Upsilon = 0.1$, $\mu = 10^{-5}$, $P = 0.1$ and $F = 10^{-6}$.

48

Figure 10b displays the fitness parameter and the pink dashed line marks the limit Υ
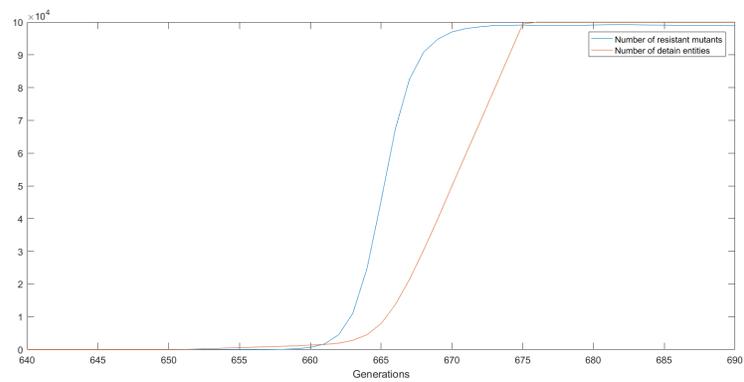

Figure 10c displays the number of resistant mutants and the number of cooperation entities
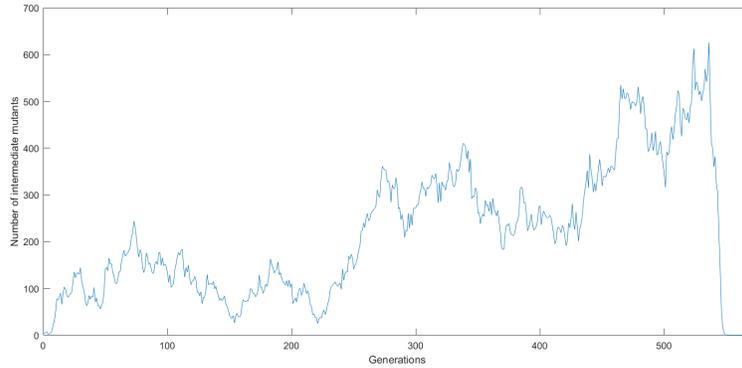
Figure 11a displays the number of intermediate mutants

Figure 11 High activation rate of cooperation entities and coexistence of resistant mutants and wild-type individuals
A mutant appears in the population at generation $t = 0$ and generates a lineage of mutants that survive long enough such that the fitness parameter reaches the limit Υ. When the fitness parameter reaches this limit, the population of mutants has already generated four resistant individuals, which are advantageous as long as the number of cooperation entities are lower than the population size, N. However, since the activation rate of the cooperation entities is very high, the number of cooperation entities reaches N when the number of resistant mutants is approximately 4×10^4. Since the wild-type individuals and resistant mutants become neutral variants, the number of each type is expected to remain constant.
The parameter sizes are: $N = 10^5$, $\Upsilon = 1$, $\mu = 10^{-5}$, $P = 1$ and $F = 10^{-8}$.
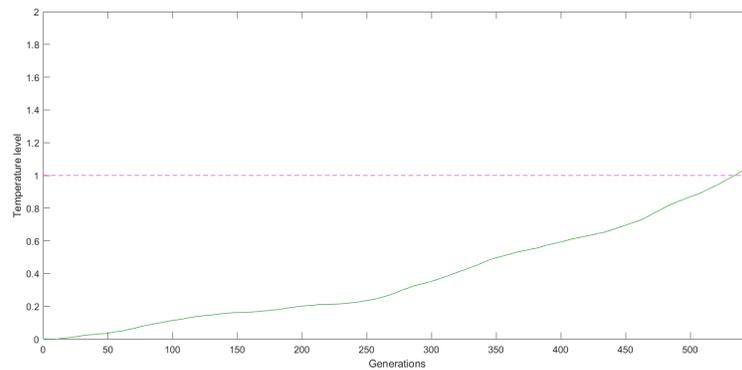

Figure 11b displays the fitness parameter and the pink dashed line marks the limit Υ
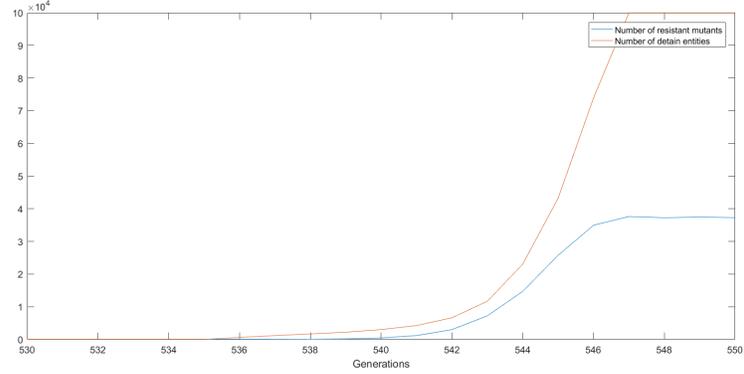
50

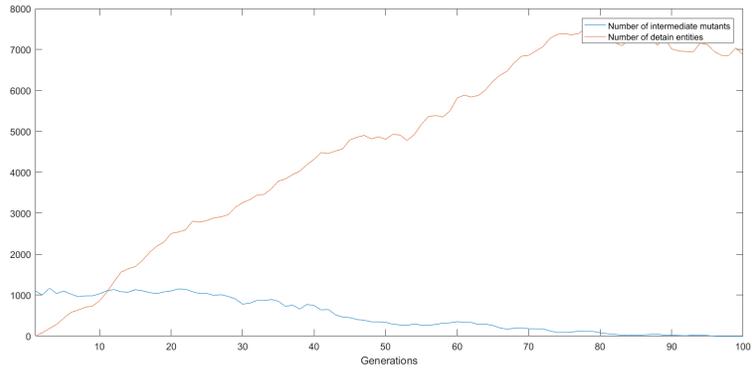Figure 11c displays the number of resistant mutants and the number of cooperation entities

Figure 12a displays the number of intermediate mutants and the number of cooperation entities before the resistant mutants invade the population

Figure 12 Low activation rate of cooperation entities and delayed invasion of resistant mutants

The fitness parameter reaches the limit Υ at generation $t = 0$, and all the mutants are non-resistant from generation $t = 0$ to generation $t = 432$. Thus, these mutants become increasingly disadvantageous as the number of cooperation entities grows. Since the diffusion rate, F, equals zero, the fitness parameter remains above Υ even though the number of mutants decreases towards zero. A resistant mutant is generated at generation $t = 433$, and as long as the number cooperation entities is less than the population size, N, the resistant mutants are advantageous and grow exponentially. Since the activation rate is relatively low, the population of resistant mutants beats the cooperation entities in the race towards N. The initial number of intermediate mutants is $i = 558$. The other parameter sizes are: $N = 10^5$, $\mu = 10^{-5}$, $P = 0.01$ and $F = 0$.
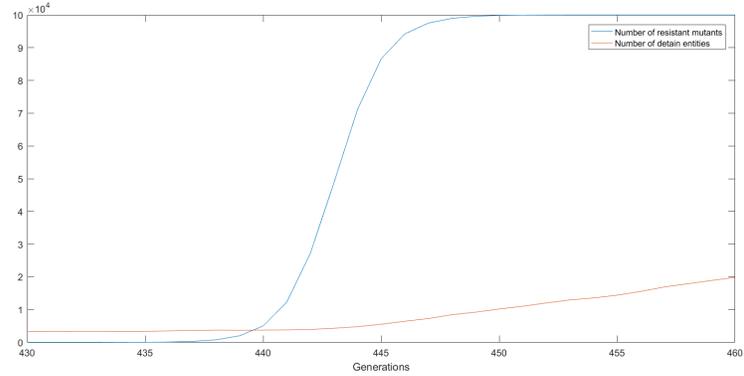
51

Figure 12b displays the number of resistant mutants and the number of cooperation entities as the resistant mutants invade the whole population
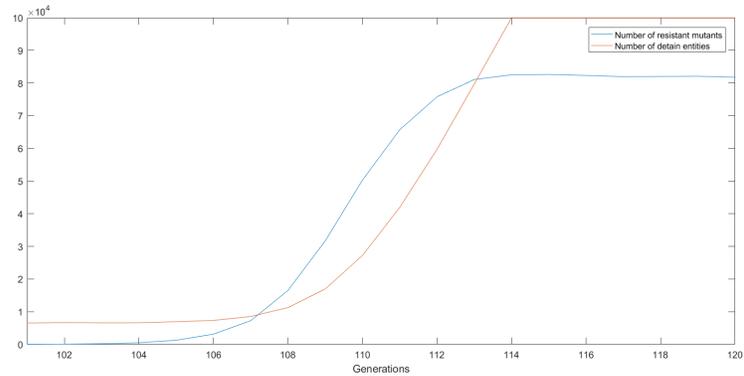

Figure 13a displays the number of intermediate mutants and the number of cooperation entities before the first resistant mutant is generated

Figure 13 Moderate activation rate of cooperation entities and co-existence of resistant mutants and wild-type individuals
The fitness parameter reaches the limit Υ at generation $t = 0$, and all the mutants are non-resistant from generation $t = 0$ to generation $t = 101$. Thus, these mutants become increasingly disadvantageous as the number of cooperation entities grows. Since the diffusion rate, F, equals zero, the fitness parameter remains above Υ even though the number of mutants decreases towards zero. A resistant mutant is generated at generation $t = 101$, and as long as the number cooperation entities is less than the population size, N, the resistant mutants are advantageous and grow exponentially. However, since the activation rate of the cooperation entities is sufficiently high, the number of cooperation entities reaches N when the number of resistant mutants is approximately 8×10^4. Since the wild-type individuals and resistant mutants become neutral variants, the number of each type is expected to remain constant. The initial number of intermediate mutants is $i = 1107$. The other parameter sizes are: $N = 10^5$, $\Upsilon = 0.1$, $\mu = 10^{-4}$, $P = 0.25$ and $F = 0$.
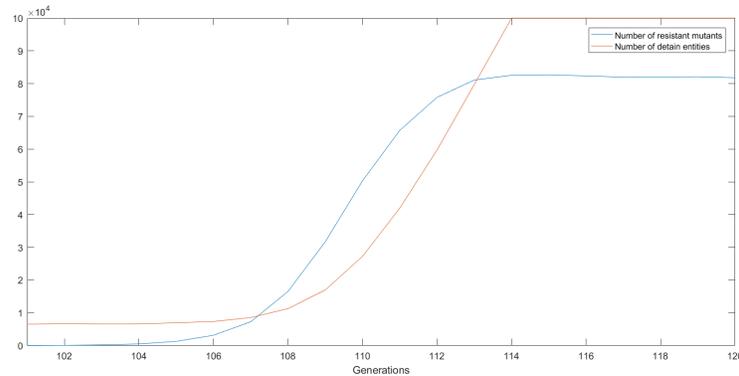
Figure 13 b displays the number of resistant mutants and the number of cooperation entities

References

[1] Moran, P. Random processes in genetics. Mathematical Proceedings of the Cambridge Philosophical Society 54, 60–71 (1958).

[2] Nowak, M. Evolutionary dynamics. Havard University Press (2006).

[3] Wodarz, D. and Komarova, N. L. Computational biology of cancer: lecture notes and mathematical modeling. World Scientific (2005).

[4] Ashcroft, P., Michor, F. and Galla, T. Stochastic tunneling and metastable states during the somatic evolution of cancer. Genetics 199, 1213–1228 (2015).

[5] Komarova, N. L. Spatial stochastic models for cancer initiation and progression. Bulletin of Mathematical Biology 68, 1573–1599 (2006).

[6] Michor, F., Iwasa, Y. and Nowak, M. Dynamics of cancer progression. Nature Reviews 4, 197–205 (2004).

[7] Nowak, M., Komarova, N., Sengupta, A., Jallepalli, P.V., et al. The role of chromosomal instability in tumor initiation. PNAS 99, 16226–16231 (2002).

[8] Breivik, J. Don't stop for repairs in a war zone: Darwinian evolution unites genes and environment in cancer development. PNAS 98, 5379–5381 (2001).

[9] Smith, J.M. The theory of games and the evolution of animal conflict. Journal of Theoretical Biology 47, 209–221 (1974).

[10] Axelrod, R. and Hamilton, W.D. The evolution of cooperation. Science 27, 1390–1396 (1981).

[11] Samuelson, L. Evolution of game theory. Journal of Economic Perspective 16, 47–66 (2002).

[12] Antal, T., Traulsen, A., Ohtsuki, H., Tarnita, C.E., and Nowak, M.A. Mutation-selection equilibrium in games with multiple strategies. Journal of Theoretical Biology 258 (2009).

[13] Nash, J. Non-cooperative games. Annals of Mathematics 54, 286–295 (1951).

[14] Kaveh, K., Veller, C. and Nowak, M. Games of multicellularity. Journal of Theoretical Biology 403, 143–158 (2016).

[15] Sompayrac, L. How the immune system works. Wiley-Blackwell, 5th edition (2015).

[16] Koebel, C.M., Vermi, W., Swann, J.B., Zerafa, N., Rodig, S.J., Old, L.J., et al. Adaptive immunity maintains occult cancer in an equilibrium state. Nature 450, 903–710 (2007).

[17] Haabeth, O.A., Tveita, A.A., Fauskanger, M., Schjesvold, F., Lorvik, K.B., et al. How do CD4+ T cells detect and eliminate tumor cells that either lack or express MHC class II molecules? Front Immunol. 5, Article 174 (2014).

[18] Klenerman, P. The immune system: a very short introduction. Oxford University Press (2017).

[19] Archetti, M. Evolutionary dynamics of the Warburg effect: glycolysis as a collective action problem among cancer cells. Journal of Theoretical Biology 342, 1–8 (2014).

[20] Bernards, R. and Weinberg, R.A. A progression puzzle. Nature 22 (2002).

[21] Archetti, M. Heterogeneity and proliferation of invasive cancer subclones in game theory models of the Warburg effect. Cell Proliferation 48 (2015).

[22] Robertson-Tessi, M., Gillies, R.J., Gatenby, R.A. and Anderson, A.R.A. Non-linear tumor-immune interactions arising from spatial metabolic heterogeneity. bioRxiv 038273 (2016).

[23] Gravenmier, C.A., Siddique, M. and Gatenby, R.A. Adaptation to stochastic temporal variations in intratumoral blood flow: the Warburg effect as a bet hedging strategy. Bulletin of Mathematical Biology (2017).

[24] Kaznatcheev, A., Vander Velde, R., Scott, J.G. and Basanta, D. Cancer treatment scheduling and dynamic heterogeneity in social dilemmas of tumour acidity and vasculature. British Journal of Cancer 116, 785–792 (2017).

[25] Norris, J.R, Markov chains. Cambridge University Press (1998).

[26] Ephraim, J.F. and Matzinger, P. Is cancer dangerous to the immune system? Seminars in Immunology 8, 271280 (1996).

[27] Lambert, G., Vyawahare, S. and Austin, R.H. Bacteria and game theory: the rise and fall of cooperation in spatially heterogeneous environments. Interface Focus 4 (2014).

[28] Turner, P. and Chao, L. Prisoner's dilemma in an RNA virus. Nature 398, 441–443 (1999).

[29] Nowak, M. and Sigmund, K. Phage-lift for game theory. Nature 389, 367–368 (1999).

[30] Marx, K. and Engels, F. Manifesto of the Comunist party. Progress Publishers (1848).

[31] Måløy, M. Frp bør legge innvandrings– og integreringsministeren på bordet. Minerva (15.12. 2017). Retrieved from: https://www.minervanett.no/frp-bor-legge-innvandrings-integreringsministeren-pa-bordet/

[32] Bell, R., Mieth, L. and Buchner, A. Separating conditional and unconditional cooperation in sequential prisoners dilemma game. PLoS ONE 12 (2017).

[33] Dunn, G.P., Bruce, A.T., Ikeda, H., Old, L.J. and Schreiber, R.D. Cancer immunoediting: from immunosurveillance to tumor escape. Nature Immunology 3, 991998 (2002).

[34] Wodarz, D. Stem cell regulation and the development of blast crisis in chronic myeloid leukemia: implications for the outcome of Imatinib treatment and discontinuation. Medical Hypotheses 70, 128-136 (2008).

[35] Høyem, M.R., Måløy, F., Jakobsen, P. and Brandsdal, B.O. Stem cell regulation: implications when differentiated cells regulate symmetric stem cell division. Journal of Theoretical Biology 380, 203–219 (2015).

[36] Måløy, M., Måløy, F., Jakobsen, P. and Brandsdal, B.O. Dynamic self-organisation of haematopoiesis and (a)symmetric cell division. Journal of Theoretical Biology 414, 147–164 (2017).

[37] Meyerson, R.B. Game theory: analysis of conflict. Harvard University Press (1997).

[38] Angus, I. Marx and Engels... and Darwin? International Socialist Review 65 (2009). Retrieved from: https://isreview.org/issue/65/marx-and-engelsand-darwin

[39] Lindberg, A. Sylvi Listhaug ar Norges Trump. Aftonbladet (30.11. 2017). Retrieved from: https://www.aftonbladet.se/ledare/a/B8bAl/sylvi-listhaug-ar-norges-trump

[40] Rustad, H. Erna cruiser inn til seier. Document.no (29.11. 2017). Retrieved from: https://www.document.no/2017/08/29/erna-cruiser-inn-til-seier/

[41] Pinker, S., Nowak, M.A. and Lee, J.J. The logic of indirect speech. PNAS 105, 833–838 (2007).

[42] Brown, P. and Levinson, S.C. Politeness: some universals in language usage. New York: Cambridge University Press (1987).

[43] Kolbeinstveit, L. Selv om det er Sylvi Listhaug som sier det, er stram innvandringspolitikk fortsatt viktig for integreringen. Minerva (31.08. 2017). Retrieved from: https://www.minervanett.no/sylvi-listhaug-sier-stram-innvandringspolitikk-fortsatt-viktig-integreringen/

[44] Sylvi Listhaug resigns as Norway's justice minister. The Local (20.03.2018) Retrieved from: https://www.thelocal.no/20180320/sylvi-listhaug-resigns-as-norways-justice-minister.

[45] Bugge, S., Holmes, M. and Thanem, T. Fersk måling: Frp fosser frem etter mistillitsforslaget mot Listhaug. Dagbladet (16.03.2018) Retrieved from: https://www.vg.no/nyheter/innenriks/i/qn79Kz/fersk-maaling-frp-fosser-frem-etter-mistillitsforslaget-mot-listhaug.

[46] Dønvold-Myhre, L. Innvandringsdebatten på internett økte under Listhaug-saken. NRK (22.03.2018) Retrieved from: https://www.nrk.no/norge/innvandringsdebatten-pa-internett-okte-under-Listhaug-saken.

Liste over publikasjoner i ph.d.-avhandling

Vitenskapelige publikasjoner

Marthe Måløy, Frode Måløy, Rafael Lahoz-Beltra, Juan Carlos Nuno og Antonio Bru (2018) An extended Moran process that captures the struggle for fitness. Under review

Arbeidsfordeling: Rafael Lahoz-Beltra, Juan Carlos Nuno, Antonio Bru og Marthe Måløy: Utvikling av modell.

Marthe Måløy: Skrive artikkel, programmering av modell, korrespondanse med journal.

Frode Måløy: Korrektur lesning, programmering av modell

Marthe Måløy, Frode Måløy, Per Jakobsen, Bjørn Olav Brandsdal (2017) Dynamic self-organisation of haematopoiesis and (a)symmetric cell division. *Journal of Theoretical Biology*.

Marthe Måløy: Utvikling av modell, skrive artikkel, programmering av modell, korrespondanse med journal.

Frode Måløy: Korrektur lesning, programmering av modell

Per Jakobsen: Samtaler om modellering og forklaringsnivå.

Marte Rørvik Høyem, Frode Måløy, Per Jakobsen, Bjørn Olav Brandsdal (2015) Stem cell regulation: Implications when differentiated cells regulate symmetric stem cell division. *Journal of Theoretical Biology.*

Marthe Måløy: Utvikling av modell, skrive artikkel, programmering av modell, korrespondanse med journal.

Frode Måløy: Korrektur lesning, programmering av modell

Per Jakobsen: Samtaler om modellering og forklaringsnivå.

Marthe Måløy

Per Jakobsen