School of Mathematics of Statistics

Newcastle University

Modelling the Proliferation of Human Stem Cells

MMath Report

By Lesley Elliot

Supervisors: Dr. Nick Parker and Prof. Anvar Shukurov

April 28, 2015

Contents

1 Introduction **1**
- 1.1 Stem Cells . . . 1
- 1.2 Stem Cell Experiments . . . 2
- 1.3 Goals of the Project . . . 3

2 Constructing a Mathematical Model **4**
- 2.1 Key Features of Stem Cells . . . 4
- 2.2 Types of Mathematical Model . . . 7
- 2.3 Agent-Based Models . . . 8

3 Modelling Stem Cell Migration **10**
- 3.1 One-Dimensional Random Walk . . . 10
- 3.2 Two-Dimensional Random Walk . . . 12
- 3.3 Two-Dimensional Biased Random Walk . . . 15

4 The Experimental Migration of Stem Cells **18**
- 4.1 Parameters . . . 18
- 4.2 Diffusion of Stem Cells . . . 20
- 4.3 Mutual Attraction of Stem Cells . . . 21

5 Growth of Stem Cell Colonies **24**
- 5.1 Initial conditions . . . 24
- 5.2 Cell Division . . . 25
  - 5.2.1 How Does a Colony Accommodate Daughter Cells? . . . 25
  - 5.2.2 A Model of a Growing Colony . . . 26
- 5.3 Quantifying the Growth of Colonies . . . 28
  - 5.3.1 The Cell Number . . . 28
  - 5.3.2 The Colony Area . . . 29
- 5.4 Stem Cell Lineages . . . 30
- 5.5 Including the Cell Cycle . . . 33
- 5.6 Including Differentiation . . . 35

6 Conclusions **38**
- 6.1 Further Developments . . . 39

7 Bibliography **41**

Abstract

Human induced pluripotent stem cells (iPSCs) are produced *in vitro* from specialised cells, resolving the ethical issue of using human embryonic stem cells in biomedical experiments. Their production is, however, currently very slow, costly and inefficient.

I model the behaviours of iPSCs with an agent-based approach (in Matlab) to gain insight into the mechanisms of their evolution in specific experiments, both published and currently conducted in the Institute of Genetic Medicine. I have developed models of isotropic and biased random walks in 2D that accurately reproduce experiments on the movements of single and several iPSCs. When the cells are more than 150 μm apart their motion is close to an isotropic random walk. Cells within 70 μm move in a biased way, being strongly attracted at separations less than 30 μm; these behaviours are successfully modelled as a strongly biased random walk with attraction. The model is then extended to describe the growth of iPSC colonies with allowance for cell division and spontaneous differentiation. I identify three aspects of the iPSC behaviour which require experimental clarification: how a colony rearranges to accommodate cells produced by division, if the division rate of cells is uniform across a colony and the rate of cell differentiation within a colony. The model is further developed to include these features so comparison against experimental data can answer these questions and the efficiency of future experiments can be optimised.

Chapter 1

Introduction

1.1 Stem Cells

Every organism is made up of **specialised cells**, which have specific functions, and **unspecialised cells** with no specific functional purpose. Examples of specialised cells are red blood cells or liver cells in humans. Unspecialised cells are called **stem cells**, and although they do not have a specific function, they do have the ability to **differentiate** into (turn into) specialised cells. Every cell (specialised or unspecialised) contains the same genes but some are turned off and some are active. The combination of on and off genes gives a specialised cell its specific function. A stem cell has enough genes turned on that it can differentiate into a specialised cell by turning particular genes off. Until a stem cell differentiates, it will continue to renew itself (divide) to give rise to daughter stem cells [24].

The number of specialised cells that a stem cell can differentiate into is referred to as its **potency**, and is determined by the number of active genes it has available to turn on and off. There are varying levels of potency [12, 6], with the highest level being totipotency. Totipotent cells have the ability to give rise to any specialised cells in an organism; examples are zygotes or spores. The next lower level is **pluripotency**. Pluripotent cells are able to give rise to many specialised cells but not all of those which totipotent cells can. There is a spectrum ranging from completely pluripotent cells, which have the ability to differentiate into any cell that an embryonic cell can, to partly pluripotent cells which are still at least able to give rise to all cells of the three germ layers [12]. Levels of potency below this decrease in the number of specialised cells they are able to give rise to, with the lowest being that of unipotent cells which are only able to differentiate into one cell lineage [7] .

Stem cells can occur naturally, for example embryonic stem cells, or can be artificially produced by inducing pluripotency. Induced pluripotent stem cells (**iPSCs**) are created *in vitro* by introducing specific genes into **somatic** cells (adult specialised cells) which give them similar properties to those of embryonic stem cells. Pluripotency can be induced in a few ways and there are different types of somatic cell that can be used [6, 11, 25]; in Section 1.2 I will discuss the specific method used in the experiment I aim to model. As the induced stem cells are pluripotent, they have the same potency as human embryonic stem cells (**hESCs**) so they can differentiate into any specialised cell an embryonic stem cell can.

The scientific benefit of stem cells is huge. They can be used in drug development and to model the spread of disease [6] due to the highly controlled and clean environment they

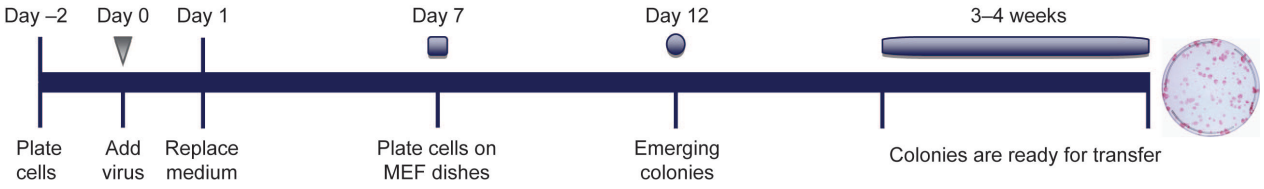


Figure 1.1: The experiment time-line of the CytoTune^{TM}-iPS reprogramming experiment [22].

provide for *in vitro* testing and their ability to regenerate indefinitely [20]. The great advantage of using iPSCs over hESCs is that this resolves the ethical controversy of extracting cells from human embryos by providing an equally potent *in vitro* alternative. Another exciting potential of iPSCs is using their potent ability in regenerative medicine to produce immune matched organs for patients in need of transplants, although there is still a lot more research needed in this area [6]. This could be achieved by extracting somatic cells from the patient, turning them into iPSCs *in vitro* and then influencing the iPSCs to differentiate into the specific specialised cell(s) required for the organ needed. Immune rejection and the transmission of infections and viruses can be mitigated by using a patient's own cells [21].

1.2 Stem Cell Experiments

As previously mentioned, there are various ways of inducing pluripotency in cells. The method I will focus on, and develop a mathematical model of, is that used at the Institute of Genetic Medicine based at the International Centre for Life in Newcastle-upon-Tyne.

The experiments are based on the CytoTune^{TM}-iPS Reprogramming Kit [22]. This method reprograms somatic **fibroblast** cells, found in the connective tissues of humans, into induced pluripotent stem cells. The time-line for this process is shown in Figure 1.1. Initial fibroblast cells are grown in culture for 2–3 days until they have 80% confluency (coverage) in the well. Unlike stem cells, fibroblasts exhibit **contact inhibition**. This means that their division rate decreases if they are in close contact with each other. Therefore, 80% confluency is the optimal compromise between the number of cells and their growth rate. At this point, **transduction** takes place: the fibroblast cells are transduced (treated) with 4 specific strains of the Sendai virus (SeV) which contain the 4 necessary transcription factors, Oct4, Sox2, Klf4, and cMyc (which hold the necessary genes), to induce pluripotency. The cells are then cultured and grown in wells with the first colonies of iPSCs appearing around 12 days after transduction.

Stem cells are different to fibroblasts in the way they grow. Firstly, they do not exhibit the issues of contact inhibition. In fact, their growth is aided by being in contact with other stem cells, so they grow in colonies. They have a much shorter division time of around 14–16 hours, compared to 26–30 hours for fibroblast cells [9]. About 3–4 weeks after transduction the confluency of iPSCs in the wells is large enough that **passaging** can take place. This involves transferring small sections of a stem cell colony from a highly confluent well into new wells to allow new colonies to grow. Consequently, the number of iPSCs produced increases.

The morphology (appearance) of stem cells differs from that of fibroblasts. Stem cells are small (usually round), with a diameter of approximately 10 μm, and their nuclei occupy most

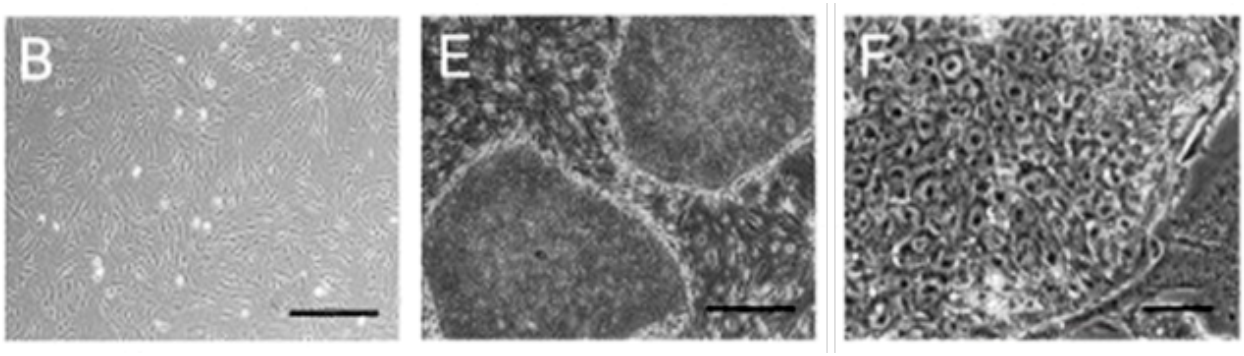


Figure 1.2: The difference in morphology of iPSCs and fibroblast cells. Panel B shows fibroblast cells at the correct confluency for transduction (80%), panel E shows an established iPSC colony and panel F shows an image of iPSC with high magnification [20].

of the cell volume with small amounts of cytoplasm between it and the cell wall. Their shapes are, however, variable and can change to become more square or rectangular as they aim to be in complete adhesion with neighbouring stem cells (to aid their growth). In contrast, fibroblasts have an elongated shape, and although they vary in length, are bigger due to the larger amounts of cytoplasm surrounding their nuclei. Figure 1.2 illustrates these different properties. Figure 1.2(B) shows the elongated shape of fibroblasts and their contact inhibition by their spread. Figure 1.2(E) shows how stem cells form compact colonies and Figure 1.2(F) shows the small circular morphology of the individual stem cells under high magnification. The white circular cells in the images are dead cells.

Other experiments have been carried out at the Centre of Life as a consequence of the work done in this project to allow testing of the mathematical models and clarification of questions that arose from the modelling. These will be discussed throughout the report in the appropriate place.

1.3 Goals of the Project

At the moment the process of producing iPSCs is slow, costly and inefficient, with the iPSCs produced being of highly variable qualities of pluripotency [15]. It costs in excess of £1000 to run the experiment at the Centre for Life which induces pluripotent stem cells and currently less than 1% of the fibroblast cells transduced become completely pluripotent stem cells, so the efficiency is $< 1\%$.

The main aim of my project is to model the production of iPSCs at the Centre for Life in sufficient detail to improve the efficiency of their experiments. This involves gaining an understanding of how to stop or control the death and spontaneous differentiation of cells, and predict the best time and parts of colonies to passage in order to only grow new completely pluripotent colonies. Currently, there is very little fundamental understanding of the mechanisms of the processes that control the iPSC behaviour, which before now had never been mathematically modelled before.

Chapter 2

Constructing a Mathematical Model

In order to create a model which will give us adequate understanding of an experiment we want to incorporate all the important characteristic behaviours of stem cells into it. It is therefore essential to outline what these are before deciding upon a model. The distinguishing features which were identified are now outlined and explained [9].

2.1 Key Features of Stem Cells

Pluripotency/Differentiation – As previously explained, pluripotency is a continuum of states determined by the number of specialised cells a stem cell can give rise to. In the experiments described in Section 1.2, the pluripotency of stem cell colonies is tested and only those exhibiting completely pluripotent characteristics are considered as stem cells.

Figure 2.1 shows an iPSC colony whose nuclei have been stained to show the cells' varying levels of pluripotency. The concentration of the pink indicator shows the level of pluripotency; the paler the pink in the nuclei the less pluripotent the cell is. We can see that the pluripotency varies within the colony. Roughly, the most pluripotent cells seem to occur at the centre of the colony with potency decreasing radially from the centre. This is characteristic of all colonies [9].

Around the edge of the colony in Figure 2.1, the nuclei become sparse and pale because stem cells differentiate into specialised cells. The specialised cells are larger with a large amount of cytoplasm surrounding the nucleus, which the staining does not show, hence the disperse appearance of the stained nuclei. The paleness is due to the fact that the specialised cells are not pluripotent. Our collaborates at the Institute of Genetic Medicine, Dr Majlinda Lako and Dr Irina Neganova, are not sure why this decreasing level of pluripotency occurs in the colony or why cells spontaneously differentiate around the edge. We do, however, know that, on a population level, we expect approximately 10% of the colony to completely lose pluripotency and differentiate. In fact it is thought to be beneficial to maintain this ratio as it appears to give rise to the highest proportion of stem cells that are completely pluripotent [9].

Division/Proliferation – It takes an iPSC between 14 and 16 hours to complete its cell development cycle and divide to give rise to a daughter cell [9]. The stages of the cell

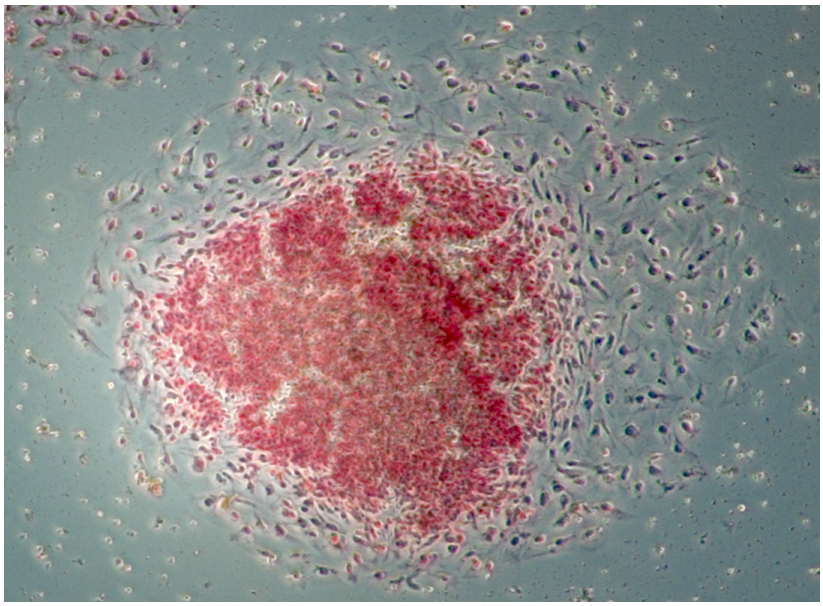


Figure 2.1: A colony of iPSC's whose nuclei have been stained to show their pluripotency. The cells which are more pluripotent have brighter red nuclei. The microscopic image was taken by the author under $\times 10$ magnification. The typical size of a stem cell is about 10μm.

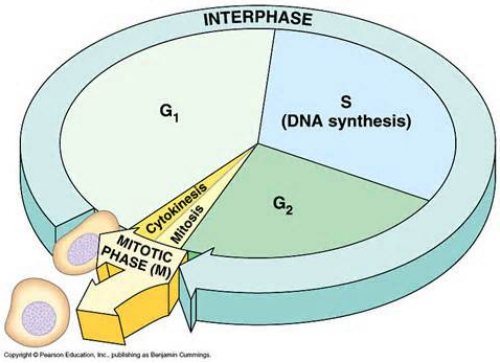


Figure 2.2: A diagram of the four stages of the cell development cycle of an iPSC [1].

cycle, shown in Figure 2.2, are G_1, S, G_2 and finally M. In the G_1 phase, also known as gap 1, the cell is growing; experiments show that this stage typically takes 2.5 hours for an iPSC [9]. The next stage is S, or synthesis, which takes between 6 and 8 hours for an iPSC [9]. In this phase the cell replicates its exact DNA in preparation for division. Next is the G_2, or gap 2, phase where the cell continues to grow before division; this takes between 3.5 and 4 hours for an iPSC. The final stage, M, is mitosis. Here the cell splits to produce a daughter cell which contains the exact copy of its DNA produced in the S phase; this takes around 1.5 hours for an iPSC [9].

Death – Transduction causes a high death rate of cells, which become too 'stressed out'. Previous experiments done at the Centre for Life show that around 50% of the cells die on the day of transduction [9]. In addition, cells die throughout the experiment for various reasons. The process of a cell dying is referred to as **apoptosis** (programmed cell death). When a cell enters this irreversible state, it takes about a day for the cell to completely die. Experiments at the Centre for Life show that between 10% and 14% of cells are under-going apoptosis at any time [9].

Cells can also die from **senescence** (old age). However, as stem cells give rise to daughter cells they are maintained in a nonageing state [19], so senescence is less important when modelling iPSCs.

When a cell dies, it detaches from the colony which is fixed to the well by an underlying matrix, and floats to the top of the medium; on photographs they are identified by white spots, as seen in Figure 1.2(B). The dead cells are then manually removed from the medium. The vacancy the detached cell leaves in the colony of iPSCs is always seen to be filled. This is likely to be a consequence of adjacent stem cells wanting to be in complete contact with each other in order to aid their proliferation.

Cell Migration – Experiments assessing the motility of hESCs have shown that their migration is a fundamental feature of their behaviour [10]. We will assume that iPSCs have similar mobility abilities due to their vast similarities to hESCs. Individual cells are not stationary but can move, apparently involved in a **random walk**. Experiments performed with two cells show that migration becomes directional if two cells are within a threshold distance [10]. Thus there is a **mutual attraction** between stem cells within a certain distance of each other. This is thought to be caused by the release of chemical signals by each cell which indicate its existence to neighbouring cells.

The existence of mutual attraction is consistent with the unique property of stem cells in their preference to grow in full adhesion. Stem cells release growth factors as they proliferate, which provide neighbour cells with essential chemicals to also proliferate [9]. A single cell with no neighbours will often die before it has even undergone a single cell division.

Experiments on hESC motility [10] have been described in sufficient detail to allow their modelling. The migration of single stem cells and stem cells separated by more than $150\,\mu$m mimics that of a particle undergoing random walk. However, if the separation distance is less than $70\,\mu$m, the movement of the cells is far more systematic, with an almost straight line path at separations less than $30\,\mu$m. This behaviour is modelled in Chapter 4.

2.2 Types of Mathematical Model

Currently there are a few different models which could capture some of the characteristics described, however none could encompass them all. I will now consider some of the different types of model available before deciding on the most appropriate way forward and beginning to develop a model of my own.

Continuous, Deterministic Models

Models of this kind include various types of dynamical systems: sets of first order ODEs whose solutions are continuous functions of time and deterministic. Examples of such basic one-dimensional dynamical systems would be the exponential-growth and the logistic model [18]. Both would be suitable to monitor the increasing number of cells as they divided and the additional parameter of the logistic model would also make it suitable to model the property of contact inhibition that fibroblasts exhibit; however we knows this is not necessary for stem cells.

Higher-dimensional dynamical systems were also considered in similar applications. In [17], a three-dimensional model of cell growth is discussed which has three states of cells: proliferating, quiescent and senescent. Proliferating cells are those which are dividing. Quiescent cells are considered as being 'frozen'; they are not dead, but are not dividing either. Quiescent is a reversible state from which cells re-enter the proliferating state at a constant rate. Senescent cells are those which are no longer able to divide, which we consider as dead; this is an irreversible state. In [17], a fourth dimension is added to model the stress of cells, which is considered to be the cause of transferring from the proliferating to quiescent state. However, this model is not suitable for stem cells, as they do not have a quiescent state and they do not become senescent [9].

A more relevant use of a high-dimensional dynamical system could be to model the movement between the phases of the cell cycle of iPSCs. A four-dimensional dynamical system is used in [3] to model the transition between the phases of the cell cycle of human tumour cells and their reactions to radiotherapy and chemotherapy, which can make the cell cycle reversible. Although this is a more applicable model it still would not allow us to capture any features of stem cells other than proliferation.

Discrete, Deterministic Models

A model currently used to describe the growth of stem cells is the Sherley model [5]:

$$N = N_0 \left[0.5 + \frac{1-(2\alpha)^{j+1}}{2(1-2\alpha)}\right] - B, \qquad j = \frac{t}{D_\mathrm{t}}, \tag{2.2.1}$$

with N the number of proliferating stem cells, N_0 the initial number of stem cells, B the cumulative number of dead cells (from mortality or migration), D_t the division time of the cells and α the mitotic fraction. The mitotic fraction is assumed to be constant and reflects the presence of nondividing cells. These can arise from cells that can produce both nondividing and dividing cells [5].

The Sherley model predicts the current number of proliferating cells in terms of j (the number of divisions), a dimensionless quantity. This discretisation of time seems sensible as the division time of an iPSC is 14–16 hours, so modelling time continuously is superfluous. Although this model encompasses death and division, it has no way of calculating B other than by counting, so fails to adequately model any characteristic other than proliferation.

Continuous, Stochastic Models

The deterministic nature of the models considered so far allows no randomness when evolving the cell states. This means the same initial condition would give the same solution each time, which is not realistic in this application. Any two experiments, with even just 5 stem cells, never produces the same result [9]. For this reason, a stochastic approach is far more suitable. However, there are disadvantages to making a stochastic model continuous when representing cell growth [8]. A discrete model is more realistic, especially for a small number of cells, as the number of cells will not be a smooth function of time but instead changes step-wise each time there is a division or death.

Discrete and Stochastic

When discretising time we can consider the model to evolve in time steps of arbitrary length. By making the time step small enough, the probability of more than one event happening at a given time becomes negligible. Models of this type can also have spatial dependence allowing us to include the migration of cells. We will therefore consider agent-based models, which include cellular automaton models evolving in discrete time and space. A cellular automaton model was applied to the migration of glioma cells in two dimensions, including cell-cell attraction through communication [2], behaviours similar to the mutual attraction properties of stem cells discussed in Section 2.1.

2.3 Agent-Based Models

Agent-based modelling (ABM) is a computer simulation with spatial (continuous or discrete) and temporal dependence which is becoming increasingly popular [23]. ABM provides insight into the characteristics of the system by modelling the behaviour of its individual components, called ‘agents’. The agents act according to a set of rules which aim to describe how they react to their environment and surrounding agents and which can be modified and developed as the model is refined. In our application, the agents are the iPSCs and the rules they adhere to would be the characteristic behaviours described in Section 2.1. For an agent-based model to be suitable, the agents must abide by certain rules. An agent must be autonomous (have the freedom to act independently), move to achieve its goals, be reactive (adapt its behaviour when its environment changes) and be situated at a definite location in its environment. Stem cells exhibit all these properties making an agent-based model ideal.

Cellular automaton models are a type of agent-based model with the restriction that agents move in discrete space; the cells migrate around a grid according to a set of rules. This will be the type of model I use in two dimensions with the square grid shown in Figure 2.3.

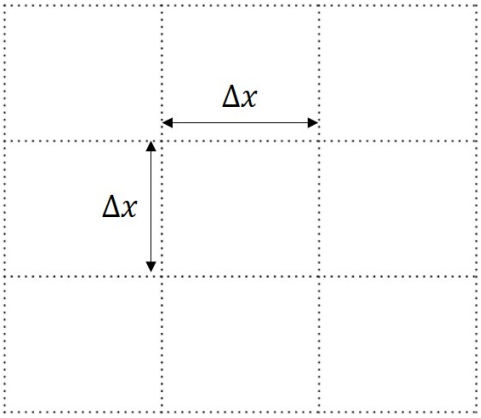


Figure 2.3: The discrete space grid on which the agents can move in a cellular automaton model.

The iPSCs in the experiments occupy a layer one-cell thick; so, their colonies are genuinely two-dimensional. Each grid point can be occupied by a maximum of one cell, so is either occupied or unoccupied at any time. There are many advantages to this method, including the fact that it incorporates spatial dependence. Firstly, the model can initially be made dimensionless, using a time step τ and grid size Δx. This means verifying the relevance of physical measurements by changing parameters in the model is much simpler, which will be extremely useful in our case when modelling experimental data. The other benefit is that we can add complexity stage by stage, as trying to include rules which describe all the characteristics given in Section 2.1 at once would be incredibly complex. Each adjustment to the model can then be tested iteratively to identify the cause of problems more easily, with further levels of complexity only being added to a coherent model. This way of modelling also allows us to determine which characteristics are important and which may be unnecessary.

Chapter 3

Modelling Stem Cell Migration

In this chapter I will explain the development of my cellular automaton model for the migration of stem cells which is built in Matlab (code included on CD). This includes the random walk behaviour of single cells and cells separated by more than 150 μm and the attractive behaviour of cells within a threshold distance. No other characteristics, including birth and death, will be included in the model yet. The number of cells is denoted N, which is constant as there are no demographic events, and the number of time steps that have evolved is denoted M. The model is formulated in terms of dimensionless time $\tau = 1$ and distance $\Delta x = 1$, but it will be tested against experimental results [10] with dimensional parameters specified in Chapter 4.

3.1 One-Dimensional Random Walk

I start by developing a model to simulate the simplest case of migration: a one-dimensional random walk which models the diffusion of cells in one dimension [4, 13]. The model is run with 1000 cells to determine the average behaviour. For simplicity, all the cells start at the origin, so $x_0 = 0$, and at each time step, τ, move one step, Δx, either left or right with equal probability. This means that the movement is isotropic.

It can be shown [4, 13] that after M time steps, the spatial position of any cell, x, can be written as

$$x = 2r - M, \tag{3.1.1}$$

where r is the number of steps taken to the right. This is because if M steps are taken in total, and r of those are to the right, then $M - r$ steps are taken to the left. Hence, the displacement from the origin can be written as $x = r - (M - r) = 2r - M$. Equation (3.1.1) explains that, if M is even only even positions can be occupied (as x can only take even values), and vice versa if M is odd. Figure 3.1 shows this by the alternating white and black bars, especially visible in the top two panels, which represent unoccupied and occupied gridpoints respectively. The histograms also give a qualitative check of the diffusion of cells by the increasing spread of occupied x values along with the increasing maximum distance specified for each simulation.

The distribution of the cells position can also be modelled by a Gaussian curve [4, 13]. The probability of a position, x, being occupied after M steps can be approximated with

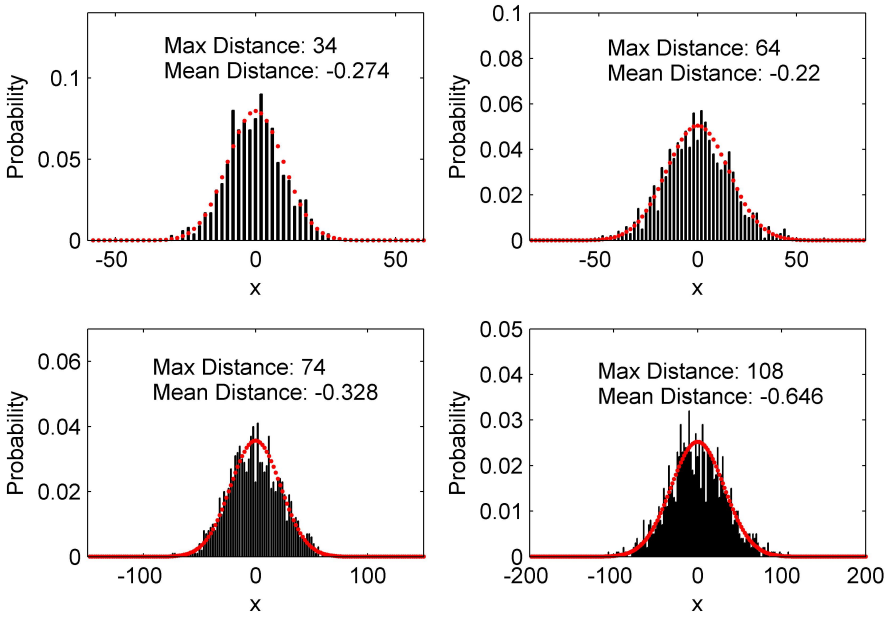


Figure 3.1: Displacement histograms of 1000 cells after 100 time steps (top left), 200 time steps (top right), 500 time steps (bottom left) and 1000 time steps (bottom right). All cells start at the origin and move isotropically. The Gaussian given in Equation (3.1.1) is overlaid in red and the mean and maximum distance of each simulation is given in the legends.

great accuracy by the Gaussian [13],

$$p(x, M) = \sqrt{\frac{2}{M\pi}} \exp\left(-x^2/2M\right). \tag{3.1.2}$$

The mean displacement of the cells at any time should be zero, due to their isotropic movement and initial position of $x_0 = 0$, and the variance of the Gaussian should increase with time, hence Equation (3.1.2) is a Gaussian with $\mu = 0$ and $\sigma_x^2 = M$. The Gaussian given in Equation (3.1.2) has also been multiplied by a factor of 2 to account for the fact only odd or even values of x can be occupied at any time. The red dotted lines in Figure 3.1 represent the values from Equation (3.1.2), evaluated at only odd or even values of x as appropriate, and appear to fit the histograms well. This shows that the model generates data according to Equation (3.1.2), so is working as desired. The mean displacement in each simulation is also given and is another confirmation of the assumption that $\mu = 0$.

If the model is working correctly the cells should diffuse according to [4, 13]

$$\sigma_x = \sqrt{M} \approx 2\sqrt{\kappa t} = 2\sqrt{\frac{(\Delta x)^2}{2\tau} t} = x\sqrt{2t}, \tag{3.1.3}$$

where κ is the diffusion coefficient (recall that $\tau = \Delta x = 1$). Figure 3.2 shows that the simulated points are fitted to a line of best fit with gradient 0.5, which confirms the square root relationship between σ_x and t, and hence confirms that the model works correctly for the diffusion of cells in one dimension. In order to quantify diffusion, we can calculate the

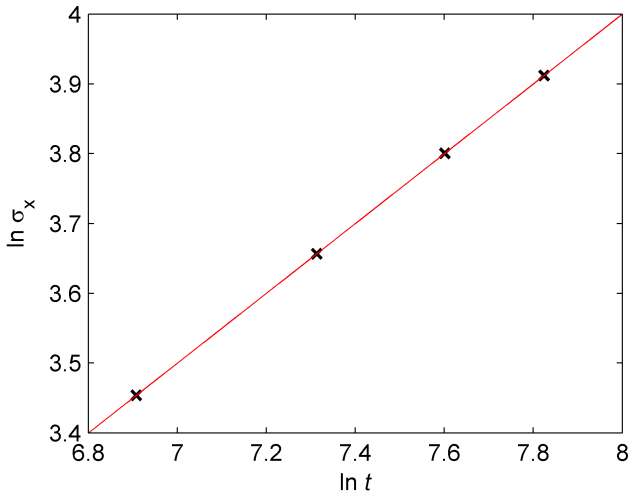


Figure 3.2: The power law dependence of σ_x on t. In black are the model values and in red is the line of best fit with equation $\ln \sigma_x = 0.5 \ln t - 3 \times 10^{-13}$ found using Matlab data fitting.

constant diffusion coefficient, κ, which tells us how efficiently the particles disperse from a high to low density. To find the value of κ, which would change if τ or Δx changed, we can equate the y-intercept of the line of best fit in Figure 3.2 to $\ln 2 + 0.5 \ln \kappa$ and rearrange for κ. With $\tau = \Delta x = 1$, the simulation gives $\kappa = 0.25$.

3.2 Two-Dimensional Random Walk

The model can now be developed to model two-dimensional diffusion, which is more realistic in application to the *in vitro* growth of stem cells in a well. We recall that the stem cells are attached to the well by an underlying matrix, and colonies grow only one cell thick. At each time step movement is now possible in four directions; up, down, left and right. Again, the probability of moving in each direction is the same so the movement is still random and isotropic, as Figure 3.3 shows. The possibility of moving in eight directions, so as to include diagonal movements, was considered but decided to be an unnecessary complication as a horizontal step followed by a vertical step gives the equivalent displacement. Restricting movement to be horizontal and vertical only also adds the simplification that each step has the same length, Δx. In developing the model both Δx and τ are kept dimensionless and equal to unity as before. The simulation is again run for 1000 cells which all have initial position at the origin, $\vec{x}_0 = (0, 0)$.

In order to test the model of two-dimensional diffusion we must consider the Rice Distribution. This gives the probability of the displacement of cells in two dimensions, Δr, from their initial position, ν, with the 'spread' quantified by σ. Figure 3.4 illustrates these parameters and the diffusion we would expect to see with the intensity of blue representing the density of cells. The probability density function is given by

$$\varphi(\Delta r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2 + \nu^2}{2\sigma^2}\right) I_0\left(\frac{r\nu}{\sigma^2}\right) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \tag{3.2.1}$$

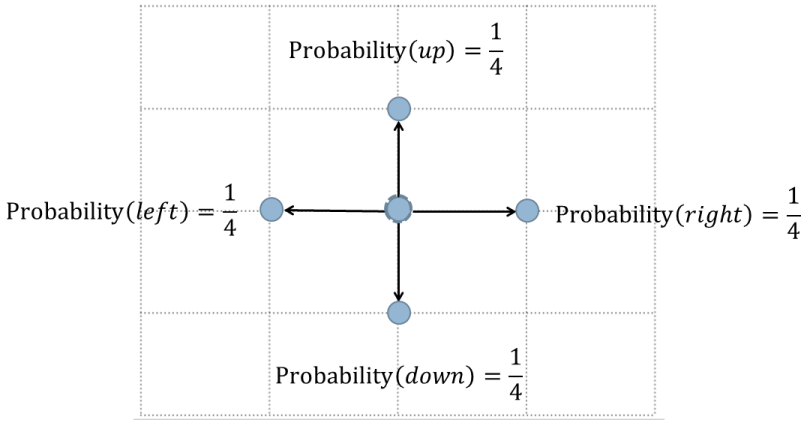


Figure 3.3: The possible movements of the isotropic two-dimensional random walk on the discrete grid of size Δx.

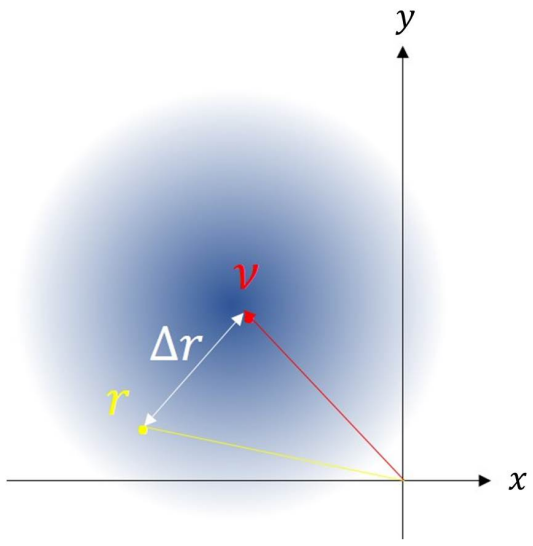
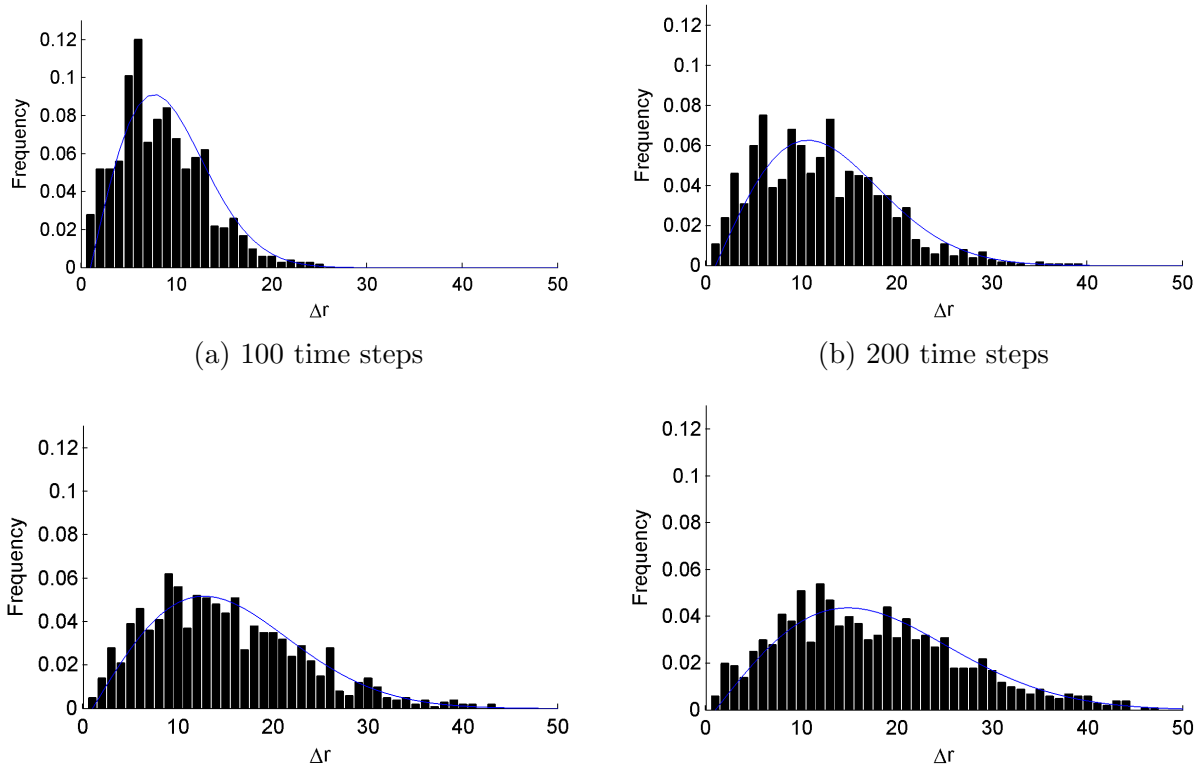


Figure 3.4: The parameters of the Rice distribution and the diffusion of cells around their initial position, ν. The intensity of blue represents the expected density of cells.



(a) 100 time steps

(b) 200 time steps

(c) 300 time steps

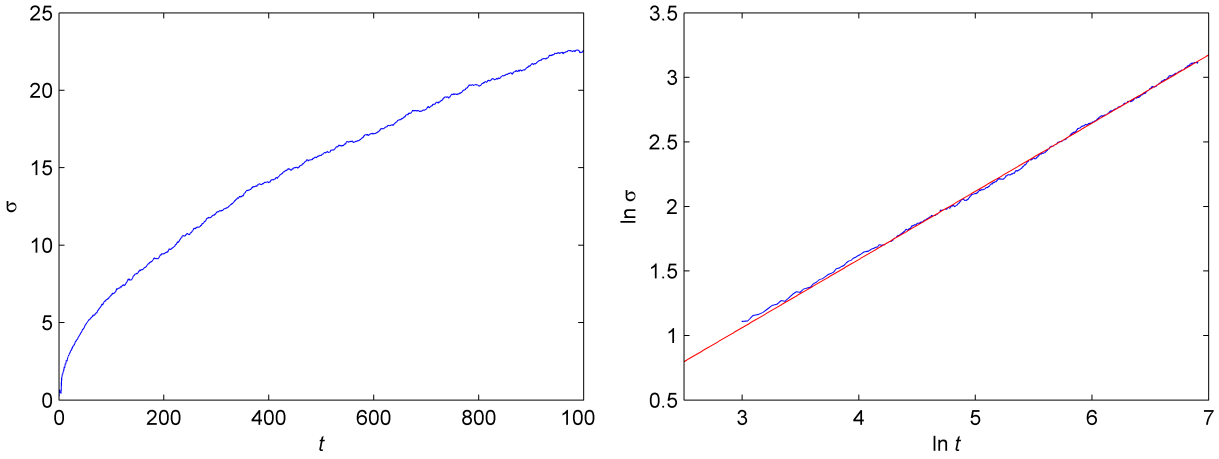
(d) 400 time steps

Figure 3.5: The displacement histograms (black) and the fitted Rice distribution (blue), (a) after 100 time steps, (b) after 200 time steps, (c) after 300 time steps and (d) after 400 time steps.

with the simplification due to $\nu = 0$, as all cells start at the origin, and $I_0(0) = 1$ for the modified Bessel function of the first kind. The value of σ, however, is unknown and should increase with time. To estimate it, we can fit the Rice density function, Equation (3.2.1), to the displacement histograms of the cells at each time step by minimising the sum of residuals squared between the two. Figure 3.5 illustrates this after 100, 200, 300 and 400 time steps. It is evident that, as time evolves, the cells spread further from their initial position $\nu = 0$.

For a quantitative check of diffusion, we can plot the estimated values of σ against t as shown in Figure 3.6(a). The diffusion of cells should again follow the diffusion law $\sigma = 2\sqrt{\kappa t}$. Figure 3.6(b) shows that, on a log-log scale, the line of best fit has a gradient of 0.52, which gives good agreement with the square root relationship we expect. When calculating the equation of the line of best fit, the first 20 estimated values of σ were excluded to increase accuracy, as the relationship given in Equation (3.2.1) is asymptotic. This confirmation means that the model adequately describes diffusion in two dimensions and confirms that excluding diagonal movement along the grid is not damaging.

Again, the diffusion coefficient can be estimated by equating the y-intercept of the line of best fit in Figure 3.6(b) with $\ln 2 + 0.5 \ln \kappa$. For the simulation shown in Figure 3.6 we find $\kappa \approx 0.087$. As expected, this is smaller than the diffusion coefficient in one dimension ($\kappa = 0.25$) as allowing movement in four directions means on average, $\Delta r < \Delta x$ after the same number of steps.



(a) The estimated values of σ versus t

(b) σ versus t on log-log scale

Figure 3.6: The diffusive spread illustrated by the dependence of σ on t. The left-hand panel shows the estimated values of σ as time evolves. The right-hand panel shows a log-log plot of this dependence to identify the power law and has line of best fit $\ln \sigma_x = 0.52 \ln t - 0.53$ found using Matlab data fitting tools.

At this point, extending the model to three dimensions was considered but deemed unnecessary, as stem cell colonies *in vitro* grow one cell thick keeping constant contact with the underlying surface of the well [9].

3.3 Two-Dimensional Biased Random Walk

The next element of migration to include in the model is the mutual attraction between cells which are close enough to interact (Section 2.1). It is not likely that their migration suddenly becomes purely attractive at a certain distance with no isotropic element. It is more plausible that the relative importance of biased and isotropic elements in the cells movement varies with attraction becoming dominant when the distance between cells reduces beyond some threshold. However, for simplicity, we will consider the attraction of cells to be binary. I will refer to the distance between two cells as D and the threshold distance where cells switch between isotropic and attractive migration as T (a parameter of the model). The angle between the horizontal and shortest distance between Cell 1 and Cell 2 is denoted θ. The initial positions of the two cells are arbitrary. Figure 3.7 illustrates this setup. If $D > T$, the two cells are completely unaware of each other and they perform independent two-dimensional random walks as in Chapter 3.2. However, their movement becomes biased if $D \leq T$ by mutual attraction.

When $D \leq T$, the probabilities of moving in each direction depend on θ. In the case shown in Figure 3.7, the probability of Cell 1 moving right is proportional to $\cos\theta$ and moving up is proportional to $\sin\theta$. The probability it moves down or left is 0 as the model only allows attractive motion. If we denote, for example, the probability of Cell 1 moving up as $\mathrm{P}_1(\mathrm{up})$, then the probabilities of motion of Cell 2 are given by $\mathrm{P}_1(\mathrm{up}) = \mathrm{P}_2(\mathrm{down})$ and vice versa and $\mathrm{P}_1(\mathrm{left}) = \mathrm{P}_2(\mathrm{right})$ and vice versa. Although the probabilities of movement of the two cells

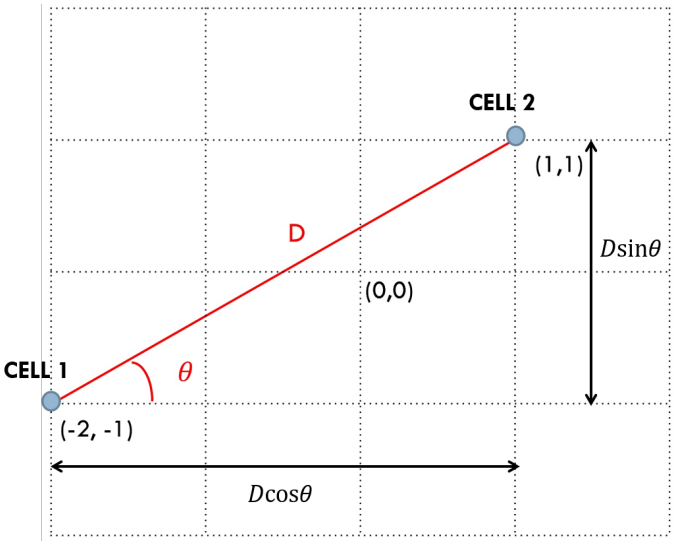


Figure 3.7: The parameters in the biased random walk model. The shortest distance between the two cells is D and the angle between the straight line joining cells 1 and 2 makes and the horizontal is θ.

are related, the direction of motion of each cell is chosen randomly (with the probabilities specified) and independently at each step. After each step, the model recalculates D and θ, and hence the probabilities of motion for each cell, which always sum to unity. The model continues to run until either the cells occupy adjacent grid points or a specified length of time has evolved.

Figure 3.8 shows two examples of the model in action. In both cases the cells have initial positions of $(x, y) = (-5, -5)$ and $(5, 5)$ and T is made large enough such that $D = 10\sqrt{2} < T$ intially. The migration of cells follows a biased random walk where the separation between the cells reduces at each step. The simulation stops when the two cells are next to each other.

If the initial positions of the two cells are such that $D > T$, the cells will either migrate further apart from each other and never come within T, or, will randomly step within the threshold distance and then perform the biased random walk to become adhesive. This behaviour is shown in Figure 3.9 with two cells which have initial positions $(x, y) = (-10, -10)$ and $(10, 10)$ and threshold distance $T = 25$, so initially $D = 20\sqrt{2} > T$ but the values are similar. We see that initially the cells move randomly around their initial positions but when D becomes smaller than T they migrate towards each other as in a biased random walk. With confidence that the complete migration model is working as we expect, we can begin to test it with specific parameter values against an *in vitro* experiment.

Currently the model only works with two cells since the experiment we discuss in Chapter 4 focuses on cell pairs. Further development would generalise this to model the migration of N cells at various separation distances. This would involve calculating the distance between all cells and whether each cell had neighbours within the threshold distance or not. Cells which had no neighbouring cells within the threshold distance would perform the isotropic random walk. However, for cells with neighbours within the threshold distance, T, the overall probability from all surrounding cells within a circle of radius T would have to be calculated for moving in each direction and then normalised to ensure they sum to unity for each cell.

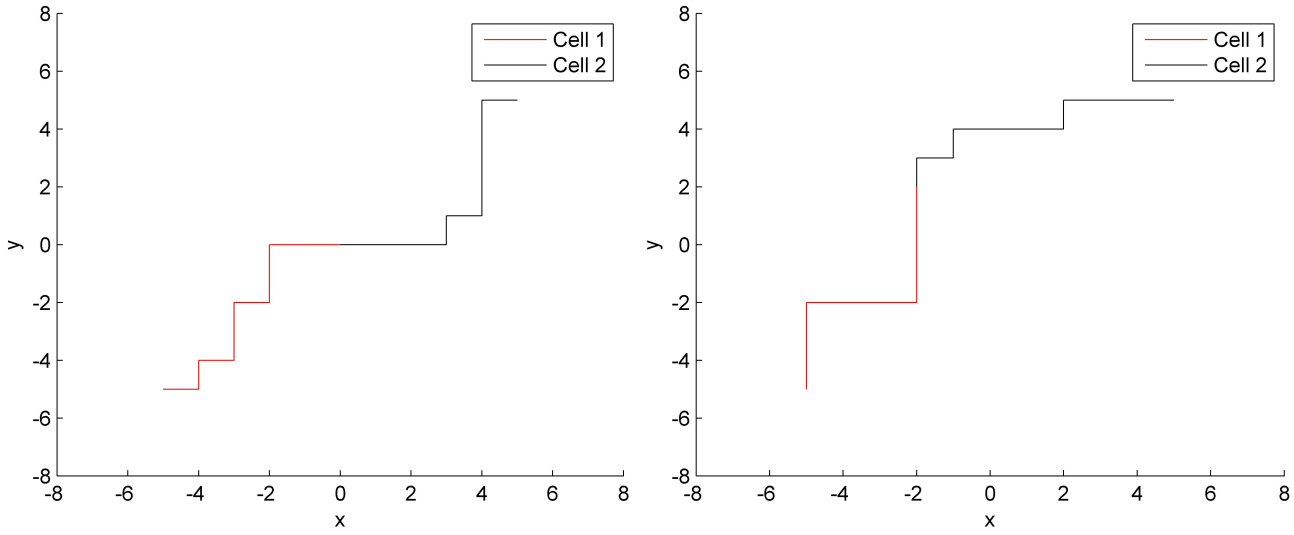


Figure 3.8: Biased, converging random walks of two cells that start at $(x, y) = (-5, -5)$ and $(5, 5)$ within the attraction distance, $D = 10 < T$, illustrated with their trajectories in the (x, y)-plane.

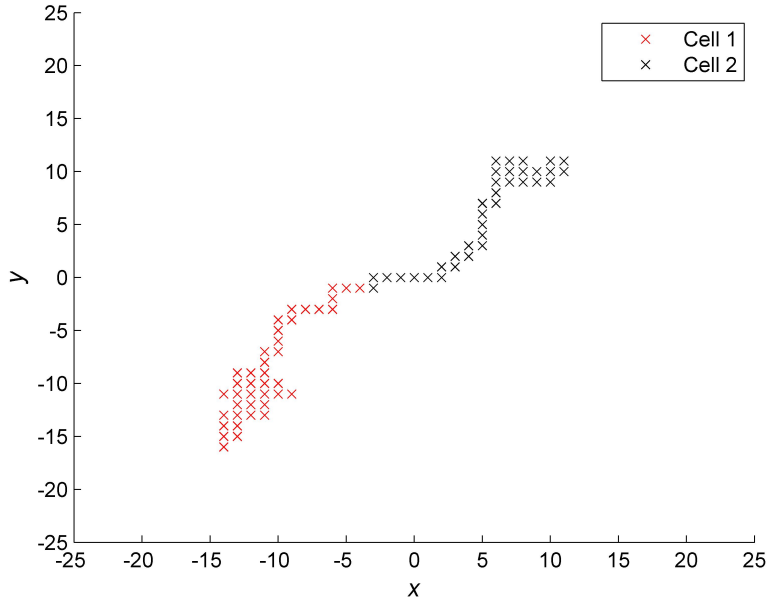


Figure 3.9: The trajectories of two cells that start at $(x, y) = (-10, -10)$ and $(10, 10)$ outside the threshold distance, with $D = 20\sqrt{2}$ and $T = 25$.

Chapter 4

The Experimental Migration of Stem Cells

Experiments and statistical analysis of the migration of single hESCs and collective hESC movement are described by Li et al. [10]. The results show these to be fundamental properties of hESCs, however, a mathematical model for the migration of stem cells was not attempted by Li et al. [10]. In this chapter I will apply our agent-based model for the migration of stem cells to their results shown in Figure 4.1.

Two experiments are discussed in [10]. One measures the migration of a single cell in the absence of neighbouring cells and the second measures the attraction of cells where neighbouring cells are present. In experiments with more than one cell, the experiment was stopped at 10 hours if the cells had not already merged. My model is built to reflect this.

4.1 Parameters

In this section I will discuss how I choose the model parameters, which now need to be dimensional, and explain some of the quantities used to quantify the migration of the hESCs.

Li et al. [10] introduce two so far undefined quantities: directionality and travel radius. The directionality of cells is defined as the ratio of the shortest distance between the start and end positions, L, to the total distance traversed, L_{trav}. In my model, L can be calculated at any time by

$$L = \sqrt{[x(\tau) - x_0]^2 + [y(\tau) - y_0]^2}, \tag{4.1.1}$$

where (x_0, y_0) is the initial position of the cell. Since each cell moves either along x, or along y, by the same distance Δx at each time step,

$$L_{\text{trav}} = M\Delta x, \tag{4.1.2}$$

where M is the number of steps. Hence, the directionality is given by

$$\frac{L}{L_{\text{trav}}} = \frac{\sqrt{[x(\tau) - x_0]^2 + [y(\tau) - y_0]^2}}{M\Delta x}. \tag{4.1.3}$$

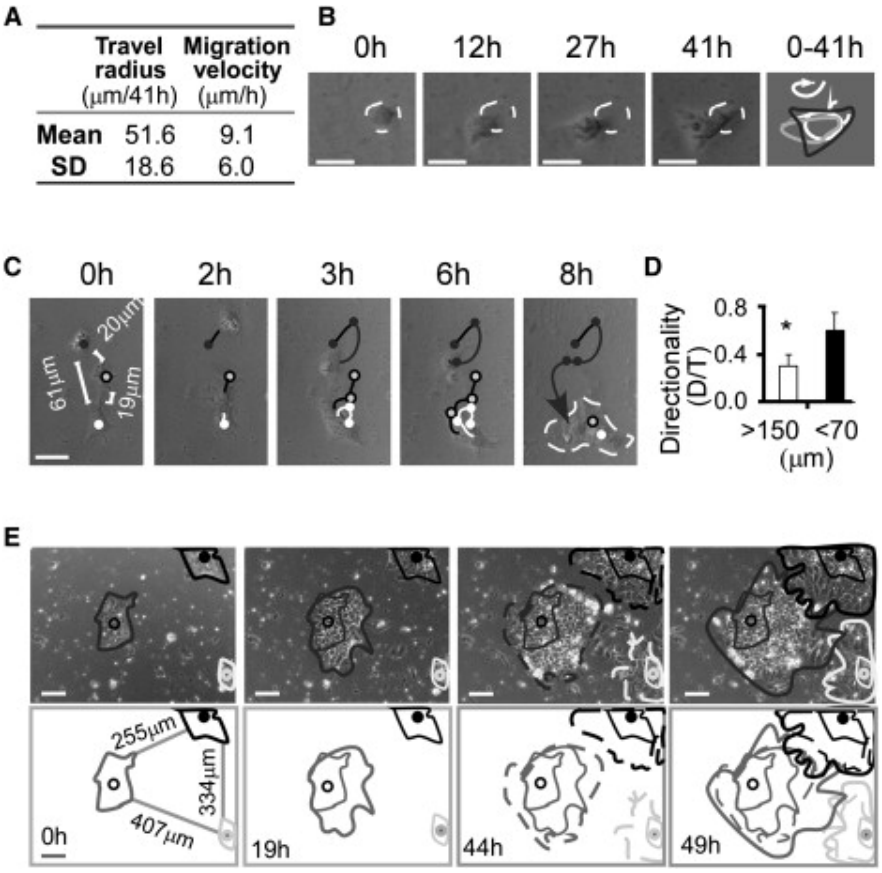


Figure 4.1: Experiments on the individual and collective migration of hESCs [10]. (A) Qualitative data on the migration of a single hESC with no neighbouring cells. Migration velocity (μm/h) quantifies the motility of the cells and the travel radius is the furthest point reached by a cell in 41 hours, referred to as $L_{\max}$. Data is from 87 cells in four independent experiments. (B) Images of the characteristic movement of a single hESC with no neighbouring cells. In the right-hand panel lines represent cell trajectories with the arrows showing the direction of movement. The bars length are $20\,\mu$m. (C) Images of the characteristic migration of single hESCs with a neighbouring cell less than $70\,\mu$m away, with the bar length $50\,\mu$m. (D) Qualitative data on the directionality of a single hESC with separation distances from a neighbour of more than $150\,\mu$m and less than $70\,\mu$m. The directionality is the ratio of L and L_{trav} as explained in Section 4.1. Data came from four independent experiments and the bars represent the standard deviation based on more than 150 cells per group. (E) Images showing the asymmetric movement and colony expansion of hESCs. Bars are $100\,\mu$m.

The travel radius of a cell is defined as the maximum linear distance from its initial position reached within 41 hours. This farthest point is not necessarily its root mean square displacement at 41 hours. I will call this parameter $L_{\max}$. This is not the most mathematically useful parameter and my model had to be adjusted to capture this information. If the cells migration is based on a tailed distribution (for example, the Rice or Gaussian distribution) then the maximum distance traveled could be uncharacteristically large with a small probability. Recording the mean distance traveled at each time step would quantify the data more appropriately. However, the analysis in Figure 4.1 uses the travel radius to quantify cell migration so it is necessary to develop my model so it can be tested against these results.

Experiments which were carried out for two cells separated by a minimum distance of $150\,\mu$m yielded results analogous to a particle undertaking a random walk. However, when the separation distance was less than $70\,\mu$m; movement became directional. For this reason, the threshold distance in my model, T, will be taken as $70\,\mu$m.

The final step before running the simulation is to establish the magnitudes of τ and Δx. Assuming, as before, that one stem cell occupies a grid point and that adjacent cells are in adhesion, we can use the approximate diameter of a stem cell as Δx, so $\Delta x = 10\ \mu$m. Then Equation (4.1.2) simplifies to $L_{\text{trav}} = 10M\,\mu$m. To calculate the corresponding value of τ we can use the mean migration velocity of a single cell given in Figure 4.1, $v = 9.1\,\mu$m/h, with

$$\Delta x = v\tau, \tag{4.1.4}$$

to find $\tau \approx 1.1$ hours.

4.2 Diffusion of Stem Cells

This section includes the migration of a single cell or two cells separated by a distance greater than $150\,\mu$m. The experiment finds the travel radius of single hESCs in the absence of neighbours as $L_{\max} = 51.6 \pm 18.6\,\mu$m, with a maximum of $88\,\mu$m, and the migration velocity as $v = 9.1 \pm 6.0\,\mu$m/h (Figure 4.1(A)). The change in position of a single cell in a well with snapshots every few hours is shown in Figure 4.1(B), with the white lines representing $20\,\mu$m for scale.

The model described in Section 3.2, with the parameters given in Section 4.1, for 1000 cells produces the average $L_{\max}$ and L in Figure 4.2. We would expect $\langle L_{\max} \rangle$ to approximately go through the point $(t, L_{\max}) = (41, 51.6)$ if the model replicates the experiments. From Figure 4.2 we can see that at 40.7 hours the model predicts the travel radius to be $67.98\,\mu$m, which is within 1 standard deviation of the mean travel radius shown in Figure 4.1(A), so gives good agreement. Some of the ambiguity in this result may be down to the precision of the experiment, especially when dealing with such small quantities, but also due to the approximation of taking $\tau = 1.1$ hours. The value of τ was calculated using Equation (4.1.4) with the mean migration velocity $9.1\,\mu$m/h. However, the migration velocity has a standard deviation of $6.0\,\mu$m/h. This is relatively very large (about 65% of the average), so τ has a standard deviation of $\dfrac{\Delta x}{v^2}\sigma_v \approx 45$ minutes. The standard deviation of the 1000 simulated travel radius values was also calculated at 40.7 hours and found to be $21.6\,\mu$m (3d.p). This also gives very good agreement with the experimental standard deviation of the travel radius of $18.6\,\mu$m after 41 hours.

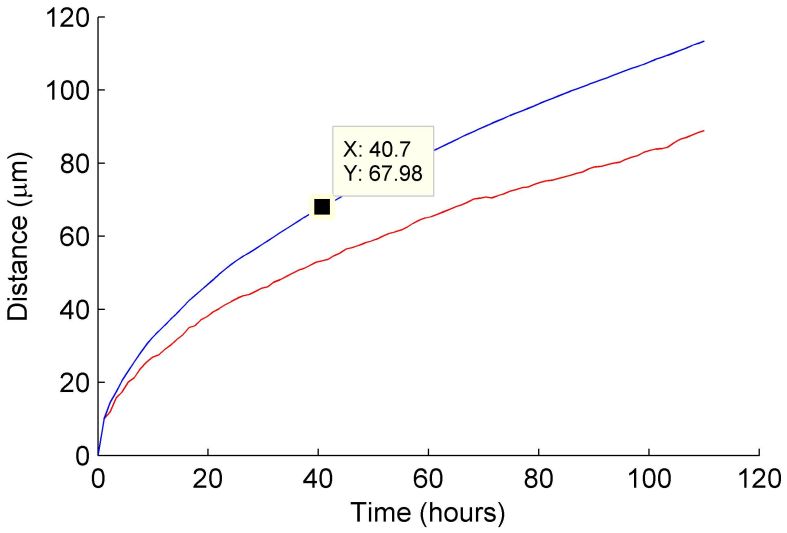


Figure 4.2: The average travel radius, $\langle L_{\max} \rangle$, in blue and average shortest distance from the cells' initial position, $\langle L \rangle$, in red in μm. Values are averaged over 1000 cells.

We can also quantify the directionality of the cells in this model, which represents the migration of single cells and also of cells separated by a distance larger than 150 μm. This is done by running the simulation for 1000 cells and determining the average directionality as $\langle L \rangle / \langle L_{\text{trav}} \rangle$. The results are shown in Figure 4.3. At 41 hours the directionality is approximately 0.15 and then appears to settle on approximately 0.1. If we now consider Figure 4.1(D), we can see that the mean directionality for cells at a distance greater than 150 μm is about 0.15, so again my model reflects the experimental values.

The agreement between my model and the experimental data for the migration of a single stem cell and cells separated by more than 150 μm strongly suggests that they follow an unbiased random walk with the parameters established by my model.

4.3 Mutual Attraction of Stem Cells

When the separation of cells is less than 70 μm the hESCs movement becomes directional, and with separations of less than 30 μm they move along a nearly straight line. Figure 4.1(C) shows this with photographs taken from a well with three cells all within 70 μm of each other. The directionality of cell migration in the model of Section 3.3, with $\Delta x = 10\,\mu$m and $\tau = 1.1$, is shown in Figure 4.4. The duration of the simulation is approximately 10 hours as the experiment continued until 10 hours had elapsed or the two cells had joined. We can see that the directionality settles at 0.7 quite quickly. Comparing this with Figure 4.1(D) for a separation less than 70 μm, we see that 0.7 is not the average value of the experimentally measured directionality, but roughly the maximum. This, however, could be expected as the model of Section 3.3 has only attractive behaviour with no isotropic element in the random walk. In this sense, the random walk model of Section 3.3 may be too strongly biased by the cell attraction. Hence we would expect the *maximum* experimentally measured directionality to be around the directionality in Figure 4.4 (≈ 0.7), as cell attraction cannot be any stronger

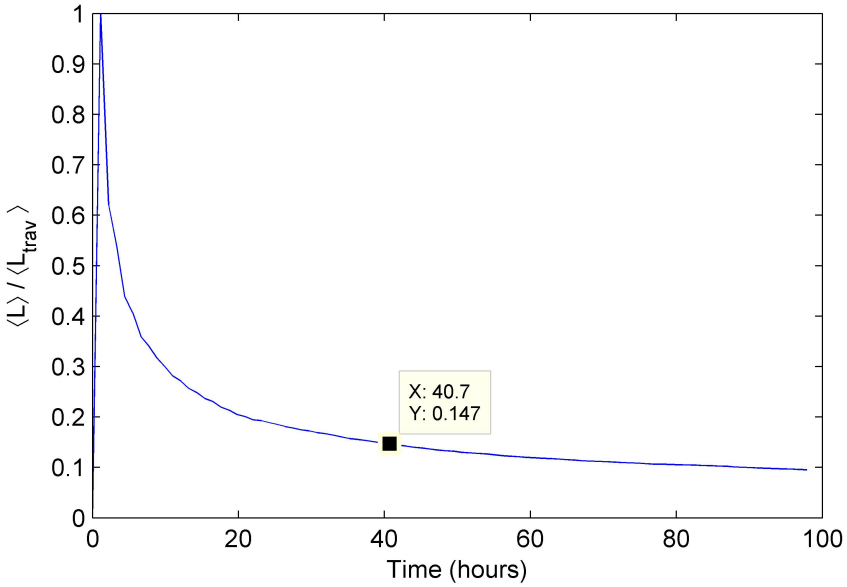


Figure 4.3: The directionality of cells (averaged over 1000 cells) in the isotropic random walk of single cells and cells separated by more than 150 μm.

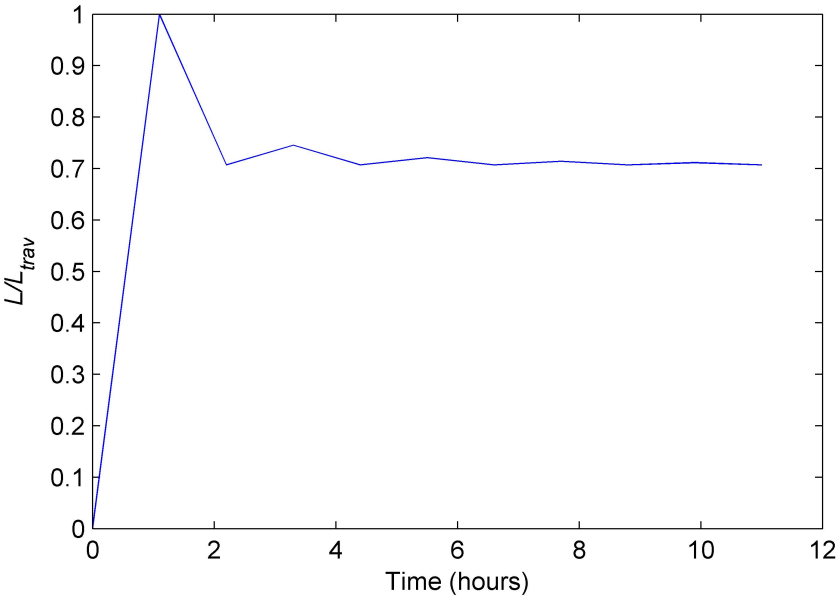


Figure 4.4: The directionality of the attractive biased random walk for the migration of cells separated by less than 70 μm.

than in our model! This completely mutual attractive migration is more visible when the cell separation is less than $30\,\mu$m [10] , so the attraction model would be more suitable for cells in this case.

The model gives good agreement for the migration of single cells, cells separated by distances greater than $150\,\mu$m and cells separated by less than $30\,\mu$m. However, the behaviour of cells with a separation distance between $70\,\mu$m and $30\,\mu$m is not modelled with great accuracy. This is probably because my model for attractive motion does not have an isotropic element in the random walk which is distance-dependent, as discussed in Section 3.3. If developing the model further this would be my next adaptation. Then, when modelling the migration of cells with separation distances between $70\,\mu$m and $30\,\mu$m, the isotropic element would decrease in intensity, vanishing when $D = 30\,\mu$m. The difficulty would be in quantifying the ratio of unbiased motion to attraction between these distances and deciding how the attraction intensity increases as separation decreases. New experiments would be required to develop the model in this direction as it would involve further parameters that need to be estimated experimentally. However, we decided to proceed to models more directly related to the experiments conducted by our colleagues at the Institute of Genetic Medicine who are more interested in the evolution of colonies of stem cells.

Chapter 5

Growth of Stem Cell Colonies

In this chapter I will discuss the development and testing of another agent-based model which was implemented in Matlab (the code for this is given on the CD). The focus will now be on modelling a colony of stem cells and how it grows. I will expand the capabilities of the previous model to include more of the stem cell features introduced in Section 2.1. This will include the division of cells and the cell development cycle for a more insightful model of cell proliferation. Finally, cell differentiation will be included. The migration model discussed in Chapter 3 will be refined and adapted to deal with a slightly different scenario.

The model still evolves in discrete time and space on a square grid in steps of duration τ and length Δx. The model now includes an 'information' matrix which records for each cell its number, two-dimensional position, age, division time (how long it takes to complete the cell cycle) and how long ago it last divided. This matrix is updated each time step, with all previous information still stored, to provide a clear picture of the how the colony evolves.

5.1 Initial conditions

An isolated stem cell rarely survives more than 15 hours (approximately one division time) due to its lack of contact with neighbouring cells and the consequent lack of chemical growth factors [9]. Therefore, each simulation is now started with a small colony of 5 cells at adjacent grid points as shown in Figure 5.1, i.e. $N_0 = 5$. This corresponds to the experiments, where iPSCs are passaged as small colonies of approximately this size [9]. Each cell is assumed to fill a grid square, i.e. to be $\Delta x \times \Delta x$ in size, with $\Delta x = 1$ in terms of dimensionless length.

At the start of a simulation each cell is given the same division time, $D_\mathrm{t} = 1500\tau$. Experiments suggest this is 14–16 hours, so stochastic, but for simplicity we will assume it is deterministic. Each cell in the initial colony is then given a random age, C, such that $1 \leq C \leq D_\mathrm{t}$ to avoid complete synchronisation in the division of the colony. This random starting age for each cell in the initial colony is also considered as the time since the cell last divided, denoted T_d.

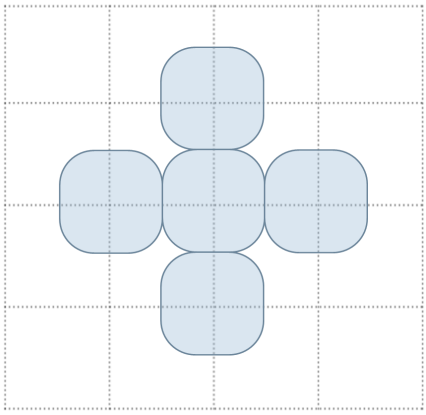


Figure 5.1: The initial colony of stem cells, shown by the blue regions, on the discrete grid.

5.2 Cell Division

The model works by increasing the age, C, and time since last division, T_d, of each cell by τ at each step. Then, when in the course of the simulation, $T_\mathrm{d} = D_\mathrm{t}$ for a cell, it splits. The daughter cell is given age zero but a value of T_d is chosen randomly between 0 and $0.07D_\mathrm{t}$ to avoid all the cells descended from the same initial parent dividing at exactly the same time. Allocating T_d at random has the same effect as making the division time scale, D_t, random. The daughter cell then occupies the grid position of the cell which has just divided, and the parent cell moves to an adjacent unoccupied grid point to avoid the layer of cells becoming more than one cell thick. This is simple when the parent cell is on the boundary of the colony: we determine which of its adjacent grid points are unoccupied and move the parent cell to one of these, with equal probabilities. However, as the colony grows, a parent cell may be in the middle of a colony with no adjacent grid points unoccupied. Thus, we must consider how the cells rearrange in the colony to make room for the new daughter cell.

5.2.1 How Does a Colony Accommodate Daughter Cells?

When we identified this problem we approached our collaborates at the Centre for Life for their opinion, as well as doing our own literature search. It quickly became apparent that this is a question biologists have never asked, or been asked, before and that there was no conclusive answer.

We thought that there might be three possibilities. Firstly, that central cells may not divide at all. This is a bold assumption to make without any experimental evidence, so this idea was not adopted here. However, a way of determining whether this is the case is discussed in Section 5.4. Secondly, the mechanism can be purely mechanical, so that when a new cell needs to create space it pushes the surrounding cells outwards by applying pressure as it grows. However, it is difficult to accept that a single cell would be able to create a pressure large enough to move a large number of other cells, especially given that the colony cells are attached to the well via a chemical matrix [9]. The alternative we came up with is that there can be a complex communication between the cells which alert them of the fact a cell is going to divide. Consequently, the surrounding cells create space by moving outwards.

The communication could be via the release of chemical signals, which we already know iPSCs are capable of doing from their interaction when migrating (Chapter 4). Thus, our mathematical model has helped to identify a new feature of stem cells behaviour which needs to be explored and which could provide fundamental understanding of their interactions.

As a result, a new experiment was launched in the Institute of Genetic Medicine to address questions arising from our model. In that experiment, about 50 small colonies of iPSCs are grown, and their microscopic photo images are taken every 15 seconds to allow trajectories of individual stem cells to be determined. Analysing this new experimental data is my priority for a summer research project in June–July 2015.

5.2.2 A Model of a Growing Colony

Assuming that stem cells are able to communicate in an appropriate manner, I use ideas from my migration model in Section 3.3 to build a mechanism which creates space for daughter cells produced inside the colony. We first calculate where the boundary of the colony is and which cells form it, by identifying the cell numbers occupying the (x, y) positions of the boundary. Then, the shortest distances between the splitting cell and all the boundary cells are calculated. If there are multiple boundary cells at this shortest distance, say N_b , one of them is chosen at random with equal probability $1/N_b$.

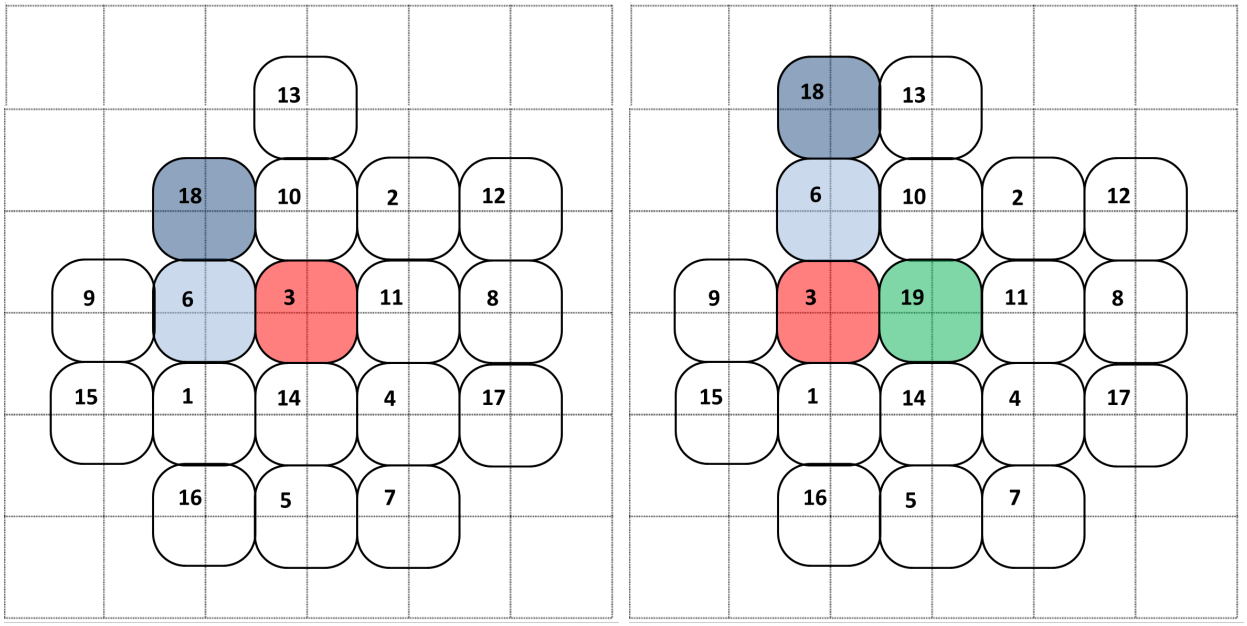
At this point the model of cell movement described in Section 3.3 is employed. Consider Cell 1 to be the splitting parent cell and Cell 2 to be the boundary cell which has been selected to move. The migration model now only includes attraction, (an infinite threshold distance), and the position of Cell 2, the boundary cell, is temporarily fixed. Then we can determine a shortest path in the (x, y)-plane between the dividing cell and the selected boundary cell.

The (x, y) coordinates of all the grid nodes on the shortest path are calculated so we can identify the cell numbers occupying these positions and move each cell one grid point towards Cell 2 along x or y as appropriate. This includes the parent cell, so the newly born daughter cell can occupy the grid point that the parent cell previously occupied. The boundary cell is then moved to one of its adjacent unoccupied grid points with equal probabilities.

This process is illustrated in Figure 5.2 with an example of an 18 cell colony where central Cell 3 (red) divides. The model creates space for the daughter Cell 19 (green) as follows:

- Cell 19 occupies the previous position of Cell 3.
- Cells 2, 9 and 18 are all $2\Delta x$ from Cell 3 (the shortest distance between Cell 3 and the boundary) so one is chosen with probability 1/3, in this case Cell 18 (dark blue).
- Cell 18 now occupies one of its adjacent unoccupied grid points.
- The chain of cells, including the parent one, between Cell 19 and the new space at the boundary all move in recurrence. The path (3,6,18) is shown by the coloured cells in the left-hand panel of Figure 5.2. Cell 3 now occupies the previous position of Cell 6, and Cell 6 now occupies the vacancy made by Cell 18.
- All other cells (white) stay in the same position.

Thus, the model selects a path of least resistance by moving only the cells along the path to the nearest boundary cell.



(a) before division (b) after division

Figure 5.2: The rearrangement of a stem cell colony when an interior cell divides. Cell 3 (red) splits producing daughter Cell 19 shown in green. The daughter cell occupies the grid point of its parent cell. Cell 18 was chosen as the nearest boundary cell to Cell 3, and moves to an unoccupied adjacent grid point. Cell 6 moves to the previous position of Cell 18 and Cell 3 moves to the previous position of Cell 6. This creates space for the new daughter cell.

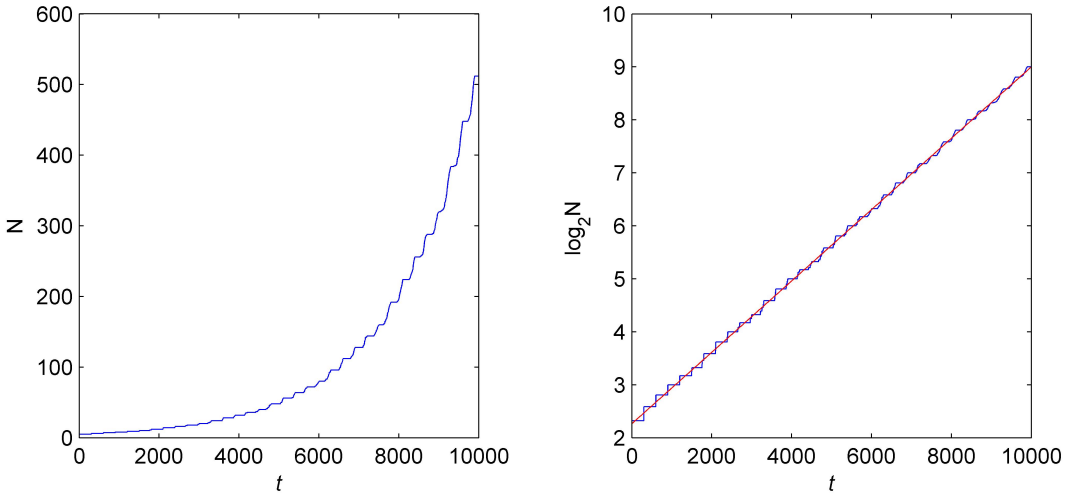


Figure 5.3: The left-hand panel shows the number of cells in the simulated colony as a function of time. The right-hand panel is the log-linear plot of this that demonstrates the power law dependence between the variables. The red line is a fitted line of best fit found with Matlab fitting tools and has equation $\log_2 N = 0.00067t + 2.3$. The division time of each cell in the simulation is $D_\mathrm{t} = 1500$, with $N_0 = 5$.

5.3 Quantifying the Growth of Colonies

It is important at this stage to check if the model describes the growth of colonies as we would expect. To do this, we need a way of quantifying the growth of the simulated colonies. This can be done by tracking the number of cells in the colony as a function of time and/or by monitoring the area of the colony as a function of time. Both will be discussed here taking $\Delta x = \tau = 1$, meaning that the division time of each cell, D_t, is 1500.

5.3.1 The Cell Number

The number of cells in a growing colony is a function of time, $N(t)$. When one cell divides, it results in two cells, the parent and daughter, so

$$N(t) = N_0 2^{t/D_\mathrm{t}} \tag{5.3.1}$$

should describe the growth. The simulation results are shown in Figure 5.3 together with the dependence of Equation (5.3.1). The line of best fit in the right-hand panel of Figure 5.3 has gradient $0.00067 \approx 1/D_\mathrm{t} = 1/1500$ and y-intercept $2.3 \approx \log_2 N_0 = \log_2 5$. This shows that the model is working as expected and gives us a way of comparing the simulation to experimental data.

Comparison to Experimental Data

In experiments presented by Ohmine et al. [14], images of a growing iPSC colony were taken every 12 hours and the number of cells in each colony counted. These images are shown in Figure 5.4. Assuming that the division time of iPSCs is 15 hours and noting that the initial

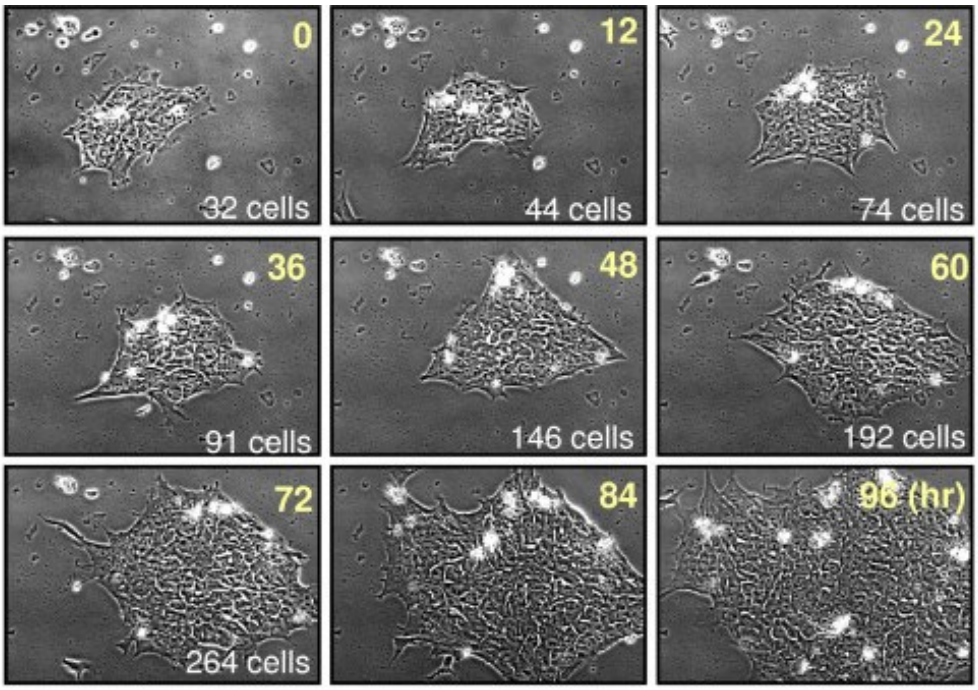


Figure 5.4: Long-term time-lapse images of an iPSC colony [14]. Time in hours is shown in the upper right corner, and the cell count is shown in the bottom right corner of each panel.

colony in the experiment had 32 cells, Equation (5.3.1) reduces to

$$N(t) = 32\left(2^{t/15\,\text{hours}}\right). \tag{5.3.2}$$

Figure 5.5 shows the experimental data on a log-linear scale to check if it follows the same exponential behaviour as Equation (5.3.2). Figure 5.5 confirms that it does, as the data fits well a straight line with positive gradient. The gradient of the line of best fit is $0.043\,\text{hour}^{-1}$ which suggests the division time of iPSCs is $1/0.043\,\text{hour}^{-1} \approx 22\,\text{hours}$. This difference from $D_\text{t} = 15\,\text{hours}$ adopted above could be for a number of reasons. Firstly, my model for cell growth does not include differentiation or death which would decrease the number of stem cells estimated by Equation (5.3.2) at each division. Figure 5.4 shows quite a few white dots in each image, which indicate dead cells, so this is certainly a contributing factor to the difference in the living cell number. Another reason which may contribute to the difference is that my model assumes that the growth throughout the colony is uniform. However, we are not sure if this is the case as the central cells may divide at a lower rate.

5.3.2 The Colony Area

The growth of the colony can also be quantified by its area, $A(t)$, also a function of time. The area of the colony at any time is given by

$$A(t) = N(t)A_\text{c} = N_0 2^{t/D_\text{t}} A_\text{c}, \tag{5.3.3}$$

using Equation (5.3.1), where A_c is the area occupied by a single stem cell. In terms of dimensionless variables, $A_\text{c} = 1$ and then Equation (5.3.3) simplifies to $A(t) = N(t)$.

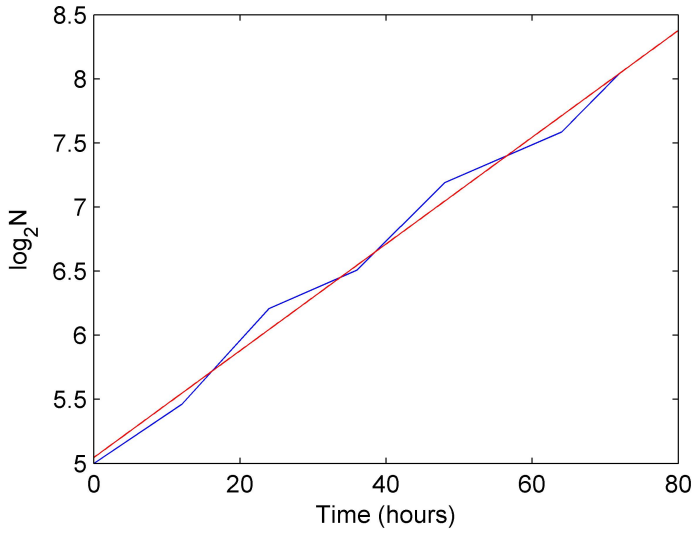


Figure 5.5: Testing the accuracy of the model of colony growth (in the absence of differentiation and death) against experimental data [14]. The line of best fit shown in red is $\log_2 N = 0.043t + 5$ hours.

Experiments show that colonies always form an approximate circle as they grow [9]. So, it is more appropriate to consider the colony radius, $r(t)$ and area $A(t) = \pi r(t)^2$. However, the colonies are not perfect circles so do not have an exact radius. Therefore, at each time step, the distance between the centre of the colony and every boundary cell is calculated and the average taken to be the average radius r_a of the colony. Figure 5.6 shows how this quantity evolves. The area of the simulated colony can then be calculated as

$$A(t) = \pi r_a^2. \tag{5.3.4}$$

Figure 5.7 shows how the area of the colony from Equation (5.3.4) evolves with time. The right-hand panel shows that the colony size also grows exponentially as the gradient of the log-linear best-fit is $0.0007 \approx 1/1500$. This is less accurate than the gradient of the line of best fit in Figure 5.3, but this could be expected due to the approximation of using r_a, and agreement to four decimal places is still satisfactory.

We can now check the accuracy of r_a as a measure of colony size. Using Equations (5.3.3) and (5.3.4), we get $N/r_a^2 \approx \pi$. So, we expect this ratio to tend to π asymptotically, because with a small number of cells the shape of the colony will be very irregular. Figure 5.8 shows that this is the case. This was important to test as it gives another validation of our assumptions of how a colony rearranges to accommodate daughter cells, as explained in Section 5.2.1. It appears that the model assumptions do not affect the behaviour of the growth of stem-cell colonies.

5.4 Stem Cell Lineages

We discussed with the biologists at the Centre for Life how to gain insight into the precise mechanisms of the rearrangement of a colony when cells are dividing. We also need to check if the growth of the colony is uniform. Do the central cells, which are smaller in appearance

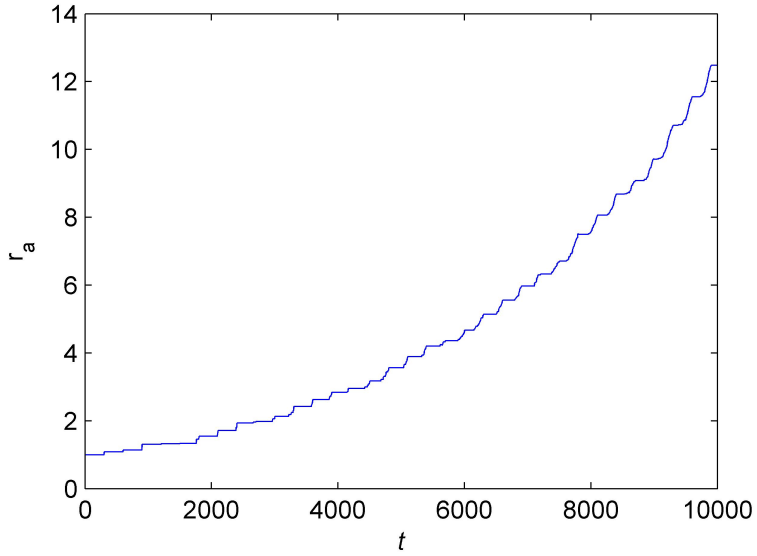


Figure 5.6: The average radius of the colony versus time.

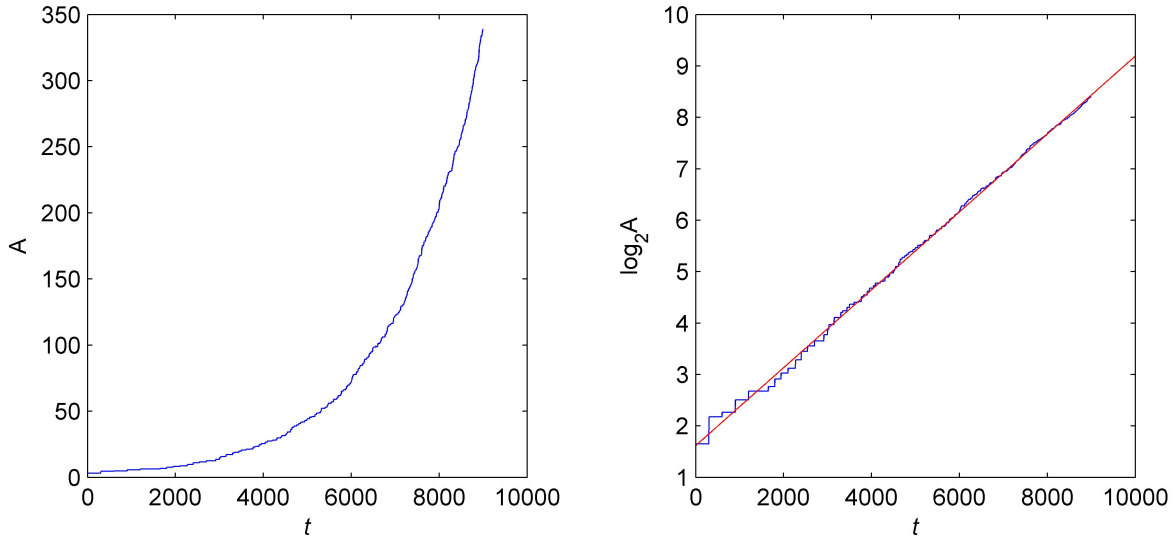


Figure 5.7: The area of the colony increases in time as in Equation (5.3.4): The right-hand panel shows the log-linear plot in blue, so the power-law variation can be confirmed. The red line is a line of the best fit found using Matlab data fitting, $\log_2 A = 0.0007t + 1.6$.

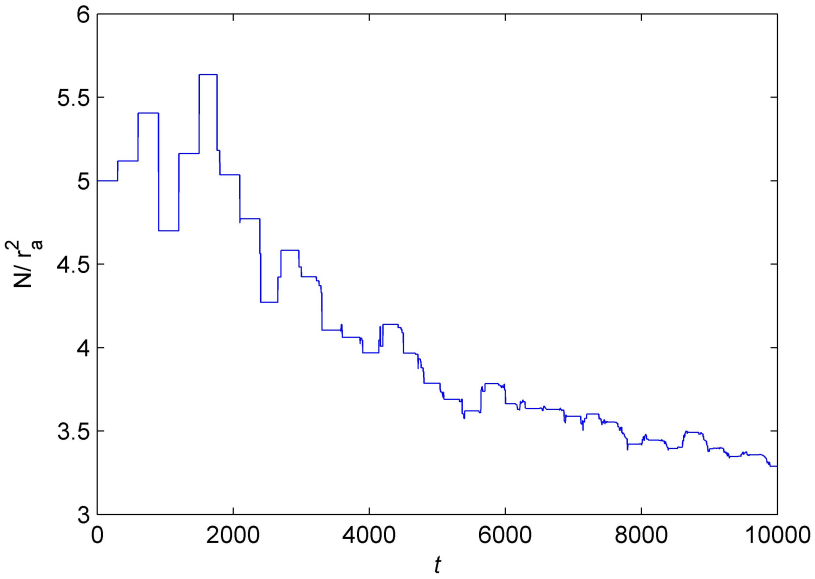


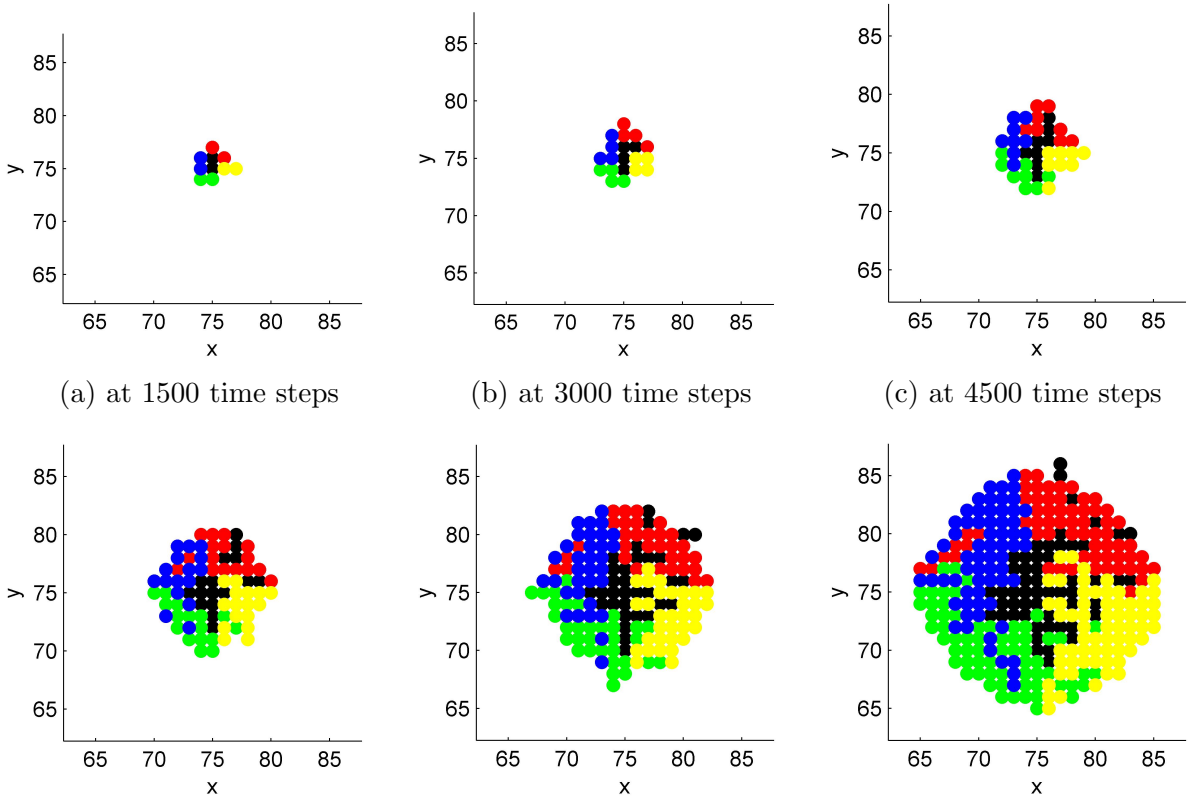
Figure 5.8: Testing the change in the colony shape as it grows. The ratio of the number of cells to the colony area plotted against time should tend to π asymptotically.

(perhaps due to being more 'squashed'), proliferate at the same rate as cells nearer the edge of a colony? A decision was made to carry out specialised experiments which start with a small initial colony where each cell is stained with a different colour [9]. This allows the migration of each cell to be tracked and the daughter cells it produced to be tracked as well. A clear picture of the growth of a colony can also be determined by comparing the relative sizes of differently coloured areas of the colony. Although time and resources have not allowed me to analyse the results of the experiment yet, I will explain the adaptation of the model that would allow the method of Section 5.2.1 to be applied to the experiment.

The 'information' matrix described in Section 5 now has another column which assigns a different number between 1 and N_0 to each cell in the initial colony. Any daughter cell produced by the initial cells or their offspring is then given the same number. By associating a different colour to each number, plots and videos can then be made which show the proliferation of the colony and its rearrangement by colour-coded cell paths.

Figure 5.9 shows an example starting with an initial colony of size $N_0 = 5$. Each initial cell is coloured separately: yellow, black, red, blue and green. The cell division parameter is $D_\mathrm{t} = 1500$, so each plot shows the colony at successive times that are multiples of D_t. Hence, the number of each colour cell doubles in each plot. This is a consequence of the weak stochasticity in the model, as the time taken for a cell to split is nearly constant. We recall that, in the model all cells have the same division time but a daughter cell is given a random value of T_d (between 0 and $0.07D_\mathrm{t}$). If the division time of each cell was made completely random we would not see the exact doubling of the colony size every 1500 time steps.

Figure 5.9 shows that initially the offspring tend to stay more or less in the same family area. Exceptions to this start to arise in panel (d) where some cells have become separated as the colony has had to rearrange itself to create space for the division of interior cells. Once a 'family' has become dispersed the cells continue to grow in different sections of the colony



(a) at 1500 time steps (b) at 3000 time steps (c) at 4500 time steps

(d) at 6000 time steps (e) at 7500 time steps (f) at 9000 time steps

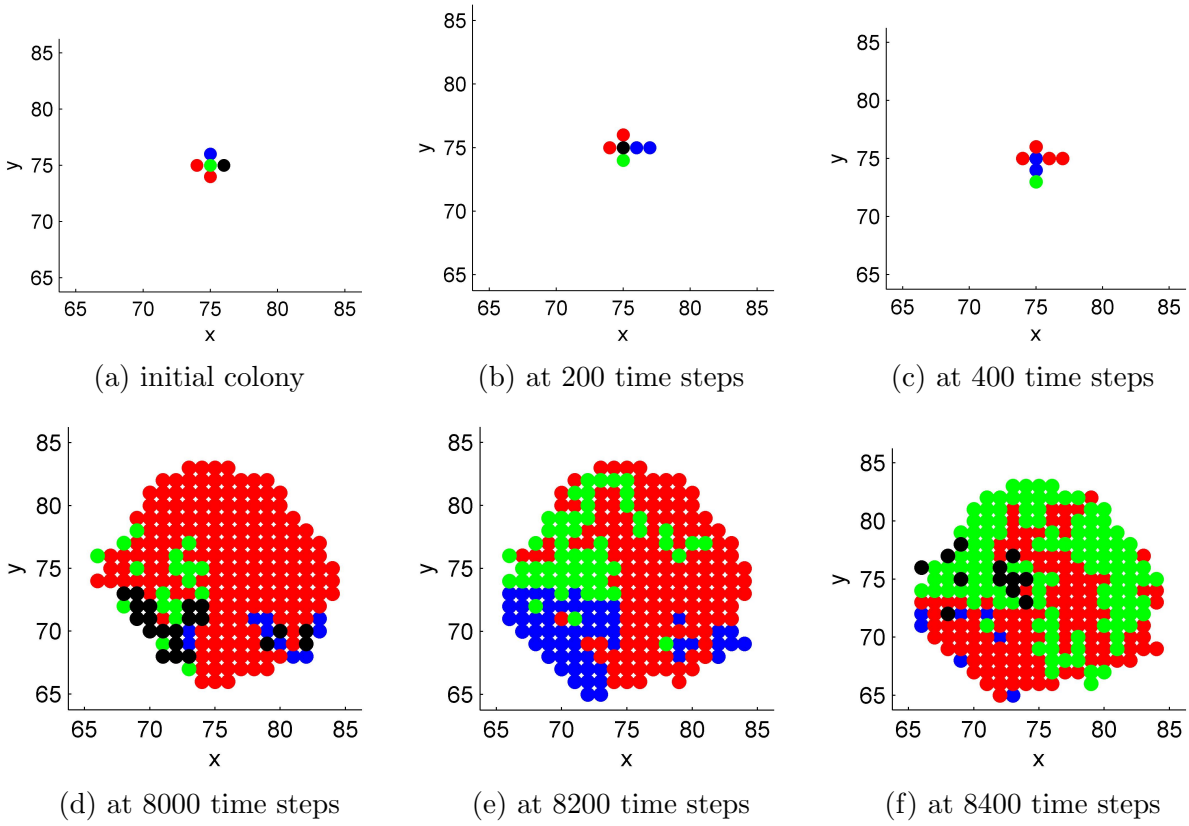
Figure 5.9: Colour-coded plots of the simulated colony at a varying number of time steps. Colour coding allows us to track parent cells and their offspring, which have the same colour. These results can be directly compared to experimental data.

as we can see in panel (f) where the colours are more mixed, especially the black cells which have dispersed the most. This is most likely because the central cell in the initial colony was black, so they have to move around the most to make room for their daughter cells, as they rarely have an unoccupied adjacent grid point when splitting. Although the mixing appears to increase with time, even after six divisions the colony is still clustered in clear lineage sections.

The spread and relative size of each cell lineage section in the colony can be used to test whether growth is uniform across a colony and to clarify whether the interior cells divide as often as the outer cells. Figure 5.9 has an equal number of cells of each lineage at all times due to the assumptions that growth is uniform and all cells have the same division time.

5.5 Including the Cell Cycle

The next development is to add the stages of the cell development cycle of an iPSC to the model. This is for two reasons. Firstly, it describes the proliferation process in more detail, and secondly, it will allow more accurate modelling of processes such as differentiation. Only stem cells in G_1 phase of the cell cycle can differentiate [9], so the model needs this level



(a) initial colony (b) at 200 time steps (c) at 400 time steps

(d) at 8000 time steps (e) at 8200 time steps (f) at 8400 time steps

Figure 5.10: Colour-coded images showing cells transition through the cell development cycle. Blue represents a cell in G_1 phase, red a cell in S phase, green a cell in G_2 phase and black a cell in M phase. After completing the cycle, a cell splits.

of detail to capture this. The time spent in each phase is recorded for each cell in the 'information' matrix, and when the duration of a phase is completed, the cell either moves to the next phase, or splits if it was in M.

The stages and durations of the cell development cycle of an iPSC are given in Section 2.1. Using the proportions of the division time spent in each phase we can determine the length of G_1 as 250τ, S as 700τ, G_2 as 400τ and M as 150τ in our model, as $D_t = 1500\tau$. These are constant for each cell for simplicity, although, a slight variation in division time is accounted for by giving daughter cells an initial age in G_1 between 0 and 100τ.

Using the relative duration of each phase we can give the initial colony a realistic proportion of cells in each phase: 16% in G_1 phase, 50% in S phase, 25% in G_2 and 9% in M phase. We also expect these ratios to be maintained as the colony grows. Once the phase of each initial cell is determined, a random age (between 1 and the phase duration) is given to each cell. Colour-coded plots and videos can then be made which show the growth of the colony in more detail, with each colour representing a stage of the cell cycle.

Figure 5.10 shows an example. In the simulation blue represents a cell in G_1 phase, red a cell in S phase, green a cell in G_2 phase and black a cell in M phase. Comparing panels (a), (b) and (c) shows clearly how the model cycles each cell through G_1, then S, then G_2, then M phase, before dividing. Comparing panels (d), (e) and (f) shows this on a larger scale and

also that the cells in each phase are in collective areas. This is a consequence of the behaviour seen in Section 5.4: that offspring of parents cells grow in the same area. Therefore, as all cells have the same constant phase durations, apart from the weak stochasity in the initial G_1 age of a daughter cell, each cell lineage will move through the cell cycle at a very similar rate. If the duration of each cell cycle stage was made stochastic this collective behaviour may be less obvious.

Comparing panels (d) and (e) we see that the black area of cells at the bottom left of the colony at 8000 time steps has increased in size and become blue by 8200 time steps. This is due to the short duration of M phase (150τ) and the fact that after completing this stage the cells divide. Comparing panels (d) and (f) we see that the few green cells on the left side of the colony at 8000 time steps have become black, i.e. transitioned from G_2 to M, after another 400 time steps. This is due to the longer duration of G_2 (400τ). We can be confident that the model is working as desired and is giving a more detailed description of the division process which can be compared to experimental data.

5.6 Including Differentiation

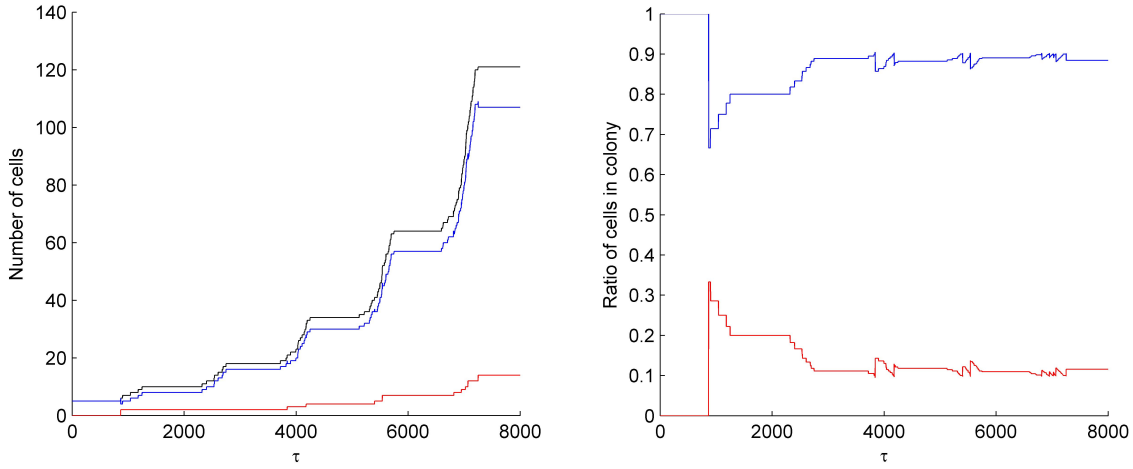
The final modification of the model discussed in this report is to include differentiation, which as explained in Section 2.1, happens around the edge of a colony. Stem cells can only differentiate when in G_1 phase of their cell cycle and, on average, we would expect approximately 10% of cells in a colony to be differentiated at any time [9]. Therefore, to incorporate this phenomenon into the model each boundary cell, which is also a stem cell in G_1 phase, is given a probability of differentiating each time step of

$$\frac{0.1N - N_{\mathrm{d}}}{\alpha}. \tag{5.6.1}$$

Here N is the number of cells as given in Equation (5.3.1), N_{d} is the number of differentiated cells (a function of time) and α is a normalisation constant to ensure the probabilities sum to unity. The normalisation constant, α, is equal to the number of cells which could possibly differentiate at each time step. Hence, it can be estimated by the proportion of cells in phase G_1 multiplied by the approximate circumference of the colony of stem cells, that is $\alpha \approx 0.16 \times 2\pi r_a = 0.32\pi r_a$.

The 'information' matrix is developed to note which cells are stem cells and which are differentiated at all times. For simplicity in the model, differentiated cells are considered as having the same morphology as stem cells, i.e. being of area 1, and do not divide. Modification of the model would have differentiated cells occupying 2 or more grid points, to account for their relative larger size, and give them a division time of 2400τ, so they divide on a comparable scale to stem cells. Using the probability in Equation (5.6.1) within our model the number of differentiated cells can be monitored as a function of time and compared to the number of stem cells. If Equation (5.6.1) is valid we expect to see proportions of 90% stem cells and 10% differentiated.

Figure 5.11 shows the number of differentiated cells and stem cells of a simulation initially with $N_0 = 5$ and $N_{\mathrm{d}} = 0$. Panel (a) shows the step wise increasing values of cell numbers of each type and panel (b) shows the relative proportions of the colony that are stem cells



(a) The number of cells vs. τ.

(b) The proportions of cells.

Figure 5.11: The number of differentiated cells and stem cells in the growing colony. In the left-hand panel the black, blue and red lines represent all cells, stem cells and differentiated cells respectively. In the right-hand panel the blue line shows the proportion of stem cells in the colony and the red line shows the proportion of differentiated cells in the colony. The desired ratios of each type of cell are seen.

and differentiated cells. It appears to take roughly 2500 time steps for the colony to reach the expected ratios of 10% differentiated cells and 90% stem cells and beyond this these proportions are maintained. This gives confirmation that the model is working as expected, and that the assumed probability of a stem cell differentiating each time step (Equation (5.6.1)) gives the desired result.

In an unpublished paper [16] experiments suggest that around the edge of a colony there is a band of constant width of differentiated cells. This band is thought to be three differentiated cells thick at all times, even as the colony continues to grow. This means that at any time we can approximate the number of differentiated cells by

$$\frac{N_{\mathrm{d}}}{2\pi r_a} \approx 3, \tag{5.6.2}$$

where r_a is the radius of the colony of stem cells, disregarding the boundary of differentiated cells. Figure 5.12 shows the number of differentiated cells in the model against $2\pi r_a$, where the probability of differentiation was given by Equation (5.6.1). It shows that the ratio is not constant, but instead an increasing step-wise function. The model was run with a smaller and larger probability of stem cells differentiating than in Equation (5.6.1) and the ratio given in Equation (5.6.2) was never seen to settle on a constant value, let alone 3.

The model shows good agreement with the proportions of stem cells and differentiated cells expected in a colony, however, for these to be maintained it is not feasible for the ratio in Equation (5.6.2) to be constant. Further experiments would need to be done to investigate whether it is the case that the ratios are 10% differentiated cells and 90% stem cells in a colony or whether there is a constant width band around the edge, however, my model has shown that both are not simultaneously possible.

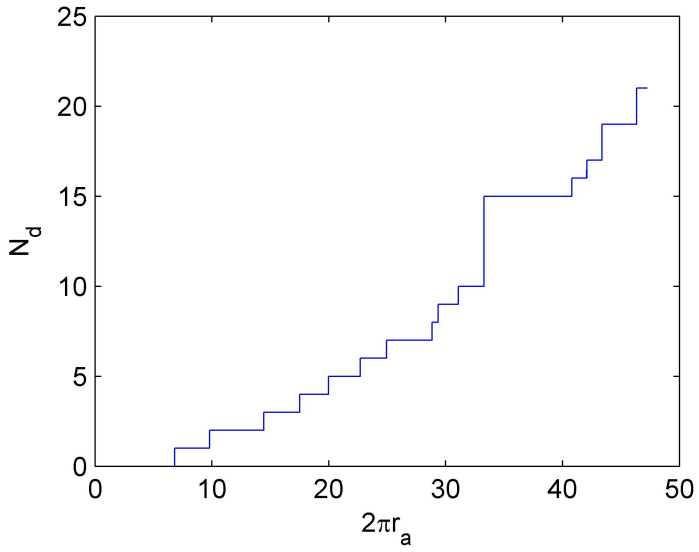


Figure 5.12: Quantifying Equation (5.6.2) using the probability of differentiation at each time step given in Equation (5.6.1). We see an increasing step-wise function, not a constant value.

Chapter 6

Conclusions

The key features of iPSCs behaviour which I initially identified as important were their division process (including cell cycle), migration, pluripotency spectrum, death and their ability to differentiate (Section 2.1). These are the primary elements of a useful mathematical model.

Stem cells are autonomous, reactive to their environment and migrate to be closer to other stem cells. An agent-based model is an adequate tool to describe such behaviours. Because stem cells divide on a large time scale of 14–16 hours [9] and have a small migration velocity of approximately 9.1 μm/h [10], it is reasonable to describe their behaviour in discrete time and space. This means we can use a particular type of agent-based mode: a cellular automaton model.

The model is based on a grid of step size Δx, where only a single cell can occupy each grid point, and evolves in time steps of duration τ. The value of τ can be made small enough that only one event happens at each time step but large enough that it is not necessary to make the model continuous in time to capture the behaviour. A cellular automaton model is ideal for including further levels of complexity step by step, particularly helpful when including the many complex characteristic behaviours of stem cells. Thus, the model can be tested at each stage for its suitability before any further development is made.

The analysis of my agent-based model for the migration of stem cells discussed in Chapters 3 and 4 assumes that iPSCs and hESCs move in the same way. It shows that it is appropriate to model the migration of a single hESC and multiple hESCs separated by a distance greater than 150 μm by a two-dimensional isotropic random walk along a grid. The mean and standard deviation of the travel radius of 1000 simulated cells obtained from the models agree well with experimental results [10].

At separation distances less than 70 μm, the migration of stem cells becomes directional, so we can consider 70 μm as a threshold distance where the cells are able to communicate and are attracted to each other. When cells are within 30 μm of each other, it is appropriate to model them according to a strongly biased random walk with attraction. This was done by comparing the directionality of cells within the threshold distance in my model with the experimental directionality [10]. For separations between 70 μm and 30 μm, the attraction between the stem cells increases from 0% to 100%, so my migration model could be improved by having a distance-dependent attraction.

The second agent-based model developed in this report describes the growth of stem cell

colonies. Upon developing the model, I identified two other features of stem cell behaviour which were not initially considered, but appear to be important for the growth of colonies. Firstly, it is unclear how a compact colony of cells, that are in close contact with each other, creates space for interior daughter cells. Secondly, it remains to be understood if the division rate of cells is uniform across a colony. A way of modelling both of these features was created by adapting my model to track the cell lineages in a colony. This would allow for comparison against relevant experimental data that has been obtained in an experiment designed at the Institute of Genetic Medicine to address these questions.

Cell division alone leads to an exponential growth of their number in time. The lack of death and differentiation in this case leads to a discrepancy in the estimated cell division time. Including the cell development cycle in the model gave more detailed insight into the proliferation of cells in the colony. This showed that equal phase durations in the model meant that large areas of a colony evolved through the cycle in almost perfect synchronisation. This needs to be tested against experimental data.

When including the differentiation of stem cells I computed the probability of differentiation that reproduces the experimentally measured proportion of differentiated cells in a colony, 10% [9]. I also explored a suggestion from another experiment that a colony has a band of differentiated cells of constant width 3 cells around the edge. My model showed that this is inconsistent with a fixed fraction of cells differentiated at any time. Both cannot be simultaneously true. Further experiments need to be carried out to determine the exact nature of the differentiation within a colony of iPSCs.

6.1 Further Developments

Although the models already give useful results, there are clear opportunities for further improvements to increase the level of insight into the behaviour of iPSCs.

Cell Migration

My model for migration inside the threshold distance, T, currently only works for simulations of two cells. It needs to be generalised to an arbitrary number of cells at various separation distances. For cells with neighbours within T, the overall migration probability depending on all surrounding cells within a circle of radius T would need to be calculated.

The other necessary development of this model is to make the attraction distance-dependent when $D < T$. This would mean adding an element of isotropic random walk, say $w(D)$, to the probability of moving in each direction which decreases with D. The probability of moving in each direction for each cell would then need to be normalised to unity. The manner in which $w(D)$ decreases would need to be investigated, however, from the experiments [10], we can assume that $w(70\,\mu\text{m}) \approx 1$ and $w(30\,\mu\text{m}) \approx 0$.

Growing Colonies

The first modification to this model would be to give each cell an independent random division time between 14 and 16 hours. This randomness would then be carried through when the

cell development cycle is added by making the duration of each phase random for each cell.

The next step would be to clarify how a colony accommodates daughter cells and whether the growth is uniform across the colony. This could be done by comparing the new experimental data on stained colonies to the model discussed in Section 5.4.

The characteristics of differentiated cells included in the model are currently very basic. The model firstly needs to distinguish between the morphologies of stem cells and differentiated cells. This may involve a differentiated cell occupying two or three grid points per cell compared to one grid point per stem cell to reflect their distinct sizes. With further development it may even become helpful to generalise the cellular automaton model to an agent-based model with continuous spatial dependence to reflect this.

The model for differentiation also needs to be adapted to include the division of differentiated cells (around 24 hours for differentiated cells compared to 14–16 hours for stem cells). If the proportion of differentiated cells is to be maintained at 10% (as experiments suggest it should [9]), including their division would mean decreasing the rate at which stem cells differentiate. Further experimental analysis needs to be done to quantify the rate of differentiation in the colony as there are currently two conflicting ideas. Once this has been done, the model will be more accurate and can be used to make suggestions of how to prevent or control differentiation within an iPSC colony.

Finally, cell death needs to be added to the model. This could initially be done in a similar way to differentiation, by giving each cell a probability of starting the process of apoptosis each time step. Then 24 hours later the cell would be considered as dead and removed from the simulated colony. The rate of death (or cells entering apoptosis) could be an arbitrary parameter to be determined from comparison of the model to experimental data.

Acknowledgments

I would like to thank my supervisors Prof. Anvar Shukurov and Dr. Nick Parker for their time, expert advice and encouragement throughout this project.

I also wish to give thanks to Dr. Irina Neganova (Institute of Genetic Medicine) for her time and effort in helping me understand the fascinating work she and the other biologists do at the Centre for Life. My deep appreciation also goes to Majlinda Lako (Institute of Genetic Medicine) for her time and help on the project.

Finally, I wish to thank my parents for their support.

Chapter 7

Bibliography

[1] Jones Biology: Cell Growth and Division. `http://www.jonesbiology.com/biology-1/chapter-10---cell-growth-and-division`. Accessed: 2015-03-26.

[2] M Aubert, M Badoual, S Fereol, C Christov, and B Grammaticos. A cellular automaton model for the migration of glioma cells. *Physical Biology*, 3(2):93, 2006.

[3] B Basse, B C Baguley, E S Marshall, W R Joseph, B Van Brunt, G Wake, and D JN Wall. A mathematical model for analysis of the cell cycle in cell lines derived from human tumors. *Journal of Mathematical Biology*, 47(4):295–312, 2003.

[4] T Czárán. *Spatiotemporal Models of Population and Community Dynamics*. Springer, 1998.

[5] B M Deasy, R J Jankowski, T R Payne, B Cao, J P Goff, J S Greenberger, and J Huard. Modeling stem cell population growth: Incorporating terms for proliferative heterogeneity. *Stem Cells*, 21(5):536–545, 2003.

[6] C A Goldthwaite. The promise of induced pluripotent stem cells (iPSCs). *National Institute of Health*, 2008.

[7] K Hochedlinger and K Plath. Epigenetic reprogramming and induced pluripotency. *Development*, 136(4):509–523, 2009.

[8] J Horowitz, M D Normand, M G Corradini, and M Peleg. Probabilistic model of microbial cell growth, division, and mortality. *Applied and Environmental Microbiology*, 76(1):230–242, 2010.

[9] Drs Majlinda Lako and Irina Neganova. Private Communications, Institute of Genetic Medicine, Centre for Life, Newcastle-upon-Tyne, 2014–15.

[10] L Li, B. H. Wang, S. Wang, L. Moalim-Nour, K. Mohib, D. Lohnes, and L. Wang. Individual cell movement, asymmetric colony expansion, rho-associated kinase, and e-cadherin impact the clonogenicity of human embryonic stem cells. *Biophysical Journal*, 98(11):2442 – 2451, 2010.

[11] T Ma, M Xie, T Laurent, and S Ding. Progress in the reprogramming of somatic cells. *Circulation Research*, 112(3):562–574, 2013.

[12] S Mitalipov and D Wolf. Totipotency, pluripotency and nuclear reprogramming. In *Engineering of Stem Cells*, pages 185–199. Springer, 2009.

[13] J.D. Murray. *Mathematical Biology: I. An Introduction*. Springer, 2002.

[14] S Ohmine, A Dietz, M C Deeds, K A Hartjes, D R Miller, T Thatava, T Sakuma, Y C Kudva, and Y Ikeda. Induced pluripotent stem cells from gmp-grade hematopoietic progenitor cells and mononuclear myeloid cells. *Stem Cell Res Ther*, 2(6):46, 2011.

[15] M Rao and N Malik. Assessing iPSC reprogramming methods for their suitability in translational medicine. *Journal of Cellular Biochemistry*, 113(10):3061–3068, 2012.

[16] K Rosowski, A Mertz, S Norcross, and E Dufresne. hESC differentiation and colony edge. Edges of human embryonic stem cell colonies display distinct mechanical properties and differentiation potential.

[17] S Schäuble, K Klement, S Marthandan, S Münch, I Heiland, S Schuster, P Hemmerich, and S Diekmann. Quantitative model of cell cycle arrest and cellular senescence in primary human fibroblasts. *PloS ONE*, 7(8):e42150, 2012.

[18] S H Strogatz. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Westview Press, 2001.

[19] M A Tabatabai, Z Bursac, W M Eby, and K P Singh. Mathematical modeling of stem cell proliferation. *Medical & Biological Engineering & Computing*, 49(3):253–262, 2011.

[20] K Takahashi, K Tanabe, M Ohnuki, Mi Narita, T Ichisaka, K Tomoda, and S Yamanaka. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, 131(5):861–872, 2007.

[21] K Takahashi and S Yamanaka. Induced pluripotent stem cells in medicine and biology. *Development*, 140(12):2457–2461, 2013.

[22] Invitrogen Life Technologies. *The CytoTune^{TM}-iPS Reprogramming Kit User Guide*. Catalog Numbers: A16517, A16518, Publication Number: MAN0009378, Revision 1.0.

[23] G Wurzer, K Kowarik, and H Reschreiter. *Agent-Based Modeling and Simulation in Archaeology*. Springer, 2013.

[24] Q Ying, J Nichols, I Chambers, and A Smith. Bmp induction of id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with stat3. *Cell*, 115(3):281–292, 2003.

[25] J Yu, M A Vodyanik, K Smuga-Otto, J Antosiewicz-Bourget, J Frane, S Tian, J Nie, G Jonsdottir, V Ruotti, R Stewart, et al. Induced pluripotent stem cell lines derived from human somatic cells. *Science*, 318(5858):1917–1920, 2007.