Investigating the interplay between cellular mechanics and decision-making in the *C. elegans* germ line

Kathryn Atwell
Oriel College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2016

Abstract

The behaviour of individual cells must be carefully coordinated across a tissue to achieve correct function. In particular, proliferation and differentiation decisions must be precisely regulated throughout development, tissue maintenance, and repair. A better understanding of how these processes are controlled would have implications for human health; cancer is, after all, dysregulated proliferation, while regenerative medicine relies on being able to influence cell decisions accurately. To investigate such fundamental biological processes, it is common practice to use an experimentally tractable model organism. Here, we focus on the germ line of the nematode worm *C. elegans*, which provides opportunities to study organogenesis, tissue maintenance, and ageing effects. Despite the advantages of this organism as a biological model, certain questions about germ cell behaviour and coordination remain challenging to address in the lab. There is therefore a need for computational models of the germ line to complement experimental approaches.

In this thesis, we develop a new *in silico* model of the *C. elegans* germ line. Novel aspects include working in three dimensions, covering the late larval period, and integrating a logical model of germ cell behaviour into a wider cell mechanics simulation. Our model produces a reasonable fit to wild-type germline behaviour, and provides the first cell tracking and labelling predictions for the larval period. It also suggests two new biological hypotheses: 1) that "stretching" growth plays a significant role in gonadogenesis, and 2) that a feedback mechanism acts on the germ cell cycle to prevent overproliferation. Having introduced the full model, we address some technical questions arising from our work, namely: what is the effect of applying a more physically realistic force law?; and can simulation performance be improved by changing the numerical scheme? Finally, we use *in silico* modelling to compare a number of hypothesised germ line maintenance mechanisms. There, our results support a model with functionally equivalent germ cells undergoing at most infrequent, transient cell cycle arrests.

Acknowledgements

If only I had a well-thought-out, beautifully written acknowledgments section to put here. Sadly I don't, mainly on account of never believing this day would come. So instead we'll have to make do with whatever I can cook up in the final day or so before submission.

Firstly, this thesis would not have been possible without the help of my embarrassingly large supervisory team. Prof. David Gavaghan has been a voice of reason throughout, and has been invaluable in navigating the bureaucracy of an Oxford DPhil. Dr. James Osborne has been great fun to work with, and a continual source of good ideas. I'm astonished by how much time he had for this project, even after moving to Australia and becoming father to twins. I owe him at least two beers. Lastly, Dr. Hillel Kugler has been a knowledgeable and encouraging supervisor. I was sorry to see him go when he relocated to Israel in my second year.

To lose one supervisor is a misfortune. To lose two begins to look like a problem with me. Therefore I'm particularly indebted to Dr. Sara-Jane Dunn and Prof. Ruth Baker for agreeing to step in as replacement supervisors. I'd like to thank SJ for her positive mental attitude and for managing to correct hundreds of pages of thesis with eerie speed. And I thank Ruth for her patience, understanding and comments. I will never forget her strong views on comma placement and on the usage of ”which” vs. ”that” (which I have yet to master).

I am grateful to our fantastic experimental collaborators from New York University, Prof. Jane Hubbard and Dr. Zhao Qin. Their unexpected enthusiasm for this project and their willingness to test our crazy hypotheses has been wonderful. In particular, Jane's expertise and determination played a huge part in getting Chapter 4 of this thesis published. I also wish to acknowledge the generous support of the Biological Computation group at Microsoft Research Cambridge; all the more generous given that I came for 12 weeks and ended up staying for three years! Thanks for making me feel like part of the team.

It is certainly true that one of the best things about academia is the people. So I must put in a good word for a lovely bunch of eccentrics from Oxford, namely Beth McMillan, Paul Taylor, Will Smith and Anna Muszkiewicz. Thanks for all the... interesting conversations guys. You will forever have a table in the cosy pub of my heart. The same goes for other friends who have lived through this DPhil with me: Steve, Richard, Zhen. You now know more about the *C. elegans* reproductive system than you ever thought possible. You're welcome.

My parents David and Gillian have been incredibly supportive throughout my studies, always ready to welcome me home with open arms. They (almost) never made fun of my topic, and in return I never made them proof read my thesis. Thanks for everything guys.

Finally, to my husband Saran: I love you, you have the patience of a saint, and I promise never to do this again!

Contents

Glossary

AC

Anchor Cell; a key cell directing vulval development in *C. elegans*.

AMT

Absolute Movement Threshold; the maximum distance a cell is allowed to travel in one time step of a Chaste simulation. Used to highlight spurious numerical results.

ARD

Adult Reproductive Diapause; a starvation response of young adult worms, in which most of the germ line dies with the exception of a small distal cell population. The germ line can be repopulated on feeding.

CD

Cell Diameters; a measure of distance from the DTC based on counting rows of germ cells.

DIC

Differential Interference Contrast; a microscopy technique that uses light with two different polarisations to boost the contrast when viewing transparent tissues.

DTC

Distal Tip Cell; the somatic cell at the tip of each gonad arm that acts as a leader cell during gonadogenesis and signals for germ cell proliferation.

FSM

Finite State Machine; a modelling formalism based around system states and the allowed transitions between them.

GLS

Generalised Linear Spring; a force law defined by Equation (3.10).

hph

Hours post-hatching.

JKR

Johnson Kendall Roberts; a model of elastic contact incorporating adhesion.

LSC

Live Sequence Chart; a modelling formalism that extends MSCs.

MSC

Message Sequence Chart; a visual modelling formalism based on message passing between entities.

MSM

Meta State Machine; a C++ library for statecharts work.

MSP

Major Sperm Protein; a protein that signals the presence of sperm in the spermatheca, leading to oocyte maturation and ovulation.

ODE

Ordinary Differential Equation.

PDE

Partial Differential Equation.

SBML

Systems Biology Markup Language; a domain specific language for expressing the details of biological models.

UML

Universal Modelling Language; a collection of formalisms for modelling and system design.

VPC

Vulval Precursor Cell; a cell type important in *C. elegans* vulval development that can adopt either a primary, secondary or tertiary fate.

Chapter 1

Introduction

The precise regulation of cell proliferation and fate decisions is fundamental to life. It enables the development of complex metazoan body plans, as well as the regeneration and repair of adult tissues. It also impacts upon human health: dysregulated cell proliferation underlies cancer, while the promise of regenerative medicine depends on a deep understanding of how stem cells execute directed differentiation, forming and maintaining the structures of the body. In short, the control of proliferation and fate is a key topic in biology that plays a crucial role in health. Since cell division and fate decisions occur in all sufficiently complex organisms, and since regulatory mechanisms can be conserved over evolutionary time scales, it is worth studying these processes in a simple model organism for which experimental data is readily available. In this thesis, we investigate the proliferation and fate decisions made by germ cells in the nematode worm *C. elegans*.

C. elegans has been used as a model organism since the 1970s, when it was popularised by Sydney Brenner [32]. Since then, three Nobel prizes have been awarded for work using *C. elegans*, ranging from fundamental research into development [234], to the discovery and refinement of new experimental techniques (GFP labelling and RNA interference [38, 75]). The worm possesses many advantages that have led to it becoming a prominent model organism. First, it has a short generation time of around 3 days [149], allowing large numbers of animals to be bred quickly and allowing phenotype to be studied over several generations. Second, the worm is hardy enough to be stored conveniently; young larvae can survive freezing and even recover after years of preservation at -70°C [149]. Third, because *C. elegans* exists in both hermaphrodite and male forms, desirable genetic traits can be propagated by self-fertilisation, while diversity can be introduced via sexual reproduction. Above all, *C. elegans* is amenable to the use of microscopy techniques to study all its cells. The worm is transparent, with a manageable number of cells for a metazoan, and for the most part these cells follow a fixed lineage, adopting predictable fates. It is this property of *C. elegans* that allowed Sulston *et al.* [234] to trace the complete lineage of all the organism's somatic cells, and which has since led to the publication of its complete neuronal network or "connectome" [265].

The lifespan of *C. elegans* is short: only 3 weeks in replete laboratory conditions [137]. As such, most of its tissues do not need to renew or repair themselves, and once development is over, few proliferative cells remain in the body. The reproductive system is the exception. The *C. elegans* germ line contains a permanent population of dividing cells at the tip of each gonad arm, maintained by signals from a specialised somatic cell [144]. These signals create a stem cell niche-like environment. As germ cells travel through the gonad toward the uterus, they eventually switch from proliferation into meiosis. The germ line therefore exemplifies controlled proliferation and differentiation, leading to tissue maintenance. Germ cells also display a carefully regulated sexual fate decision. In the larva, all germ cells develop into spermatazoa, while in the adult hermaphrodite the entire germ line switches to oocyte production [109]. The *C. elegans* germ line is therefore an excellent system for studying diverse fate decisions. Moreover, germ cells are influenced by the surrounding environment of the gonad, an organ that develops during the larval life cycle stages [144]. The wider reproductive system therefore affords opportunities to study organogenesis.

Despite the advantages of *C. elegans* as a model organism, there remain obstacles associated with experimental work on the germ line. Germ cells do not follow a fixed pattern of timed, oriented divisions of the kind seen somatically. Instead, the pattern of germ cell divisions varies from worm to worm, and the resulting cells appear almost identical under light microscopy [141]. As a result, tracking individual germ cells is challenging, and no comprehensive description of germ cell lineages exists. The only way to distinguish germline cells is by using fluorescent labels. The challenge then becomes achieving reliable, continuous expression of a fluorescent marker in the germ line. Unfortunately, germ cells act effectively to silence foreign DNA [238], and, while recent advances allowing single copy insertions go some way toward addressing this issue [275], techniques for single cell tracking and clonal labelling remain under development. Meanwhile, the extreme sensitivity of the adult reproductive system to stress prevents it being live imaged over long periods. A final problem is that certain interesting properties of germ cells are difficult to measure at all; such as the activity of certain intracellular signalling molecules, or whether the cell is in mitotic or meiotic S phase [114]. Even when live measurements are unnecessary and the germ line can be removed and fixed, sometimes the antibodies do not exist to mark the proteins or cells of interest.

In light of these difficulties, there is a need for alternative approaches to complement experimental work on the germ line. *In silico* models are a good candidate; able to generate predictions where the corresponding experiments are time consuming or technically infeasible. For example, visualising the abundance of a particular molecule in a germ line simulation is trivial, as is single cell tracking, labelling and lineaging. Working hypotheses about the movement and behaviour of germ cells can be developed using a computational model, and, once the necessary experimental techniques *do* become available, models will aid in interpreting labelling data [60]. Moreover, multiscale models, which cover both intracellular and intercellular interactions, can help to extrapolate from a hypothesis about individual cell behaviour to its implications for the entire germ line. Simulations using different germ cell behaviour may result in subtly different whole organ properties. This could then allow the underlying hypothesis to be experimentally tested, by looking only at easy-to-measure, large scale properties.

For an *in silico* germ line model to be applicable to a wide range of problems, it must incorporate realistic cell mechanics as well as reasonable single cell behaviour. The mechanical interactions between germ cells determine their movement, which, in turn, determines the neighbours and signalling environment they are exposed to over time. A hypothesis about tissue renewal that initially seems plausible may fail once the movement of cells is taken into account. The mechanical properties of germ cells also determine how they pack together, which influences commonly measured germline attributes, as well as the forces individual cells experience. There have been reports in many other biological systems of cells altering their behaviour in response to mechanical forces and cell-cell contact. For example, the osteocytes in bone respond to stresses placed on the skeleton, while epithelial cells grown in culture experience contact inhibition of proliferation, ceasing to divide once the available space is filled [133, 276]. It will be of interest to the community to explore whether similar mechanical feedback might exist on *C. elegans* germ cells. If so, *C. elegans* could be used as a new model system to study these poorly understood physical mechanisms.

Finally, an *in silico* germ line model should ideally be three-dimensional (3D), allowing it to capture cell movement in the gonad tip and turn regions in full, rather than following only a cross-section or cell movement down the long axis. While it would be interesting to construct a male germ line model, here we focus only on the hermaphrodite, for which much more quantitative data is available. In this thesis, we set out to develop multiscale models of the hermaphrodite germ line, and to use them to investigate hypotheses about germ cell behaviour in larvae and adults.

1.1 Thesis overview

The structure of this thesis is as follows. Chapters 2 and 3 both review pre-existing work. In Chapter 2, we summarise our current biological knowledge of the *C. elegans* germ line; covering its development, maintenance and response to perturbations. We also introduce several open questions, which the rest of the thesis goes on to explore. Chapter 3, meanwhile, covers cell-based modelling tools and techniques. In this third chapter, we discuss which approaches might be suitable for representing the germ line, as well as the tools and techniques employed in later chapters (Statecharts and the Chaste modelling library). We also review the few existing computational models of this system.

Chapter 4 presents our own 3D model of the *C. elegans* germ line, which covers the organ's development during the L3 and L4 larval stages (see Chapter 2), as well as its adult maintenance. We assess our model's fit to a range of wild-type germline properties reported in the literature. We then discuss three new biological insights that were obtained during the modelling process, namely: an important role for "stretching" growth during gonadogenesis; the requirement for a feedback mechanism to prevent germline overproliferation; and predicted germ cell tracking and labelling behaviour.

Chapter 5 examines some technical ways in which our cell-based model might be improved. We consider two main topics: whether a more detailed, physically realistic force law could change the outcome of our simulations; and whether an alternative numerical scheme could improve simulation accuracy and performance. We examine an alternative force law proposed by Buske *et al.* [33], and find that it does not significantly change the result of distal germ line simulations, in fact behaving similarly to the simple spring-based model used in Chapter 4. Meanwhile, switching from a forward Euler to a Runge-Kutta 4 method for updating cell positions *can* improve performance and accuracy, as can adaptive time-stepping (with some caveats). Chapter 5 ends with a short section proposing an alternative type of intracellular model, cell logic modules, which seeks to address some of the challenges associated with our earlier statecharts approach. Cell logic modules are used throughout the investigations in Chapter 6.

Chapter 6 considers a different kind of change to our model: altering the way in which the proliferative zone is maintained at the cellular level. We focus on the distal germ line and compare three different proliferative zone maintenance hypotheses, some of which include a distinct stem cell population, while others rely on functionally equivalent germ cells. We assess how well each maintenance mechanism fits known germline properties, and also model two perturbations to normal function: sudden withdrawal of the niche signal; and ageing.

Finally, Chapter 7 discusses the main conclusions of our work and suggests avenues for future investigation.

1.2 Thesis aims

The aims of this thesis can be summarised as follows. Firstly, we aim to develop a 3D model of the germ line that incorporates both cell movement and behaviour, while also covering the late larval development period. Secondly, we wish to explore a number of open questions about the germ line, involving the interplay between mechanics/movement and decision-making. These open questions are introduced in Section 2.9. Briefly, we ask: how are germ line expansion and gonadogenesis related?; how are cell numbers kept stable in the young adult?; and how is the proliferative zone renewed? Chapter 4 contributes toward the first two of these questions, while Chapter 6 focuses on the third. Chapter 5, meanwhile, aims to address some more general technical questions relating to the effect of force law choice and the efficiency of different numerical schemes in a cell-based modelling context.

1.3 Experimental collaboration

This thesis contains new experimental data collected by our collaborators at New York University. Throughout, where we refer to our collaborators, we refer to the lab of Prof. Jane Hubbard at the Skirball Institute of Biomolecular Medicine, NYU Langone Medical Center. Specifically, all microscopy images and data were obtained by Dr. Zhao Qin unless otherwise noted, while Prof. Hubbard provided many useful discussions on *C. elegans* germline biology.

Chapter 2

Biological literature review

In this chapter, we review current biological knowledge about the *C. elegans* hermaphrodite germ line; covering its larval development, adult structure and function. Normal germ line maintenance is described, a process that relies on coordinated cell proliferation, meiosis, sex determination and apoptosis. We also detail the disruptive effects on the germ line of starvation, ageing and certain experimental perturbations. The review concludes by introducing several open questions in this field that we claim can be investigated *in silico*.

2.1 Orientation

Figure 2.1 introduces the basic anatomy of an adult *C. elegans*. The major organs are labelled, and the terms dorsal, ventral, anterior and posterior are used to refer to different regions of the body. The gonad itself is found in the centre of the worm, wrapped around the intestine and taking up a considerable proportion of the body cavity. Indeed, the adult gonad contains approximately 2500 cell nuclei, while the rest of the body contains only 810 [109, 234]. The gonad consists of a pair of U-shaped arms that meet at the uterus (Figure 2.1, cyan). The end of each arm closest to the uterus is referred to as *proximal*, while the end furthest from the uterus is *distal*.

In hermaphrodite individuals, the gonad produces both sperm and oocytes over the course of a life cycle, enabling self-fertilisation and allowing these animals to reproduce alone. However, a small percentage of *C. elegans* are males (about 0.1% of the offspring of a hermaphrodite [112]). Males arise by non-disjunction, when an embryo receives one X chromosome instead of two [112]. The

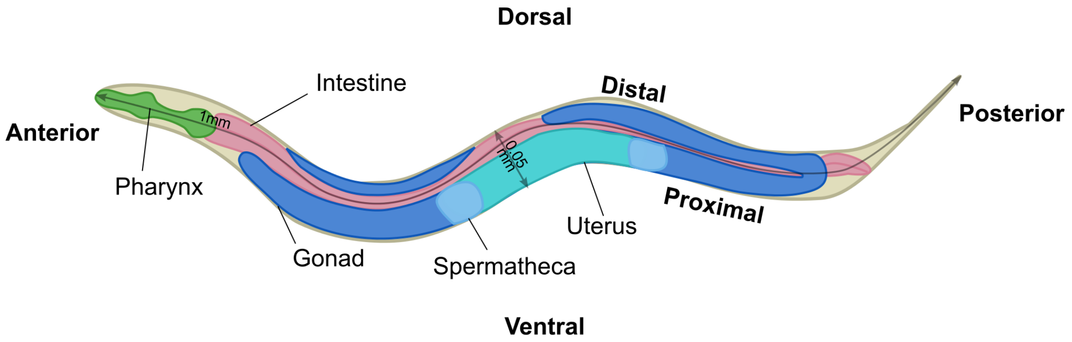


Figure 2.1: Basic hermaphrodite anatomy. A cartoon of the interior of a *C. elegans* hermaphrodite. For clarity, we omit the musculature, nervous system and waxy cuticle. The terms *dorsal* and *ventral* refer to the back and front of the animal respectively, while *anterior* and *posterior* refer to the head and the tail. In the context of a gonad arm, *proximal* refers to the end closest to the uterus, while *distal* refers to the end furthest from the uterus. Cartoon based on [9].

male gonad produces just sperm, and males must mate with hermaphrodites to reproduce. The primary differences in male anatomy compared to Figure 2.1 are the addition of a tail used in mating, and a straightened out posterior gonad that connects to the tail. Throughout the rest of this work, we will focus on *C. elegans* hermaphrodites.

2.2 Larval reproductive development

The *C. elegans* reproductive system develops during the larval part of the life cycle, which lasts around 1.5 days between hatching and maturity [149]. The larval period is subdivided into four developmental stages, referred to as L1, L2, L3 and L4 (in chronological order; see Figure 2.2). Each developmental stage ends in a moult, when the worm's cuticle is shed.

Immediately after hatching, at the start of L1, the reproductive system consists of a small membrane in the middle of the ventral surface of the worm, containing four primordial gonadal cells (Figure 2.2, L1) [109]. The two innermost cells (denoted Z2 and Z3) are the germ cell precursors, which go on to produce all of the animal's gametes. Meanwhile, the outer pair of cells (Z1 and Z4) give rise to the somatic tissues of the gonad [141]. These four cells slowly divide inside the gonad membrane until the L2 larval stage, when a rearrangement takes place. Most somatic cells cluster in the centre of the developing organ, splitting the germ cells into two pools and creating an anterior and a posterior gonad arm (Figure 2.2, L2) [141]. A single somatic cell remains at the tip of each arm and is referred to as a Distal Tip Cell, or DTC. The DTCs play two crucial roles in the hermaphrodite reproductive system: first, they act as leader cells during gonad development, and second, they produce a signal that keeps nearby germ cells proliferative [141, 144].

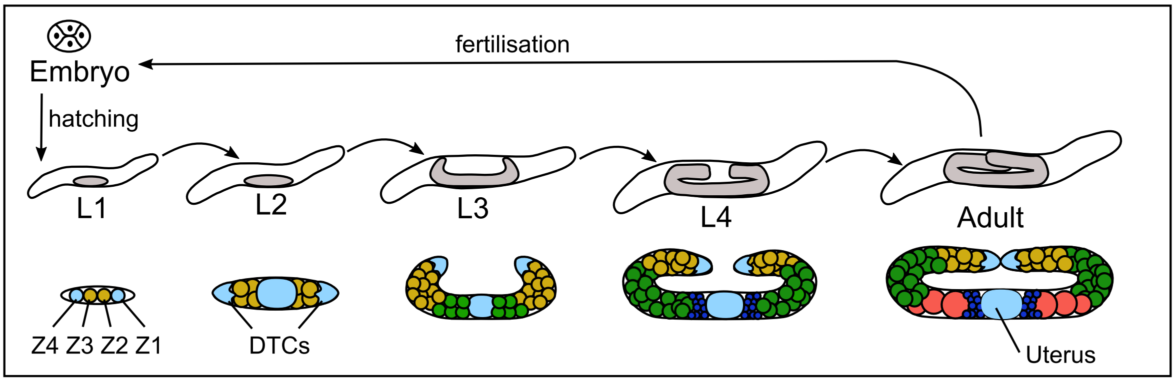


Figure 2.2: Germ line development in the larva. The top row of images show the *C. elegans* life cycle, with the shape of the growing gonad indicated by a grey silhouette at each stage (not to scale). The bottom row of images show the development of the germ line inside the gonad membrane. Cells coloured pale blue are somatic; these include the distal tip cells as well as the cells composing the uterus and vulva (represented by an oval). All other cells are germ cells, colour coded as follows: yellow = proliferative; green = meiotic; dark blue = sperm; pink = oocyte. In L1, we note the labels given to the somatic gonad cells and primordial germ cells. Germ cell numbers are underrepresented from L2 onward in this cartoon, and the sheath cells that surround the gonad are not depicted.

Reproductive development during the L3 larval stage is largely characterised by rapid gonad growth and expansion of the germ cell pool. The DTCs actively migrate toward the anterior and posterior ends of the worm, causing the gonad arms to extend [109, 144]. This migration requires the expression of two genes that encode metalloproteases, *gon-1* and *mig-6*. Since metalloproteases are enzymes that break down the extracellular matrix, it is thought the DTCs remodel surrounding tissue to make progress [29, 135].

Late in the L3 larval stage each of the DTCs executes a turn, moving from the ventral surface onto the dorsal surface, then changing direction to travel back into the centre of the worm. Turning requires the netrin receptors UNC-5 and UNC-40 along with their substrate UNC-6 [279]. UNC-5 expression is activated at the start of the turn, and it is hypothesised that the DTCs follow an UNC-6 gradient to travel from the ventral to the dorsal side [233, 258]. The integrin system also seems to play a role in turning, possibly by altering the adhesive properties of the DTC [268]. The molecular pathways involved in DTC migration and pathfinding are reviewed in [268]. Migration only halts with the adult moult, by which point the DTCs have begun to pass each other on the dorsal surface (see Adult, Figure 2.2). The complete sequence of gonad development during L3 and L4 is illustrated by Figure 2.3, which shows how a gonad arm grows from a short tube into an elongated U-shape.

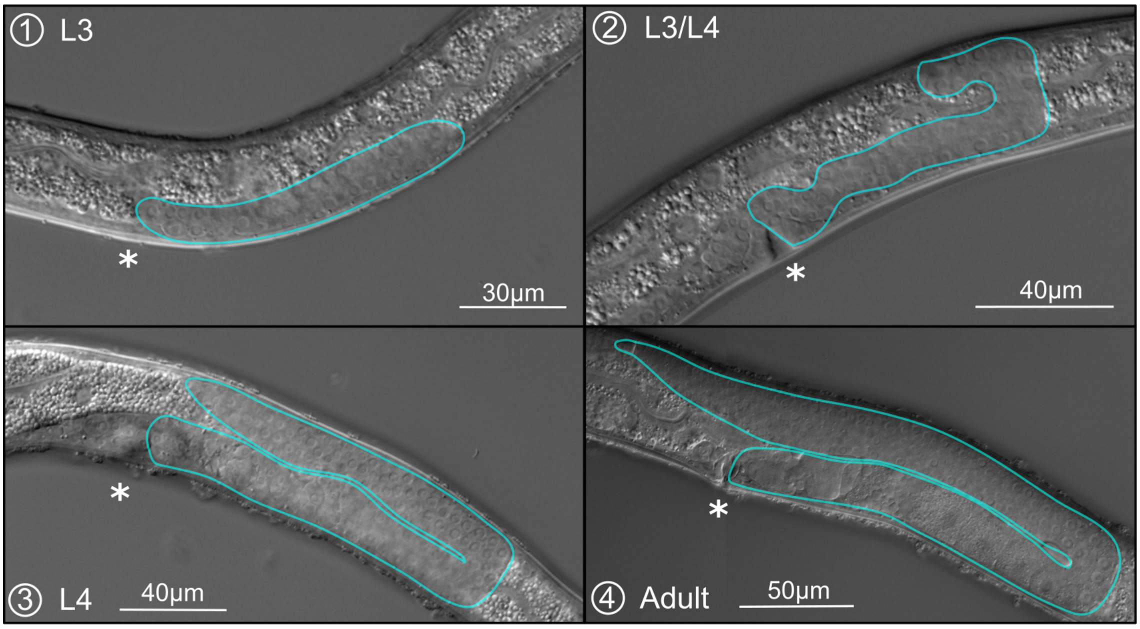


Figure 2.3: Gonad arm growth during L3 and L4. A sequence of Differential Interference Contrast (DIC) microscopy images, showing the growth of a single gonad arm during L3 and L4. The arm has been outlined in blue, and a white star indicates the position of the developing vulva. These and all other microscopy images in this thesis were taken by Dr. Z. Qin (see Section 1.3).

As the gonad lengthens, the germ cells inside rapidly proliferate to fill the available space [109]. The pressure generated by germ cell divisions inside the organ is thought to contribute to the anterior/posterior extension of the gonad arms, by accelerating DTC migration. This idea is supported by the observation that underproliferation mutants with few germ cells also have a shorter adult gonad [139].

Over the course of L3 and L4, other important changes take place in the germ cell population itself. The DTCs provide a signal that keeps nearby germ cells proliferative and inhibits meiosis [144]. As the DTCs move further away from the centre of the animal, proximal germ cells fall out of range of this signal, and thus begin meiosis (Figure 2.2, L3) [109]. Meiotic entry can be discerned by a characteristic crescent-shaped nucleus (due to chromosome condensation) as well as increased cell size [57, 109]. By L4, the proximal-most meiotic cells have begun to develop into sperm [109] (Figure 2.2, L4). Each spermatogenic germ cell grows into a large primary spermatocyte, before dividing twice to give four spermatids and a residual body (see Figure 2.4) [143, 263]. The residual body is broken down and resorbed, while the four spermatids go on to become mature, motile sperm. Sperm and spermatids collect in the proximal end of the gonad and are pushed into a flexible chamber called the spermatheca by the first oocyte as it is ovulated. The location of the spermatheca can be seen in Figures 2.1 and 2.5.

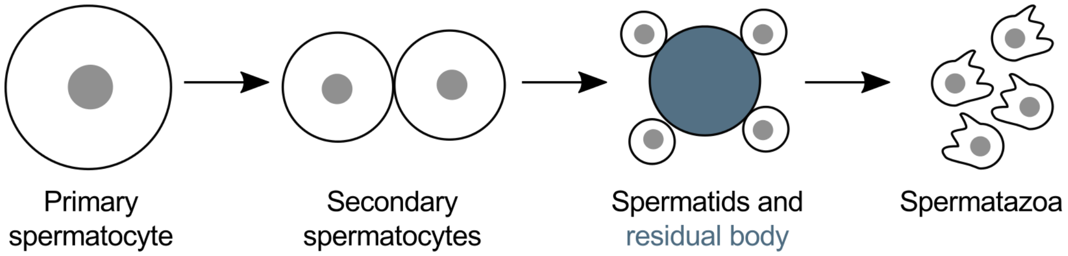


Figure 2.4: Spermatogenesis. The sequence of events during sperm production in L4. A large primary spermatocyte divides into four small spermatids and a residual body. While the residual body is eventually resorbed, the four spermatids undergo further changes, becoming motile spermatazoa. Based on [235].

Sperm production halts after the adult moult, at which point reproductive development is complete. Thereafter, all germ cells entering meiosis either develop into oocytes or undergo apoptosis in the gonad turn [95, 109]. At around this time, the germ cell division rate also slows, marking a shift from germ line expansion to maintenance [147]. The molecular pathways that regulate germ cell proliferation, meiosis, sex determination and apoptosis are all discussed in more detail below, as these processes continue in the adult.

2.3 Structure and function of the adult germ line

Figure 2.5 shows the organisation of the young adult germ line, at around 35.5 hours post-hatching (hph) [109]. Since the sole function of the adult germ line is to provide a constant supply of oocytes, the organ is configured similarly to a production line. At the distal end of the gonad, close to the DTC, is a proliferative zone containing dividing cells. Further along the gonad, germ cells cease proliferating and enter meiosis [102]. On reaching the turn, a number of germ cells undergo apoptosis, while the survivors continue to develop into oocytes. Finally, in the proximal arm, we see a queue of increasingly mature oocytes awaiting ovulation [211].

Ovulation is the last step in this process, where a mature oocyte is ejected into the spermatheca. There it is fertilised before entering the uterus. It is thought that mitotic pressure (generated by the distal proliferative population) drives cells most of the way though the gonad, while the final push into the spermatheca is provided by muscular sheath cells that surround the proximal region [46, 163]. Ovulation not only requires force, but also a chemical signal indicating the presence of sperm. The Major Sperm Protein (MSP) promotes oocyte maturation as well as prompting the sheath cells to contract [169, 170]. After around 5 days of adulthood, a hermaphrodite exhausts its supply of self-sperm [117]. It must then mate with a male to produce further offspring.

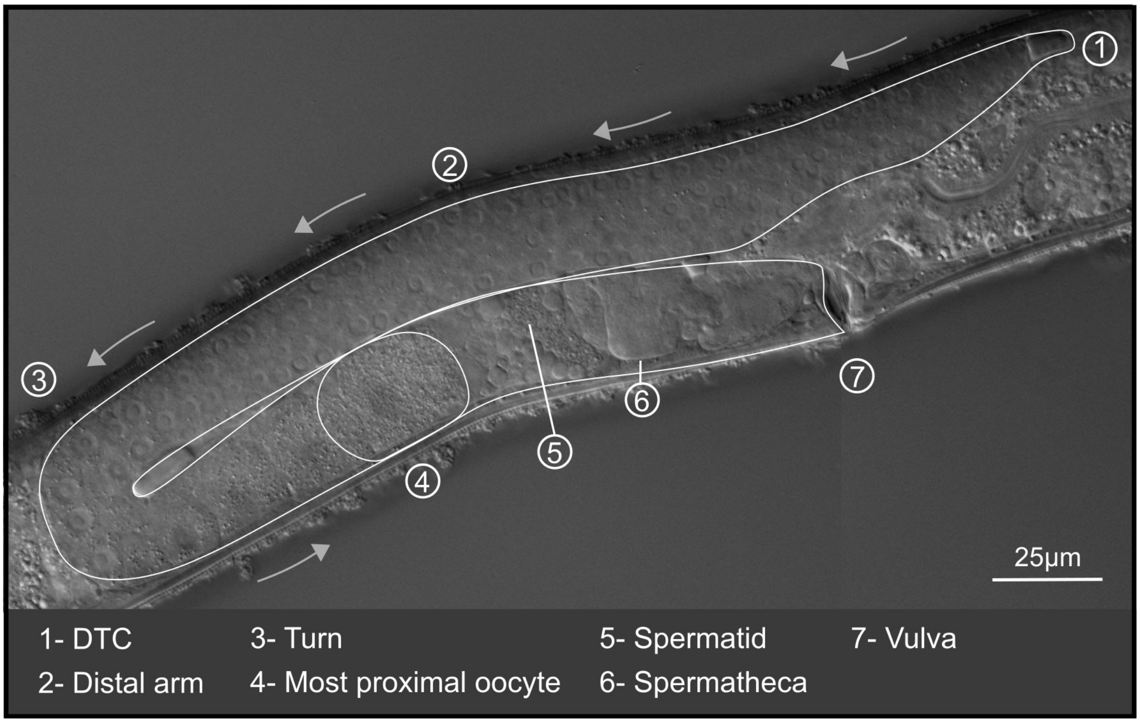


Figure 2.5: The adult germ line. DIC micrograph of a single young adult gonad arm. Arrows indicate the dominant direction of germ cell movement. The arm is outlined in white, as is the proximal-most oocyte. When this oocyte is ovulated, it pushes spermatids into place within the spermatheca.

The production of large oocytes requires a considerable amount of cytoplasmic material. Certain anatomical features of the reproductive system help ensure enough cytoplasm is available for oocyte growth. Firstly, nutrient rich yolk proteins are synthesised in the intestine [142]. Since the gonad wraps around the intestine, these proteins can pass readily into the germ line and accumulate in oocytes. Secondly, the gonad has a cytoplasm reservoir running down its centre called the rachis [1] (see Figure 2.6). Distal germ cells have an opening in their membranes leading to the rachis that may allow them to contribute material to this reservoir [267]. More proximally, narrow projections from the rachis attach to maturing oocytes and stream cytoplasm into them, supporting their growth [267]. As a result of their connections to the rachis, the majority of germ “cells” are actually syncytial, that is, they are individual nuclei within a shared cytoplasm. Oocytes do not complete their cell membranes until they bud off from the rachis in the proximal gonad [109]. Despite their syncytial nature, distal germ cells are not known to share cytoplasmic components with each other, and their cell cycles are not synchronised. Therefore, the field refers to all germ nuclei by convention as germ cells [114].

With the rachis streaming cytoplasm and the intestine providing yolk proteins, three oocytes can be produced per gonad arm per hour [162]. However, the majority of germ cells still do not

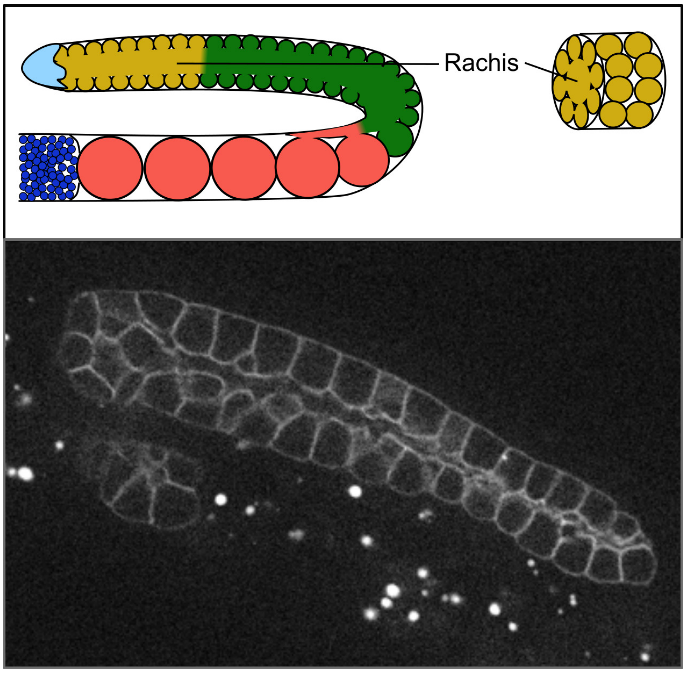


Figure 2.6: The rachis. Top: a cartoon showing the rachis, the cytoplasm reservoir that runs down the centre of the distal gonad and turn. Most germ cells open onto the rachis, with mature oocytes budding off in the proximal gonad. Taking a section across the distal arm, the rachis appears as a circle surrounded by 8-12 germ cells [109]. Bottom: a lengthwise section through an early L4 gonad arm. The rachis is visible as a narrow gap separating the two lines of germ cells. This is a fluorescence image of *xnSi1* [43], a strain of *C. elegans* containing a GFP marker targeted to germ cell membranes.

develop into mature oocytes. Instead, most undergo apoptosis in the gonad turn. Estimates of the ratio of germ cell deaths to ovulations suggest that between 2 and 30 cells die for each oocyte produced [114, 123]. Such a high death rate even in healthy individuals suggests that apoptosis performs an essential function. It has been proposed that dying germ cells may act as nurse cells, donating their cytoplasm to surviving oocytes and thereby providing them with further resources for growth [95].

In the next few sections, we will describe in greater detail the processes involved in normal germ line function; including proliferation, meiosis, and apoptosis. We begin with a discussion of germ cell proliferation in the distal gonad.

2.4 The proliferative zone

A proliferative germ cell population is maintained at the distal end of the adult gonad, consisting of around 230 cells [103]. When viewed under the microscope, these cells appear to be arranged in rows, giving rise to the common practice in the field of quoting measurements in *cell rows* or *cell diameters* (CD) [114]. Formally, the proliferative zone is defined as: *the region stretching from the DTC to the first cell row containing two or more meiotic cells* [50] (see Figure 2.7). Cells in meiosis have a characteristic crescent-shaped nucleus, which is used to identify them for this purpose. The requirement that the last row must contain more than one meiotic cell guards against outliers, since isolated meiotic cells do sometimes appear in a majority proliferative region.

The length of the proliferative zone and the number of cells it contains are important measurements in the field, because they indicate the germ line's capacity for self-renewal. In wild-type animals, the proliferative zone typically measures 20-25 CD, with dividing cells completely absent beyond 30 CD from the distal tip [51, 154]. The length of the proliferative zone in microns is somewhere between 50µm and 90µm, but these units are rarely used in the literature, with CDs preferred [102, 154].

Proliferative zone maintenance by the DTC

It has long been known that the DTC is key to proliferative zone maintenance. If the DTC is ablated, all germ cells enter meiosis and the proliferative zone is lost [144]. The biochemical pathway responsible for this behaviour is increasingly well-understood, and is summarised in

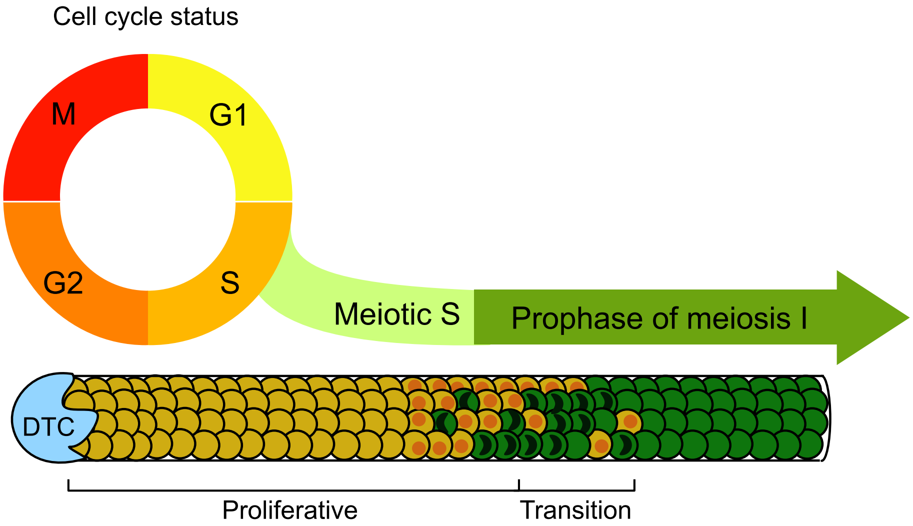


Figure 2.7: The proliferative zone. Along the bottom of this figure is a cartoon of the distal germ line with accurate cell row counts. It shows the extent of the proliferative zone, which terminates at the first row containing two or more meiotic cells. Following convention in the field, we depict the DTC surrounding the distal-most cell rows. However, in reality, the DTC possesses longer cell processes that may enable interactions with germ cells over greater distances. The existance of a diffusible molecule produced by the DTC also cannot be ruled out, as an additional mechanism alongside direct cell-cell interactions. Along the top of this figure is a graphic indicating progress through mitosis/meiosis. Distal proliferative zone cells are mitotic and cycle through G1, S, G2 and M phases. Further along the zone cells are increasingly found in meiotic S phase, a specialised S phase in preparation for meiosis [123].

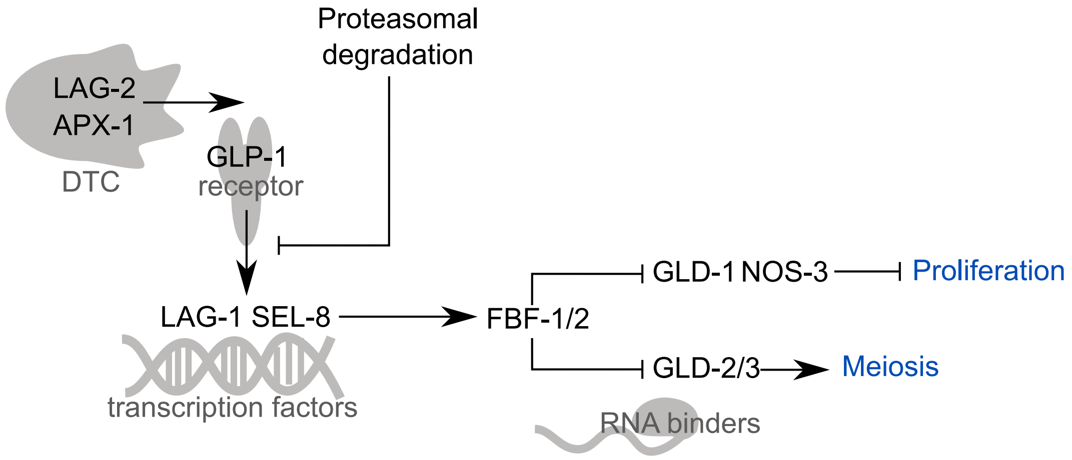


Figure 2.8: Delta/Notch-like DTC signalling maintains the proliferative zone. LAG-2/APX-1 displayed by the DTC binds GLP-1 on germ cells. The GLP-1 intracellular domain then disassociates and activates the LAG-1 transcription factor. This signal is propagated through a network of RNA binding proteins, including FBF and two parallel GLD pathways. Ultimately, proliferation is promoted and meiosis inhibited in cells close to the DTC. See text for references.

Figure 2.8. The DTC displays two membrane bound ligands, LAG-2 and APX-1, which bind the GLP-1 receptor on nearby germ cells [16, 51, 108, 180]. The GLP-1 intracellular domain then disassociates, enters the germ cell nucleus, and activates the transcription factor LAG-1 (together with cofactor SEL-8) [44, 61]. LAG-1 alters the expression of a number of RNA binding proteins, notably FBF-1/2, a pair of translational inhibitors [49, 278].

At least two distinct pathways lie downstream of FBF: a GLD-1/NOS-3 pathway and a GLD-2/3 pathway [69, 83, 104, 131]. These are thought to act in parallel, since the *gld-1*/*gld-2* double mutant displays serious abnormalities, while both single mutants are similar to wild type [131]. Antibody staining reveals a GLD-1 gradient along the length of the gonad, with the GLD-1 concentration rising further from the distal tip [104]. It is therefore proposed that DTC signalling inhibits GLD expression in the distal-most germ cells, with the effect of preventing meiosis and promoting continued proliferation. The DTC signalling pathway as a whole strongly resembles the Delta/Notch pathway in flies; indeed, LAG-2 is a Delta homologue, GLP-1 a Notch homologue, and LAG-1 is homologous to Su(H) [44, 239, 271].

The downstream details of DTC signalling remain somewhat unclear, since the RNA binding proteins involved have many targets, and it is not known which are important for proliferation control [130, 138]. In particular, it is unclear how exactly this pathway influences the cell cycle. Recent evidence suggests that CYE-1 and CDK-2 may be involved (the *C. elegans* homologues of cyclin E and cdk2). Depletion of CYE-1/CDK-2 causes germ cells to enter meiosis prematurely, a phenotype that requires the presence of GLD-1 [125]. GLD-1, meanwhile, has been shown to inhibit CYE-1/CDK-2 activity in the *C. elegans* embryo [28]. Taken together, these facts could point to a mutual repression between CYE-1/CDK-2 and GLD-1, which may form part of a "toggle switch" between proliferation and meiosis [125].

In addition to the question of how DTC signalling influences the cell cycle, there also exist problematic mutant phenotypes that do not make sense in the context of the simple pathway described above. For example, the phenotype of a *fbf-1/fbf-2* double mutant is less extreme than that of a *glp-1* mutant [49], implying that other pathways operate alongside the one shown in Figure 2.8. As such, the search continues for other factors controlling germ cell proliferation and meiotic entry. Figure 2.9 presents some of these additional putative regulators.

Clearly DTC ligands play an important role in proliferative zone maintenance. Recently, however, it has been discovered that the DTC can interact with germ cells in other ways; through processes (thin cell membrane projections), and gap junctions [35, 229]. Byrd *et al.* (2014) [35] demonstrated that the DTC extends thin processes between the distal-most germ cells. These processes are collectively known as the DTC plexus, and may supply signalling factors or anchor a stem cell population in place. Meanwhile, Starich *et al.* (2014) [229] showed that gap junctions exist between the DTC and distal germ cells, formed by the innexins INX-8/9 and INX-14/21. The role of these gap junctions is unknown, but mutants lacking them show severely reduced larval germ cell proliferation, and their germ lines often become necrotic in the adult.

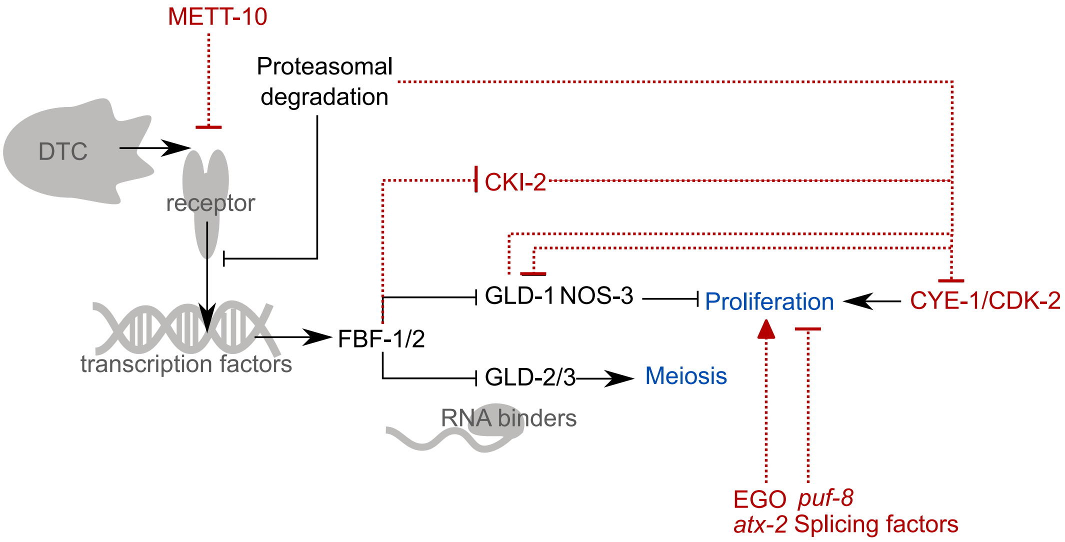


Figure 2.9: A more detailed picture of proliferation control. Shown in red are other factors thought to influence germ cell proliferation and meiotic entry, alongside the classical Delta/Notch pathway. In particular, mutual repression between GLD-1 and CYE-1/CDK-2 may form a toggle switch and provide a link to the cell cycle. These details are reviewed in [103].

Spatial trends within the proliferative zone

An obvious question raised by the above discussion is whether the DTC provides a niche environment that confers proliferative capacity on germ cells, or whether there exists a distinct stem cell population occupying the distal gonad. The issue of specialised stem cells versus functionally equivalent germ cells continues to be controversial in the field, with conclusive studies prevented by the lack of appropriate tracking and labelling techniques. This has prompted some researchers to try to infer information about the proliferative germ cell population by looking at spatial trends along the gonad; in particular, spatial trends in cell cycle behaviour.

The first quantitative studies of the proliferative zone were undertaken in 2006, by Maciejowski *et al.* and Crittenden *et al.* [50, 154]. These investigations involved capturing microscopy images of hundreds of *C. elegans* gonad arms, then processing the images to identify all cells currently dividing (i.e those in M phase). A quantity called the *mitotic index* can then be calculated for each proliferative zone row, defined as:

$$\text{Mitotic index of row } i = \frac{\text{Total number of dividing cells observed in row } i}{\text{Total number of cells observed in row } i}. \tag{2.1}$$

Assuming that the length of M phase stays roughly constant, the mitotic index can be used as a proxy for the germ cell division rate.

Figure 2.10A shows the observed mitotic index profile along the proliferative zone. Significant spatial variation is evident, with the distal-most germ cell rows cycling relatively slowly, and a peak in mitotic activity occurring around 7CD away from the distal tip. Thereafter, mitotic index decays rapidly with distance from the DTC [154]. This observation *could* indicate the presence of a distinct, slow cycling cell population occupying the distal part of the proliferative zone. We return to mitotic index profiles and what they can tell us about the germ line in Chapter 6.

In addition to studying the mitotic index, it is possible to analyse the percentage of germ cells found in each cell cycle phase. The established technique for estimating these percentages involves time course experiments in which germ cells are labelled in multiple ways. For example, brief exposure to the nucleotide analogue EdU will label only cells in S phase. However, a longer period of EdU labelling, followed by PH3 staining for M phase cells, can provide an estimate of the length of G2. Figure 2.10B shows estimated cell cycle phase percentages for the adult proliferative zone obtained in this manner [82]. An alternative approach uses software (such as Irises [255]) to estimate the amount of DNA in stained germ cell nuclei. Diploid DNA content would indicate a cell in G1,

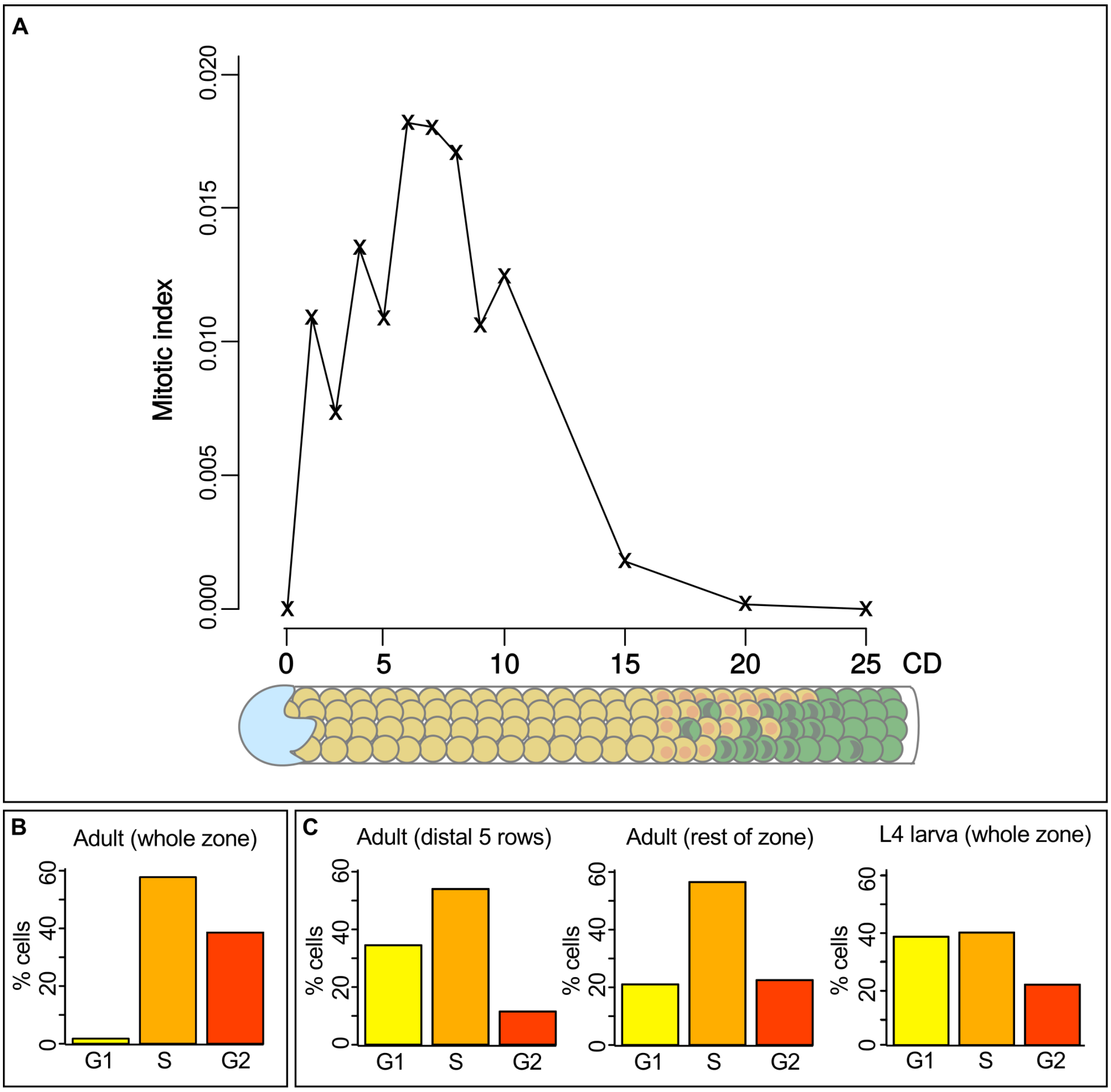


Figure 2.10: Spatial trends in cell cycle behaviour. A) Mitotic index along the proliferative zone, based on data in [154]. The rate of cell division increases with distance from the DTC up to a peak around 7CD away, and decays thereafter. B) A bar graph showing the estimated percentage of germ cells in each cell cycle phase, based on labelling experiments reported in [82]. C) A separate set of estimated cell cycle phase percentages based on DNA content analysis. Data provided by Dr. D. Michaelson (Hubbard lab, see Section 1.3). Reading left to right, these graphs relate to the distal-most five rows of the adult germ line, the rest of the adult proliferative zone, and the larval proliferative zone, respectively.

twice that DNA content would indicate a cell in G2, and a measurement somewhere in between would indicate an S phase cell (still in the process of duplicating its DNA). DNA content analysis points to subtle spatial trends in cell cycle phase occupancy, as well as to differences between the larval and adult proliferative zones (Figure 2.10C). However, the results are inconsistent with the traditional labelling approach, particularly as regards the length of G1.

In summary, the proliferative zone is a frequently studied region of the germ line because it is the site of tissue renewal. Proliferation is maintained by the DTC, primarily through Delta/Notch-like signalling; although other factors also contribute, including physical connections and gap junctions. While the issue of whether separate germline stem cells exist remains unresolved, certain groups of germ cells do seem to display different cell cycle behaviour, depending on their position in the zone and the age of the animal. These are trends that an *in silico* model of proliferative zone maintenance should seek to reproduce and explain.

We now turn to the next step in the oocyte production process: meiotic entry, followed by gradual progress through meiosis I.

2.5 Meiosis and the meiotic entry decision

Germ cells eventually stop proliferating and undergo meiosis, a process that reduces their DNA content to a single copy of the genetic material. This is a necessary step in the production of gametes. Meiosis takes place over two cell divisions: meiosis I and meiosis II. A specialised meiotic S phase is now widely considered to be the first step in the process [31, 123]. During meiotic S the DNA is duplicated, and links are established between sister chromatids. Prophase of meiosis I follows, during which genetic material is exchanged between homologous chromosomes, generating variation. Prophase of meiosis I can be further subdivided into leptotene, zygotene, pachytene, diplotene and diakinesis; the precise events that take place during each stage are illustrated in Figure 2.11. Next, the meiosis I division occurs, producing two diploid cells with reshuffled chromosomes. Finally, the meiosis II division happens without DNA duplication, giving rise to four haploid daughters.

Figure 2.11 shows where the different stages of meiosis occur in the hermaphrodite germ line [123]. For most of their journey through the gonad, germ cells are slowly progressing through prophase of meiosis I, reaching diakinesis as they enter the proximal region. Meiosis then pauses, and oocytes do not undergo the two meiotic divisions until *after* ovulation [7]. The molecular regulation of meiosis is outside the scope of this thesis. However, multiple layers of regulation exist to ensure homologous pairing, exactly one crossover per chromosome, the repair of extraneous double strand breaks, and correct DNA segregation into daughter cells (see review in [151]).

An interesting question is when exactly germ cells become committed to meiotic entry. The first visible sign of meiosis is a crescent-shaped nucleus, but by the time that appears cells are already in the leptotene/zygotene stages of meiotic prophase. Presumably, commitment to meiosis actu-

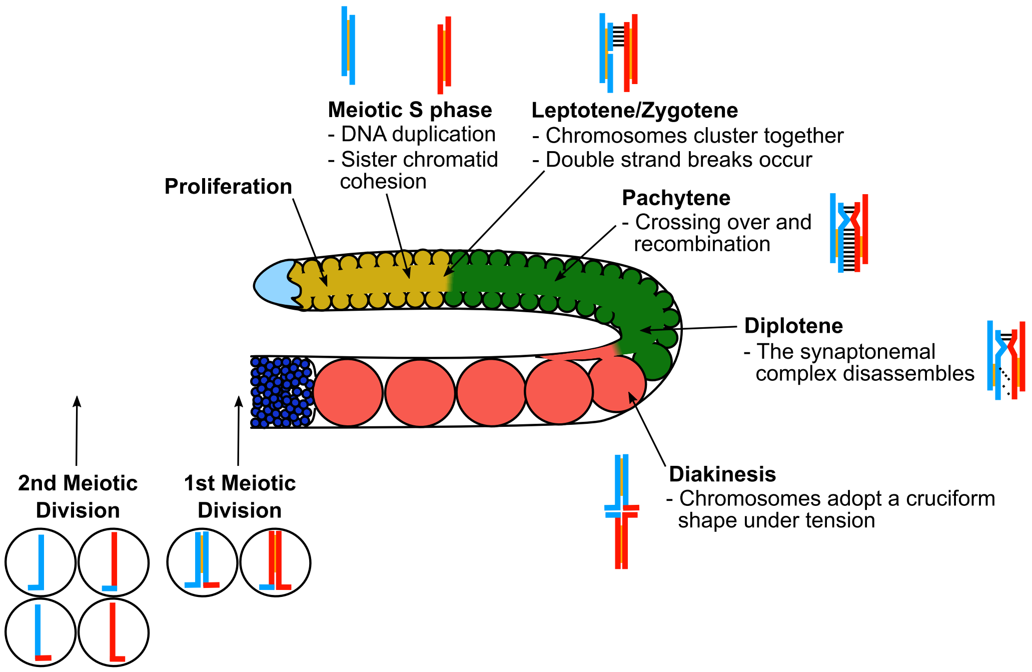


Figure 2.11: Progress through meiosis. A diagram showing the stages of meiosis and where in the gonad they occur. Most time is spent in prophase of meiosis I, during which genetic material is exchanged between homologous chromosome pairs. The two meiotic divisions do not take place until after ovulation. Based on figures in [123, 151].

ally takes place earlier, while cells are still in the "proliferative" zone. In fact, as many as 30% of proliferative zone cells are thought to be in meiotic S phase, and as such no longer actively dividing [81, 82]. Unfortunately, meiotic and mitotic S phase cannot be distinguished using current labelling techniques. As a result, crescent-shaped nuclei continue to be used as a convenient marker for the proliferation/meiosis transition. The region containing a mixture of crescent-shaped and normal nuclei is called the *transition zone*. In larvae it is only 1CD long, while in adults it measures around 5CD [102].

At some point during their passage through the gonad, adult germ cells make a further key decision: they commit to becoming oocytes. This is significant because, while adult germ cells always develop into female gametes, larval germ cells develop into sperm. Therefore, hermaphrodite fertility depends on a switch in germ cell sexual fate taking place. How this is achieved is the topic of the next section.

2.6 Sex determination

Sex determination is an important germ cell decision that must be properly regulated to ensure fertility. In general, germ cells that enter meiosis during the L3 larval stage become sperm-fated, while cells that enter meiosis at later times become oocyte-fated [274]. In the adult, the eventual sex of germ cells matches the chromosomal sex of the organism, but in the larva, sperm are transiently produced in an XX animal. This temporary switch to sperm production is achieved largely through post-transcriptional regulation, with mRNA binding factors modifying the default sex determination pathway.

For the interested reader, Figure 2.12 summarises the molecular regulation involved in sex determination. Briefly, in adult hermaphrodites, a high dosage of the X chromosome genes *sex-1* and *fox-1* triggers a cascade of negative interactions [36, 184] that ultimately represses the expression of the *fem-1/2/3* genes [100, 111, 165]. Levels of the TRA-1 transcription factor are raised as a result [111, 218, 230], and this causes female somatic development while preventing spermatogenesis [21, 40, 72, 113, 127]. In larvae, FOG-2 acts together with the mRNA binding protein GLD-1

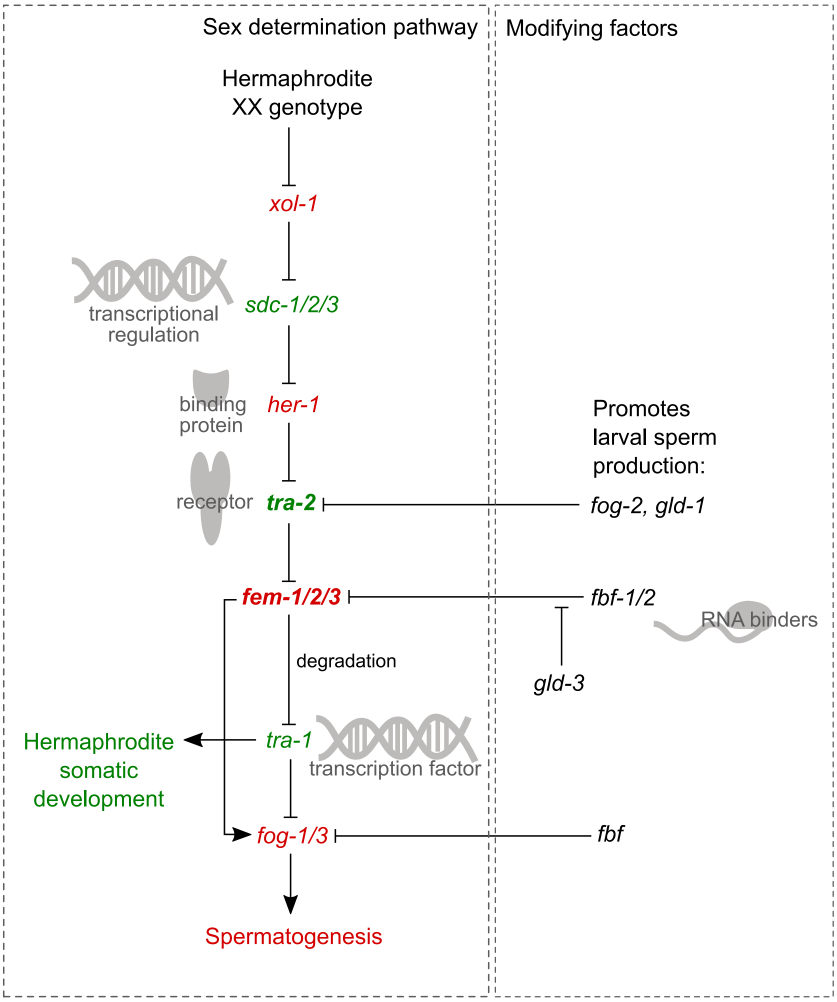


Figure 2.12: The molecular regulation of sex determination. Left: the cascade of negative interactions leading from an XX genotype to hermaphrodite development and suppression of spermatogenesis. Genes/processes that are inhibited in hermaphrodites are coloured red, while genes/processes that are activated are coloured green. Right: some of the genes that allow temporary sperm production to occur in the larva, by altering the balance between *tra-2* and *fem-3*.

to allow a transient period of sperm production to occur [48, 58, 122]. Several other regulators are also involved in hermaphrodite spermatogenesis. However, in Figure 2.12 we focus particularly on those molecules that are also involved in proliferation control, and which therefore create an interesting link between the two major germ cell decisions.

As with meiotic entry, a key question is when and where sex determination takes place. The fact that various stimuli can alter germ cell sex has been put forward as evidence that at least some adult germ cells are sexually labile [175, 189, 210]. Barton and Kimble [21] have studied the timing of sex determination using a temperature sensitive mutation in the *fog-1* gene (Figure 2.12), which was triggered at different time points during larval development. They found that *fog-1* is required specifically between 22 and 32 hph for spermatogenesis to occur. Since the beginning of this period coincides with first meiosis in the germ line, it was proposed that germ cells commit to a sexual fate when they enter meiosis [21]. However, more recent work has challenged this view, highlighting subtle differences between the proliferative zones of male and hermaphrodite animals: different cell cycle lengths; different rates of cell movement; and different levels of sex related proteins [176, 230, 247]. These differences suggest that sex specific traits are present even among the distal-most proliferative cells. As such, the exact timing of sex determination remains unclear.

So far, we have discussed how a proliferative germ cell enters meiosis and becomes committed to an oocyte fate. However, not all germ cells develop into mature oocytes; a large number undergo apoptosis in the gonad turn. The next section looks at germline apoptosis and its possible role in adult *C. elegans*.

2.7 Germline apoptosis

Cell death occurs frequently in the adult hermaphrodite: between one and five cell corpses can be found in the germ line at any given time [95]. Dying germ cells first separate from the rachis, then undergo apoptosis, before finally being engulfed and removed by the sheath cells that surround the gonad [95]. Apoptosis can only occur in, and just prior to, the gonad turn; a restriction that is thought to be enforced by regulatory proteins displaying spatial concentration gradients (for example, GLD-1) [104, 215]. Apoptosis also does not occur in the larval or the male germ line, implying that only oocyte-fated germ cells undergo programmed cell death [89, 95, 124].

There appear to be at least two different types of germline apoptosis. One type is triggered in response to DNA damage or the failure of a meiosis checkpoint [89, 124], while the other occurs at a high rate even in healthy individuals. Since the second type of apoptosis is common in healthy *C. elegans*, it is assumed to play a role in normal germ line function and is therefore known as *physiological apoptosis* [95]. Both processes share a core pathway involving the release of the protein CED-4 from complex with CED-9 (see Figure 2.13) [89, 95]. However, apoptosis in response to damage can be blocked by a *ced-9* gain of function mutation, while physiological apoptosis continues [89, 95]. Damage-induced apoptosis also requires EGL-1 and CED-13; physiological apoptosis does not [95, 216]. As such, there are clear molecular differences between the two processes.

While the advantage of removing genetically damaged cells is obvious, there are a number of possible explanations for physiological apoptosis. Dying germ cells may act as nurse cells, donating their cytoplasm to developing oocytes [95]. Alternatively, apoptosis in the turn may serve to cull excess cells after they have contributed proteins to the rachis, since insufficient resources are available to turn them all into mature oocytes [19]. These hypotheses are consistent with the fact that physiological apoptosis only occurs in oogenic germ lines; they are also supported by the finding that mutants lacking apoptosis produce smaller, less viable oocytes [11]. A final possibility is that germ cell death may be required for tissue homeostasis. A fixed germ cell count could be maintained in the gonad either by varying the cell death rate, modulating proliferation, or altering the laying rate. However, *ced-3* mutants (which cannot undergo physiological apoptosis) do not contain an abnormal number of cells, and their laying rates are largely unchanged [223].

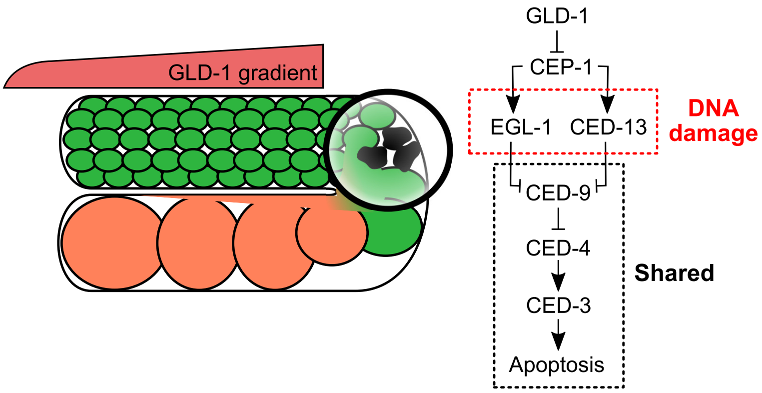


Figure 2.13: Types of apoptosis in the germ line. Germ cell apoptosis comes in two forms: one is triggered by DNA damage, the other (physiological apoptosis) occurs in healthy animals. Both apoptosis types share a pathway involving CED-9, CED-4 and CED-3. However, only damage-induced apoptosis requires EGL-1 and CED-13, and only it can be disabled by *ced-9* gain of function.

As such, variation in the proliferation rate may be a more important control mechanism for tissue homeostasis [19].

We have now covered the processes involved in normal germ line function. In the next section, we look at some key perturbations and how they disrupt the germ line.

2.8 Disruptions to normal germ line function

Here, we describe ways in which normal *C. elegans* germ line behaviour can be disrupted. We begin by looking at natural triggers of germline change (starvation and ageing), before discussing some relevant experimental perturbations and what was learned from them.

Starvation

Starvation is perhaps the most studied disruptor of germ line function. Moreover, dietary restriction is of great interest to *C. elegans* researchers generally, because of its effects on the ageing process [146, 259]. Therefore, while this topic is not central to the thesis, we give a brief overview of the germline response to food scarcity.

The effect of starvation on the germ line depends strongly on the age of the worm. In general, when food is scarce, animals invest fewer resources in reproduction to focus on survival until conditions improve. Young larvae do this by temporarily arresting their development (L1 larvae undergo L1 diapause [129], while L2 larvae adopt a hardy *dauer* form [37]). Germ cells in these animals remain arrested in G2 phase until refeeding occurs [183]. Older L3 worms cannot initiate diapause, so they instead scale back the expansion of their germ line, resulting in fewer eventual offspring [55, 147, 168]. Adult animals, meanwhile, respond to starvation by “bagging”: eggs are retained inside the uterus and hatch there [219]. Bagging kills the mother, but provides protection and food for the progeny. Finally, starving young adult animals that do not yet contain viable embryos can avoid bagging and instead undergo Adult Reproductive Diapause (ARD) [12, 219]. During ARD, the majority of the germ line dies, with only 30-35 cells preserved [12]. Oocyte production continues at a greatly reduced rate, and the germ line can be repopulated by the surviving cells on refeeding [12, 219].

Ageing

Like many other organs, the germ line deteriorates with age. Perhaps the most obvious change in its function occurs 5 days after L4, when the supply of self-sperm becomes depleted [117]. In the absence of sperm, oocytes fail to mature and begin to stack up inside the proximal arm and uterus [11, 152, 162]. Sperm count is therefore the limiting factor governing the reproductive lifespan of a hermaphrodite. However, it is not the only bound on fertility. Mated individuals have sufficient sperm but still stop reproducing around 10 days post-L4 [117]. Oocyte quality seems to be the other important factor in reproductive ageing. Older animals produce smaller oocytes that lack the resources for embryonic development [11, 152]. They also produce more male offspring and contain abnormal looking germ nuclei, suggesting an increase in chromosome disjunction errors [152, 203].

Other age related changes occur at the opposite end of the germ line, in the proliferative zone. Notably, the total number of proliferative germ cells declines after young adulthood, falling from around 200 cells to 50 over a period of 12 days (see Figure 2.14A) [88, 197]. Interestingly, the extent of proliferative cell loss with age can be modified. Qin and Hubbard [197] recorded the proliferative cell count of 12 day old animals with a range of genetic mutations. They found that a reduction of function mutation in the *daf-2* insulin receptor gene protects against proliferative cell loss (Figure 2.14B). The effect appears to be mediated by the transcription factor DAF-16, indicating involvement of canonical insulin signalling.

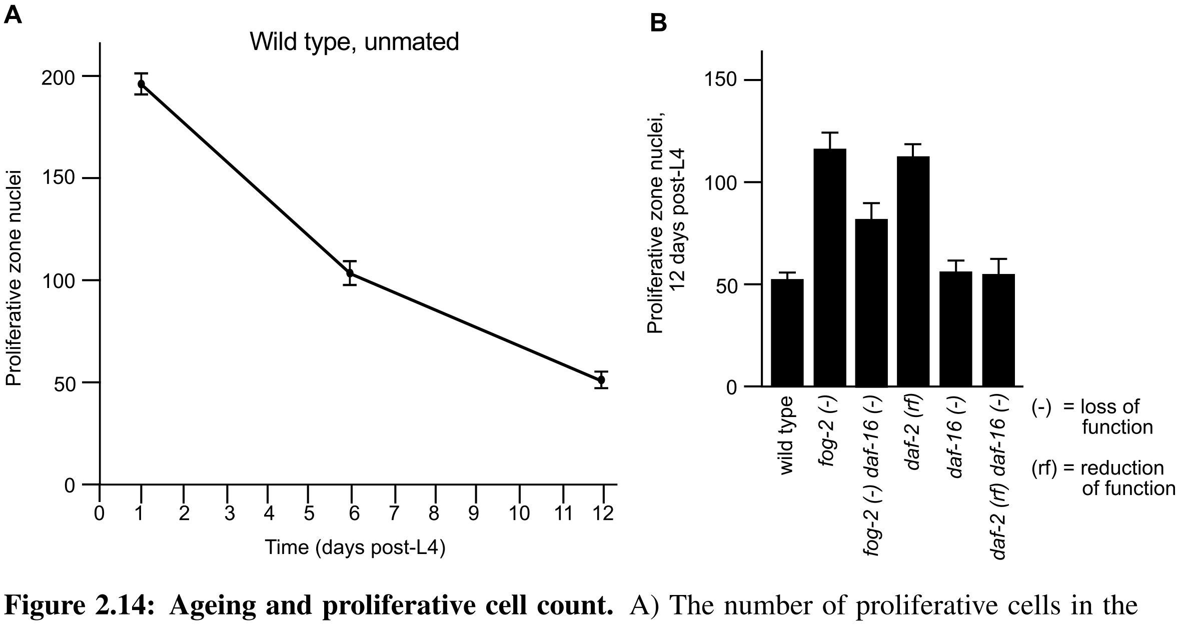


Figure 2.14: Ageing and proliferative cell count. A) The number of proliferative cells in the wild-type germ line declines with age, roughly halving every 6 days. B) Various mutations alter the extent of proliferative cell loss. *fog-2(-)* mutants, which have no sperm and exhibit slower germ cell flux through the gonad, show improved proliferative cell maintenance. The effect is partially mediated by the insulin pathway protein DAF-16. More generally, other genes involved in insulin signalling affect proliferative cell loss, such as *daf-2*. Figures adapted with permission from [197].

Moreover, a number of different mutations that slow down germ cell movement through the gonad also protect against proliferative cell loss. For example, *fog-2* mutants have no sperm and a low ovulation rate, resulting in slow germ cell progress through the gonad. *fog-2* mutants maintain their proliferative cell count roughly twice as well as wild-type animals (Figure 2.14B) [197]. Taken together, the results in [197] point to a positive relationship between high germ cell flux and proliferative cell loss with age. The mechanism underlying this relationship is not fully understood, although Qin and Hubbard report that it is partly mediated by *daf-16* [197]. In Chapter 6, we will return briefly to ageing effects and consider whether a particular proliferative zone maintenance hypothesis can reproduce the observed cell count decline.

Disrupting the DTC proliferation signal

As we saw in Section 2.4, the DTC is crucial for continued germ cell proliferation. Here, we describe two important experimental studies, both of which used a temperature sensitive mutation to disrupt DTC signalling in an adult worm. The results have implications for how the proliferative zone is maintained on a cellular level.

In 2010, Cinquin *et al.* used a temperature sensitive mutation in the gene *glp-1* to disable the receptor that senses DTC ligands [47]. Since DTC signalling is necessary for proliferative zone maintenance, the expected outcome of this mutation is for all germ cells to eventually stop proliferating and enter meiosis. Cinquin *et al.* triggered the mutation by raising temperature, waited several hours, then recorded which germ cell rows contained meiotic cells. By varying the waiting period, they were able to build up a picture of how the proliferative zone enters meiosis over time. In this investigation, meiotic germ cells were identified by their crescent-shaped nuclei.

Cinquin *et al.* observed that after DTC signalling disruption, germ cells enter meiosis in a proximal to distal wave (see Figure 2.15A). In the first 5 hours after temperature upshift, they saw the end of the proliferative zone retreat from 13CD to 5CD. The five distal-most cell rows were then reported to enter meiosis near-simultaneously over the course of a further hour [47]. These findings were interpreted as evidence for a distinct stem cell population occupying the first five rows of the gonad [47]. The distal-most germ cells were proposed to be immature, leading them to enter meiosis last, and to share a common stem identity, causing them to enter meiosis simultaneously. In addition, the proximal to distal wave of differentiation was attributed to the presence of transit amplifying cells [47]. Proximal transit amplifying cells would have undergone more divisions than distal ones, and would therefore enter meiosis earlier.

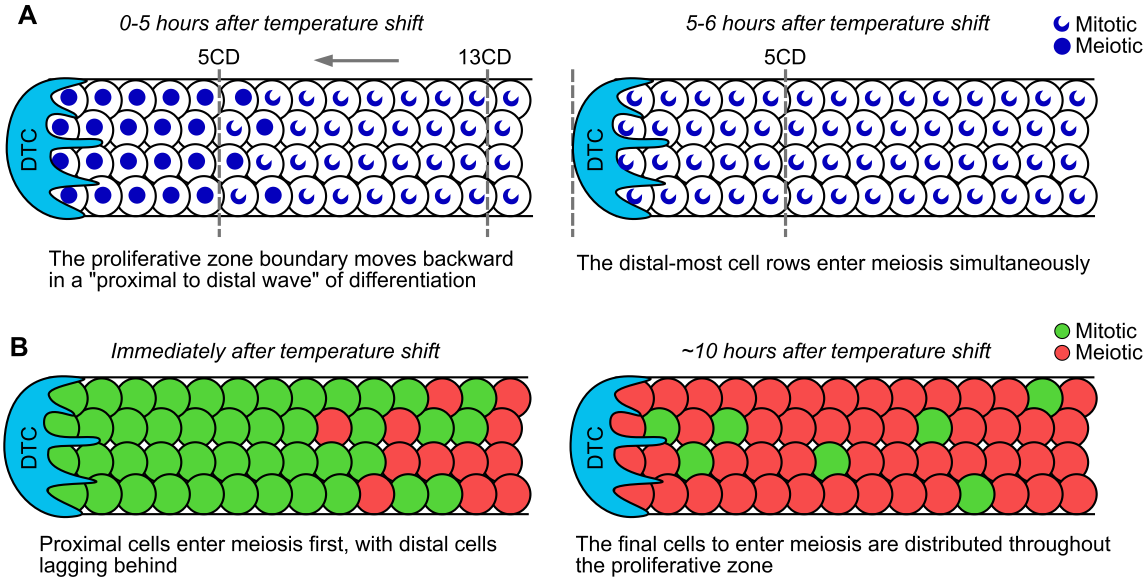


Figure 2.15: Meiotic entry following DTC signal disruption. A) The results reported by Cinquin *et al.* [47]. For the first 5 hours, germ cells enter meiosis in a proximal to distal wave. Then, over the next hour, the remaining cell rows enter meiosis almost simultaneously. B) The results reported by Fox and Schedl [81]. Initially, meiosis events are confined to the proximal part of the proliferative zone. However, 10 hours later, the last cells to enter meiosis are distributed throughout the zone, not concentrated near the distal tip.

More recently, in 2015, Fox and Schedl performed a very similar experiment to the one described above [81]. They again used a temperature sensitive *glp-1* mutation to disrupt DTC signalling, and studied the manner in which the proliferative zone entered meiosis. However, in this newer investigation, antibody staining was used to distinguish between proliferative and meiotic cells. Antibodies against the chromatid cohesion proteins REC-8 and HIM-3 [191, 277] can act as markers under suitable fixation conditions [102], with REC-8 staining proliferative cells and HIM-3 meiotic ones. Fox and Schedl claim that this technique is more sensitive than assessing nuclear appearance, and that it can detect meiosis at an earlier stage [81].

Following temperature upshift, Fox and Schedl observed a spatial trend in meiotic entry. They reported seeing proximal germ cells enter meiosis first, with distal germ cells lagging behind. However, they did not describe the pattern as constituting a wave of differentiation. Moreover, Fox and Schedl reported that the final germ cells to enter meiosis are distributed throughout the proliferative zone, not concentrated near the distal tip (Figure 2.15B). Based on this finding, along with experiments that involved labelling for specific cell cycle phases, Fox and Schedl concluded that the distal-most germ cells are *not* functionally distinct, but acquire proliferative capacity from the DTC. They also asserted that when the DTC signal is removed, germ cells simply complete their current cell cycle then enter meiosis, without intervening transit amplifying divisions.

Prior to 2015, germline stem cells were frequently mentioned in the literature [13, 114, 260] and the experiments of Cinquin *et al.* appeared to confirm the existence of distinct, immature germ cells. Fox and Schedl's work presents a new and significant challenge to that viewpoint, which we consider further in Chapter 6.

2.9 Open questions

Despite all that is known about germ cell behaviour, many open questions remain regarding germ line development and maintenance. Often, these questions concern the tissue or whole organ scale; we wish to understand how the behaviour of a germ cell population is coordinated, in the context of the gonad, to ensure correct function.

For example, there is evidence to suggest that gonad growth influences the developing germ line, while germ cells, in turn, influence gonadogenesis. Underproliferation mutants have a smaller adult gonad because they lack the pressure of germ cell divisions to drive the DTC forward at a normal rate. *Whether this is the full extent of mechanical interaction between the germ line and the gonad is unknown*, and the topic is difficult to study experimentally since germ line and gonad development are inextricably linked *in vivo*. Certainly it is essential that gonad growth and germ line expansion occur at well matched rates; otherwise inappropriate contacts can arise between germ cells and the somatic gonad, leading to tumour formation [164]. An *in silico* model would allow different aspects of gonadogenesis and germ cell proliferation to be investigated in isolation, making it easier to study their separate contributions.

Another question that remains controversial is *whether a distinct population of germline stem cells exists, or whether all germ cells are functionally equivalent.* Apart from the experiments of Cinquin *et al.* described in Section 2.8 [47], other evidence in favour of germline stem cells includes the preservation of a small cell population during ARD [12] and the slower cell cycle of the distal-most cell rows [154], a property frequently associated with stemness. Nevertheless, all of these observations could equally well be explained by functionally equivalent germ cells, provided the distal-most cells acquire special properties as a result of close contact with the DTC. This second perspective is currently gaining ground, in light of the work of Fox and Schedl (2015) [81], as well as the discovery of processes and gap junctions linking the DTC to distal germ cells [35, 229]. Unfortunately, the experimental techniques needed to detect a niche-bound stem cell population (such as individual cell tracking, clonal labelling and long term imaging) remain unavailable in the *C. elegans* germ line. However, in a computational model these aspects of the system can be much more easily explored. By simulating a variety of different maintenance mechanisms, it should be possible to compare the resulting whole germ line behaviour with existing measurements, and perhaps find subtle differences between these hypotheses.

Finally, there is the question of *how germ cell numbers are homeostatically maintained in the young adult gonad*. While proliferative cell count does decline in older animals, it is stably maintained for a day or so in wild-type *C. elegans*, and for even longer in individuals with certain protective mutations. Moreover, when the rate of germ cell removal is reduced (either by disabling apoptosis or preventing ovulation) significant overcrowding of the gonad does not occur. This suggests that the proliferation rate can be adjusted to compensate for slower germ cell removal, but the exact mechanism and whether it is biochemical or mechanical in nature is unknown. A model incorporating germ cell mechanics as well as intracellular signalling should be capable of studying the effect of mechanical feedback on proliferation rates.

2.10 Summary

In this chapter, we reviewed the biology of the *C. elegans* hermaphrodite germ line, covering its larval development, adult maintenance, and perturbations to germ line function. We also introduced several open questions regarding the coordination of germ cell behaviour. We explore these questions further in later chapters; Chapter 4 comments on the relationship between germ line and gonad development and on germ line homeostasis, while Chapter 6 looks at potential mechanisms of proliferative zone maintenance.

A realistic, predictive model of the germ line has the potential to: 1) enable different aspects of germ line development to be studied in isolation; 2) generate predictions for experiments that are currently infeasible; and 3) identify subtle measurable differences between complex hypotheses. We therefore set out to construct such a model. Since many interesting germline properties ultimately depend on germ cell movement and packing, it is important that our model incorporates reasonable cell movement as well as individual cell behaviour. The next chapter therefore discusses existing approaches for modelling an interacting cell population, and for representing intracellular decision-making.

Chapter 3

Review of existing techniques, tools and models

This review chapter is divided into three main parts. In the first, we introduce a variety of approaches for modelling an interacting population of cells. We also discuss available software tools for modelling cells and tissues, with a particular emphasis on Chaste: an open source C++ library used throughout the rest of this thesis [171]. In the second part, we cover techniques for modelling intracellular decision-making, focussing on discrete approaches and particularly statechart models, which are used in Chapter 4. Finally, in the third part, we review pre-existing models of the *C. elegans* germ line: the techniques they employed, the insights they provided, and possible areas for improvement.

3.1 Representing a population of cells

In this section, we consider different techniques for modelling a population of cells. We examine the choices a modeller can make, beginning with high-level decisions and becoming increasingly specific. For each approach mentioned, we briefly consider its advantages and disadvantages in the context of *C. elegans* germ line modelling. Finally, we discuss existing software tools for cell-based work, focussing on the Chaste library [171].

3.1.1 Continuum vs. agent based models

At a high level, there are two different approaches to modelling a tissue or population of cells. The first, the continuum approach, does not consider individual cells but rather deals with smoothly varying quantities, such as cell density. The change in cell density over time can be modelled using a differential equation, either an Ordinary Differential Equation (ODE), or a Partial Differential Equation (PDE) if the precise spatial distribution of cells is of interest. A classic example is the Keller-Segel model of chemotaxis, in which two equations govern the time evolution of cell density (n) and chemoattractant concentration (v) (Figure 3.1A) [136]. Cells are assumed to diffuse at a rate k_1, move up a chemoattractant gradient at rate k_2, and divide at rate k_3. Meanwhile, the chemoattractant diffuses, is consumed by cells at rate k_4, and is replenished at rate k_5. This gives the following pair of equations

$$\begin{aligned}\frac{\partial n}{dt} &= \nabla \cdot (k_1 \nabla n - k_2 n \nabla v) + k_3(n,v),\\ \frac{\partial v}{dt} &= \nabla^2 v - k_4(n,v)n + k_5(n,v).\end{aligned}$$

The above model was used to explore the conditions under which a homogeneous distribution of amoebae becomes unstable and forms aggregates. Similar continuum approaches have since been used to model tumour growth and wound healing [202, 225]. An advantage of such PDE models is that they are amenable to analysis using well-established mathematical techniques, which may provide an exact solution in some limit, or allow results to be proven about system behaviour.

Continuum approaches can also be used to model the deformation of a tissue due to applied force. Given certain mechanical properties of the tissue, a constitutive equation can be formulated describing the relationship between stress and strain [97]. Ultimately, PDEs for displacement and pressure fields can then be solved over the tissue using, for example, a finite element method. These continuum mechanics models are appropriate whenever the behaviour of individual cells is not of interest; for instance, in cancer modelling to represent the healthy tissue surrounding a tumour, or in embryonic modelling to study the forces needed to induce morphological change [41, 140, 187].

The alternative to continuum approaches is agent based modelling, in which individual cells are explicitly represented (reviewed in [8], see Figures 3.1B and C). Each cell has a position that updates over time, and cells may also have an internal state or program determining their actions.

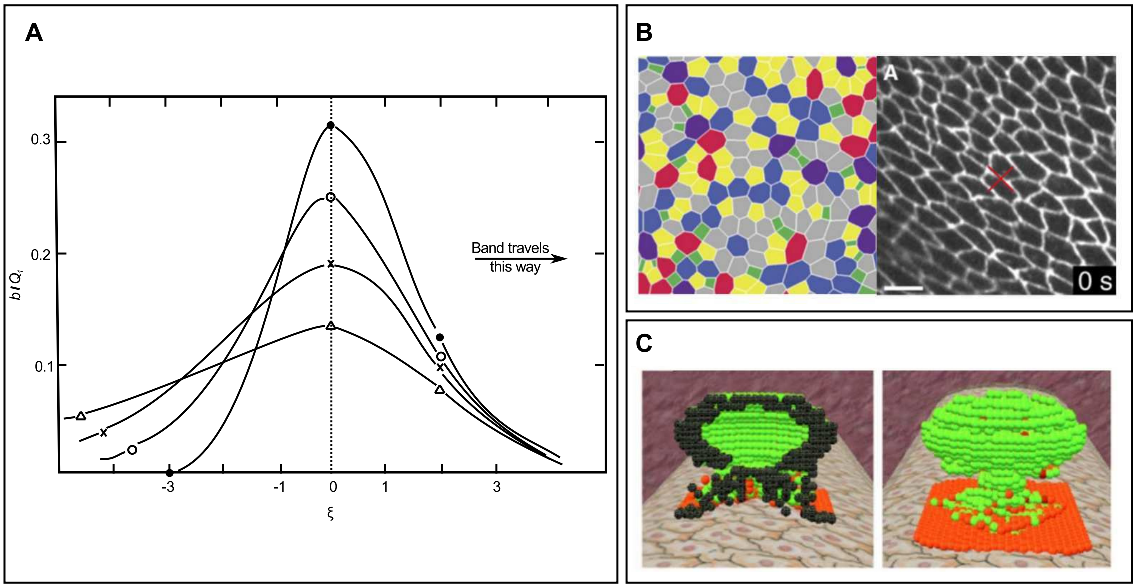


Figure 3.1: Examples of continuous and agent based models. A) Evolution of the cell density profile over time in a Keller-Segel model of chemotaxis. B) Left: cells in an agent based computer model of the *Drosophila* wing, colour coded according to number of sides. Right: an actual section of *Drosophila* wing with cell membranes fluorescently labelled, for comparison. The red cross indicates a target for laser cell ablation. C) A 3D agent based model of pancreatic development, in which individual cells move on a lattice according to certain rules. Figures reproduced with permission from: Keller and Segel (1971) [136]; Farhadifar *et al.* (2007) [73]; and Setty *et al.* (2008) [220] [Image copyright (2008) National Academy of Sciences], respectively.

Agent based models are suited for the investigation of cell-cell signalling, as well as the emergent properties that arise from individual cell behaviour. Tissue growth can be easily represented by adding more agents, and simulation output can be directly compared with single cell experimental data, such as fluorescence microscopy images (Figure 3.1B). As a result, agent based models are frequently used in developmental biology to study the mechanisms of growth and patterning; for example, in the *Drosophila* wing disc [80], colonic crypt formation [34], and *C. elegans* vulval development [185].

Moving forward, we concentrate on agent based models, since many interesting questions about the *C. elegans* germ line concern individual cell behaviours, such as meiotic entry, cell cycle progression and fate decisions. As discussed in Section 2.3, most germ cells are technically syncytial, but they are considered in the field to be separate entities, meaning that agent based techniques are appropriate for representing the germ line.

3.1.2 On-lattice vs. off-lattice models

Within agent based modelling there are two possible approaches to represent cell positions: on-lattice and off-lattice. In the on-lattice approach, space is divided up into a discrete grid of lattice sites. A common type of on-lattice model is the cellular automaton [256], in which each cell occupies a single lattice site and moves to a new site at each time step according to a set of update rules (Figure 3.2A). Position update rules typically take into account the number and type of neighbouring cells, and can also depend on other information associated with lattice sites, such as nutrient or signalling factor concentrations. Hybrid models often use PDE or Ordinary Differential Equation (ODE) systems to update a separate grid containing chemical concentrations [201]. Position update rules can also be stochastic, causing cells to move according to a biased random walk (see [221] and space-jump models [188, 248]). Cellular automata have been used to model tissue maintenance by stem cells in a variety of biological contexts [221, 227].

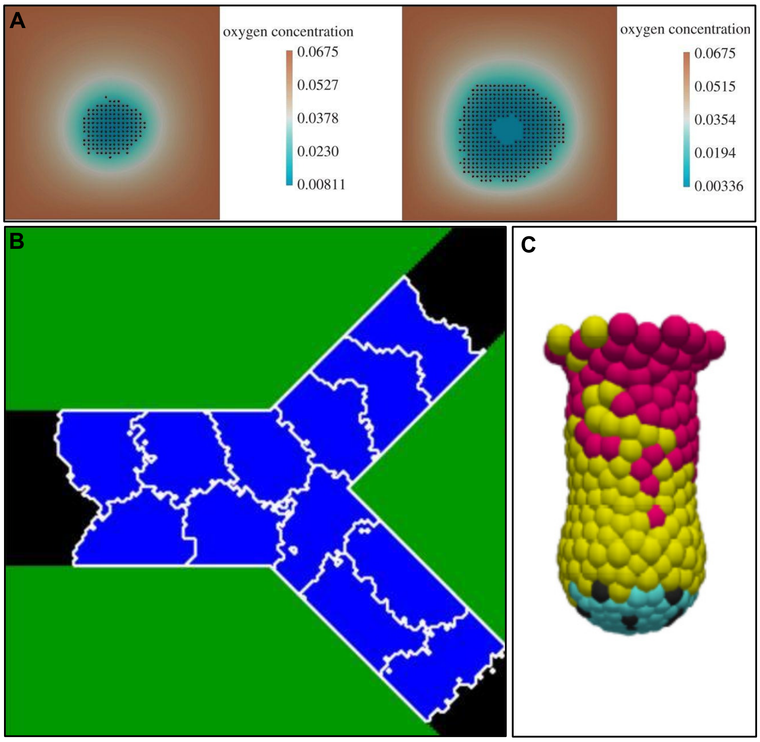


Figure 3.2: Examples of on-lattice and off-lattice models. A) A hybrid cellular automaton model of tumour growth. Cells occupy lattice sites in a discrete grid, and proliferate at a rate dependant on the local oxygen concentration. B) A cellular Potts model of lumen formation in a branched blood vessel. Cells are shown in blue with their borders highlighted in white. As the simulation progresses, cell-cell repulsion forces produce changes in cell shape. C) An off-lattice, overlapping spheres model of the colonic crypt. Figures reproduced with permission from: Figueredo *et al.* (2013) [74]; Boas and Merks (2014) [30]; and Dunn *et al.* (2016) [68], respectively.

A different on-lattice technique is the cellular Potts model (Figure 3.2B). Potts models were originally developed for problems in statistical physics, where they were used to simulate the changes in alignment of neighbouring lattice spins with temperature [195]. However, they have since been used in biology to study cell sorting and tumour invasion [93, 250]. In a cellular Potts model, each cell is represented by a collection of lattice sites sharing a common identity (ID). Cellular shape changes can therefore be captured as well as movement by allowing individual site IDs to update at each time step. The form of the update rule is also different: a Hamiltonian function is defined that allows the energy of the entire system to be calculated. Randomly generated updates are then accepted with a probability that depends on the associated energy change (decreases in energy being more likely).

The main drawback of on-lattice models is that the presence of an underlying grid may lead to artefacts in cell movement, such as the instantaneous shifting of an entire cell row to make room for a new daughter cell [166]. Moreover, cellular automaton update rules need not have any physical basis and can introduce unrealistic effects. For example, using a biased random walk to update positions makes an assumption about the dominant direction of movement that could be incorrect.

The alternative is to take an off-lattice approach, allowing cells to move freely to any point in the domain rather than only between lattice sites (Figure 3.2C). Typically, a force law is chosen to describe the interaction between any pair of cells. The net force on a cell can then be calculated by summing over all its interactions with its neighbours (see Section 3.1.3), then adding any frictional forces due to the surrounding medium. Cell positions may then be updated in accordance with Newton's second law of motion $F = ma$. The most general position update equation has the form:

$$\sum_j \mathbf{F}_{ij} - \eta_i \frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} + K\sqrt{\mathrm{d}t}\, r = m_i \frac{\mathrm{d}^2\mathbf{r}_i}{\mathrm{d}t^2}, \tag{3.1}$$

where $\mathbf{F}_{ij}$ is the force on cell i due to cell j, $\mathbf{r}_i$ is the location of cell i, m_i is its mass and η_i is its drag coefficient. $K\sqrt{\mathrm{d}t}\, r$ is an optional stochastic term based on Brownian motion, where r is Gaussian distributed noise and K is proportional to the Boltzmann constant. Frequently, an overdamping assumption is made, meaning that drag dominates over acceleration. The deterministic form of the update equation can then be further simplified to

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \frac{1}{\eta_i} \sum_j \mathbf{F}_{ij}. \tag{3.2}$$

Off-lattice models have been successfully applied to studying tissue homeostasis [67], epithelial layers [86] and tumour growth [208].

In this thesis we focus on off-lattice models. The primary motivation is to avoid artefacts in cell movement due either to the lattice or unphysical update rules. While cellular Potts models do use update rules with a physical inspiration, they do not capture mechanical interactions. Moreover, cellular Potts models have no intrinsic time scale, only generations, which makes it difficult to couple them to an intracellular model involving time delays. Therefore, we will now go on to discuss some important modelling decisions associated with off-lattice approaches: choice of cell representation, and choice of force law.

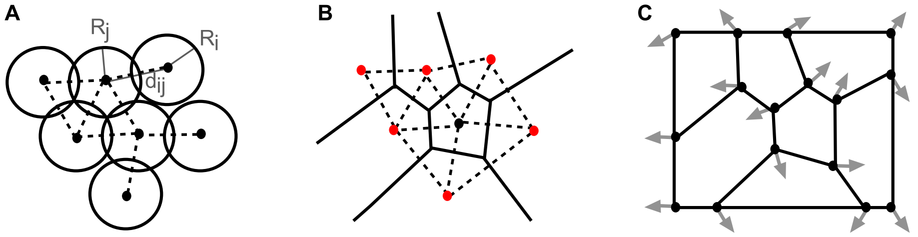
3.1.3 Cell centre vs. vertex-based models

There are at least three different ways to define cells and cell-cell contacts in an off-lattice model. The simplest is to represent cells by circles or spheres of some radius (Figure 3.3A). This method provides a natural cutoff distance beyond which cells are too far apart to influence each other. A pair of cells can interact only if they overlap, i.e. if

$$d_{ij} < R_i + R_j, \tag{3.3}$$

where R_i and R_j are cell radii, and d_{ij} is the distance between the cell centres. Overlapping sphere models have been used to study tissue renewal in the colonic crypt, and to model epithelial populations more generally [33, 67, 86].

A slightly different approach also uses a collection of centre points to define cells, but cell boundaries are determined by a Voronoi tessellation (see Figure 3.3B; [65, 166, 174]). Given a set of points, a Voronoi tessellation divides space into regions that are closer to a particular point than any other. As a result, cells appear realistically packed in visualisations and do not interpenetrate (cf. Figures 3.3A and B). In a Voronoi model, cells exert force on each other only if their associated points are joined by an edge in the Delaunay triangulation (the dual mesh to the Voronoi tessellation; see dashed lines, Figure 3.3B). This method of defining cell-cell connectivity can produce different behaviour from an overlapping spheres model under certain circumstances. For example, during large compression deformations different cell rearrangements occur, resulting in more variable stress in the overlapping spheres case [192]. An important implementation point to



Overlapping spheres

Voronoi/Delaunay tessellation

Vertex-based

Figure 3.3: Off-lattice definitions of cells and connectivities. A) A 2D overlapping spheres model, in which cells are represented by circles that exert force on each other when they overlap. B) A Voronoi tessellation model. Cells are defined by Voronoi regions, which are closer to a particular cell point than any other. Forces act along the edges of the dual Delaunay triangulation (dashed lines). Here, red points are associated with unbounded “ghost” cells. C) A vertex-based model, in which cells are polygons defined by the positions of their boundary vertices. Forces act on these vertices directly.

consider is that the boundary cells in a Voronoi tessellation extend to infinity. To combat this, a layer of “ghost” cells is usually added surrounding the population of interest (Figure 3.3B, red). Ghost cells can then be excluded from visualisations and calculations.

A third approach is to use a vertex-based model, which defines cells as polygons and explicitly specifies the vertices making up each cell boundary. Vertex positions are then updated directly (Figure 3.3C). Positions can be updated based on the net force each vertex experiences (a function of properties such as cell area and perimeter). Alternatively, vertex positions can be chosen to minimise some energy function over the entire system [73, 181, 264]. Additional rules are required to determine when a large cell should be divided, and to decide how small cells and element overlaps should be resolved. Vertex-based models are commonly used to study epithelial layers, because the cell shapes, junctions and rearrangements in the model can be usefully compared to those in microscopy images (Figure 3.1B) [73, 199].

We will use a cell centre model to represent the *C. elegans* germ line, for the following reasons. Vertex-based models are computationally expensive and difficult to implement in three dimensions. As we are not interested in the precise shape or intercalation of germ cells, the extra detail captured by a vertex-based model is unnecessary. With regard to cell-cell connectivity, previous work has shown that tissues composed of overlapping spheres behave similarly to those composed of Voronoi cells under many circumstances, including the application of elastic compression, tension and shear forces [192]. As a result, the potential benefits of using a Voronoi tessellation do not justify the computational effort required to produce the tessellation in three dimensions. As such, we focus on cell centre models with spherical cells from this point on.

3.1.4 Force laws

Cell centre models require a force law to describe the interaction between any pair of cells i and j. Force acts along the vector $\mathbf{r}_{ij}$, directed from the centre of cell i to the centre of j. The force experienced by cell i can therefore be written as

$$\mathbf{F}_{ij} = f_{ij}\hat{\mathbf{r}}_{ij}, \tag{3.4}$$

where $\hat{\mathbf{r}}_{ij} = \mathbf{r}_{ij}/|\mathbf{r}_{ij}|$. Typically, f_{ij} is a function of cell separation or contact area that includes parameters representing mechanical properties. This section gives a general overview of the types of force law that exist. The issue of choosing a force law for our *C. elegans* germ line model is discussed in Sections 3.1.6, 4.1.1 and 5.1.

Simple repulsion forces

It is reasonable to expect that a pair of overlapping cells should repel each other due to elastic forces. The simplest choice of f_{ij} that achieves this is a linear function. If v_{ij} is the length of the cell-cell overlap ($v_{ij} = R_i + R_j - d_{ij}$) then

$$f_{ij} = -kv_{ij}, \tag{3.5}$$

produces a repulsion force. As the amount of cell-cell overlap increases, the magnitude of the repulsion grows up to a maximum of $k(R_i + R_j)$. Often, a cutoff distance is imposed beyond which cells cannot influence each other; a restriction that is both realistic and reduces computational cost. A natural choice of cutoff is $R_i + R_j$, meaning that only overlapping cells interact. Figure 3.4A shows a linear repulsion force with this cutoff applied.

An alternative way to introduce repulsion is to scale force strength according to the contact area, A_{ij}, between cells, giving

$$f_{ij} = -kA_{ij}. \tag{3.6}$$

Geometric arguments can be used to find A_{ij} in terms of the cell radii and separation:

$$A_{ij} = A_{ji} = \pi \left[R_i^2 - \frac{(R_i^2 - R_j^2 + d_{ij}^2)^2}{4d_{ij}^2} \right]. \tag{3.7}$$

Figure 3.4B illustrates the difference this alternative scaling makes.

Hard core potentials

The force laws described above enable a pair of cells to overlap completely, which is not always desirable. A maximum overlap, v^*, can be enforced by allowing the repulsion strength to grow

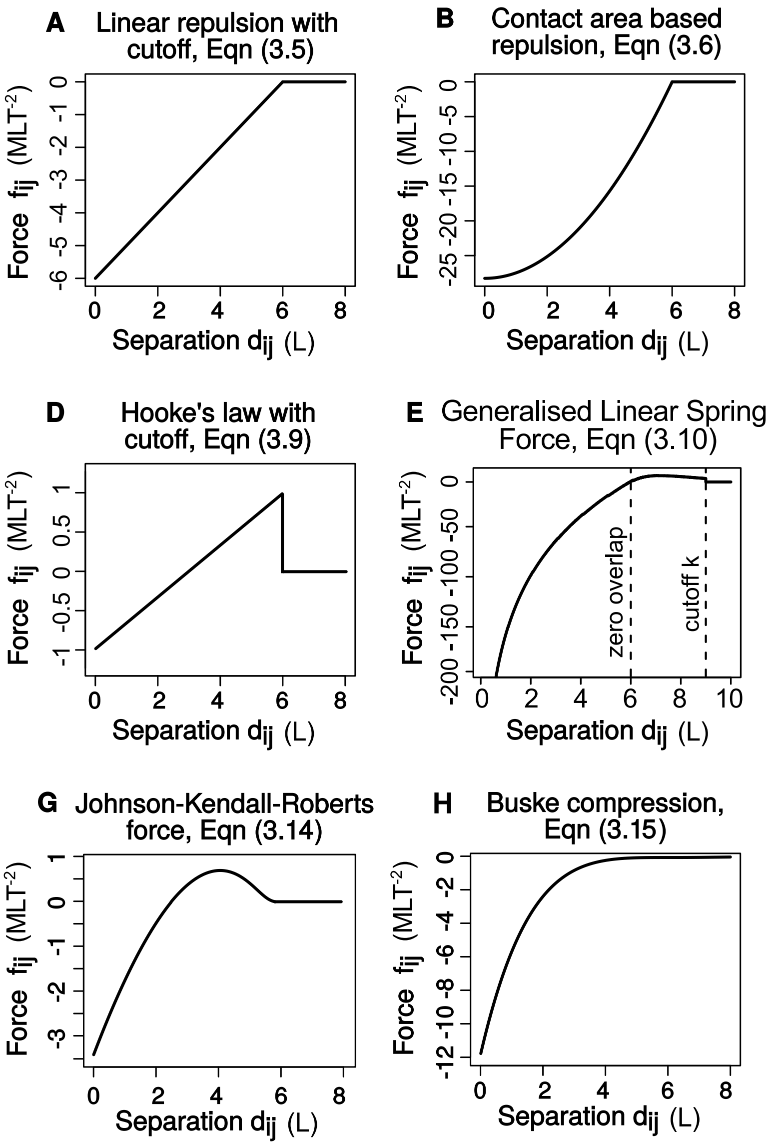


Figure 3.4: Force as a function of cell-cell separation for a number of force laws. Forces were calculated for two cells of radius 3 length units (L), using the following parameter values. In A) $k = 1(\mathrm{MT}^{-2})$ and in B) $k = 1(\mathrm{MT}^{-2}\mathrm{L}^{-1})$, where M = mass units and T = time units. In C) the exponential function has $\alpha = 10$, $v^* = 3(\mathrm{L})$ and $k = 1(\mathrm{MT}^{-2})$. In D) the spring has rest length $d_{rest} = 3(\mathrm{L})$ and strength coefficient $\mu = 1(\mathrm{MLT}^{-2})$. In E) $\alpha = 5$, $k = 9(\mathrm{L})$ and $\mu = 15(\mathrm{MT}^{-2})$. In F), G) and H) the Young's modulus is taken to be $E = 1(\mathrm{MT}^{-2}\mathrm{L}^{-1})$ and the Poisson ratio is $\nu = 1/3$. In G) $\sigma = 0.1$ and in H) $K = 1(\mathrm{ML}^2\mathrm{T}^{-2})$.

exponentially for overlaps $v_{ij} > v^*$. For example,

$$f_{ij} = -kv^* \exp\left[\alpha\left(\frac{v_{ij}}{v^*} - 1\right)\right], \tag{3.8}$$

where α is a constant (see Figure 3.4C). This is known as a hard core potential, since cells behave as if they have an incompressible core. The force needed to make a cell pair overlap by much more than v^* is so great it is essentially never encountered, particularly when α is large.

Hooke's law

Force laws can also take into account cell-cell adhesion, which opposes repulsion and may cause cells to have a non-zero steady state overlap. A simple way to incorporate adhesion is to treat cell centres as if they were joined by springs. The rest length of the spring is the preferred cell-cell separation, and cells further apart than this distance will move toward each other. Hooke's law gives

$$f_{ij} = \mu\left(\frac{d_{ij} - d_{rest}}{d_{rest}}\right), \tag{3.9}$$

where μ is a spring strength constant and d_{rest} is the rest length. A cutoff distance must be imposed in this case to limit the range of adhesion (see Figure 3.4D).

Generalised spring force laws

Using Hooke's law can lead to unstable cell-based simulations. Under Hooke's law the repulsion force between cells is relatively weak even at large overlaps. Cells can therefore overlap to the extent that they experience undesirable and unrealistic interactions with their next-nearest neighbours. The Generalised Linear Spring (GLS) force law aims to combat this problem, and originates from work by Pathmanathan *et al.* in 2009 [192]. The GLS force on cell i due to cell j is given by:

$$f_{ij} = \begin{cases} \mu\,(R_i + R_j)\log\left(1 - \dfrac{v_{ij}}{R_i + R_j}\right), & \text{if } v_{ij} > 0 \text{ and } d_{ij} < k, \\ \mu\, v_{ij}\,\exp\left(\dfrac{\alpha\, v_{ij}}{R_i + R_j}\right), & \text{if } v_{ij} \leq 0 \text{ and } d_{ij} < k, \\ 0, & \text{if } d_{ij} \geq k. \end{cases} \tag{3.10}$$

For two overlapping cells the GLS force is equivalent to Hooke's law to first order, providing spring-like repulsion with μ denoting the spring strength. However, the GLS force law produces stronger repulsion at large overlaps, similar to a hard core potential. This improves simulation stability and prevents indirect cell-cell interactions. In addition, the GLS force law incorporates weak attraction between non-overlapping cells, which decays exponentially with distance at a rate determined by the parameter α. This weak attraction can be interpreted as cell-cell adhesion, and also ensures that the force law is smooth at $d_{ij} = R_i + R_j$. Finally, a cutoff distance k is imposed beyond which cells cannot interact. Figure 3.4E illustrates the size of the GLS force as a function of cell separation.

A slight variation on the GLS force law is a pure repulsion force, in which only overlapping cells are permitted to interact. As a result, no weak attraction between cells occurs:

$$f_{ij} = \begin{cases} \mu\,(R_i + R_j)\log\left(1 - \dfrac{v_{ij}}{R_i + R_j}\right), & \text{if } v_{ij} > 0, \\ 0 & \text{if } v_{ij} \leq 0. \end{cases} \tag{3.11}$$

Hertz and Johnson-Kendall-Roberts contact forces

So far, we have made heuristic assumptions about cell-cell forces: that they should be linear or spring-like. However, it is possible to formulate force laws based on more in-depth physical arguments. Prominent examples come from the Hertz theory of contact and Johnson-Kendall-Roberts (JKR) theory [128, 192, 194]. The Hertz approach neglects adhesion and considers the deformation that occurs when two spheres share a circular contact surface. It is possible to solve for the pressure field required to generate such a deformation, in terms of the Young's modulus, E, and Poisson ratio, ν, of a cell. Integrating pressure over the contact area gives the following expression for the force between cells i and j:

$$f_{ij} = -\frac{2E}{3(1-\nu^2)}\sqrt{\frac{R_i R_j}{R_i + R_j}}\,v_{ij}^{\frac{3}{2}}, \tag{3.12}$$

(subject to the assumption that the deformation is small) [194]. Figure 3.4F illustrates the profile of a Hertz contact force. JKR theory extends this reasoning by reintroducing adhesion via a surface energy term. The result is an equation relating v_{ij} to contact surface area, along with a different expression for the force:

$$v_{ij} = \frac{A_{ij}}{\pi}\frac{R_i R_j}{R_i + R_j} - \sqrt{\frac{4\pi\sigma(1-\nu^2)}{E}}\left(\frac{A_{ij}}{\pi}\right)^{\frac{1}{4}}, \tag{3.13}$$

$$f_{ij} = -\frac{2E(R_i + R_j)}{3R_iR_j(1-\nu^2)}\left(\frac{A_{ij}}{\pi}\right)^{\frac{3}{2}} + \left(\frac{8\pi\sigma E}{2(1-\nu^2)}\right)^{\frac{1}{2}}\left(\frac{A_{ij}}{\pi}\right)^{\frac{3}{4}}, \tag{3.14}$$

where σ relates to the strength of adhesion. Figure 3.4G illustrates the JKR force as a function of cell separation.

More detailed force laws

Finally, a few force laws take into account an even wider range of cell-cell interactions. One example is the model of Buske *et al.* (2011), which includes repulsion, adhesion and uniform compression [33]. We discuss this force law in detail in Section 5.1. Here, for illustrative purposes, we show only the compression component (Figure 3.4H), which penalises deviations from a cell's target volume, V^T:

$$f_{ij} = -\frac{K\pi}{3V^T}\left(V^T - V\right)\left(2(R_i - x_{ij})(2R_i - x_{ij}) + (R_i - x_{ij})^2(2R_i - x_{ij})\right)\left(1 - \frac{x_{ij}}{d_{ij}}\right). \tag{3.15}$$

V above is the cell's compressed volume, calculated based on its current radius. The remaining terms account for any volume lost due to the cell-cell overlap. x_{ij} is the distance from the centre of cell i to its contact surface with j, given by $(R_i^2 - R_j^2 + d_{ij}^2)/2d_{ij}$, and K is the energy of uniform compression [33]. Detailed biophysical force laws like Equation (3.15) come at the cost of needing to solve a larger, coupled system of equations each time step, as well as having more parameter values to select. There is therefore a trade-off to be made between realism on the one hand and speed and simplicity on the other.

3.1.5 Cell-based modelling software

Some of the cell-based models described in this section take significant time to implement, meaning it is neither practical nor efficient for every scientist to write their own code from scratch. As a result, numerous software tools have been made available for developing and running cell-based models.

A variety of applications exist for on-lattice modelling. There are tools available for running cellular automata, of which the best maintained appears to be Golly [249]. However, biology-specific cellular automaton tools are rare, perhaps because these models are simple enough to write from scratch in Mathematica or MATLAB [14, 266]. A more substantial on-lattice project is CompuCell3D, which runs cellular Potts models of tissues [237]. A user interface is provided to simplify simulation design, and various plugins can be added to modify the energy function governing cell movement. Intracellular decision-making can also be incorporated by linking to the program Bionetsolver [10]. CompuCell3D has been successfully used to study gastrulation, angiogenesis and cell migration [217, 226, 253]. A similar software tool is Morpheus, which also supports the development of cellular Potts models [231]. The main distinction appears to be that Morpheus separates model design from implementation details. Intracellular models are expressed in a high-level, domain specific language, without the need to explain how they will be simulated. Morpheus has good support for ODE models and reaction diffusion equations, which has led to its use in studying vascular development and pattern formation [17, 18].

For off-lattice modelling, the existing tools are often tailored toward a specific biological domain. For instance, VirtualLeaf is a software package for modelling plant growth that takes into account the cell wall and strong cell-cell adhesion typical of plant tissues [167]. It uses a vertex-based approach, with vertex position updates generated stochastically and either accepted or rejected based on their effect on the system-wide energy level. VirtualLeaf aims to make model development accessible to plant scientists and has been used to study root growth and meristem patterning [177, 257]. Another example is CellModeller, which is primarily used to simulate bacterial colonies [204]. It represents bacteria by rigid capsules and updates cell positions at each time step so as to remove overlaps arising from division or growth. Nutrient fields can also be applied and updated by solving PDEs. CellModeller is written in Python and maintains excellent performance through GPU acceleration and parallelism.

A more general purpose modelling library is Chaste (Cancer Heart and Soft Tissue Environment) [171]. Chaste provides a collection of C++ classes useful for two main tasks: simulating cardiac electrophysiology, and running cell-based models. The cell-based code supports vertex-based, cell centre and Voronoi tessellation models, as well as on-lattice simulations (cellular automata and Potts models). This flexibility has enabled comparative studies of different modelling approaches [192], in addition to work on colonic crypt renewal and epithelial wound healing [67, 79]. Further details about Chaste are given in Section 3.1.6.

For our *C. elegans* germ line work the decision was made to use Chaste, as it provides the best support for general purpose off-lattice modelling. Moreover, expertise is readily available at Oxford, and Chaste was designed according to good software engineering practices such as test driven development, version control, and thorough documentation. The code is released under an open source licence and is reasonably easy to use and build upon.

3.1.6 Chaste: Cancer Heart and Soft Tissue Environment

Chaste is more akin to a library than a modelling tool, in that it does not provide a user interface for designing and running simulations. Instead, users write C++ programs called test files that describe simulations at a high level. Listing A.1 in Appendix A shows a typical example.

Most test files begin by creating a cell population. A range of population types are available, but in this thesis we primarily use node based cell populations, which are off-lattice and based on overlapping spherical cells. Certain details about each cell must be specified when the population is created. For example, every cell must have an associated cell cycle model, responsible for regulating its division and other intracellular properties. Once the cell population is set up, a simulation object can be created. Features such as force laws and boundary conditions are then added to the simulation in a modular fashion.

When an off-lattice simulation runs, Chaste enters a time-stepping loop that repeats the following operations:

1. Remove dead cells;
2. Carry out divisions;
3. Update cell cycle models;
4. Update cell positions;
5. Apply boundary conditions;
6. Apply modifiers;
7. Output data.

We will look briefly at each of these steps.

Cell removal and division

Users can add a “cell killer” to a simulation that either removes certain cells immediately or marks them as apoptotic. For instance, all cells above a certain plane, or all cells with fewer than six neighbours could be targeted. Cells marked for apoptosis shrink rapidly, before being removed from the simulation. Similarly, cell division is a matter of looping over the population and identifying cells marked as “ready to divide”. A daughter cell is then created that inherits the properties of its parent. The two resulting cells are placed offset from the parent’s position along a division vector with a randomly chosen direction (see Figure 3.5).

Updating cell cycle models

In Chaste, every cell has an associated cell cycle model, which is responsible for updating its current phase and deciding when division should occur. Mitotic cells cycle through phases G1, S, G2 and M, and are marked “ready to divide” at the start of M phase. Cells that are not proliferative remain permanently in G0.

All cell cycle models are prompted to update at each time step. The particular form of the update depends on the type of model. A *fixed duration* cell cycle model will check how much time has been spent in the current phase. If enough time has passed, the phase is advanced. More complex *generation based* models also take into account how many times the cell has previously divided. After a certain number of divisions, cells enter G0 and become quiescent. The most detailed cell cycle models are based on a system of ODEs that must be solved at each time step to determine the current phase [5].

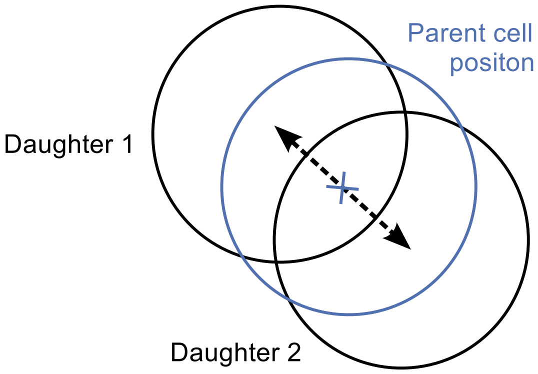


Figure 3.5: Cell division in Chaste. A pair of daughter cells are placed equidistant from the parent’s position, along a vector with a randomly chosen direction.

Updating cell positions

Cell positions are updated according to Newton's second law with overdamping (Section 3.1.2). Chaste solves Equation (3.2) using the forward Euler method with a fixed time step, dt:

$$\mathbf{r}_i^{N+1} = \mathbf{r}_i^N + \frac{dt}{\eta_i} \sum_j \mathbf{F}_{ij}. \tag{3.16}$$

As for the force law specifying $\mathbf{F}_{ij}$, a number of options are implemented in Chaste and the user can select any force compatible with their cell population model. In practice, however, many cell centre simulations use a Generalised Linear Spring (GLS) force, given by Equation 3.10 and described in full in Section 3.1.4.

Table 3.1 lists Chaste's default parameters for the GLS force law. The defaults were chosen based on colonic crypt modelling work; they produce a realistic rate of cell movement in crypt simulations, comparable to the rates in Meineke *et al.* (2001) [166]. Notably, the important quantity that emerges is the ratio μ/η, so non-dimensionalisation is appropriate and η can be held constant at unity provided μ is suitably scaled.

Applying boundary conditions

Once cell positions have been updated, boundary conditions are applied. Cells that have passed through a boundary are identified, then moved a minimal distance so they lie within the allowed region again. All boundary conditions are rechecked at the end of the time step to ensure they remain satisfied. This validation step helps detect incompatible boundary conditions where two or more boundaries are in use.

Parameter	Description	Value	Reference
μ	Spring strength	$15\,\mathrm{kg\,h^{-2}}$	Based on movement rates in [166]
α	Adhesion decay rate	5	Arbitrarily chosen to smooth the force law
k	Interaction cutoff	$3R_{max}\,\mu\mathrm{m}$	Require $k > 2R_{max}$ so that overlapping cells interact
η	Drag coefficient	$1\,\mathrm{kg\,h^{-1}}$	Can be scaled to 1 given an appropriate μ

Table 3.1: Default parameters for the GLS force law. R_{max} is the maximum cell radius in a simulation.

Modifiers and data output

Modifiers are applied at the end of a time step. These include any pieces of custom code the user wants to insert. The most common use cases are: 1) updating a cell property unrelated to the cell cycle, and 2) writing data to a file. At specified intervals, Chaste also saves a snapshot of the entire cell population that can be visualised in Paraview [4].

We have now covered the existing tools and techniques for modelling a cell population, as well as giving an overview of the Chaste library. However, a germ line model should also represent the complex decision-making processes that occur within individual cells. Therefore, we will now review a range intracellular modelling techniques, commenting on their suitability for representing germ cell behaviour.

3.2 Representing intracellular logic

Cells take input from their environment and transform it into appropriate actions. A useful way of thinking about this transformation is as a logical program; though cellular programs differ from their man-made counterparts in terms of flexibility, stochasticity, and robustness. A wide variety of approaches exist for modelling the internal logic of a cell. We begin by making a distinction between continuous and discrete models, before focussing in on a small selection of techniques inspired by computer science, including Petri nets, Boolean networks and finite state machines (FSM).

3.2.1 Continuous vs. discrete models

Intracellular models can be broadly classified as either continuous or discrete. Continuous models typically consist of a set of equations, such as a system of coupled ODEs. The variables may represent amounts of protein, ions or mRNA, either in the whole cell or in a given compartment, activation state or complex. These variables evolve over time, and since proteins are ultimately responsible for most cellular functions, a set of protein concentrations can mapped to an observable behaviour.

A good example of this approach can be seen in mathematical models of the cell cycle. These include variables representing the amount of different cyclins, kinases and other regulators in particular complexes and phosphorylation states [52]. Dynamical systems theory can be used to show that the system settles into a repeating limit cycle, with a threshold concentration of some factor interpreted as the trigger for division. Indeed, coupled differential equations describing protein levels can generate a range of useful behaviours, including switching, hysteresis, oscillations, and adaptation [222, 251]. Continuous models have proven useful in almost every area of biology, from circadian rhythms [212], to axon potential generation [110], to metabolomics [126].

By contrast, discrete models often look more like a set of instructions for evolving the cell in time. They typically feature a representation of the cell's state, an indication of which states could be entered into next, and some form of update rule to be applied at each time step. An advantage of discrete approaches is that they allow high-level hypotheses about cell behaviour to be tested, without the need to speculate about the molecular details. Some discrete models are also amenable to model-checking; a technique that uses automated tools to explore model behaviour systematically and exhaustively [272]. Model-checking can determine whether a model always produces biologically realistic outcomes, or whether certain special cases exist giving unexpected results. A number of discrete modelling techniques are discussed in this section, including Petri nets, Boolean networks and statecharts. For further details, see [39, 76, 153, 262].

We chose to focus on discrete approaches in this thesis because the molecular mechanisms that coordinate germ cell behaviour are insufficiently well characterised for detailed mathematical modelling of the biochemistry. While the initial steps of DTC signalling are clear, the pathway expands downstream into a poorly understood network of mRNA binding proteins, with potentially hundreds of targets. Even where the direct protein-protein interactions at work are known, estimates of the binding, unbinding and decay rates are not. These quantities would require a concerted effort to measure *in vitro*, and they are more difficult to estimate than the parameter values of a more abstract model, which can be related to observable cell properties. In general, the open questions described in Chapter 2 are concerned with whole cell and whole organ behaviours. Given that we are still trying to distinguish between high-level hypotheses, a discrete approach to modelling germ cell behaviour is appropriate.

3.2.2 Petri nets

The first intracellular modelling formalism we consider is the Petri net, which is similar in spirit to an ODE system and is often used to simulate sets of chemical reactions. A Petri net consists of a collection of *places* (represented by circles), each of which contains a certain number of *tokens* at any given time (see Figure 3.6) [39]. In the case of a chemical reaction network, places correspond to chemical species, while tokens show how many molecules of each one are present. Places are connected to each other via *transitions*, which are rectangular boxes joined to their incoming and outgoing places by arcs. Each transition requires a certain number of input tokens to be available before it can fire, and on firing it generates a certain number of output tokens. In a chemical reaction system, transitions represent reactions while the number of input and output tokens specifies the stoichiometry. Extensions to standard Petri nets have been developed that feature values attached to tokens (coloured Petri nets), exponentially distributed delays before transition firing (stochastic Petri nets), and places marked with real values instead of tokens (hybrid Petri nets) [39].

In computer science, Petri nets are used to study asynchronous and concurrent systems, since they do not specify the order in which transitions fire, only the conditions under which they are enabled. Non-biological applications include the design of resource sharing protocols and communication protocols [178]. Mathematical analyses exist that can determine the set of reachable token distributions for a Petri net, or put bounds on the number of tokens in a given place [178]. In biology, Petri nets have been successfully used to model metabolic and gene regulatory networks; including the pentose phosphate pathway in red blood cells [200], the lambda phage genetic switch [159], and Delta-Notch mediated fate determination in *Drosophila* [172]. Though Petri nets have been used for more abstract models, they naturally lend themselves to tracking counts of individual molecules. As such, they could be considered unnecessarily detailed for the questions at hand in

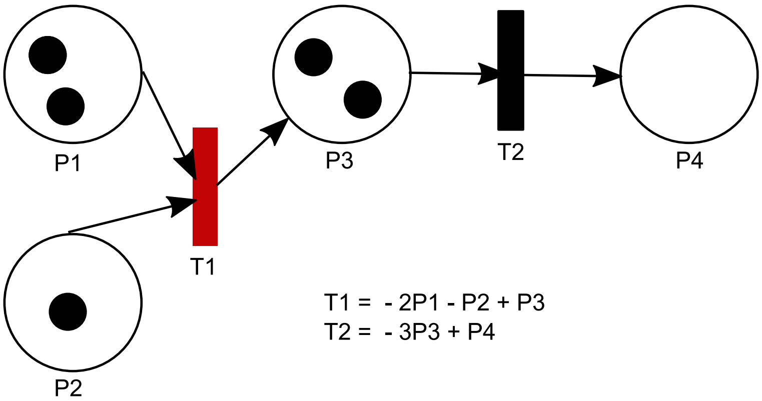


Figure 3.6: A Petri net. Small black circles represent tokens, large circles represent places, and rectangles represent transitions. The definitions of T1 and T2 specify how many tokens they consume from their input places and how many tokens they generate. In the state or *marking* shown, sufficient tokens are available for T1 to fire but not T2, which we indicate by colouring the T1 rectangle red.

the germ line. A Petri net model would likely require knowledge of the molecular interactions governing germ cell behaviour and their stoichiometry. We therefore move on to consider alternative modelling techniques.

3.2.3 Boolean and qualitative networks

Boolean networks are a commonly used model type, particularly well-suited to studying the time evolution of gene expression in large genetic networks [134, 246]. A Boolean network consists of a number of nodes representing genes, each of which has two possible states: ON (expressed) or OFF (silent). Each node is connected by directed arcs to a set of input nodes (see Figure 3.7). To determine whether a node should be ON or OFF at time T+1, a logical function is applied to the states of its inputs at time T. If the function evaluates to TRUE the node turns ON, and if not it is switched to OFF. For instance, if node A has inputs B and C together with the logical function AND, then A will be ON at T+1 only if B and C were ON at T.

The use of Boolean networks to represent genetic circuits was first proposed in [134], and they have since been applied to a wide range of biological systems [26, 66, 193, 261]. There are two main uses of Boolean network models. Where the underlying genetic network is already known, a model can make predictions about mutant phenotypes and essential genes [261]. *Essential* here could mean "loss of function mutations disrupt a particular signalling pathway" or "loss of function mutations are lethal" depending on the context. Steady state patterns of activity can be determined, which may correspond to biologically meaningful states such as "health" or "disease".

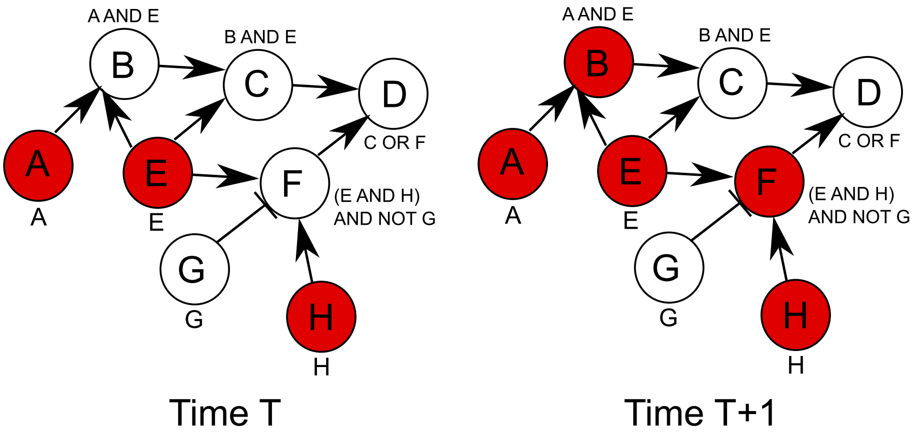


Figure 3.7: One update of a Boolean network. Each circle represents a gene and those coloured red are currently ON. Beside each gene is a logical condition on its input nodes that must be met for it to be expressed. At time T+1, gene activities are updated based on activities at T. Thus, in the above example, B activates at T+1 because A and E are ON at T, and F activates at T+1 because E and H (but not G) are ON at T. C and D do not activate until the next time step (T+2).

Perturbations can also be identified that produce a change of state [6, 205]. Where the underlying genetic network is *not* well-characterised, techniques from verification and model-checking can be used to construct a Boolean network that reproduces experimentally measured expression levels [66, 206, 272, 273]. Therefore, Boolean networks can be used both to reverse engineer genetic networks from data and to make biologically relevant predictions.

Various extensions of Boolean networks exist that aim to introduce extra detail into the model. For example, asynchronous Boolean networks allow for the fact that some activations and inactivations occur faster than others. Updates are processed one gene at a time, with new activity levels depending on the current state of the system rather than the state at time T. Explicit delays may be introduced to account for substantial rate differences [190, 213]. In developmental biology, spatial detail has been introduced by considering multiple interacting Boolean networks associated with different domains or cells [190, 193]. Finally, some extensions of Boolean networks move away from ON/OFF gene activities altogether and allow multiple levels of activation [209, 213]. In the qualitative network formalism, each gene can take on a finite range of integer expression levels. There remains an underlying network with sets of input nodes, and logical functions determine whether each gene receives a positive or a negative stimulus at a given time step (Figure 3.8) [26, 209]. Qualitative networks also permit formal analysis of model behaviour and identification of steady states.

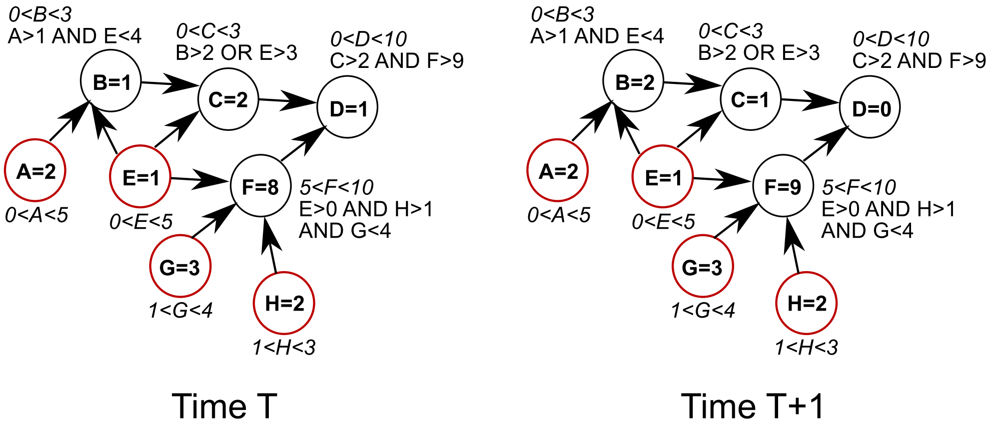


Figure 3.8: One update of a qualitative network. Again, circles represent genes, but now each gene has an associated integer value giving its activation level within an allowed range. Values are updated according to a target function on the input nodes at time T. In this example, we add +1 to the activation level of all genes where the condition is satisfied, and -1 whenever the condition is not satisfied. Red circles indicate genes with no input nodes, whose activity levels are set by the modeller.

Boolean and qualitative networks are powerful methods for studying large systems of interacting genes. However, in the *C. elegans* germ line, large gene expression datasets of the kind needed to test and synthesise network models are not routinely collected. Rather, staining and reporter constructs are typically used to investigate a few genes or proteins at a time. Moreover, we are interested in properties of germ cells that are not easily linked to a pattern of gene expression at present. Therefore, we will consider a third type of discrete model, based around high-level, user-defined states and transition conditions.

3.2.4 Finite state machines, statecharts and sequence charts

A very general approach to represent intracellular logic is to model the cell as a finite state machine (FSM). An FSM is an abstract machine that can exist in a number of states represented by labelled boxes (Figure 3.9). For a cell, relevant states might include "Mitosis" or "Meiosis". Allowed transitions are shown as arrows between states, often with an attached condition that must be met for them to occur [91]. In a biological context, transition conditions will frequently depend on signals from outside the cell, such as the state of neighbouring cells or the current position. An FSM may also perform *actions*, side effects that are triggered on entry to, exit from, or while in a particular state. FSMs are a highly abstract, unstructured way to represent a program; the majority of biological modelling work actually uses various extensions of this idea such as statecharts or reactive modules [77, 105, 198, 220]. However, there are instances where pure FSMs have been used in biological research, for example, in modelling the metabolism of DNA [87].

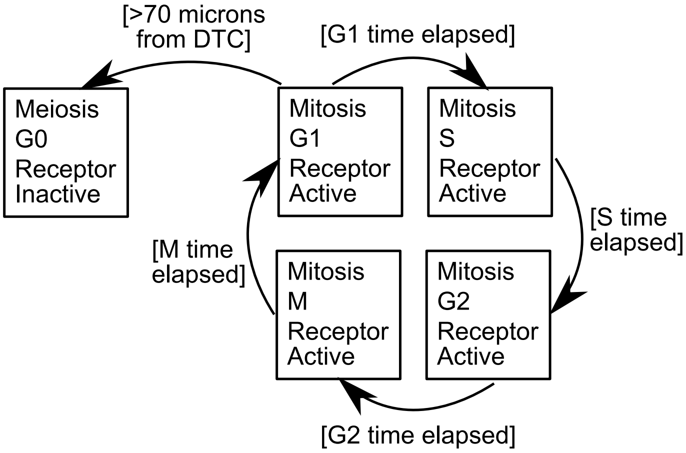


Figure 3.9: A finite state machine. Boxes represent states, which here are composed of three pieces of information. Arrows show the allowed transitions between states, and a transition condition is written alongside each one. In a biological context, transitions must often reference information external to the FSM, as well as continuous quantities, such as positions or delays.

A drawback of FSM models is that if the state consists of many pieces of information, a change in just one of them requires a new box to be added to any diagram of the system. For complex biological models, this can rapidly lead to state explosion. Statecharts are an extension of the FSM formalism, designed to allow complex systems to be represented visually in a concise, readable fashion [105]. Figure 3.10 shows an example statechart representing a cell. To prevent state explosion, *nesting* is introduced. The "Prolif" state contains four cell cycle states nested inside it: G1, S, G2 and M. If G1 is currently active, then implicitly so is Prolif and any states that contain it. Multiple different combinations (Prolif AND G1, Prolif AND S, Prolif AND G2, Prolif AND M) are thereby condensed into a single section of the chart. A stabbed arrow symbol (•↷) is introduced to indicate the default nested state to enter if no other behaviour is specified. So, in Figure 3.10, a cell entering Prolif will, by default, start in G1.

The other significant feature of statecharts is their use of *orthogonal regions*, which are sections of the chart separated by a dashed line that update in parallel. Orthogonal regions allow statecharts to be written in a modular fashion, with one region for each distinct process. Transition conditions can still depend on the state of other regions, since state information is communicated to all parts of the chart.

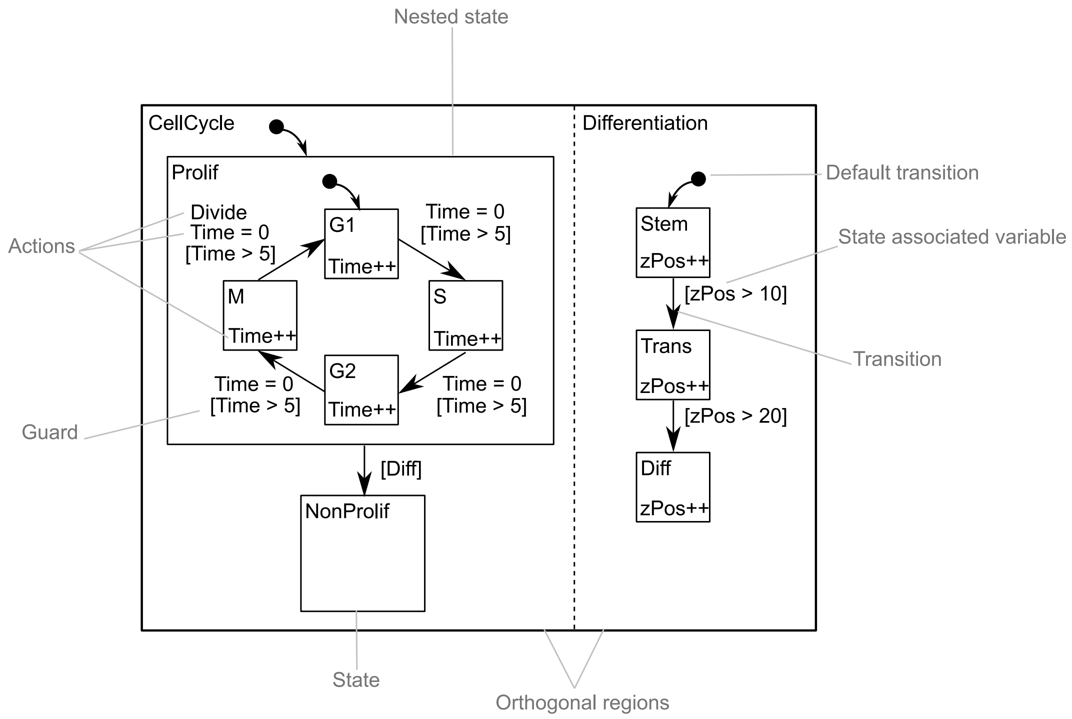


Figure 3.10: An example statechart representing a cell. Distinct orthogonal regions control the cell cycle and differentiation, with the default starting states being G1 and Stem. Transition conditions or *guards* are given in square brackets, and examples of state associated variables and actions are included.

Certain features of statecharts are particularly useful in a biological context. Their modular design reflects the underlying division of labour in the cell, as well as making diagrams easier to interpret. Some implementations allow variables to be associated with states to represent continuous quantities, such as time delays or concentrations; others allow an initial state to be chosen explicitly rather than assuming the default. Moreover, statecharts are amenable to model-checking to ensure that they match experimental observations [107], although large populations of interacting cells remain intractable. Statecharts have been widely used to model cell behaviour in contexts including pancreatic development [220], immune cell activation [71], and the *C. elegans* reproductive system [221] (see Section 3.3).

The techniques described thus far focus on the internal state of a cell and its subsystems. There is an alternative approach, Message Sequence Charts (MSCs), that concentrates on the *communication* between subsystems [92]. In an MSC, each subsystem is represented by a vertical life-line, time runs from the top to the bottom of the chart, and arrows represent messages sent between subsystems. An MSC lays out the expected behaviour of a system under a given scenario: “If this sequence of messages has been sent, this pattern of behaviour should result.” An extension, Live Sequence Charts (LSCs), is somewhat more expressive, and allows the user to specify behaviours that are forbidden as well as those that must occur (see Figure 3.11) [56]. The strength of MSCs

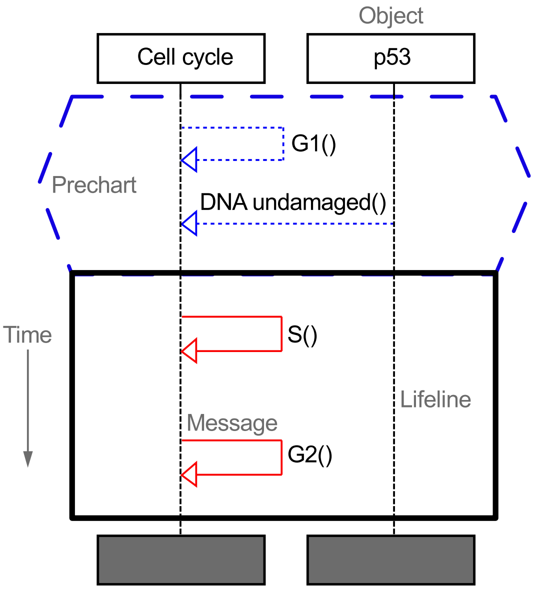


Figure 3.11: An example LSC. Above is a simple LSC representing the G1 cell cycle checkpoint in eukaryotes. The interacting objects are p53 and the cell cycle, represented by a pair of lifelines. Time runs down the diagram from top to bottom, and arrows between the two lifelines indicate messages sent between objects. Drawn in blue is a *prechart*. If the sequence of messages in the prechart occurs, this LSC specifies that the messages in the black box *must* follow. The chart therefore states that if the cell cycle has entered G1 and p53 levels indicate the DNA is undamaged, then transitions into the S and G2 phases must occur in that order.

and LSCs, with their emphasis on scenarios, is that they can be used to represent formally the qualitative patterns of behaviour observed in an experiment [132, 148].

Based on their high-level, readable representation of cell behaviour, we chose to use statecharts in our *C. elegans* germ line model. In the next section, we cover some of the available software for intracellular modelling, focussing on tools that support statecharts.

3.2.5 Intracellular modelling software

Thanks to the rapid growth of computational biology as a field, there are now many specialised tools available for intracellular modelling work. General purpose software like MATLAB continues to be used to run models based on systems of equations; though even here specialist packages exist to help with prototyping, analysis and fitting [157, 240]. For Petri net modelling, Cell Illustrator provides a mature and user friendly tool, allowing pathway diagrams to be drawn then simulated as hybrid functional Petri nets [182, 252]. Several tools also exist to support the running and analysis of Boolean and qualitative networks [23, 64, 179]. Interestingly, we are now moving toward standardised representations of models that allow them to be imported between tools and simulated in different ways. For example, the Systems Biology Markup Language (SBML) allows the essentials of a model to be described in an XML file, which can then be translated into a system of equations (either algebraic or differential), a Petri Net, or a network model [116].

Statecharts are somewhat different from other intracellular approaches in that they do not normally operate at the level of individual genes and proteins. This places them slightly outside the usual ecosystem of computational biology tools. Fortunately, many excellent statechart tools and libraries are available, thanks to their use in industry for designing critical systems [105, 118, 244]. The majority of these tools actually support the Unified Modelling Language (UML), a larger formalism of which statecharts are only a part [242].

The statecharts program with the greatest capabilities appears to be Rational Rhapsody by IBM [119]. It provides a visual environment for designing UML models that also supports simulation and verification, as well as integration with other tools [241]. The statechart implementation allows for timed expressions, actions, and overriding of statechart inheritance; all features that could be useful in a biological context. Moreover, the semantics of Rhapsody statecharts have been rigorously described [106]. The main disadvantage of Rhapsody is its licensing, with a commercial license costing upward of $1000 and a convoluted process for obtaining an academic license. The licensing situation would prevent statechart code being integrated into our cell-based modelling library, and it could discourage other scientists from building on our work.

Several other tools meet the criteria of being free, open source and updated within the past year. These include Modelio, Papyrus, Umbrello UML Modeller, and UML Designer [173, 186, 243, 245]. All four projects appear well-maintained and well-documented, with several being used in industry. However, they tend to focus on model design and validation rather than simulation. Papyrus and Umbrello both support C++ code generation, which could allow a statechart model to be integrated into Chaste, but at the cost of adding an extra dependency to the already long list of programs Chaste requires.

Finally, two statechart libraries are included in Boost, a well-respected C++ library collection [207]. A major advantage of choosing a Boost library is that Chaste already relies on Boost extensively. Using it to implement statecharts would therefore avoid extra dependencies and keep the Chaste installation procedure streamlined. The two available libraries are Boost Statechart and Boost Meta State Machine (MSM) [45, 59]. Both take an object oriented approach, with states defined as structs or classes. MSM codes in the structure of the chart as a transition table, while Boost Statechart instead templates each state on its parent and children. At the time this project began there was not much to choose between these libraries. Both approaches were attempted, and Boost Statechart proved easier to use. On the basis of cost, licensing, and ease of integration with Chaste, Boost Statechart was therefore selected for implementing our germ cell model.

Having covered techniques for both inter- and intracellular modelling, we now turn our attention to previous models of the *C. elegans* germ line. These models successfully applied the techniques we have discussed to obtain new insights about the germ line system.

3.3 Previous *C. elegans* germ line models

In this final section, we describe pre-existing models of the *C. elegans* germ line. When this project began, only two prior models were available, both from 2012 [24, 221]. However, in 2015, three new germline modelling papers were released, including a 3D model by Hall et al. [99] (published while the work in Chapter 4 of this thesis was under review; see [15]). Here, we focus on the cell-based work of Setty *et al.*, Beyer *et al.* and Hall *et al.*. Other relevant modelling work is summarised at the end of the section.

3.3.1 Setty *et al.* 2012

One of the first *C. elegans* germ line models was developed by Setty *et al.* in 2012 [221]. It implemented an on-lattice approach to germ cell movement, defining a 2D cross-section through the gonad, then dividing up the space inside into a lattice (Figure 3.12A). Over the course of a simulation, the size and shape of this lattice was regularly updated in accordance with experimental measurements, to mimic gonad growth. Thus the Setty *et al.* model was able to cover larval germ line development as well as adult maintenance. Each germ cell was represented by collection of lattice sites sharing a common ID: a 1x1 block for most cell types, ranging up to a 3x3 block for oocytes. Cells were frequently prompted to move and (less frequently) prompted to divide, with both actions requiring the availability of a neighbouring empty lattice site. Given a choice of empty sites, germ cells moved randomly during the larval part of the simulation, but were biased toward proximal motion in the adult stages.

Germ cell behaviour was governed by a statechart model running inside each cell. The statechart included orthogonal regions controlling: 1) the binding state of the GLP-1 receptor; 2) the cell cycle phase; 3) progress through differentiation; 4) choice of movement direction; and 5) the activity of various intracellular effectors (LAG-1, GLD, NOS-3, MPK-1, MEK-2, Meiotic-X, Sperm-Effector and Oocyte-Effector). Part of this statechart is reproduced in Figure 3.12B. The intracellular model was tested by simulating gain and loss of function mutations, and in each case was found to produce biologically realistic outcomes for the germ line as a whole.

Various interesting conclusions were drawn from the Setty *et al.* model. First, reducing the range of the DTC signal in simulations sometimes led to loss of the proliferative zone, suggesting a possible cause of proliferative zone instability *in vivo*. Second, the length of the cell cycle during the larval stages was found to have a lasting impact on proliferative zone size in the adult. Third, a wave of differentiation was successfully simulated using a population of identical germ cells. This is the phenomenon introduced in Section 2.8 that was originally described by Cinquin *et al.* [47], and which had been viewed as evidence for distinct germline stem cells. The Setty *et al.* model suggested that a distinct distal cell population need not exist. Finally, the average niche residence time for a germ cell was recorded, both with directed and undirected divisions.

Strengths of the Setty *et al.* model include its detailed representation of the cell cycle, containing different phases and stages of meiosis. This should allow the coupling between the cell cycle and other germ cell decisions to be explored, though the topic was not covered in any depth by the

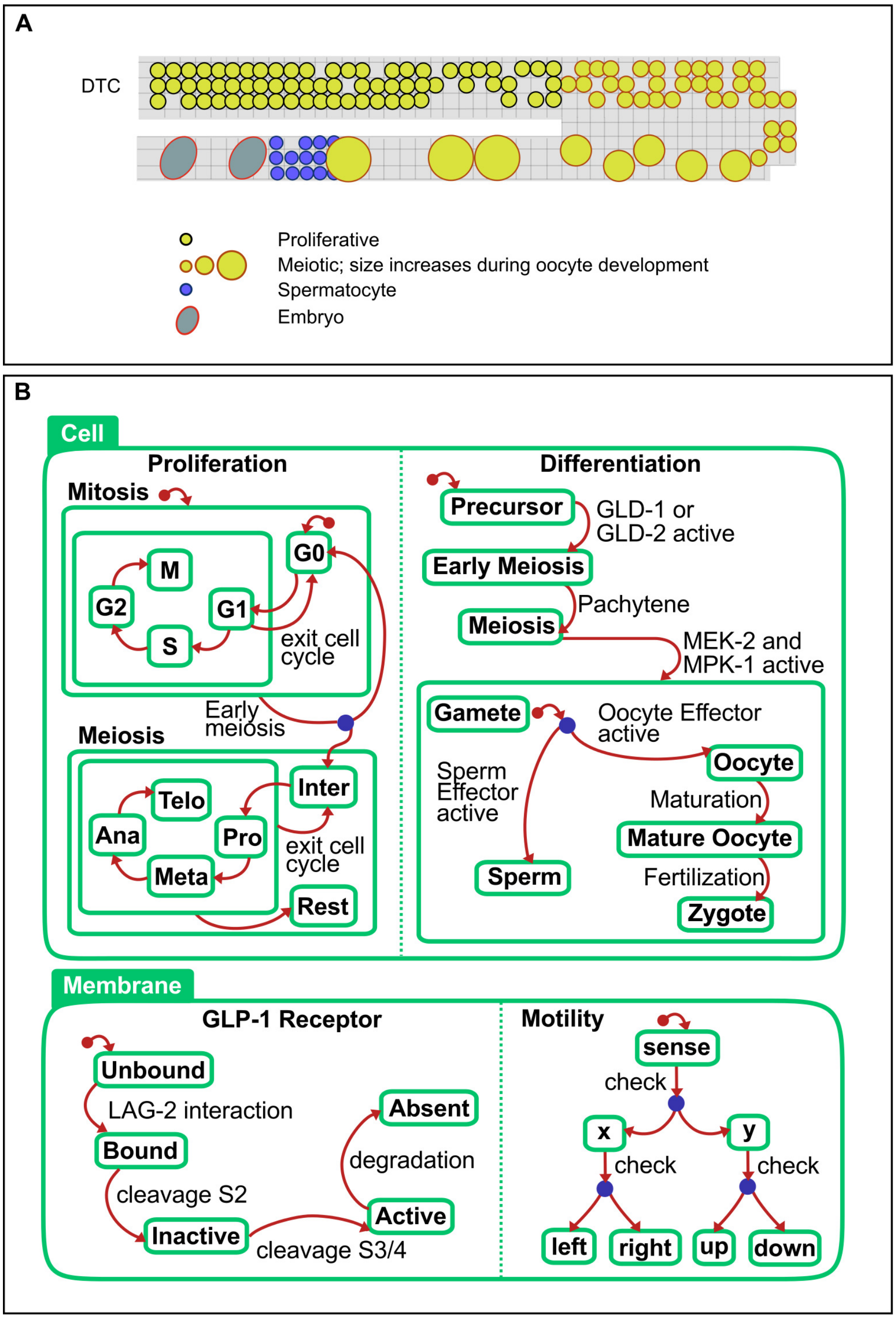


Figure 3.12: The Setty *et al.* germ line model. A) An on-lattice model of cell movement. A cross-section through the gonad is represented as a 2D lattice, which grows and changes shape during the larval stages of the simulation. Cells occupy between 1 and 9 lattice sites depending on their differentiation status, and move in a biased random walk. B) Part of the statechart model of germ cell behaviour. Pictured are the orthogonal regions controlling the cell cycle, differentiation, GLP-1 binding and movement direction. Many transitions depend on the activity level of intracellular signalling factors, which themselves turn on and off based on distance from the DTC. Images adapted with permission from [221].

authors. Since all cell behaviours, including movement, are controlled by the statechart, the Setty *et al.* model should also be amenable to model-checking. However, applying such techniques to a large interacting cell population would be challenging. A less satisfying aspect of this work is the fact that many effectors in the statechart are essentially redundant. Ultimately, most transitions are driven either by distance from the DTC or simulation time, and several effectors in the Setty et al. statechart are never commented on by the authors. It is therefore questionable whether they should have been included in the model. Finally, perhaps the weakest aspect of this work is its on-lattice approach to cell movement. Adding a bias toward proximal movement during the adult stages helps establish ovulation, but it could also affect several of the properties Setty *et al.* go on to study; notably proliferative zone stability and niche residence times. An off-lattice model based on physical laws and cell-cell repulsion would not suffer from such problems.

3.3.2 Beyer *et al.* 2012

Beyer *et al.* also published a *C. elegans* germ line model in 2012 [24], with a number of significant differences from the work of Setty. First, an off-lattice approach was taken to cell movement. Germ cells were represented by circles, which experienced a repulsion force whenever they overlapped, as well as a drag force proportional to velocity (Figure 3.13A). Rather than make an overdamping assumption, cell positions and velocities were both solved for using Verlet integration, a time-stepping technique based on a Taylor series expansion [254]. The Beyer et al. model of cell movement is therefore second order in time, corresponding to Equation 3.1 without the stochastic term.

A second point of difference is that Beyer *et al.* used a static gonad boundary condition; as such, their model only applies to the adult germ line. However, a subsequent extension did add periodic boundary conditions to the distal gonad, a change that takes the rachis into account (Figure 3.13B) [25].

Finally, Beyer *et al.* took a more *ad hoc* approach to modelling germ cell behaviour. In their model, each germ cell simply contained a collection of state variables updated by an F# program [24]. Slightly fewer effectors were included than in the Setty *et al.* statechart, with the focus being apoptosis and oocyte maturation as opposed to proliferation. Again, many variables were ultimately dependent on a cell's position in the gonad. However, Beyer *et al.* phrased this dependence in mechanistic terms, referring to different signalling environments provided by different regions of the organ. The Beyer *et al.* model also incorporated steady germ cell growth, and apoptosis was included in a "death zone" near the turn.

The Beyer *et al.* model produced cell death, ovulation and migration rates in reasonable agreement with biological estimates. The authors then tested different potential mechanisms for physiological

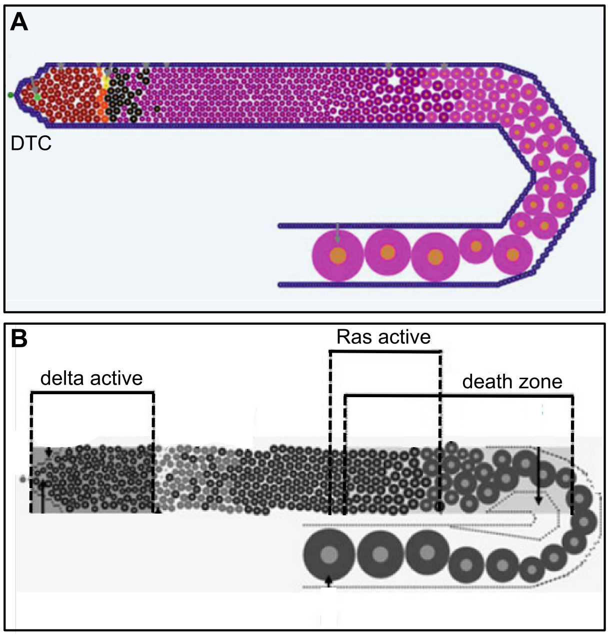


Figure 3.13: The Beyer *et al.* germ line model. A) A snapshot of the original model, reproduced with permission from [24]. Cells are confined to a 2D region representing a cross-section through the adult gonad. Within this region, germ cells move in an off-lattice fashion according to a molecular dynamics model. Cell behaviour is controlled by an F# program, which updates a cell's state according to the signalling environment it encounters. B) An extension of the model in [24] that applies periodic boundary conditions to the distal germ line in order to model the rachis. Reproduced with permission from [25].

apoptosis, finding that random cell death and death of the largest germ cells gave the most realistic-looking results. Advantages of the Beyer *et al.* model include off-lattice cell movement, and its inclusion of the rachis, which should bring simulated and actual cell counts into closer agreement. However, the papers using this model addressed a fairly limited range of biological questions, since larval development could not be modelled [24, 25]. These papers also included few new experimental measurements for comparison with the computational results.

3.3.3 Hall *et al.* 2015

A more recent model of the *C. elegans* germ line was published by Hall *et al.* in 2015, and also covered the adult life cycle stage only [99]. This was the first 3D germ line model, with cells represented by spheres confined to a region the size and shape of the adult gonad (Figure 3.14A). Cell movement was governed by the equations of Brownian dynamics [96]. Again, linear germ cell growth was assumed, division occurred after a cell had doubled in volume, and apoptosis was permitted in a fixed zone near the turn.

Hall *et al.* chose to represent germ cell logic using a qualitative network (Figure 3.14B; see also Section 3.2.4). The network incorporated a larger number of signalling factors than previous models, though with slightly more weight given to oocyte maturation and meiotic progression than to the mitotic cell cycle. Once again, a germ cell's changing position in the gonad was associated with different levels of Delta-Notch and MSP signalling, ultimately driving changes in the qualitative network's activity levels. Unfortunately, the code for the qualitative network was not made available, and the paper lacks certain details about starting states and target functions, making it difficult to comment further on the intracellular model.

Having developed a 3D, off-lattice model, Hall *et al.* were able to investigate germ cell movement

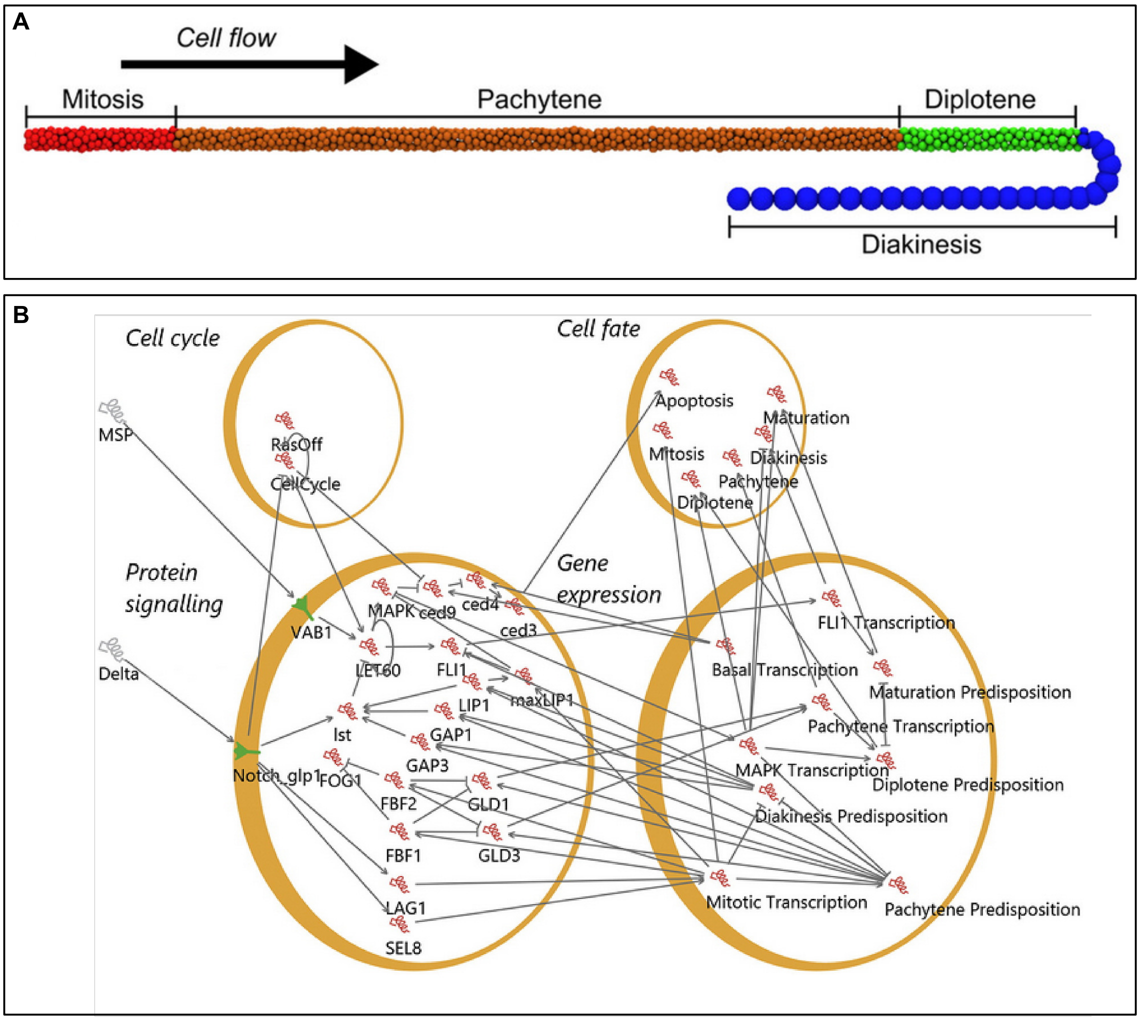


Figure 3.14: The germ line model of Hall *et al.*. A) A 3D, off-lattice model of germ cell movement. Cells are confined to a 3D region representing the adult gonad and move due to repulsion interactions with their neighbours as well as random Brownian motion. Germ cells grow linearly and divide after doubling in volume. B) A qualitative network model of germ cell behaviour. A large number of effectors are incorporated, with a particular focus on oocyte maturation. Changes in component activity are driven by the changing levels of Delta and MSP signalling that occur along the gonad. Images reproduced with permission from [99].

and packing. Their first conclusion was that distal germ cells often move laterally, mixing up the proliferative population. This is evolutionarily desirable, since it delays the takeover of the germ line by descendants of a single (potentially flawed) cell. Secondly, this model suggested some form of feedback mechanism must exist to prevent germ line overproliferation in mutants with impaired apoptosis. The authors chose to introduce negative feedback from cell crowding on division. Finally, model-checking techniques were applied to a greatly simplified version of the Hall *et al.* model, and showed that the gonad can robustly ensure correct fate decisions by channelling cells through a linear series of signalling environments.

The Hall *et al.* model extended previous work by moving to three dimensions, using off-lattice mechanics, and taking into account a large number of intracellular effectors. The decision to use a qualitative network, which is amenable to model-checking, allowed a thorough investigation of fate decisions, albeit in a simplified system. However, statecharts arguably remain a preferable formalism as they are more suited to high-level, abstract modelling and have a wider range of mature software tools available. Further improvements could include extending the Hall *et al.* model to cover larval development, or adding detail to the proliferative cell cycle.

3.3.4 Other relevant *C. elegans* models

The germ line models described thus far have all been cell-based, with a high level of detail, and applicable to a wide range of biological questions. Recently, two germ line models have been published that are much simpler, and more focussed in terms of the questions they address. The first of these, Chiang *et al.* (2015) [42], aimed to identify the evolutionary pressures influencing the rate of germ cell division. A slower germ cell cycle leads to a lower mutation rate, but a faster cell cycle allows animals to become fertile earlier. To study this trade-off, Chiang *et al.* modelled the distal gonad as a 2D lattice, with each site containing at most one cell. On exiting the distal gonad, germ cells enter a queue and ultimately either die or become gametes. The model's parameters included the dimensions of the proliferative zone and the length of the germ cell cycle at the zone's start and end. Optimal parameters values were identified for the model, subject to a range of different cost functions. The conclusion was that the *C. elegans* germ line uses a spatial gradient in cell cycle length to minimise mutations; subject to the constraint that reproductive development must not be significantly delayed.

A very different germ line model was published by Mattingly *et al.* in 2015 [160]. Their aim was to develop an equation for predicting the time a germ cell takes to travel from the distal tip to any other point in the gonad. Such an equation would allow some information about germ cell movement to be reconstructed from still images. The authors took an approach familiar from fluid dynamics: they assumed that any imbalance between proliferation and apoptosis must cause a change in germ cell flux — otherwise cells would accumulate somewhere in the gonad. The flux balance equation was manipulated to give an expression for the time taken to travel a given distance from the DTC. Parameters were then chosen based on biologically reasonable ranges, and on the estimated time taken to reach the turn. Using their model, Mattingly *et al.* were able to plot MPK-1 phosphorylation as a function of time since leaving the distal tip, revealing that this signalling pathway may be subject to timed regulation. This work is an excellent example of how a simple mathematical model can complement experimental studies, helping to overcome the limitations of current techniques.

Finally, the work discussed in this section should be viewed in the context of a wider effort to understand *C. elegans* fate decisions through logical modelling. The majority of this effort has focussed on a slightly different model system: the development of the worm vulva.

The *C. elegans* vulva forms in response to a signal from the Anchor Cell (AC), which prompts the closest Vulval Precursor Cell (VPC) to adopt a primary fate. The primary VPC then sends out a lateral signal, blocking primary fate in its neighbours. Nearby VPCs instead become secondary, while their neighbours become tertiary. The application of logical models to this system was first proposed by Kam *et al.* in 2004 [132]. They used an LSC to formalise the observed pattern of VPC fates. Subsequently, in 2005, Fisher *et al.* produced a statechart model of the behaviour of six VPCs and the AC [78]. The model was able to predict correctly the outcome of a range of genetic perturbations. It also generated a new hypothesis regarding the asynchronous timing of VPC fate decisions.

A notable difficulty with early vulval development models was the time taken to fit them to experimental data. A second paper by Fisher *et al.* addressed this issue by using model-checking techniques to automate testing and refinement [77]. Since then, logical models of other *C. elegans* systems have been developed that also use model-checking to guarantee a required set of behaviours [150]. An alternative approach for models that run quickly is to use Markov Chain Monte Carlo methods to automate fitting. This technique was employed by Sun *et al.* in their

2007 Dynamic Bayesian Networks model of vulval development [236]. Unfortunately, germ line models are not amenable to such techniques as they contain many more cells, making exhaustive model-checking impractical and run times too slow for stochastic parameter searches.

3.4 Summary

In this chapter, we reviewed existing techniques for modelling a cell population, and for representing intracellular logic. In the process, we highlighted the approaches and tools that we will use to model the *C. elegans* germ line.

In the case of a cell population, we focussed on agent based techniques and their associated decisions: on-lattice vs. off-lattice movement; choice of cell representation; and choice of force law. An off-lattice, overlapping spheres model was chosen as appropriate for the germ line, as it bases cell movement on physical laws, while avoiding unnecessary detail that could render a 3D model computationally intractable. Of the available cell-based modelling software, Chaste was chosen as a suitable library, due to its general purpose nature and good software engineering practices [171].

In the case of intracellular logic, Petri nets, Boolean networks and qualitative networks were all considered, before choosing to use statecharts in our germ line model. Statecharts provide an appropriate level of detail for studying the whole organ effects of individual cell decisions, particularly where the molecular mechanisms are not well-characterised. Boost Statechart was identified as a suitable choice of library, since it is freely available and Boost is already a Chaste dependency.

Finally, we reviewed existing models of the *C. elegans* germ line [24, 99, 221]. These models have yielded many interesting biological findings, however, areas for improvement remain. An off-lattice model would be less likely to suffer from grid artefacts and movement biases than the work of Setty *et al.*, while a 3D model would be better able to capture cell movement in the turn and near the distal tip. The work of Hall *et al.* provides a 3D off-lattice model, but it does not cover the larval period or model proliferation in detail. With these points in mind, the next chapter describes our *in silico* germ line model and presents the results of some whole organ simulations.

Chapter 4

A 3D *in silico* model of the *C. elegans* germ line

This chapter describes the 3D germ line model that we have developed, and presents the results of our whole organ simulations. This work has been published in the journal Development [15], with the theoretical work contributed by K. Atwell and the experimental work by Dr. Z. Qin. Where we reference our collaborators in this chapter, we refer to the group introduced in Section 1.3; primarily Dr. Qin.

We begin by giving an overview of the model and explaining its design, before going on to justify our choice of parameters based on a combination of new experimental measurements and literature review. We evaluate the agreement between our model's behaviour and that observed in the wild-type *C. elegans* germ line, both in qualitative and quantitative terms. Finally, we discuss some of the new biological insights we obtained by developing and running a whole germ line model, including: a role for "stretching" growth in late L4 gonadogenesis; a requirement for feedback on germ cell proliferation to regulate cell numbers; and predicted outcomes for cell tracking and labelling experiments.

4.1 Model overview

Our *C. elegans* germ line model consists of three main components: a mechanically interacting germ cell population; a growing gonad boundary condition; and a statechart model of individual cell behaviour. We will look at each of these components in turn and describe how they were designed and implemented.

4.1.1 Representing germ cells

The first component of our model is a mechanically interacting population of germ cells. We chose to represent the cell population in three dimensions, for a number of reasons. First, certain types of investigation are not possible using 1D or 2D representations of the gonad. For example, in a 2D section model like Figure 4.1A, germ cells cannot move past one another. It is therefore impossible to run a simulation in which some cells are fixed in place by adhesion without completely blocking the flow of cells through the system. Second, working in 3D gives us greater confidence that the forces and levels of crowding experienced by cells in our model are realistic (there are no artefacts arising from symmetry assumptions). Third, longitudinal sections can be taken through a 3D model that are directly comparable with microscopy z-stack images (see Figure 4.1B). Finally, the software developed as a result of this work will be applicable to a wider range of biological questions if it supports 3D simulations as well as 2D. We therefore work in full 3D to develop a model that is highly generalisable.

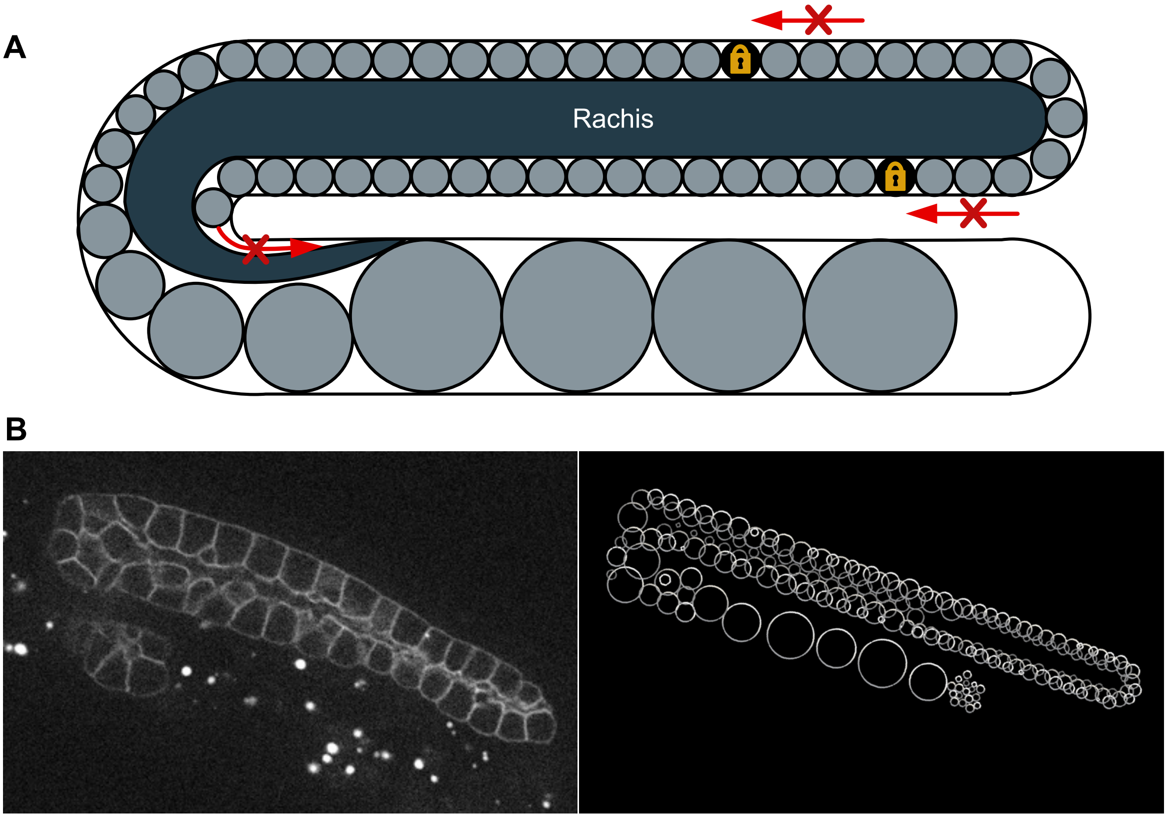


Figure 4.1: Modelling the germ line in two and three dimensions. A) Illustration of a germ line model that represents a 2D section through the gonad. If certain cells are anchored in place by adhesion (indicated by a padlock) then neighbouring cells are unable to move past. The rachis can also potentially obstruct the flow of cells. B) On the left is a z-stack microscopy image showing a slice through a late L4 *C. elegans* germ line. On the right is a slice through our 3D model, showing how the two images are comparable.

We represent the germ cell population using a cell centre model with spherical cells, as described in Section 3.1.3, since this approach is both computationally tractable in three dimensions, and sufficiently detailed to study the tissue scale effects emerging from individual cell behaviour. The germ cell population model was implemented using Chaste (Section 3.1.6) [171].

For the initial model, we chose to apply a relatively simple repulsion force between germ cells. This is identical to the GLS force described in Section 3.1.4, but with the additional constraint that only overlapping cells may interact. The decision to neglect attractive cell-cell forces at this stage reduces the number of difficult-to-determine force parameters in the model. Moreover, when cells overlap heavily in an overlapping spheres model, any attraction forces arising from Equation (3.10) are likely to be between next-nearest neighbours and as such spurious. The force on germ cell i due to germ cell j is therefore given by Equation (3.11),

$$\mathbf{F}_{ij} = \begin{cases} \mu\,(R_i + R_j)\log\left(1 - \dfrac{v_{ij}}{R_i + R_j}\right)\hat{\mathbf{r}}_{ij}, & (\text{if } v_{ij} > 0), \\ \mathbf{0}, & (\text{if } v_{ij} \leq 0), \end{cases}$$

where μ is a spring strength constant and v_{ij} is the length of the cell-cell overlap. Cell positions are updated based on the net force on each cell, in accordance with the overdamped form of Newton's second law (Equation (3.2)).

The *C. elegans* germ line contains a wide range of cell sizes, from small spermatazoa to mature oocytes that fill the entire proximal gonad tube. It would be unphysical for cells of such different sizes to move by the same amount in response to an identical force. We therefore made a minor adjustment to the usual position update equation by scaling the drag coefficient of each cell according to its radius, R_i,

$$\eta_i = \overline{\eta}\left(\frac{R_i}{5}\right). \tag{4.1}$$

Here, $\overline{\eta}$ is the default drag coefficient used in the Chaste library ($\overline{\eta} = 1\,\mathrm{kg\,h^{-1}}$) [67]. Scaling η according to cell radius implicitly assumes that germ cells experience Stokes' drag. However, this assumption is valid for small spheres moving through a viscous medium, and so is both appropriate in a biological context and consistent with our overdamping assumption [22].

The default drag coefficient in the Chaste library is intended to be suitable for modelling large cells: namely, human crypt epithelium cells, which are about 5μm in radius. The main effect of the scaling in Equation 4.1 is therefore to make smaller proliferative germ cells move more freely. If this scaling is removed, proliferative germ cells become slow to move apart and to fill the available space. Cells pack together in the distal gonad to a much greater extent, and the organ is incompletely filled with cells during the developmental part of a simulation. As we will see later in the chapter, this has an impact on the adult organ in our model, with the fully developed gonad being shorter and fewer spermatocytes produced.

4.1.2 Defining a growing gonad boundary

The second key component of our model is its boundary condition. The boundary must confine germ cells to lie inside the gonad membrane but outside the central cytoplasm reservoir or rachis. It must also evolve over the course of a simulation, to reflect the growth and development that take place during the L3 and L4 larval stages. Chaste already has a built-in mechanism for applying boundary conditions (Section 3.1.6). However, in previous work the boundary has usually been static or subject only to a specific class of deformations [63, 67]. Therefore, the main task is to implement a boundary condition capable of dramatically changing shape over time.

We model gonadogenesis from the beginning of the L3 larval stage onward, starting at around 18.5 hours post-hatching (hph) [269]. By this point, the reproductive system has already developed posterior-anterior symmetry [141]. Starting simulations from early L3 therefore means we can omit the complex cell rearrangements involved in establishing distinct gonad arms. It also lets us focus on a single U-shaped arm as representative of the entire organ.

The decision was made to model the DTC explicitly as a specialised cell, since the DTC acts as a leader and determines gonad shape by its migration. *In vivo*, DTC pathfinding is thought to be controlled by a combination of signalling gradients laid down in the worm body, and changing receptor expression patterns in the DTC itself over the course of development (Section 2.2). In our model, the DTC's position is updated at each time step according to a leader cell program, rather than by solving the usual equations of motion (Equation (3.2)) with the forward Euler method. The leader cell program prescribes the direction of DTC movement, and includes certain changes known to occur at specific times. Initially, it instructs the DTC to travel in a straight line, mimicking the outward migration along the ventral surface of the body that takes place during L3 (see Figure 4.2A). At 23 hph (around the time of the L3/L4 moult) the DTC is instructed to turn in a semicircle, representing the move from the ventral to the dorsal surface of the worm's interior. Finally, the DTC is instructed to travel in a straight line back into the centre of the body, with

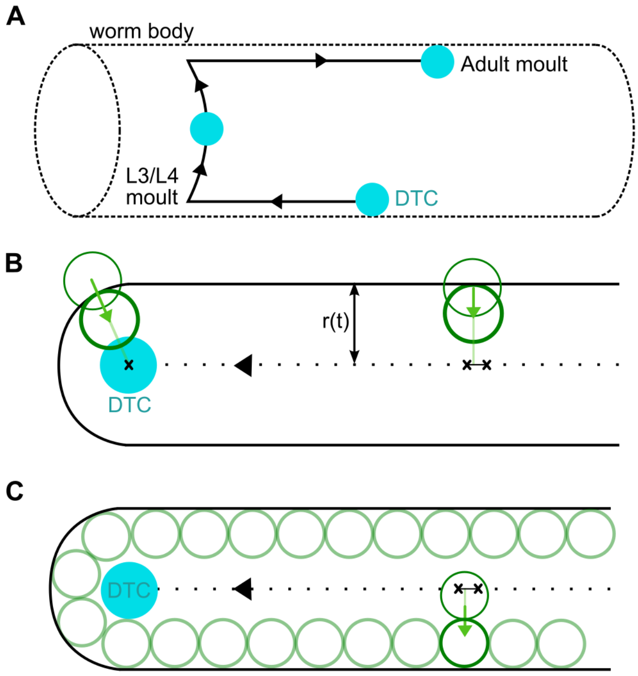


Figure 4.2: The growing gonad boundary condition. A) The movement of the DTC is explicitly modelled, and its path dictates the shape of the boundary. During L3, the DTC travels in a straight line, before turning in a semicircle at the time of the L3/L4 moult. It then migrates in a straight line back into the centre of the worm, halting at the time of the adult moult. B) Cells further than $(r(t) - R_i)$µm from the DTC path are considered outside the gonad (where $r(t)$ is the gonad radius and R_i is the cell radius). Cells outside the gonad are moved toward the closest point on the DTC path until they lie inside the boundary again. C) In the distal arm and turn, the rachis is enforced by making cell centres lie exactly $(r(t) - R_i)$µm from the DTC path. The rachis then appears as an excluded volume in the centre of the organ.

migration terminating at the time of the adult moult (35.5 hph [115]). Since the rate of DTC migration in microns per hour varies substantially over the course of larval development, we used four parameters to describe its movement speed at different stages: an early L3 rate (L_{L3E}), a late L3 rate (L_{L3L}), an early L4 rate (L_{L4E}), and a late L4 rate (L_{L4L}). During each simulation, a collection of evenly spaced points on the DTC's path are recorded, and together this set of points defines the organ midline.

The gonad boundary itself is modelled as a tube of radius $r(t)$µm surrounding the DTC path, with the gonad radius continually updated during the larval portion of the simulation. Again, because the rate of *C. elegans* gonad growth varies substantially over time, four different radial growth rate parameters are applied depending on the simulation stage: early L3 (r_{L3E}), late L3 (r_{L3L}), early L4 (r_{L4E}) or late L4 (r_{L4L}).

Germ cells further from the DTC path than $r(t) - R_i$µm are considered outside the gonad boundary (where R_i is the cell radius). The boundary is enforced by moving these cells toward the closest point on the DTC path, until they lie just inside the allowed region again (see Figure 4.2B). The result is a tube shaped gonad with hemispherical endcaps. In the distal and turn sections of the boundary, the rachis is modelled by forcing cells to lie exactly $r(t) - R_i$µm from the DTC path (Figure 4.2C). An excluded volume representing the rachis is thereby created in the middle of the organ, and this volume naturally narrows as maturing meiotic cells increase in radius. This approach of pushing cells to the surface of the gonad was preferred over specifying a solid excluded volume inside the organ, since, in preliminary tests, a fixed rachis tended to impede cell growth and movement around the turn. *In vivo*, the rachis sweeps to one side after rounding the turn. We do not model this feature, since *in silico* it can cause cells to become temporarily trapped between the rachis and the gonad membrane.

Stretching growth of the gonad

Two additional aspects of *C. elegans* gonadogenesis are taken into account in the boundary condition. Firstly, the experimental measurements made by our collaborators to help parametrise this model revealed that the proximal gonad grows significantly during late L4 (see Section 4.2.1 for details). By the late L4 stage of development, the DTC is moving in the opposite direction to proximal growth, and it is therefore impossible to attribute the elongation of the proximal gonad to DTC migration. Instead, the microscopy images suggest that the turn moves away from the center of the animal, causing the organ to “stretch”. This effect is captured in our model by allowing points on the semicircular turn section of the DTC path to shift leftward during late L4 at a fixed stretching rate of $S_{L4L}\,\mu\mathrm{m\,h}^{-1}$ (see Figure 4.3). New points are added into the resulting gap to maintain an even point spacing along the organ midline.

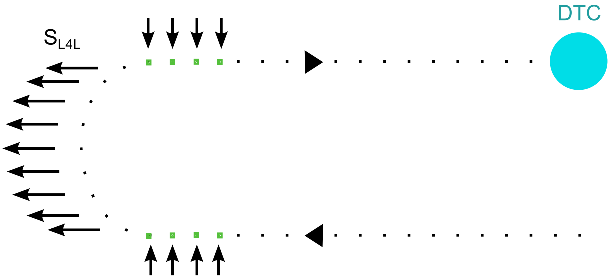


Figure 4.3: Stretching growth of the gonad. During late L4, the *C. elegans* gonad lengthens in a manner that cannot be attributed to DTC migration, but which instead seems to involve movement of the turn away from the centre of the animal. To capture this type of growth, points on the turn section of the DTC path are shifted leftward at a prescribed rate during late L4, with new points added in the gap to maintain equal spacing.

The influence of germ cell pressure on DTC migration

The second factor to take into account is that germ cell pressure inside the gonad is known to influence the migration of the DTC [139]. If fewer germ cells are present inside the organ, the DTC moves more slowly. Given the complexity of the rest of the model, we decided to represent this feedback on DTC migration in a simplified manner, through a mechanism we term "DTC halting". Within this mechanism, the DTC travels at a rate specified by the model parameters *unless* no germ cells are present within 5μm of the distal tip. If there are no germ cells within that range, the DTC halts until germ cells have proliferated to fill the gap behind it. This modelling choice also reflects the fact that the wild-type gonad should be filled with germ cells throughout development; no substantial gaps should appear.

An alternative way to implement the gonad boundary would have been to hard-code into the model the shape of the organ at a sequence of time points. Since Chaste is designed to be modular, it would still be easy to write an alternative boundary condition to replace the leader cell based boundary described above. However, we argue that explicitly modelling DTC movement brings certain advantages. For instance, it allows changes in DTC migration to realistically affect the morphology of the finished organ. If the DTC travels slowly during L3, the turn will still occur at the prescribed time and the proximal arm will be left short, as observed *in vivo* [139]. Similarly, if DTC migration is delayed during L4 the distal arm will be left short. A leader cell based boundary condition is also applicable to a wider range of scenarios, since it can represent mutant as well as wild-type behaviour. For example, the program controlling the DTC can be modified to mimic errors in pathfinding, by allowing the direction of DTC movement to change at random. The DTC will then weave around the body interior during simulations, forming an irregularly shaped adult gonad behind it.

4.1.3 A statechart model of germ cell behaviour

The final component of our model is a description of individual germ cell behaviour. Here, we represent cell behaviour using a statechart (Section 3.2.4), since statecharts offer a reasonable level of abstraction for the questions at hand, focussing on high-level processes and decision-making. Our statechart model consists of a collection of orthogonal regions handling different intracellular functions: a number of regions responsible for DTC signalling, a cell cycle region, and a sex determination region. The general design is based on work by Setty *et al.* (2012) [221],

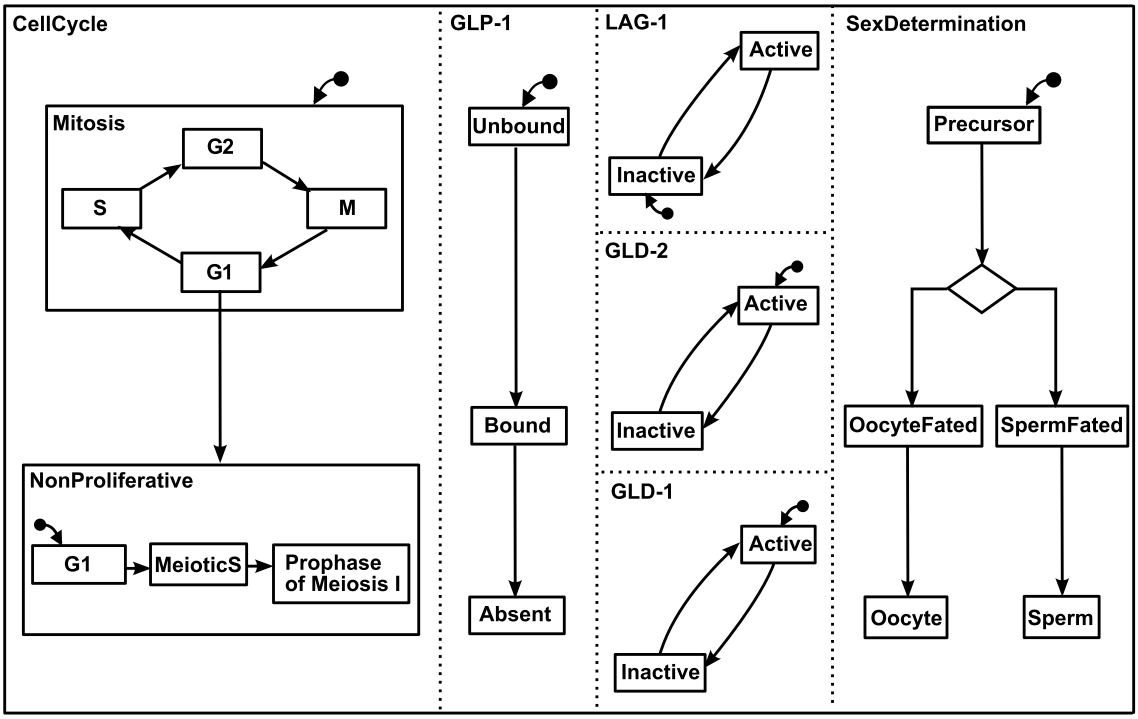


Figure 4.4: Overview of the germ cell statechart. The entire statechart is pictured above in a highly simplified form. Only states, regions, transitions and default transitions are shown, while transition conditions and actions are omitted. See Figures 4.5-4.7 for a complete diagram of each orthogonal region.

but we have streamlined the model by removing many molecular effectors that played no role in signal processing other than acting as intermediaries. Figure 4.4 gives an overview of the germ cell statechart. We will discuss each orthogonal region in turn, before briefly describing how a statechart model of cell behaviour was integrated into Chaste.

DTC signalling

The main orthogonal region responsible for responding to DTC signals is the GLP-1 region (Figure 4.5). This controls the state of a germ cell's extracellular GLP-1 receptors. All germ cells begin each simulation with GLP-1 unbound. Whenever a cell is within 35μm of the DTC, its GLP-1 receptors become bound, reflecting the interaction that occurs between GLP-1 on germ cells and LAG-2/APX-1 ligands displayed by the DTC. 35μm, or 5-6CD, is an estimate of the maximum distance from the distal tip at which germ cells can closely contact the DTC and experience binding interactions [47, 154]. As a result of GLP-1 binding, LAG-1 becomes activated and GLD-1 and GLD-2 inactivate. The inactivation of the GLD proteins goes on to affect other regions of the statechart, ultimately preventing meiosis and causing the cells close to the DTC to proliferate.

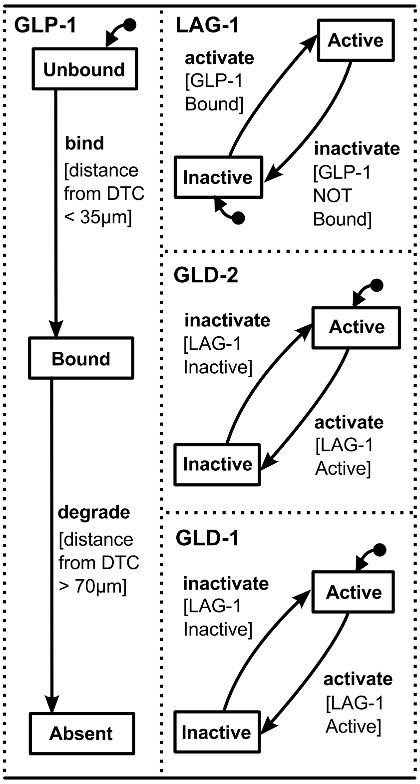


Figure 4.5: Response to DTC signalling. DTC signals are processed by four statechart regions: GLP-1, LAG-1, GLD-1 and GLD-2. The extracellular receptor GLP-1 is initially unbound, and becomes bound in cells close enough to the DTC to interact with its displayed ligands. GLP-1 binding causes LAG-1 to activate, switching off the two GLD proteins and ultimately preventing meiosis in distal cells. GLP-1 remains bound in germ cells and their daughters up to 70µm from the DTC, beyond which point the receptor is degraded. In the absence of GLP-1, LAG-1 becomes inactive, GLD-1/2 are activated, and meiosis is eventually triggered.

On division, daughter germ cells inherit the state of their parent's statechart. As a result, cells with bound GLP-1 produce daughters with bound GLP-1, even if they are more than 35µm from the DTC. This "memory effect" was found to be necessary in previous germ line models to ensure maintenance of the proliferative zone [221]. Once a germ cell moves more than 70µm from the DTC, its GLP-1 state switches to absent, representing degradation of the receptor. A 70µm distance was chosen because, during larval development, the first cells to enter meiosis do so at around 70µm from the distal tip [114, 221]. Our statechart therefore implements a simple threshold model of DTC signalling. Once GLP-1 is absent, LAG-1 becomes degraded in germ cells, GLD-1 and GLD-2 switch on, and the decision to enter meiosis is ultimately triggered.

Cell cycle model

In the cell cycle region of the chart (Figure 4.6), all germ cells start each simulation in the mitosis state. Mitotic cells cycle through the substates G1, S, G2 and M, with a division action occurring at

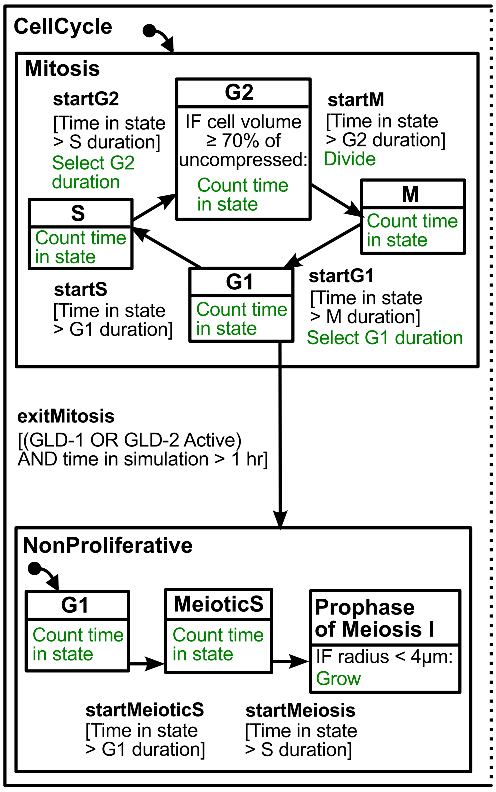


Figure 4.6: The germ cell cycle model. Germ cells begin each simulation in mitosis and cycle through the substates G1, S, G2 and M; dividing at the beginning of M phase. Cell cycle phase transitions occur after a predetermined amount of time has elapsed, with the lengths of G1 and G2 incorporating some random variation. Cells in G2 can also undergo a transient arrest in overcrowded conditions. Exit from proliferation is only possible from the G1 phase, and is triggered by the activation of GLD proteins in cells over 70µm from the DTC. Non-proliferative cells then undergo a fixed length G1 and meiotic S, before entering meiosis proper and growing up to a radius of 4µm.

the beginning of M phase. Every cell cycle phase has a predetermined length, and a transition into the next phase takes place once the appropriate amount of time has elapsed. We denote the mean durations of the four cell cycle phases by $dG1$, dS, $dG2$ and dM. Mean phase lengths increase substantially at around the time of the adult moult, taking the total cell cycle length from 3 hours to 8 hours (see Section 4.2.2 on choice of parameters). This reflects the fact that larval germ cells cycle faster than adult ones in order to expand the germ cell pool [50, 114, 145].

To prevent extensive cell cycle synchronisation, which is not supported by the data [114], a random initial phase is chosen for each germ cell at the beginning of the simulation. This is done by assigning cells an age sampled from a uniform distribution, with $0 \leq age \leq dG1+dS+dG2+dM$. A cell's age then determines its starting phase and the time already spent therein; with an age of 0 corresponding to a cell at the beginning of G1, an age of $dG1$ corresponding to a cell at the beginning of S, and so on. To further mitigate synchronisation, the lengths of the G1 and G2 phases are allowed to vary by a small amount about the mean. The G1 duration is chosen at the beginning of each G1 phase by sampling from a normal distribution with mean $dG1$ and variance $sdG1$, where s is a stochasticity parameter set to 0.1. A similar procedure is followed for G2. The precise amount of variation that occurs in *in vivo* phase lengths is unknown. We therefore introduced only a small amount of stochasticity, just enough to prevent extensive synchronisation between daughter cells in trial-and-error tests.

The final noteworthy aspect of the mitosis model is that it includes mechanical feedback on proliferation akin to contact inhibition [2], such that germ cells subject to heavy crowding undergo a temporary cell cycle arrest. Specifically, we calculate a “compressed” volume for each cell, based on its overlaps with its neighbours. This is not necessarily true compression, but in an overlapping spheres model it acts as a measure of crowding or deformation. Cells in G2 phase with a compressed volume less than 70% of their rest volume stop making progress through the cell cycle. Progress through G2 phase does not resume until levels of crowding are reduced. Here, the rest volume of cell i with radius R_i is considered to be $(4/3)\pi R_i^3$. The compressed volume, meanwhile, is $(4/3)\pi R_{ic}^3$, where R_{ic} is a “compressed” radius taking into account overlaps. Briefly, in an efficient 3D packing, spherical cells should have twelve neighbours. If cell i has twelve or more overlapping neighbours then

$$R_{ic} = \frac{1}{N}\sum_n \left[R_i - 0.5(R_i + R_n - d_{in})\right],$$

where N is the total number of neighbours, n ranges over all neighbouring cells, and d_{in} is the distance between the centres of cells i and n. For a cell with fewer than twelve neighbours, a correction term is applied to account for the extra space available

$$R_{ic} = \frac{1}{12}\left[\sum_n \left[R_i - 0.5(R_i + R_n - d_{in})\right] + (12 - N)R_i\right].$$

See also [67] and the supplementary methods section of [15] for more information. We chose to apply this temporary arrest in G2 phase, since G1 is known to be short in *C. elegans* germ cells [81, 82], and G2 arrests have already been documented under starvation conditions [85]. We discuss the motivation for introducing a contact inhibition-like mechanism in Section 4.3.2.

Exit from mitosis is modelled as a one way transition occurring in cells with active GLD-1 or GLD-2. Such cells are sufficiently far away from the DTC that they no longer experience GLP-1 signalling. Germ cells are not permitted to exit mitosis during the first hour of the simulation, in order to allow time for DTC signalling to inactivate GLD in proliferative zone cells. Otherwise, a number of germ cells would exit mitosis immediately during the first time step, as GLD proteins are active initially by default.

Exit from mitosis is further restricted to occur from the G1 state. This reflects the fact that *C. elegans* germ cells are thought to undergo a specialised S phase in preparation for meiosis [123]. Therefore, the decision to switch from mitosis into meiosis must occur *prior* to S phase, and we chose to place it in G1. Recent experimental evidence is consistent with a G1 meiosis decision [81], though the choice could also take place during the G2 phase of the previous cell cycle, with the differentiation decision being communicated to both daughter cells.

After exiting mitosis, non-proliferative germ cells undergo a G1 with fixed length $dG1$, and a meiotic S phase with fixed length dS before beginning meiosis proper. Meiotic cells then increase in radius at a linear rate g, up to a maximum of 4µm, reflecting the generally larger size of meiotic germline cells [155, 180].

Sex determination

The final, sex determination region of the chart is responsible for controlling germ cell development into mature gametes (Figure 4.7). All germ cells start each simulation in the precursor state, meaning they are uncommitted to either male or female fate. Germ cells remain precursors until they are 200µm from the DTC, at which point a sex determination decision occurs.

The exact signal responsible for triggering sex determination *in vivo* remains unclear. One proposed mechanism is that germ cells choose their fate on entry into meiosis, with cells entering meiosis before a certain time becoming sperm, while subsequent meiotic cells become oocytes. Implementing this mechanism in our model generated a mixture of sperm and oocytes all along the proximal arm, rather than distinct regions containing each type of gamete. We therefore switched to a distance-based trigger for sex determination with a 200µm threshold, which did produce sperm at the proximal end of the gonad only, in a region separate from developing oocytes. Note that we do not claim sex determination is distance-based *in vivo*, only that it is a useful first approximation for modelling purposes.

On reaching 200µm from the DTC, germ cell fate is determined by the current simulation time. Cells reaching the decision distance before 32.5 hph become sperm-fated, while cells arriving at later times become oocyte-fated. The switching time of 32.5 hph was chosen based on experi-

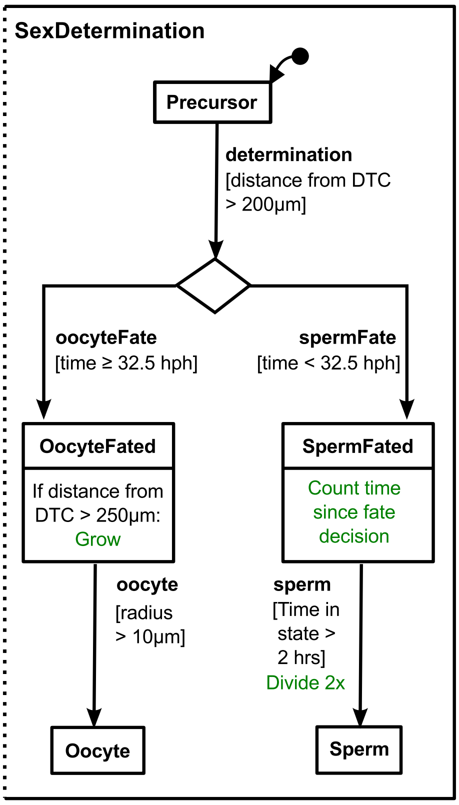


Figure 4.7: The sex determination model. Germ cells start each simulation as precursors, uncommitted to a particular fate. On reaching 200µm from the DTC, cells undergo sex determination based on the current simulation time. Prior to 32.5 hph cells become sperm-fated, while later cells become oocyte-fated. Sperm-fated cells take 2 hours to develop before dividing twice to give four mature sperm. Meanwhile, oocyte-fated cells are subject to the possibility of apoptosis (see text). Surviving oocyte-fated cells that reach 250µm from the DTC grow at a fixed rate, eventually becoming mature oocytes.

ments by Barton and Kimble [21]. These showed that the period prior to 32.5 hph is critical for spermatogenesis, and that if sperm production is blocked up to this time it cannot be recovered later in larval development.

Sperm-fated cells in our model experience a 2 hour delay before dividing twice to give four mature sperm. Oocyte-fated cells, meanwhile, grow steadily at rate g on entering the proximal arm, and are considered mature on reaching a radius of 10µm. Oocyte-fated cells are also subject to the possibility of apoptosis. Like all cell death events, this process is handled by the wider Chaste simulation and not by the intracellular statechart, so it does not appear in Figure 4.7. A cell killer is included in the simulation, which eliminates oocyte-fated germ cells with a probability $P_{death} = 0.025$ for every hour spent outside the proximal arm. The only other process that removes germ cells from the model is fertilisation. A fertilisation cell killer is included that removes one oocyte and one sperm from the system, whenever both types of gamete are present in the final 25µm of the gonad.

Integrating statecharts into Chaste

The statechart model described above was implemented in C++ using the Boost Statechart library. Normally, Chaste expects each cell to have an associated cell cycle model, responsible for determining the timing of divisions along with other intracellular behaviour. We therefore wrote a new cell cycle model class that wraps around a Boost statechart and delegates most decision-making to the chart. Instead of the cell cycle model itself executing some logic at each time step, it simply prompts the statechart to update. Conversely, when the statechart needs to perform an action, it can do so by setting variables of the cell cycle model. To allow active states to be visualised, state information is written into the existing "CellData" object, which Chaste automatically outputs in a suitable format for viewing.

The main difficulty we encountered involved copying the state of a statechart. In non-biological applications, newly created statecharts are expected to enter into their default initial state. In contrast, we wanted the state of a daughter cell to match that of its parent. This was ultimately achieved by looping through every possible state and querying whether it was active in the parent statechart. If so, the daughter statechart was forced to transition into the same state directly. Archiving (saving) state information was achieved by encoding the collection of active states as a bitset, which can then be serialized using Boost. On loading, each 1 in the bitset causes a transition to occur, forcing the loaded chart into the appropriate state.

We have now described the design of our *in silico* germ line model: its cell population model, boundary condition, and representation of cell behaviour. In the next section, we detail our chosen parameter set and how these values were selected.

4.2 Choice of parameter set

Our *in silico* model has a large number of parameters, not only in the intracellular statechart, but also in the dimensions and evolution of the boundary condition. Fortunately, many of the size and growth rate parameters can be determined from microscopy images. This section describes some new experimental measurements that were made, as well as how the remaining parameter values were chosen.

4.2.1 Measuring larval gonad dimensions and cell radii

The typical dimensions of the *C. elegans* gonad were determined based on a collection of DIC microscopy images provided by our collaborators. A number of wild-type worms were grown at 20°C in abundant food conditions. Individual animals were then selected and classified as either L3, L3/L4, mid L4 or young adult depending on the appearance of the vulva. Gonad images were subsequently captured and measured using the software ImageJ [214]. Table 4.1 summarises the results, while Figure 4.8 shows precisely how measurements were made. To ensure consistency between images, measurements were made at well-defined points in the gonad close to identifiable structures. So, in larval images, the gonad radius was measured next to the turn, while in young adult images the radius was measured level with the proximal-most oocyte (the adult gonad constricts near the turn, making that region unrepresentative of the rest of the organ).

Growth rates were calculated from the measurements in Table 4.1 by subtracting the length at the beginning of a developmental stage from the length at the end and dividing by the duration. So, for instance, the rate of radial growth between 22 hph and 26 hph was set to $0.153\,\mu\mathrm{m}\,\mathrm{h}^{-1}$. For the linear growth rate during late L4 the calculation is slightly more complex, since we want to distinguish between growth due to DTC migration and growth due to movement of the turn (see Section 4.1.2). We therefore made the assumption that any change in proximal arm length over this period is associated with turn movement, and that the distal arm "stretches" by an equal amount.

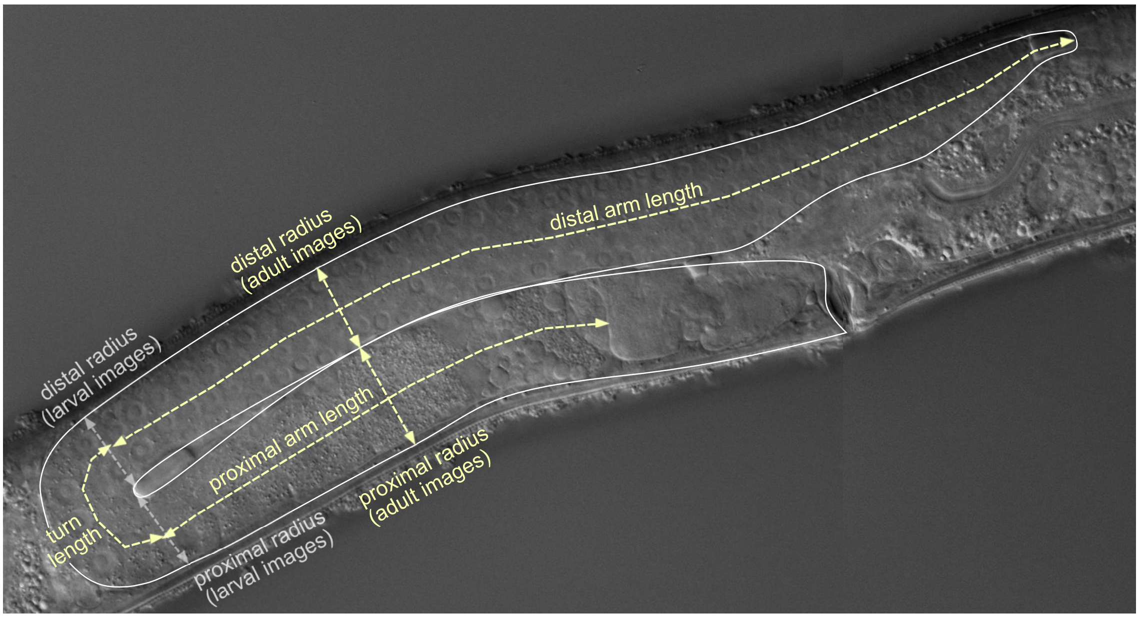


Figure 4.8: Positions at which gonad measurements were made. Corresponds to the lengths in Table 4.1. The microscopy image above was taken by our experimental collaborators (Section 1.3) and shows a young adult animal.

Measurement	**Mid L3**	**L3/L4 moult**	**Mid L4**	**Young adult**
Approximate hours post-hatching	22	26	31	35.5
Sample size	10	10	15	6
Length of proximal region (μm)	61.7(9.33)	56.1(6.16)	53.0(12.3)	126(21.5)
Length of distal region (μm)	NA	12.3(4.02)	118(17.8)	248(25.6)
Length of turn (μm)	NA	23.1(2.95)	27(2.59)	30.8(4.86)
Total length (μm)	61.7(9.33)	91.4(7.76)	198(23.3)	405(46.7)
Proximal radius (μm)	5.48(1.36)	6.37(1.02)	7.73(1.62)	12.3(3.17)
Distal radius (μm)	NA	5.80(0.939)	8.21(1.24)	10.3(1.06)
Mean radius (μm)	5.48	6.09	7.97	11.3

Table 4.1: Measurements of *C. elegans* gonad dimensions. All reported lengths are means, with the standard deviation given in brackets. Note that much of the variation is due to small age differences between worms at the same developmental stage. Approximate hours post-hatching taken from [269].

The remaining increase in gonad length is then attributed to DTC migration. This gives an L4 growth rate of 16.2μm$\,\mathrm{h}^{-1}$ due to “stretching” growth and 13.6μm$\,\mathrm{h}^{-1}$ due to movement of the DTC. Table 4.2 lists the resulting parameter values used in the model, while Figure 4.9 illustrates gonad growth rates over time.

Parameter	**Value**	**Notes**
Initial gonad length	32μm	Based on [115]
DTC migration rate: 18.5-22 hph (L_{L3E})	8.77 (2.68)μm$\,\mathrm{h}^{-1}$	
DTC migration rate: 22-26 hph (L_{L3L})	7.43 (3.05)μm$\,\mathrm{h}^{-1}$	
DTC migration rate: 26-31 hph (L_{L4E})	21.3 (4.91)μm$\,\mathrm{h}^{-1}$	
DTC migration rate: 31-35.5 hph (L_{L4L})	13.6 (12.8)μm$\,\mathrm{h}^{-1}$	
Stretching growth rate: 31-35.5 hph (S_{L4L})	16.2 (5.51)μm$\,\mathrm{h}^{-1}$	
Initial gonad radius	5.48 (1.36)μm	
Radial growth rate: 18.5-22 hph (r_{L3E})	0μm$\,\mathrm{h}^{-1}$	Estimated, based on the lower radial growth rate observed early in larval development
Radial growth rate: 22-26 hph (r_{L3L})	0.153 (0.0626)μm$\,\mathrm{h}^{-1}$	
Radial growth rate: 26-31 hph (r_{L4E})	0.367 (0.207)μm$\,\mathrm{h}^{-1}$	
Radial growth rate: 31-35.5 hph (r_{L4L})	0.74 (0.392)μm$\,\mathrm{h}^{-1}$	
Timing of the gonad turn	23 hph	Based on [115], roughly coincides with the L3/L4 moult
Radius of the semicircular gonad turn	11.5μm	Adult proximal and distal arms should almost touch

Table 4.2: Parameter values relating to the gonad boundary. Values are derived from the measurements in Table 4.1 except where otherwise stated. Standard deviation in brackets.

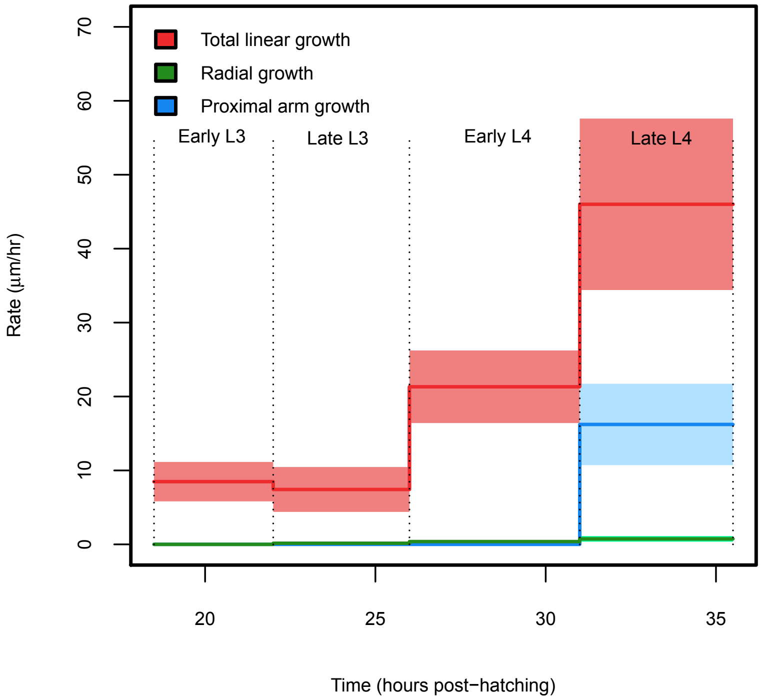


Figure 4.9: Experimentally determined rates of gonad growth over time. Shown above are the total linear growth rate (red), radial growth rate (green) and proximal arm growth rate (blue) for the *C. elegans* gonad over time, according to our measurements. Shaded regions are ±1 standard deviation.

A second set of images was used to determine the radius of a mitotic germ cell, which is an important parameter for modelling the proliferative zone. 1µm confocal z-stacks were captured of a number of live *xnSi1* animals; a strain that expresses a GFP marker on germ cell membranes [43]. In each image, a number of randomly chosen germ cells were measured within the first 10-15 rows from the DTC. Each cell's radius was measured in the image where it appeared widest, in an attempt to choose the part of the z-stack passing directly through the cell centre. Table 4.3 summarises our results. There is some imprecision involved in these measurements, since tightly packed germ cells tend to adopt a square shape, making it difficult to identify the relaxed radius. The feasible range for this parameter was determined to be 2.5µm to 3.5µm, and we chose to average all the results in Table 4.3, obtaining a mitotic germ cell radius of 2.8µm.

Measurement	L3	L3/L4	L4	Young adult
Number of animals imaged	4	3	4	2
Total cells measured	20	20	20	20
Mean radius (µm)	2.70	2.60	2.92	3.16

Table 4.3: Proliferative germ cell radius measurements. Based on measuring distal germ cells in a number of fluorescence microscopy z-stacks. A mean of 20 measurements is reported for each developmental stage, and the mean over all stages was used as a parameter in the model.

4.2.2 Other parameter choices

Other parameters in our germ line model could not be determined from microscopy images. In those cases, we relied on estimates gathered from the literature wherever possible. Table 4.4 lists the parameter values used in our wild-type germ line model, along with the source or rationale behind each choice. Parameters marked with an asterix are considered free parameters, with values that could not be fixed based on the information available.

Due to the long run time of this model (in the neighbourhood of 10 hours for a simulation up to the point of sperm depletion) it was not possible to carry out an extensive exploration of the parameter space. Stochastic approaches, such as Markov Chain Monte Carlo parameter space exploration, require thousands of simulations, which would certainly be infeasible in this case. Shorter runs, meanwhile, do not provide enough information to assess whether appropriate adult germ line behaviour is achieved. Attempts at writing an automated program to vary and evaluate parameters also failed, due to discontinuous or otherwise problematic changes in model behaviour. For example, if a moderate variation in some parameter causes no sperm to be produced, then fertilisation cannot occur, few cells are removed from the system, and the simulation will run extremely slowly as germ cells accumulate inside the gonad. Similarly, if a parameter variation leads to extensive overcrowding, large net forces and infeasibly rapid cell movement will occur, and the simulation will halt early yielding no useful data. Finally, it is difficult to make cell row measurements in an automated fashion, as that entails writing an algorithm that can reliably partition the cell population into what a human would consider rows.

For these reasons, we instead varied the values of the free parameters manually, and attempted to obtain as close an agreement as possible with available quantitative data on the germ line. Section 4.3.2 discusses the fit of our current model to this data, while Section 4.3.6 explores the local behaviour of the model about our chosen parameter set. In general, this model aims to be as quantitative and predictive as the available data will allow, and we would argue that it achieves those aims at least as well as other modelling work on the same system [25, 99, 221]. Many parameters (such as gonad and cell dimensions) *are* heavily constrained, thanks to the availability of high quality experimental results. However, until fully quantitative studies become the norm in this field, there is perhaps not yet enough information to justify stochastic parameter space exploration and fitting. As such, the role of our model is to help further simplify the system by identifying the most important free parameters to focus on (via local sensitivity analysis). It should also allow biological hypotheses to be tested quantitatively, highlighting any fundamental problems.

4.2.3 Initial conditions

We have already stated many of the parameter values relating to the beginning of a simulation. However, for clarity, we now summarise the initial conditions in full. At the beginning of each simulation the gonad boundary is given by a cylinder with hemispherical endcaps of length 32μm and radius 5.48μm. 16 germ cells are present arranged in a row along the midline (Figure 4.10). The initially active states of each cell's statechart are marked by a default transition symbol (•↷) in Figures 4.5-4.7. Since all germ cells start within 35μm of the DTC, all germ cells are initially proliferative.

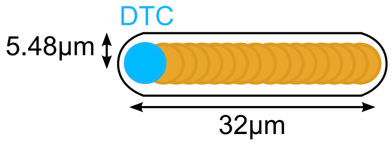


Figure 4.10: The initial state of a simulation. The starting gonad boundary dimensions are based on our new measurements, and 16 germ cells are present arranged in a row along the midline.

Parameter	**Value**	**Source**	**Notes**
Cells			
Initial germ cell count	16	[115]	
Meiotic cell radius	4µm	[155, 180]	Based on images in these publications.
Oocyte radius	10µm		Oocytes should almost fill the proximal arm.
Sperm radius	1.5µm	[224]	Based on images in this publication.
Cell growth rate (g)	1µm	*	
Cell cycle			
Larval cell cycle duration	3h	[145] *	Choice also influenced by larval cell counts reported in [115].
Total adult cell cycle duration	8h	[82]	
Cell cycle phase breakdown	$dG1$ = 2% of total dS = 57% of total $dG2$ = 39% of total dM = 2% of total	[82]	
Stochasticity (s)	0.1		Arbitrary, prevents synchronised divisions.
Timing of the switch from larval to adult cell cycle	31-35.5h	[147]	After mid L4 but prior to the adult moult.
Threshold volume for contact inhibition	70% of rest volume		Arbitrary threshold chosen to test a general contact inhibition hypothesis.
Force law			
Baseline drag coefficient (η)	1kg h^{-1}	[67]	
GLS force law spring coefficient (μ)	50kg h^{-2}	*	
Other values			
Distance from the DTC at which the GLP-1 signal becomes absent	70µm	[115]	Approximate position at which the first meiotic cell appears during development.
Distance from the DTC at which sex determination occurs	200µm	*	
Oogenesis/spermatogenesis switch time	32.5h	[21]	End of the *fog-1* temperature sensitive period.
Delay before producing mature sperm	2h	*	
Position at which oocyte growth begins	250µm		Just beyond the start of the proximal arm. Allowing oocyte growth earlier in the gonad results in large cells blocking the turn.
P_{death}	0.025 (probability h^{-1})	*	Produces sustained ovulation from young adulthood with a reasonable rate.

Table 4.4: Model parameters not determined from microscopy images. * Indicates parameters that were tuned to fit the experimental data in Section 4.3.2, Figure 4.12.

4.3 Results

Having described the model and its parameters, we now present some simulation results. First, we make qualitative comparisons between simulation output and the wild-type *C. elegans* germ line, based on general appearance and behaviour. Next, we compare quantitative simulation results with data on a range of germline properties taken from the literature. We then discuss some biological insights obtained from the model development process. Finally, we look at the effect of varying certain parameters, focussing on those with the least available data to constrain their value.

4.3.1 Example simulation output

Figure 4.11 shows snapshots taken from a representative simulation (a corresponding video is available at http://tinyurl.com/EleganGermLine). The simulation was carried out using the parameter set in Section 4.2. Certain special features discussed in Sections 4.1.2 and 4.1.3 were also incorporated, namely: DTC halting; stretching gonad growth during late L4; and contact inhibition for adult proliferative germ cells.

Qualitatively, the simulated germ line resembled the wild-type *C. elegans* organ throughout development and into adulthood. The correct high-level organisation of the germ line emerged, with the establishment of a proliferative zone, the appearance of meiotic cells at the appropriate time, and the production of gametes in roughly correct numbers, with sperm spatially separated from oocytes. In contrast to the clear separation between sperm-fated and oocyte-fated cells, the border between the proliferative and meiotic zones exhibited a mixed transition region similar to the real germ line; though in our simulations the transition zone is present from the larval stages rather than in the adult only. A transition zone emerges because the model restricts meiotic entry to a particular cell cycle phase. As a result, some germ cells can enter meiosis as soon as they are far enough away from the DTC, while others must wait until they reenter G1. The organisation of the germ line into zones was stably maintained for several simulated days, until the supply of self-sperm was exhausted. At that point, fertilisation became impossible and oocytes began backing up in the proximal arm, a problem which also occurs *in vivo* in older animals. Importantly, we were able to achieve this level of fit using biologically reasonable parameters and gonad dimensions, and without the need to introduce a bias in the direction of germ cell movement (as in previous work [221]).

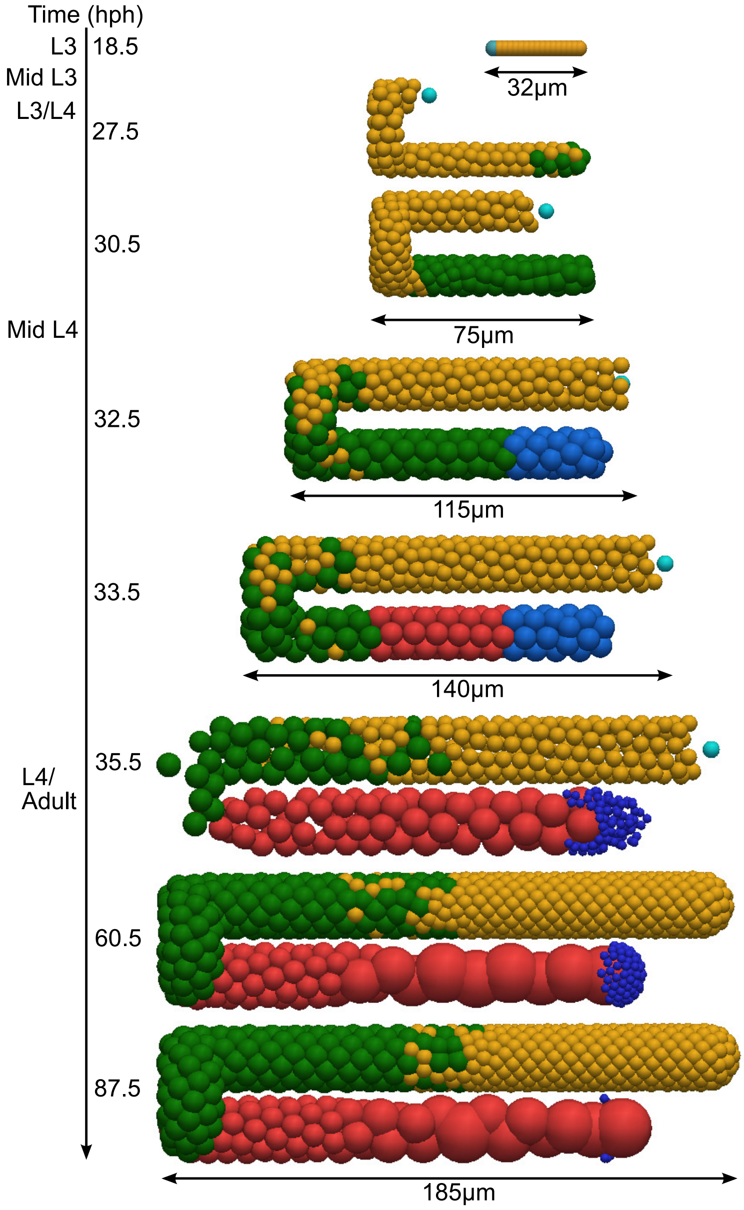


Figure 4.11: Snapshots of a representative germ line simulation, captured throughout the larval period and at two late adult time points. In the above images, germ cells in the mitosis and meiotic S states are coloured yellow, while cells that have begun meiosis proper are coloured green. Oocytes and oocyte-fated cells are pink, sperm-fated cells are pale blue, and mature sperm are dark blue. The DTC nucleus, where visible, is cyan. The parameter set is given in Section 4.2.

The main area in which our results do not yet match the *C. elegans* germ line is the gonad turn. In adulthood, the turn in our simulations appears overfilled with germ cells that are smaller than expected. This area of the simulation could perhaps be improved by beginning oocyte growth earlier, although when we tested this in preliminary simulations cells tended to get stuck in the turn. The real *C. elegans* gonad also varies in radius, becoming wider in the middle of the distal section and narrower near the turn and the DTC; a fact which we neglected for the sake of simplicity in this initial model. Since a wider rachis accounts for most of the greater width of the distal gonad, we expect this simplification will not greatly affect the accuracy of cell counts produced by our model, except perhaps in the turn and at the distal end. In short, the model represents a compromise between capturing all the details of the system, and tractability.

4.3.2 Comparison with quantitative data

To obtain a more in-depth evaluation of our model's behaviour, we recorded quantitative data relating to a number of germline properties, then compared our results with measurements available in the literature. Figure 4.12 shows the output of our model for six different properties (mean of 30 simulations, with a region ±1 standard deviation shaded). Also plotted on each graph is the experimentally observed value of the property as reported in the literature (a solid red line indicates time series data, while a dashed red line indicates a measurement at a single time point).

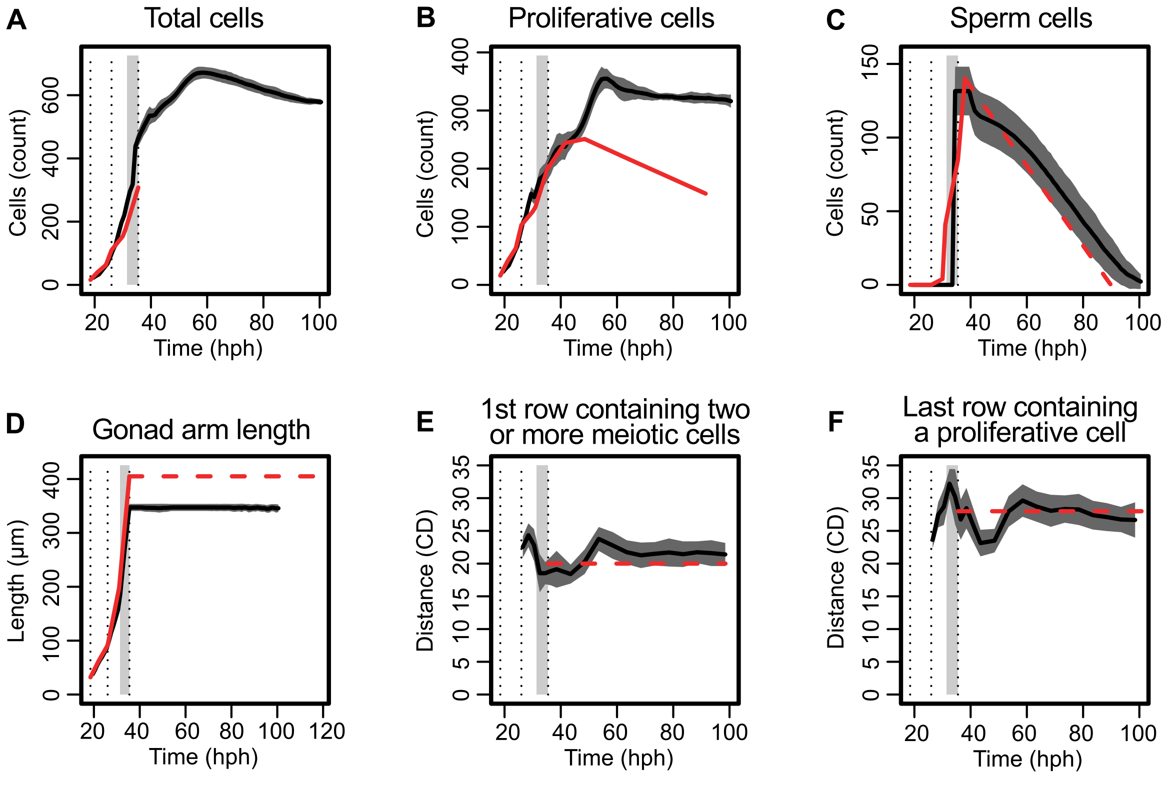


Figure 4.12: Comparing quantitative model output to literature values for a range of germline properties. In each graph, a solid black line gives the simulated result (mean of 30 runs with a region ±1 standard deviation shaded). Red lines represent experimental data, with a solid line indicating that time series data were available, and a dashed line indicating a single reported value. Vertical dotted lines mark the beginning of L3, L4 and adulthood. Finally, a light gray bar indicates the period where our model switches from larval to adult behaviour and parameters. A)-C) Counts of total germ cells, proliferative cells and sperm, respectively, against time in hph. D) Total length of the gonad arm in microns. E) Length of the proliferative zone in CD. F) Distance in CD from the DTC to the last row containing a proliferative cell. Experimental data taken from [102, 115, 139, 162].

During the larval period, a good agreement was obtained between simulated and expected cell counts (Figure 4.12A-C). The total number of cells, the number of proliferative cells, and the sperm count all increased at a realistic rate. This suggests the model is using reasonable values for the parameters governing the larval cell cycle. For the majority of the larval period, the length of the simulated gonad also matched experimental measurements well (Figure 4.12D). This need not be the case, because our model includes a DTC halting mechanism, which pauses leader cell migration if insufficient germ cells are present to fill the organ. Again, the implication is that the germ cell cycle parameters used in our model provide a sufficient proliferation rate to support normal gonadogenesis.

The precise length of the germ cell cycle is a controversial parameter in the field, with estimates of the adult cell cycle length ranging from 8 to 24 hours [50, 81, 82, 90]. In the absence of long term live imaging, these estimates are derived from pulse chase labelling experiments, and disagreement may arise from the use of different labelling techniques. During the L3 period, when all germ cells are within the proliferative zone, it *is* possible to use experimental cell counts to calculate that the L3 germ cell cycle length must be approximately 3 hours. However, once cells begin to exit the proliferative zone and differentiate, it becomes much more difficult to relate observed cell counts to the underlying cell cycle length without the help of a computational model.

Turning to the adult part of the simulation, our model maintained a stable number of proliferative cells over several days (see Figure 4.12B). This does not exactly match the experimental data, as we made no attempt to capture the effects of ageing. Qin *et al.* (2015) recently showed that ageing causes a more rapid decline in proliferative germ cell count than was previously supposed [197]. In this respect, therefore, our model behaves more like a germ line permanently in the young adult stage. The sperm count in the simulated germ line declined steadily as a result of fertilisation and ovulation, and a realistic ovulation rate was maintained until all self-sperm were depleted (Figure 4.12C). In terms of gonad length, the final adult gonad in our simulations was shorter than average, but close to the normal range for an animal at 35.5 hph (see measurements in Section 4.2.1). The model may be producing a shorter adult organ due to our decision to strictly halt gonadogenesis at a specific time point. While DTC migration is known to stop at around the time of the adult moult, it is unclear whether feedback mechanisms exist that may slightly delay or hasten the end of DTC movement to produce an organ of the appropriate length.

Finally, we also recorded the length of the proliferative zone in cell diameters (CD) from the DTC. Proliferative zone length is an important quantity in the field, since it is easy to measure and relates to the size of the stem/progenitor pool. The proliferative zone is defined as the area distal to the first row containing two or more meiotic cells. We therefore recorded the first row in which two or more meiotic cells appeared (Figure 4.12E), as well as the final row containing a proliferative cell (Figure 4.12F). Both properties were close on average to the expected values for a young adult germ line. However, some fluctuation in proliferative zone length occurred between 40 and 50hph, which may be attributed to the transient effect of switching from larval to adult model parameters.

Having compared our model to the real biological system both qualitatively and quantitatively, we go on to discuss the insights gained from running germ line simulations. In many cases, it was the process of developing the computational model that highlighted gaps in our knowledge, leading to new hypotheses. Firstly, we underline the significance of “stretching” growth during late L4 gonadogenesis, and we consider some potential mechanisms. Secondly, we discuss the need for feedback to help prevent germ line overproliferation, and we propose a contact inhibition-like mechanism as an initial hypothesis. Finally, we present the results of *in silico* cell tracking and labelling experiments; a type of investigation that is currently technically infeasible *in vivo*.

4.3.3 “Stretching” is an essential component of L4 gonadogenesis

Our model of gonadogenesis changed significantly over the course of this investigation, in light of early simulation results and new experimental data. Initially, we sought to keep our model of gonadogenesis simple, by prescribing a fixed rate of DTC migration for each larval stage, sufficient to reproduce the experimentally observed increase in length. All gonad growth occurred as a result of leader cell migration, with new gonad tube forming behind the moving DTC. We found that, in order to account for the large increase in length observed in late L4, the DTC would need to travel at a rate of $46\mu\text{m}\,\text{h}^{-1}$. However, when this rate was applied in a simulation, a biologically unrealistic gap quickly appeared behind the DTC (see Figure 4.13A). The germ cells in the proliferative zone did not divide fast enough to fill the space created by rapid DTC migration. As a result, germ cells fell behind the DTC by more than 70μm, leaving them outside the GLP-1 signalling range. Ultimately, the proliferative zone was lost.

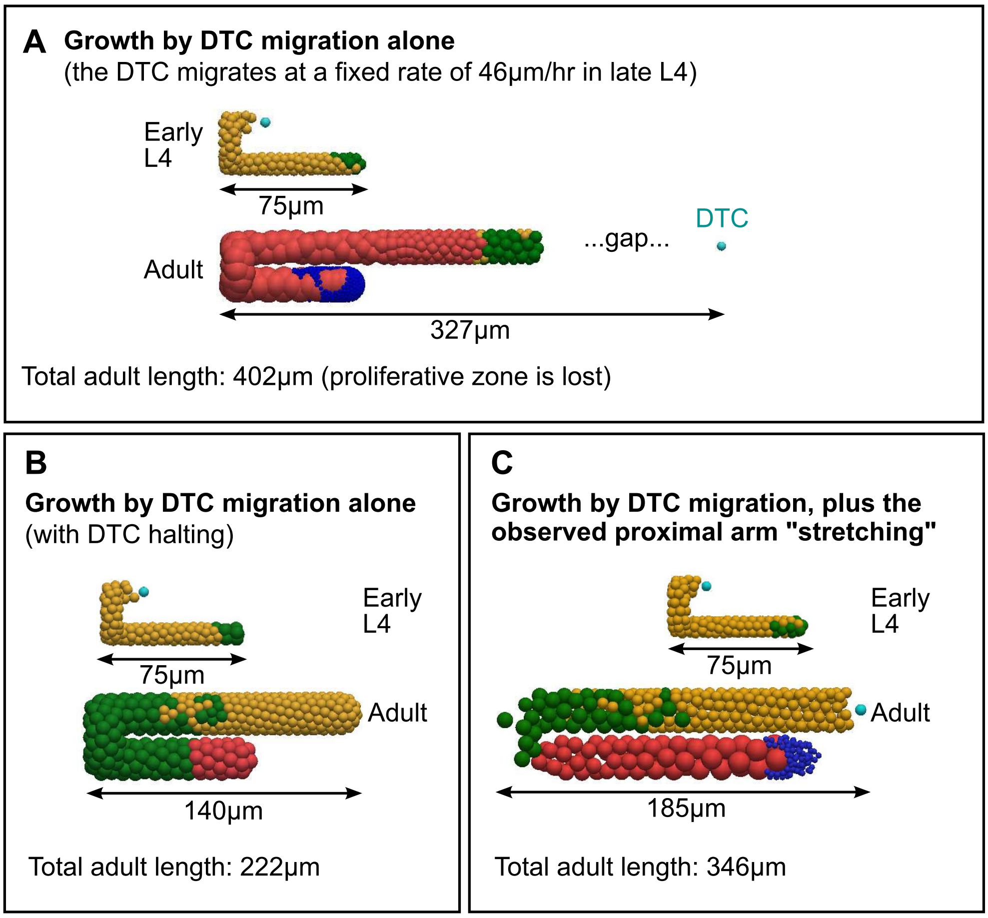


Figure 4.13: The impact of including "stretching" growth during late L4 gonadogenesis. A) A simulation in which gonad growth occurs by rapid DTC migration only (no stretching/movement of the turn). This results in proliferative zone loss, as the DTC travels faster than germ cells can proliferate to fill the available space, leaving a gap behind the leader cell. B) The same simulation as A), but with DTC halting introduced (see Section 4.1.2). The result is a shortened adult gonad, as the DTC must frequently pause until germ cells fill the available space. C) A simulation including "stretching" growth in addition to DTC migration. This results in gonad growth more consistent with experimental measurements.

One possible interpretation of this result is that the rate of larval germ cell division could be too low in our model. However, we chose a shorter larval cell cycle length than the prevailing estimate in the literature (3 hours as opposed to 4 [145]), and recent experimental evidence is also consistent with our choice of an 8 hour adult cell cycle [81, 90]. We therefore sought other explanations as to why our gonadogenesis model was failing to match the biological system. First, we tried introducing a DTC halting mechanism, in which DTC migration pauses if no germ cells are present within 5µm. Adding DTC halting is a simple way to let a lack of germ cell pressure to slow DTC movement; an effect that is observed *in vivo* [139]. It also prevents unrealistic gaps opening up behind the DTC. With halting enabled, gapping and proliferative zone loss no longer occurred. However, the DTC was forced to pause so often that it travelled far slower than its target migration rate of 46µm h^{-1}, resulting in a shorter than expected adult gonad (Figure 4.13B).

Examining our microscopy images, we found that the proximal gonad lengthens significantly during the late L4 larval stage. Proximal growth is difficult to explain in terms of DTC migration, and seems to involve movement of the gonad turn away from the centre of the animal, effectively "stretching" the organ as a whole. We added this stretching growth to the model as described in Sections 4.1.2 and 4.2.1. When both stretching and DTC halting were included, simulations produced a gonad arm that was short, but within normal ranges (Figure 4.13C). While some unrealistic gaps did appear near the gonad turn in this model, the outcome is still a better fit to the biological system. Gapping behind the DTC can ultimately lead to loss of the proliferative zone and with it loss of fertility, which cannot be allowed to occur *in vivo*. By contrast, transient low cell density elsewhere in the germ line could be tolerated, as it does not affect fertility. We conclude from these results that L4 stretching is a non-negligible component of gonadognesis; certainly for the purpose of developing a realistic *in silico* model.

Proximal gonad growth associated with outward movement of the turn has been observed before (see for example [115]). However, it has rarely been remarked upon. We speculated that the apparent stretching could be caused by the pressure of proliferating cells inside the organ, similar to previous findings regarding the impact of germ cell pressure on DTC migration [139]. Our experimental collaborators went on to test this hypothesis using animals with a temperature sensitive GLP-1 loss of function mutation. At low temperatures such mutants appear normal, but at higher temperatures the DTC signalling pathway is disrupted and all germ cell proliferation halts. Temperature upshift was carried out in early to mid L4, to see whether proximal gonad elongation and turn movement still occurred in the absence of proliferation. Loss of germ cell proliferation was found to have no significant impact on proximal growth, meaning some other explanation must exist for the phenomenon. Candidate mechanisms, which are unfortunately harder to test, include: a previously unknown genetic program affecting the gonad membrane itself; pressure provided by the growth of oocytes; or a side effect of the development of somatic reproductive structures, such as the uterus.

4.3.4 Feedback on proliferation is required for germ cell number homeostasis

While developing our germ line model, we initially had difficulty maintaining a stable number of proliferative germ cells. Using a cell cycle model without contact inhibition, and with an adult cell cycle length of 8 hours, the proliferative germ cell count quickly grew far beyond reasonable levels (Figure 4.14A). The expected behaviour is for proliferative cell numbers to be held constant in the young adult, before declining in older animals due to ageing effects. Overproliferation is not a realistic behaviour for the wild-type germ line.

The overproliferation effect persisted, at a somewhat reduced level, when the adult germ cell cycle length was increased to 24 hours; the longest estimate for this parameter found in the literature [50] (Figure 4.14B). We attempted to counterbalance proliferation by raising the germ cell death rate. However, a higher death rate did not prevent the proliferative cell count from increasing. Moreover, death rates above the level shown in Figure 4.14 caused other unrealistic model behaviour. Very high death rates led to depletion of all germ cells in the gonad turn region, resulting in an unrealistically long period without ovulation as germ cells slowly refilled the organ.

This model behaviour led us to speculate that a feedback mechanism might exist in the *C. elegans* germ line, helping to maintain a stable number of proliferative cells. In particular, we hypothesised that crowding/tight packing of adult germ cells may have a negative effect on their proliferation rate, as this is the simplest and most direct way in which overproliferation could be sensed and prevented. Such a feedback mechanism would be akin to contact inhibition, an effect originally proposed in the context of fibroblast cultures in which cells react to a lack of available space [3]. Contact inhibition-like mechanisms have been included in numerous computational biology

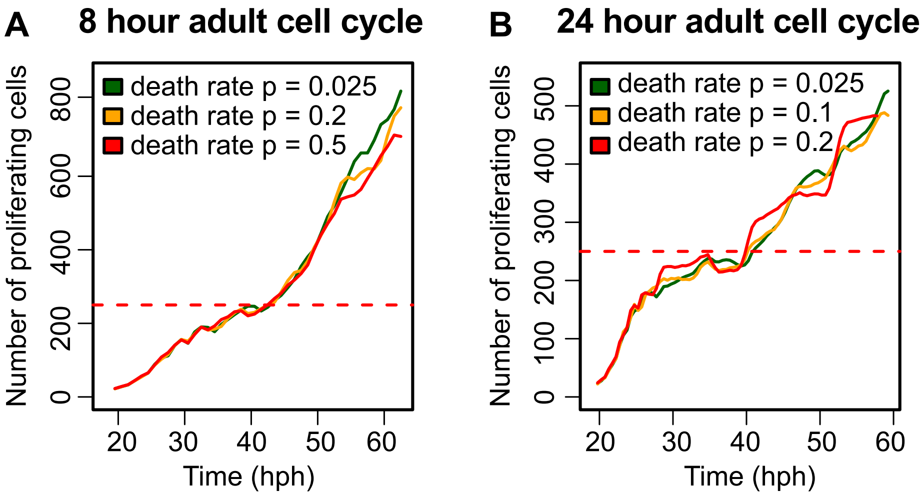


Figure 4.14: Overproliferation in the absence of contact inhibition. Both panels show attempts to achieve homeostasis in the number of proliferative germ cells in our model, by balancing rates of cell proliferation and death in the absence of contact inhibition. Each graph plots the number of proliferative cells present over time (mean of five simulations). Panel A) shows results for an adult germ cell cycle length of 8 hours, while panel B) shows results for a cell cycle length of 24 hours. Both panels display results for three different values of the death rate; higher death rates than those shown produced unrealistic outcomes (see text). A horizontal red dashed line marks the experimental proliferative cell count, which is greatly exceeded in every case.

models [33, 86]. In our case, inhibition was implemented as described in Section 4.1.3; by causing cells to transiently arrest in G2 while their "compressed" volume was less than 70% of their relaxed volume. Using a cell cycle model with contact inhibition, a stable number of proliferative cells was able to be maintained (see Figure 4.12B, Section 4.3.2).

To investigate whether contact inhibition might occur *in vivo*, our collaborators performed an experiment that aimed to induce tighter packing of germ cells, then measure any effect on the rate of division. It was decided to compare the germ lines of wild-type animals with those of *ced-3(-)* mutant worms. Since *ced-3(-)* mutants are incapable of physiological apoptosis, we reasoned that older *ced-3(-)* animals should develop a backlog of oocyte-fated cells that would normally have undergone cell death. Over time, this backlog should eventually cause tighter cell packing in the proliferative zone. The *ced-3(-)* mutation could therefore induce germ cell crowding without causing confounding changes in DTC signalling.

Figure 4.15A shows the average distance between proliferative germ cell rows in wild-type and *ced-3(-)* animals. It confirms that the *ced-3(-)* germ cells were packed significantly closer together, but only after a day or more of oogenesis. This is to be expected, since a backlog of oocyte-fated cells would take time to accumulate and affect the proliferative region. Figure 4.15B, meanwhile, shows the mitotic index (MI) of the proliferative zone for both populations. Recall that MI is the percentage of germ cells observed in the process of dividing in a still microscopy image, and that MI is used as a measure of the proliferation rate. The results indicate that both populations again had a similar MI at one day post-L4, but by 2 days post-L4 the *ced-3(-)* germ lines displayed a significantly reduced MI, and therefore slower proliferation.

The experimental results in Figure 4.15 are consistent with a contact inhibition mechanism operating in the *C. elegans* germ line. However, much more experimental work is needed to either confirm or refute a link between tight packing of adult germ cells and slower proliferation. For instance, a molecular mechanism must be identified that would allow germ cells to sense crowding. A potential candidate is the Hippo pathway, which has been implicated in contact inhibition and organ size control in other species [70]. Several Hippo pathway components are conserved in *C. elegans*, but their role has not been yet thoroughly investigated [270]. In addition, some conflicting evidence exists that would need to be addressed for a contact inhibition hypothesis to be accepted. We discuss this conflicting evidence fully in Chapter 7. Therefore, while our model suggests that moderating feedback on proliferation is necessary, it is still worth considering other, less direct

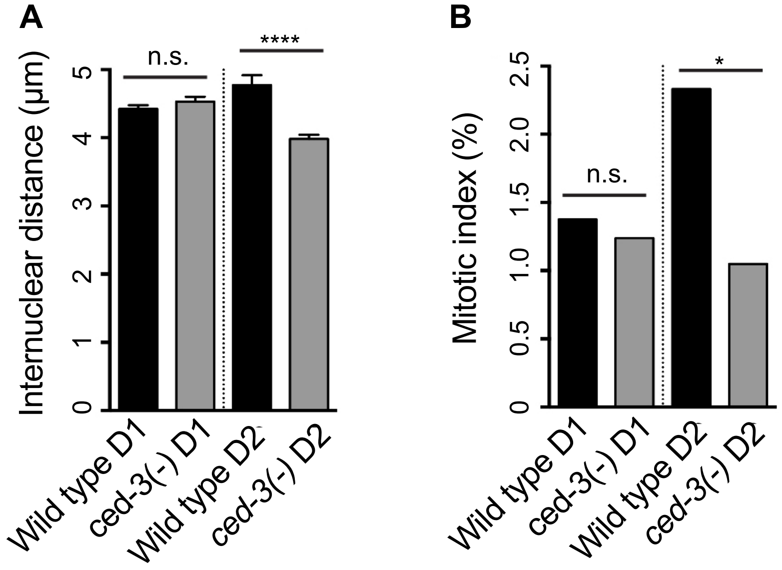


Figure 4.15: Experimental data on the relationship between impaired apoptosis, germ cell packing and proliferation rate. A) A comparison of the mean internuclear distance in the proliferative zone of wild-type animals and *ced-3(-)* mutants (specifically, *ced-3(n717))*. The time points shown are: D1, one day after the L4 moult, and D2, 2 days after. By the D2 time point, mutant animals show a significantly reduced internuclear distance, indicating tighter germ cell packing (two-tailed Student's t-test P<0.0001). B) Measurements of the proliferative zone mitotic index in the same two populations at the same time points. Again, by D2, *ced-3(-)* mutants show a significantly reduced rate of germ cell proliferation (Mann-Whitney U test, P<0.05). Impaired apoptosis is therefore associated with both tighter germ cell packing and reduced proliferation. Number of gonad arms analyzed: 33 [WT D1], 23 [*ced-3(-)* D1], 22 [WT D2], 21 [*ced-3(-)* D2]. n.s. = not significant and error bars represent standard error of the mean. Figure reproduced with permission from [15].

mechanisms that could provide that feedback, such as a biochemical signal that communicates the ovulation rate back to the proliferative zone [197]. We will return to the topic of contact inhibition in Chapter 6, where we look at its effects on the proliferative zone in simulations.

4.3.5 Germ cell ordering is roughly maintained during larval development

A significant drawback of the germ line as a model experimental system is that individual cell tracking and long term live imaging are currently infeasible. However, improved labelling and imaging techniques are under development which could make these experiments possible [90, 275]. We therefore decided to use our model to obtain predicted tracking and labelling results. When improved techniques do become available, our model could perhaps be used to interpret complex *in vivo* data.

We began by running a simulation in which three germ cells were labelled at the beginning of L3: one distal cell (red), one mid-gonad cell (green), and one proximal cell (blue). The paths of these cells and their descendants were then tracked through late larval development and into adulthood. Figure 4.16 shows the results. Some of the cell movement trends discussed below can be seen more clearly in the video at http://tinyurl.com/ElegansGermLine. Initially descendants of a single cell remain tightly grouped, and during L3 the dominant cell movement direction is toward the DTC. However, in adulthood, the dominant direction reverses, with cells moving toward the uterus (Figure 4.16, 35.5-43.5 hph). This switch in adulthood from proximal-distal to distal-proximal movement arises naturally in our model from a combination of cell mechanics and the boundary condition, whereas it must be hard-coded into on-lattice models [221].

Going into L4, the distal-most cells (red) travel further and faster than proximal cells, filling the space created by DTC migration. The distal-most cells also produce many more descendants, as they remain longer in the proliferative zone during gonadogenesis. Meanwhile, proximal cells (blue) remain pressed into the proximal end of the gonad and develop into sperm. In late L4, all clonal groups become more dispersed, with particular cell mixing happening in the mid-gonad. Some germ cell movement toward the turn also occurs over this period, as cells move to fill the space created by stretching (Figure 4.16, 33.5 hph). Tracking into adulthood showed that the descendants of the proximal L3 cells are eventually eliminated from the gonad by either apoptosis or fertilisation (note the few remaining green and blue cells at the final time point). From a biological perspective, these results indicate that germ cells in different positions at L3 are already destined to contribute to different tissues, raising the question of whether they may be specialised in any way. Certainly the fact that one specific group of L3 germ cells produces most of the proliferative zone should have implications for the spread of germline mutations.

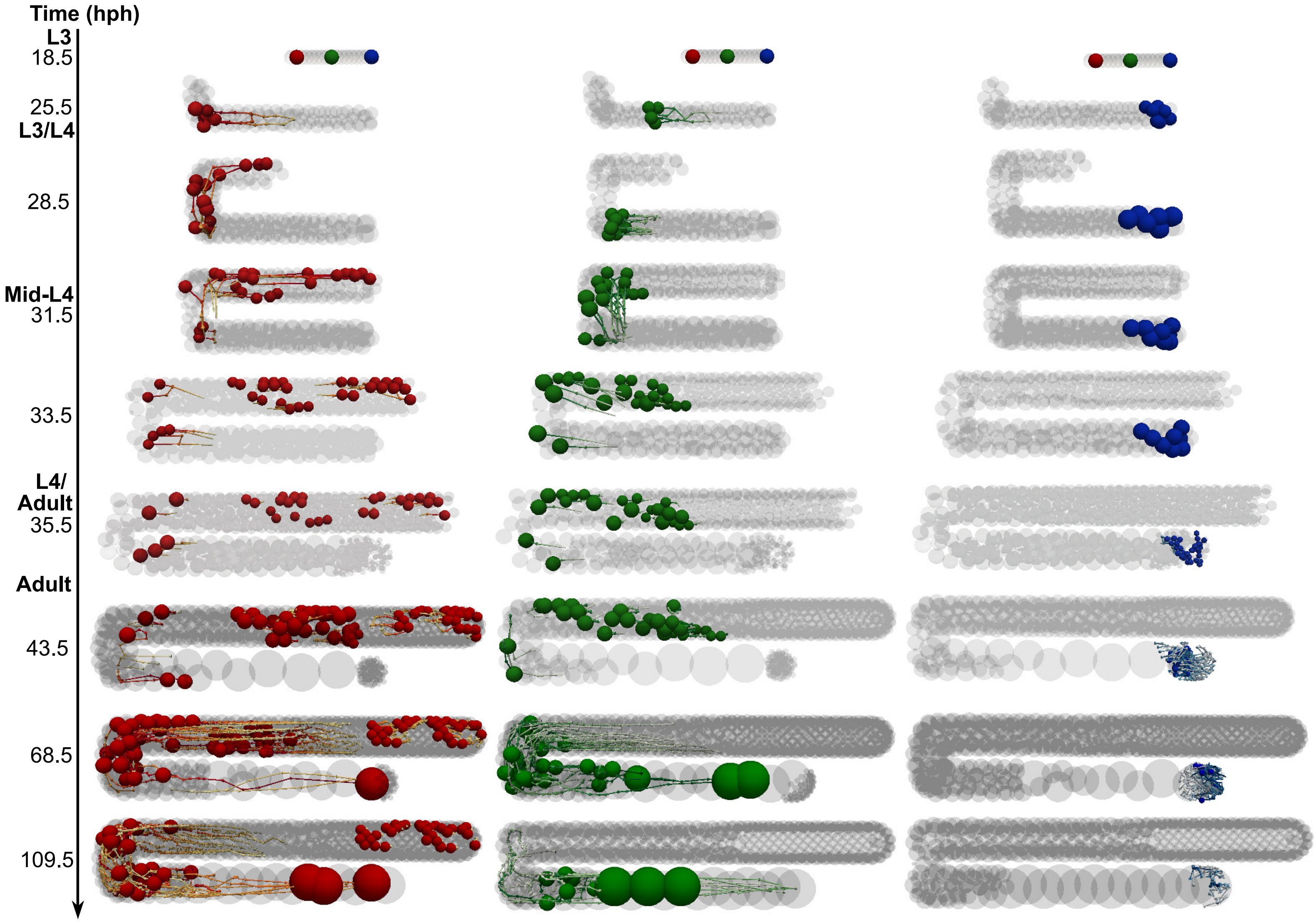


Figure 4.16: Tracking individual germ cells. In this simulation, three cells were labelled at the beginning of L3: one distal (red), one mid (green) and one proximal (blue). The columns of the figure show the paths taken by each of these cells, along with their descendants. Other germ cells are indicated in light gray. Note that some tracks do not end in a cell, owing to that cell's removal from the simulation by apoptosis or fertilisation. This is particularly apparent at 109.5 hph, as green oocytes are being ovulated.

To obtain an overview of the motion and arrangement of *all* germ cells, we ran a second simulation labelling every cell according to its L3 ancestor. Each cell present at the beginning of L3 was assigned a colour, which was then inherited by daughter cells. Figure 4.17 shows the pattern of coloured labels present at the young adult stage of the simulation. Once again, our model predicts that the proximal to distal ordering of germ cells is roughly maintained over the course of development. Cells close to the proximal end of the gonad in L3 develop into sperm, while distal germ cells at L3 contribute to the proliferative zone. Clonal groupings of cells eventually appear in the distal gonad in this simulation. However, multiple ancestor labels remain in the simulated germ line at the point when self-sperm are depleted. This observation is of interest with regard to the possibility of achieving monoclonality in the germ line; a state in which all adult cells descend from a single labelled ancestor. Monoclonality is an indicator of a neutral drift stem cell maintenance mechanism [228]. These *in silico* labelling results suggest that, to have a chance of observing monoclonality *in vivo*, ovulation would need to be prolonged beyond the point of sperm depletion; for example, by mating.

It remains to be seen whether these predictions will match experimental observations, once the techniques become available for *in vivo* cell tracking and labelling.

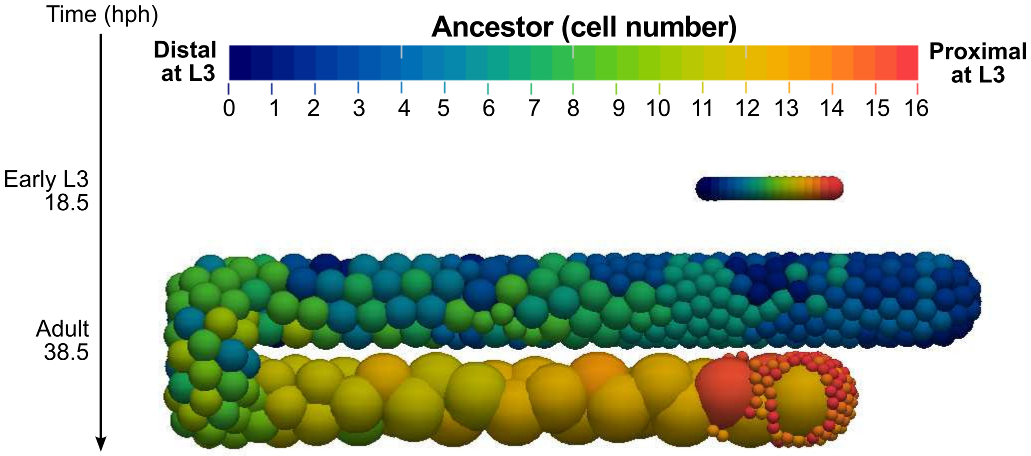


Figure 4.17: A germ line simulation with cells labelled according to their L3 ancestors. All germ cells present at the beginning of L3 were assigned a colour depending on their distal to proximal position. This colour was then inherited by daughter cells. The image above shows the resulting pattern of cell colours in the young adult germ line, providing a prediction for how germ cell ordering changes over the course of development.

4.3.6 Local parameter variation

As discussed in Section 4.2, it is not possible to conduct an exhaustive search of our model's parameter space; primarily because running a simulation for long enough to determine steady state behaviour is time consuming, precluding techniques that require thousands of runs. Instead, we chose to vary a subset of the model parameters by 25% above and below their usual values, to explore the local behaviour of the model about our wild-type parameter set. We focussed on those parameters with the least experimental data available to constrain their values: the time taken to produce mature sperm, the cell growth rate, the death rate, cell-cell repulsion strength, sex determination position, and larval germ cell cycle length (see Table 4.4). Five replicates were carried out for each parameter set, with the ARCUS-B[1] computing cluster used to complete the investigation in a reasonable timeframe. During each run, six germline properties were recorded: gonad length; number of proliferative cells; sperm count; total cell count; the position of the furthest proliferative cell from the DTC; and the position of the closest meiotic cell to the DTC. Figures 4.18-4.23 show the results (mean of all replicates, with a region ± 1 standard deviation shaded).

In several of the figures below some unusual behaviour can be seen at around the time of the L4 to adult transition; in particular, in the panel showing the position of the last proliferative cell. The distance from the DTC to the last proliferative cell peaks at the start of adulthood, before falling back and oscillating for the rest of the simulation. The peak can be attributed to the stretching component of our model; as the turn moves away from the centre of the animal, germ cells move toward the turn to fill the space created. Some proliferative cells are caught up in this movement and are carried far away from the DTC. The distance between the last proliferative cell and the DTC decreases once these cells enter meiosis. The oscillation is likely due to some amount of synchronisation of germ cell divisions. Given our contact inhibition hypothesis, it is possible that cells repeatedly become tightly packed, relax, are released from inhibition, then divide, producing tight packing once again. Alternatively, synchronisation could arise from the change in cell cycle length going from L4 into adulthood. All germ cells exit the larval cell cycle within 3 hours of each other, but then begin a much longer 8 hour cycle. As a result, cells are not evenly distributed throughout the cell cycle at the start of adulthood, but rather are clustered within the first 3 hours.

---

[1] ARC refers to the University of Oxford Advanced Research Computing facility, see http://dx.doi.org/10.5281/zenodo.22558

In terms of sensitivity to parameter variations the model was found to be insensitive to variations in a number of parameters, namely: the time taken for meiotic cells to become mature sperm; the precise cell growth rate; and the germ cell death rate (Figures 4.18-4.20). Varying the strength of the cell-cell repulsion force, μ, had only a moderate effect (Figure 4.21). As expected, if cell-cell repulsion is increased, germ cells spread out and fill space inside the gonad more effectively. As a result, germ cells were present behind the DTC for longer periods, supporting DTC migration and leading to a longer adult organ. More rapid gonad growth also affected the sperm count in our model, since germ cells reached the requisite distance from the DTC for sex determination earlier. A longer period of sperm production therefore occurred prior to the switch to oogenesis.

One of the most influential parameters tested was the position in the gonad at which sex determination occurs (Figure 4.22). When sex determination is allowed to take place closer to the DTC, fate decisions happen earlier in larval development. This results in a longer period of spermatogenesis, a higher adult sperm count, and a larger number of cells in total (since each sperm-fated cell produces four mature sperm, the number of sperm-fated cells has a disproportionate influence on total cell count). The effect of this parameter on other germline properties was less pronounced, with no clear trend visible. It is worth remembering that we assumed a simple, distance-based trigger for the sexual fate decision because it produces realistic separation of male and female gametes during development. The *in vivo* trigger for sex determination may not be distance-based. However, we can still infer from our model that the timing of germ cell sex determination is critical for producing a correctly organised germ line with appropriate numbers of gametes. As such, the precise mechanism governing this cellular decision warrants further research.

The other highly influential parameter was the length of the larval germ cell cycle (Figure 4.23), a factor identified as important in previous experimental and modelling studies [139, 221]. A faster larval cell cycle leads to more germ cells in the gonad, better supported DTC migration, and a higher adult sperm count. A substantially higher proliferative cell count also results, despite the length of the proliferative zone remaining the same, indicating that germ cells pack together more tightly as a result of frequent divisions. Although the model is indeed sensitive to the value of this parameter, we have at least one biological estimate available of the larval germ cell cycle length [145]. In addition, for the period of larval development in which all germ cells are proliferative, experimental cell counts can be used to estimate the cell cycle duration.

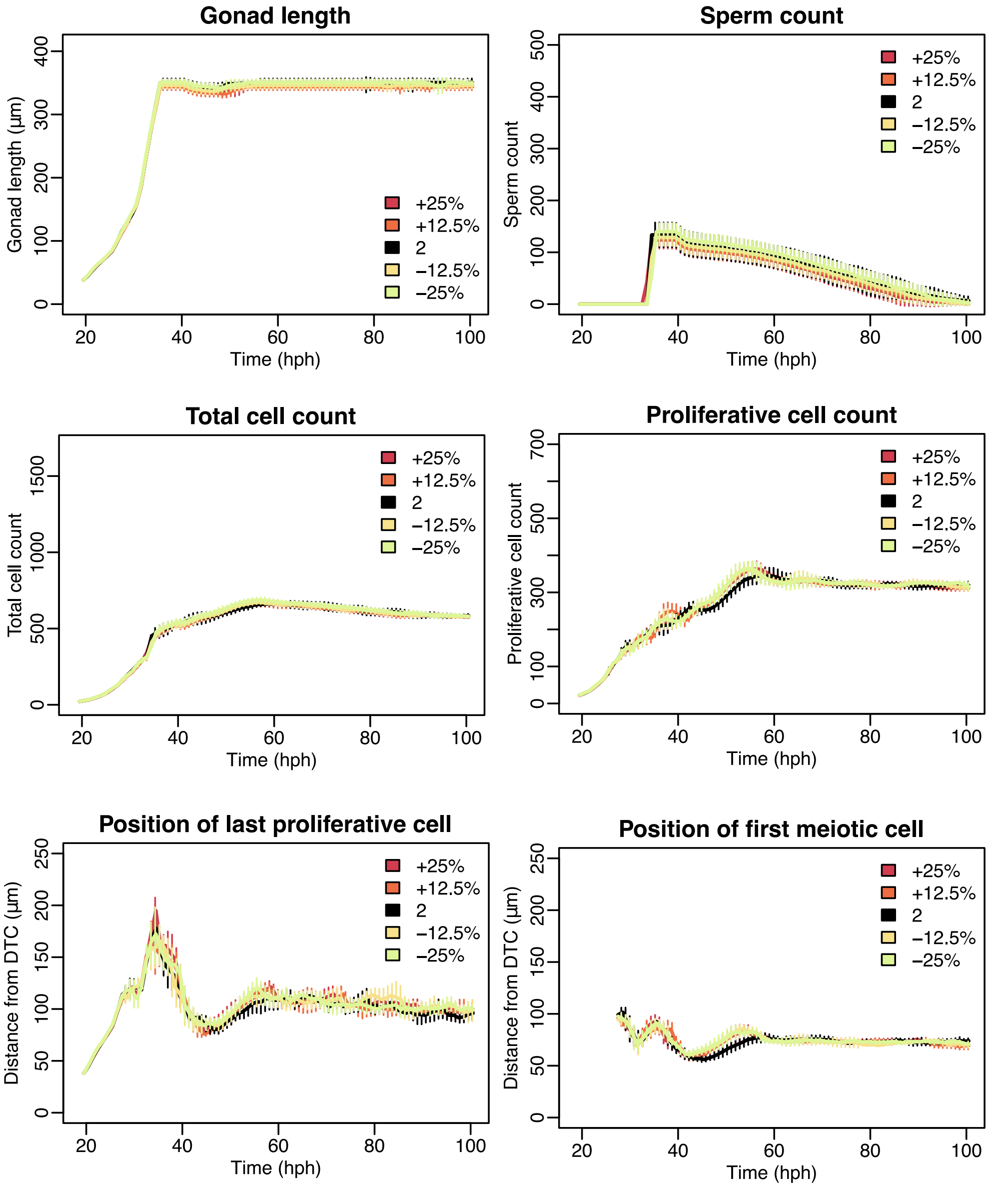


Figure 4.18: Varying the delay between meiosis and mature sperm production (hours). Each panel corresponds to a particular germline property, and we plot the simulated value of that property over time for five different parameter values. The black trace corresponds to the usual, wild-type parameter choice. The yellow and green traces show reductions in the parameter of interest, while the orange and red traces show increases (see legend). Each trace is the mean of five replicates, and bars show one standard deviation either side of the mean.

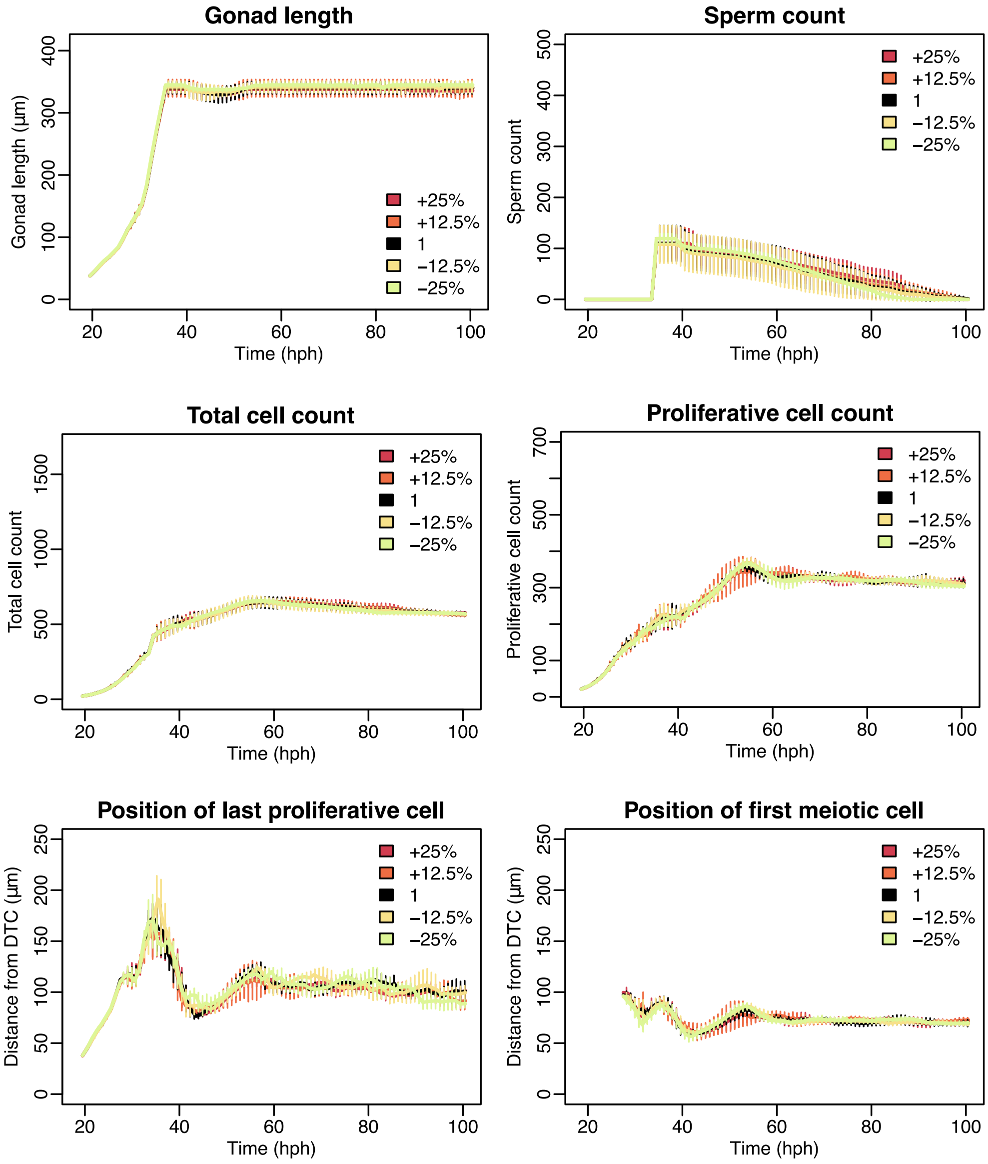


Figure 4.19: Varying the cell growth rate (μm h^{-1})**.** Each panel corresponds to a particular germline property, and we plot the simulated value of that property over time for five different parameter values. The black trace corresponds to the usual, wild-type parameter choice. The yellow and green traces show reductions in the parameter of interest, while the orange and red traces show increases (see legend). Each trace is the mean of five replicates, and bars show one standard deviation either side of the mean.

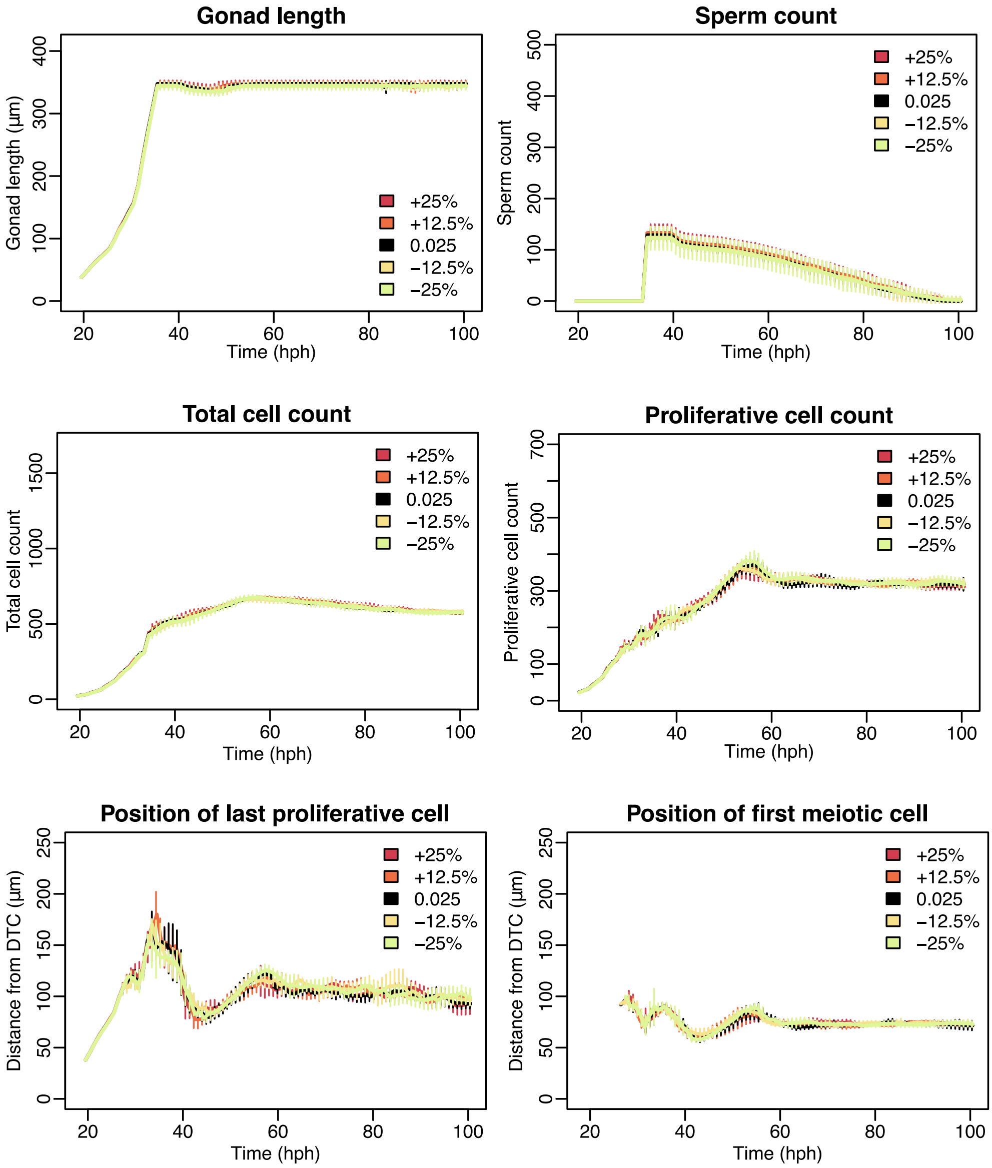


Figure 4.20: Varying the germ cell death rate (probability per hour spent outside the proximal arm). Each panel corresponds to a particular germline property, and we plot the simulated value of that property over time for five different parameter values. The black trace corresponds to the usual, wild-type parameter choice. The yellow and green traces show reductions in the parameter of interest, while the orange and red traces show increases (see legend). Each trace is the mean of five replicates, and bars show one standard deviation either side of the mean.

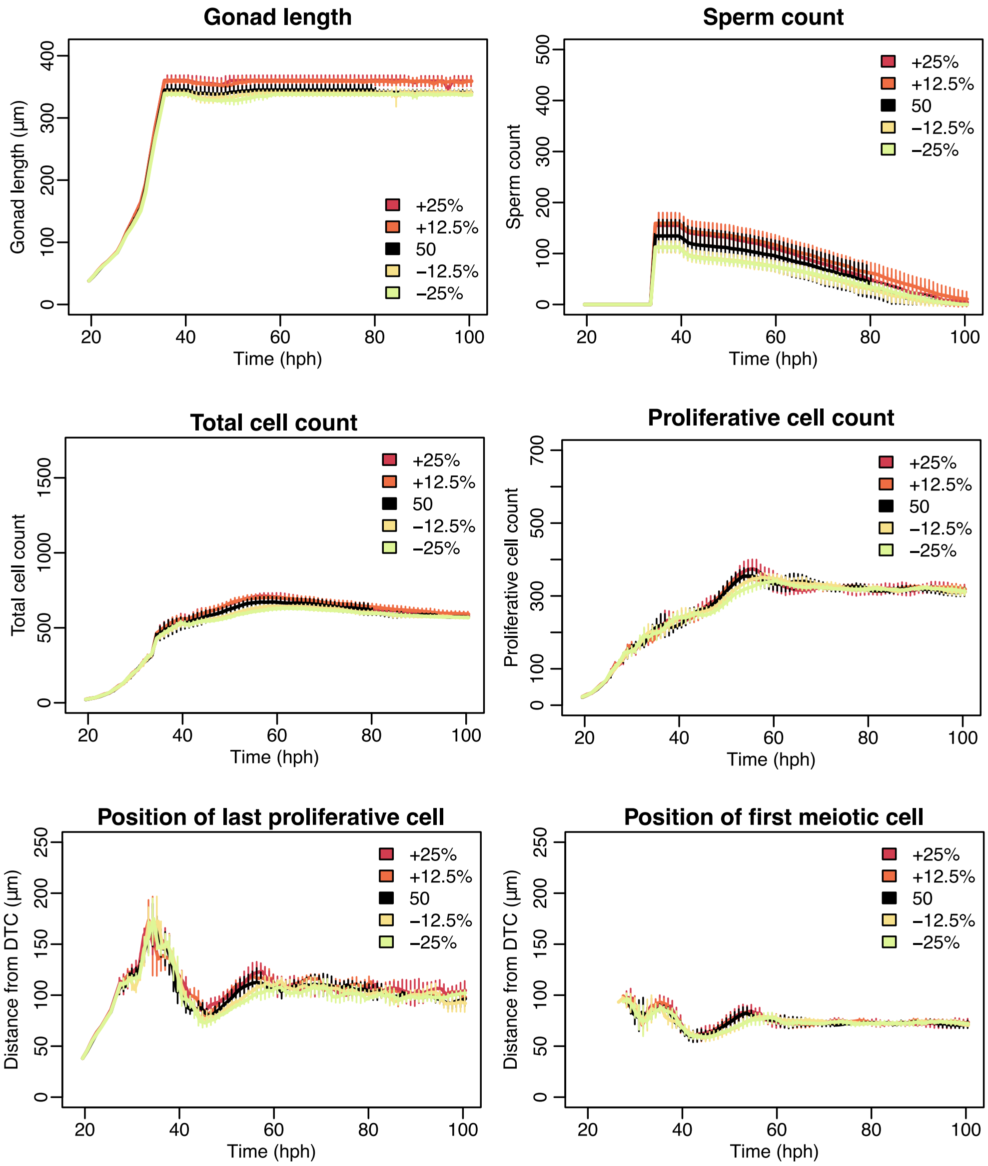


Figure 4.21: Varying the cell-cell repulsion strength ($\mathrm{h\,\mu m^{-1}}$)**.** Each panel corresponds to a particular germline property, and we plot the simulated value of that property over time for five different parameter values. The black trace corresponds to the usual, wild-type parameter choice. The yellow and green traces show reductions in the parameter of interest, while the orange and red traces show increases (see legend). Each trace is the mean of five replicates, and bars show one standard deviation either side of the mean.

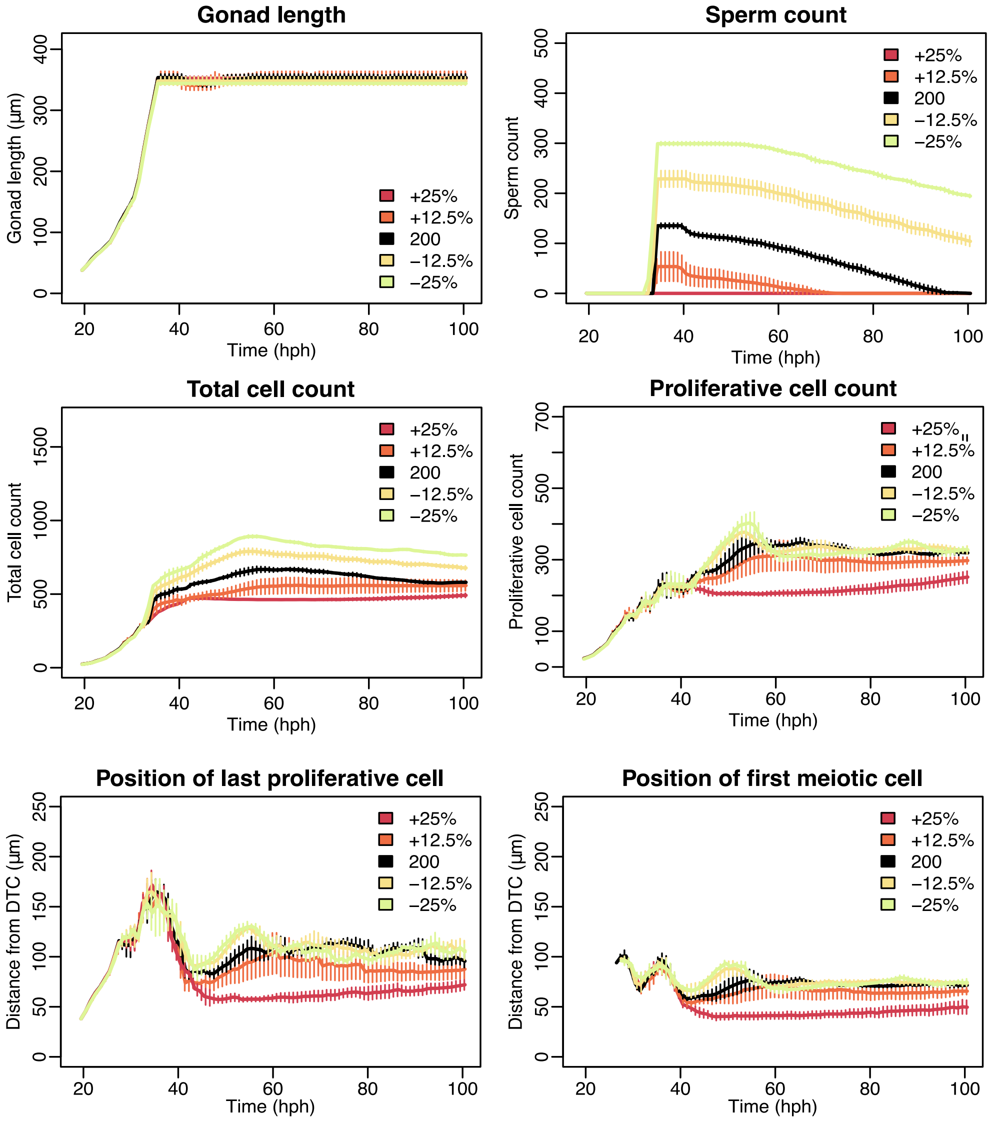


Figure 4.22: Varying the position at which sex determination occurs (µm **from the DTC).** Each panel corresponds to a particular germline property, and we plot the simulated value of that property over time for five different parameter values. The black trace corresponds to the usual, wild-type parameter choice. The yellow and green traces show reductions in the parameter of interest, while the orange and red traces show increases (see legend). Each trace is the mean of five replicates, and bars show one standard deviation either side of the mean.

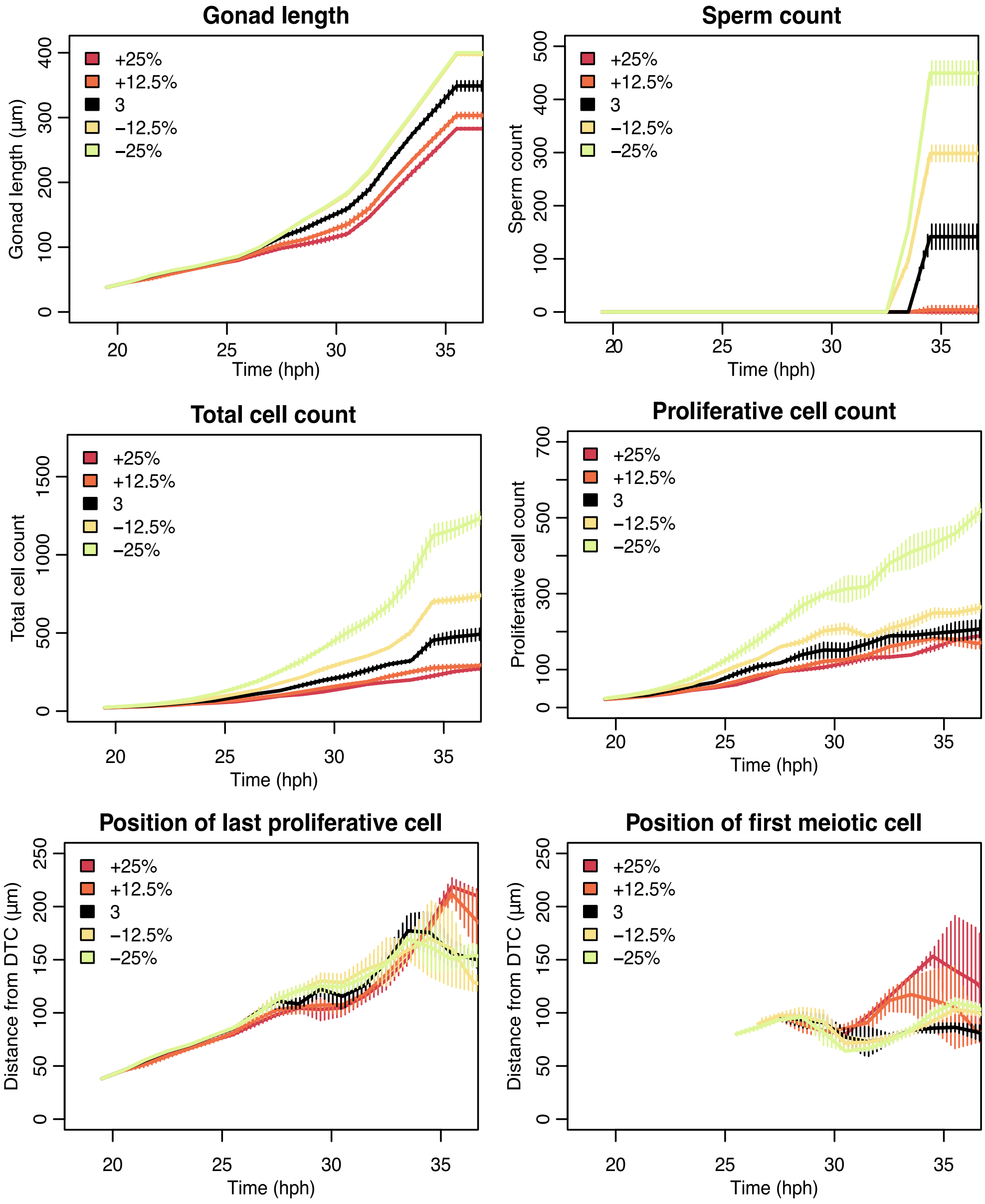


Figure 4.23: Varying the larval germ cell cycle length (hours). Each panel corresponds to a particular germline property, and we plot the simulated value of that property over time for five different parameter values. The black trace corresponds to the usual, wild-type parameter choice. The yellow and green traces show reductions in the parameter of interest, while the orange and red traces show increases (see legend). Each trace is the mean of five replicates, and bars show one standard deviation either side of the mean.

4.4 Comparison with the model of Hall *et al.* (2015)

We end this chapter with a brief comparison between our model and that of Hall *et al.*; the only other 3D germ line model currently available [99].

The clearest difference between the two models lies in the boundary condition. Whereas we use a growing gonad boundary to study the larval period, Hall *et al.* focus on the adult germ line and therefore use a fixed boundary. The Hall *et al.* figures that appear to show development actually involve a small number of initial cells proliferating to fill a gonad that is fully formed. As such, these simulations are more representative of refeeding behaviour following ARD.

In terms of mechanics, both models use a similar basic equation to update cell positions (Equation 3.2), which introduces an overdamping assumption. However, Hall *et al.* also include a Brownian motion term $\sqrt{2k_BTdt\left(\frac{\eta}{m}\right)r}$, where k_B is the Boltzmann constant, T is temperature, η is a drag coefficient, m is cell mass, and r is Gaussian distributed noise. There is a further difference in how drag is scaled with cell size: Hall *et al.* assume that drag scales with mass, whereas we scale drag according to cell radius, consistent with our overdamping assumption.

Finally, with regard to the intracellular program, Hall *et al.* use a qualitative network to model the decision-making process, whereas we use a statechart. Qualitative networks are perhaps better suited to modelling the molecular details of germ cell decision-making, as they involve a network of entities (genes or proteins) with associated activities, which are updated depending on the activities of their neighbours. This is in contrast to the higher lever approach encouraged by statechart models. Another notable difference lies in which part of the intracellular model includes most detail. Hall *et al.* focus on apoptosis and oocyte maturation, whereas we focus on the mitotic cell cycle and meiotic entry. Overall, we would argue that our intracellular model is easier for other researchers to reproduce and extend, since our code was made public and a wide range of statechart tools are available in contrast to the smaller selection of qualitative network tools.

Both models produced some unique results. For instance, Hall *et al.* drew conclusions from their simulations about the mechanism of RAS/MAPK activation in germ cells. They also performed some exhaustive verification of cell fate decisions in a simplified system. Meanwhile, our model highlighted the importance of stretching growth during gonadogenesis, and made predictions about cell movement during the L3 and L4 larval stages. We did come to some similar conclusions to Hall *et al.* (2015), for example that monoclonality is unlikely to occur within a normal reproductive lifespan, and that regulatory feedback on proliferation is necessary in some circumstances. The fact that two independently developed models agree on these points is in their favour.

4.5 Summary

In this chapter, we described our 3D *C. elegans* germ line model; its design, implementation and parameters. We then compared typical simulation output with observations of the wild-type germ line. Qualitatively, our model resembles the germ line quite closely. The correct sequence of fate decisions occur during development, a reasonable number of gametes are produced, and the organisation of the germ line is stably maintained over several days. Quantitatively, the model gives a reasonable fit in terms of gonad growth, sperm count, and proliferative zone length in cell diameters.

Weaker areas of fit include the presence of too many small oocyte-fated cells in the gonad turn. Increasing the cell death rate to combat this crowding tends to impair ovulation, particularly during the early adult stages. Potentially, future models could address this issue by making germ cell size strictly proportional to position in the gonad. While we feel this approach to be too prescriptive, it has some basis in biology (material from the rachis constantly flows into germ cells as they migrate through the organ, so more proximal cells should be larger). The result would be a smaller number of large cells in the turn. Alternatively, future models could focus on simulating the adult steady state, in which case a higher apoptosis rate might be acceptable as it would not cause problems during the larval to adult transition. In addition, during the larval portion of our model, the transition zone between proliferation and meiosis is too long, covering multiple cell rows rather than 1CD. This occurs due to a combination of desynchronised germ cell cycles, and restriction of meiotic entry to G1. It is therefore difficult to see how the fit could be improved without making fundamentally different biological assumptions, such as highly synchronised divisions during the larval stages or a meiotic entry decision decoupled from the cell cycle.

Several new hypotheses and predictions arose in the process of developing and running our model. Firstly, it was noted that the proximal germ line appears to “stretch” during late L4. This is not a point that is emphasised in the biological literature, but the effect is significant and it greatly influenced our ability to produce a realistic-looking organ *in silico*. We also carried out simulated tracking and labelling experiments, which predicted that the ordering of germ cells is roughly maintained during larval development. Proximal cells at L3 become sperm, while distal cells at L3 form the proliferative zone. Monoclonality was not observed within the normal hermaphrodite reproductive span, which may be important in terms of genetic diversity and preventing a single flawed germ cell from dominating the niche.

The most intriguing possibility raised by this model is that adult *C. elegans* germ cells may be subject to a feedback mechanism on their cell cycle, preventing overproliferation. In our model, we propose a contact inhibition-like mechanism as the simplest working hypothesis. With contact inhibition, a stable proliferative cell count is achieved, whereas without it, cell counts grow to more than twice the expected value. There is increasing discussion in the field, particularly among modellers, about the need for a mechanism to explain why mutants with impaired apoptosis or ovulation do not contain an abnormal number of germ cells [15, 19, 99, 160]. Regulatory feedback on the proliferation rate seems to be a good explanation. However, while contact inhibition is a straightforward hypothesis, it is also problematic in certain respects. We will look further at the impact of introducing contact inhibition on the proliferative zone in Chapter 6, and a full discussion of the advantages and disadvantages of this hypothesis can be found in Chapter 7.

In the next chapter, we consider two technical areas of our model which could potentially be improved. We consider applying a more physically realistic force law, and we investigate whether an alternative numerical scheme could update cell positions more accurately and efficiently.

Chapter 5

Investigating choice of force law and numerical method

In this chapter, we investigate some potential technical improvements that could be made to the cell-based models used thus far. First, we consider whether a more complex force law — that proposed by Buske *et al.* (2011) [33] — provides a more realistic representation of cellular mechanics than the GLS force implemented in Chaste. We aim to determine whether the extra detail included in the Buske force law measurably changes simulation results, and so evaluate the costs and benefits of additional complexity. Second, we examine more closely our choice of numerical method, implementing several alternative time-stepping schemes and assessing whether they improve accuracy and performance relative to the widely-used forward Euler method. Finally, we propose an alternative type of intracellular model, cell logic modules, which aims to address some of the difficulties associated with our earlier statecharts approach.

5.1 Alternative force laws

So far, we have applied a pure repulsion force law between germ cells (Equation (3.11)), meaning that overlapping cells repel each other as if their centres were joined by a nonlinear spring. This choice of force law has two limitations. Firstly, it neglects all other kinds of cell-cell interaction apart from elastic repulsion. In particular, adhesion is ignored. Secondly, the single parameter in a spring force law (spring stiffness) does not correspond to a real, measurable property of cells. This makes choosing an appropriate spring constant a matter of trial and error. If instead the force law included parameters such as the Young's modulus or Poisson ratio of a germ cell, it might be possible to determine those values experimentally, using techniques such as atomic force microscopy [121]. Where experimental measurements are lacking, sensible bounds could at least be imposed based on the properties of other materials (a cell's Young's modulus should be lower than that of a typical metal, say, and perhaps comparable to rubber).

Here, we implement a force law proposed by Buske *et al.* that includes cell elasticity, adhesion and resistance to compression [33]. The strength of these interactions depends on the material properties of the cells involved, rather than on lumped parameters such as spring stiffness. The Buske model also uses a different position update equation that takes into account cell-cell friction, and cell radii are permitted to change over time in response to applied forces. We compare the Buske force law with Chaste's existing GLS force, which includes a combination of nonlinear repulsion and weak adhesion. By enabling and disabling parts of the Buske force law, it is possible to see which components make a measurable difference to simulation results.

The structure of this section is as follows. We begin by describing the Buske force law in detail. We then present the results of a simple force comparison in the case of an overlapping pair of cells. Next, we discuss how boundary conditions can be applied to a simulation using a Buske force law, allowing more complex tests to be run. Finally, we compare the two force laws in the context of a distal germ line simulation, to see whether the behaviours observed in our whole germ line model persist regardless of force choice.

5.1.1 The GLS force and the Buske force law

The GLS force law as implemented in Chaste was introduced in Section 3.1.4. The force on cell i due to cell j is given by Equation (3.10),

$$\mathbf{F}_{ij} = \begin{cases} \mu\,(R_i + R_j)\,\log\left(1 - \dfrac{v_{ij}}{R_i + R_j}\right)\hat{\mathbf{r}}_{ij}, & \text{if } v_{ij} > 0 \text{ and } d_{ij} < k, \\ \mu\, v_{ij}\, \exp\left(\dfrac{\alpha\, v_{ij}}{R_i + R_j}\right)\hat{\mathbf{r}}_{ij}, & \text{if } v_{ij} \le 0 \text{ and } d_{ij} < k, \\ \mathbf{0}, & \text{if } d_{ij} \ge k. \end{cases}$$

where R_i and R_j are cell radii, v_{ij} is the length of the cell-cell overlap, d_{ij} is the distance between cell centres, and $\hat{\mathbf{r}}_{ij}$ is a unit vector directed from cell i to cell j. For two overlapping cells, the GLS force is identical to the pure repulsion force applied in Chapter 4, with μ denoting spring strength. However, the GLS force law also includes weak attraction between nearby cells, which decays at a rate determined by α. Finally, if cell separation exceeds a cutoff distance, k, no interaction can occur. The usual position update equation (Equation (3.2)) is used with a GLS force, which introduces a drag coefficient denoted by η. Chaste's default parameter set for this force law can be found in Table 3.1.

The force law developed by Buske *et al.* is significantly more complex. In the Buske model, cells can respond to a net force either by moving or by changing in radius [33]. In either case, the aim is for the population as a whole to reach a lower energy configuration. The total energy of the interaction between cells i and j is comprised of three terms: the deformation energy $\left(W_{ij}^D\right)$, the adhesion energy $\left(W_{ij}^A\right)$, and the uniform compression energy $\left(W_i^K\right)$.

The deformation energy is given by

$$W_{ij}^D = \sqrt{\frac{R_i R_j}{R_i + R_j}} \frac{2(R_i + R_j - d_{ij})^{\frac{5}{2}}}{5D}. \tag{5.1}$$

This expression derives from the Hertz model of contact, and the parameter D simply collects together all occurrences of the Young's modulus, E, and the Poisson ratio, ν [194]. Written in terms of these quantities, $D = 3(1-\nu^2)/2E$.

The adhesion energy between two cells is

$$W_{ij}^A = \varepsilon A_{ij}, \tag{5.2}$$

where ε denotes energy per unit contact area, and A_{ij} is the size of the contact surface. A_{ij} can be calculated as

$$A_{ij} = \pi(R_i^2 - x_{ij}^2) = \pi(R_j^2 - x_{ji}^2), \tag{5.3}$$

where x_{ij} is the distance between the contact surface and the centre of cell i

$$x_{ij} = \frac{(R_i^2 - R_j^2 + d_{ij}^2)}{2d_{ij}}. \tag{5.4}$$

Finally, the compression energy of cell i is given by

$$W_i^K = \frac{K}{2V_i^T}(V_i^T - V_i^A)^2 \tag{5.5a}$$

$$= \frac{K}{2V_i^T}\left(V_i^T - \frac{4\pi R_i^3}{3} + \sum_j \frac{\pi}{3}(R_i - x_{ij})^2(2R_i - x_{ij})\right)^2. \tag{5.5b}$$

Here, V_i^T is the cell's target volume; the volume it would occupy under zero applied force. We take this to be the volume of a sphere with the cell's initial radius. V_i^A is the cell's actual volume;

calculated based on its current radius R_i, minus any volume that overlaps with neighbouring cells. K denotes the bulk modulus, which can be written in terms of a cell's Poisson ratio and Young's modulus as $E/3(1-2\nu)$.

The forcing term for the Buske position update equation is then

$$\mathbf{F}_i = \sum_j \frac{\partial (W_{ij}^D + W_{ij}^A + W_i^K)}{\partial d_{ij}} \hat{\mathbf{r}}_{ij}, \tag{5.6}$$

and the full position update equation is

$$\eta_{BM} \frac{d\mathbf{r}_i}{dt} + \sum_j \eta_C A_{ij} \left(\frac{d\mathbf{r}_i}{dt} - \frac{d\mathbf{r}_j}{dt} \right) = \mathbf{F}_i. \tag{5.7}$$

Here, η_{BM} is a drag coefficient representing resistance to movement along a membrane or through a medium. It closely corresponds to η in the GLS update equation (3.2). By contrast, η_C is a coefficient representing cell-cell friction, found in the Buske model only.

Cells in the model of Buske *et al.* can also change their radius in response to applied forces. The forcing term for the radial update equation is

$$G_i = \sum_j \frac{\partial (W_{ij}^A - W_{ij}^D - W_i^K)}{\partial R_i}, \tag{5.8}$$

and the full radius update equation is

$$\eta_{VO} \frac{dR_i}{dt} + \sum_j \eta_C A_{ij} \left(\frac{dR_i}{dt} + \frac{dR_j}{dt} \right) = G_i, \tag{5.9}$$

where η_{VO} is a coefficient representing friction during volume changes.

The default parameter set for the Buske force law is listed in Table 5.1. Also shown are converted parameters in units of microns and hours, which are more typical for a Chaste simulation. We used these converted parameters when comparing the Buske and GLS force laws, and the original parameter set when studying the Buske model in isolation. The majority of these parameter values have some basis in experimental data. For example, E is ultimately based on optical stretcher and scanning-probe experiments; ν is based on cell manipulations using a micropipette; ε is based on cell surface receptor and ligand densities; and η_{VO} is based on the viscosity of cytoplasm. None of this experimental work involved germ cells. Rather, it used a variety of cell types commonly employed in tissue mechanics work (fibroblasts, red blood cells and endothelial cells). However,

Parameter	Value	Unit conversion	Source
E	1000 Pa	12960 $\mathrm{kg\,\mu m^{-1}\,h^{-2}}$	[86, 94, 156]
ν	0.333	0.333	[86, 156, 158]
D	1.33×10^{-3} $\mathrm{Pa^{-1}}$	1.03×10^{-4} $\mathrm{\mu m\,h^2\,kg^{-1}}$	Function of ν and E
ε	2.0×10^{-4} $\mathrm{Nm^{-1}}$	2592 $\mathrm{kg\,h^{-2}}$	[84, 86]
K	1000 Pa	12960 $\mathrm{kg\,\mu m^{-1}\,h^{-2}}$	Function of ν and E
η_{BM}	3.2 $\mathrm{Nsm^{-1}}$	11520 $\mathrm{kg\,h^{-1}}$	Fit by Buske *et al.* based on crypt simulations
η_C	5×10^{10} $\mathrm{Nsm^{-3}}$	180 $\mathrm{kg\,\mu m^{-2}\,h^{-1}}$	[86]
η_{VO}	400 $\mathrm{Nsm^{-1}}$	1.44×10^{6} $\mathrm{kg\,h^{-1}}$	[27, 86]

Table 5.1: The default Buske force law parameter set. Taken from [33]. These values are used throughout the chapter except where otherwise indicated. We also include a conversion into units more typically used in Chaste.

the original Buske *et al.* model was a representation of the colonic crypt, a system that has a very similar geometry to the distal *C. elegans* germ line.

When we started this work, Buske forces had already been partially implemented in Chaste. Cell positions could be updated by constructing a system of ODEs based on Equation (5.7), then solving this system using methods from the PETSc library [20]. However, cell radius changes were not yet enabled. We added the radial update equations (5.9), then carried out some simple tests to check the correctness of the new code. These tests are described in the next section.

5.1.2 A cell pair subject to Buske forces

To test our Buske force law implementation, we ran simulations of a single cell pair. By enabling only one type of cell-cell interaction at a time, simple tests can be created where the expected behaviour is known, allowing simulated radii and positions to be checked.

Elastic interactions only (Equation (5.1))

First, we ran a test with only elastic cell-cell interactions enabled (i.e. with $W_{ij}^A = W_i^K = 0$). If radial changes are also disallowed, then the steady state cell positions should satisfy

$$
\begin{aligned}
&\frac{\partial W_{12}^D}{\partial d_{12}} = 0 \\
\implies &-\sqrt{\frac{R_1R_2}{R_1+R_2}}\frac{(R_1+R_2-d_{12})^{\frac{3}{2}}}{D} = 0 \\
\implies &R_1 = 0, \quad \text{or} \quad R_2 = 0, \quad \text{or} \quad d_{12} = R_1+R_2.
\end{aligned}
$$

A pair of overlapping cells should therefore move apart until $d_{12} = R_1 + R_2$, eliminating the overlap. We ran the corresponding test for two cells with radii 5μm and 10μm, with their centres initially placed 10μm apart (using cells of different sizes leads to more interesting behaviour later in this section). Certain parameter values were altered in these tests in order to reach the steady state in a more reasonable time (see figure legends). Figure 5.1A demonstrates that the expected behaviour indeed occurred.

When instead the two cell centres are fixed in place and only radial changes are permitted, the steady state radius of cell 1 should satisfy

$$\begin{aligned} &\frac{\partial W_{12}^D}{\partial R_1} = 0 \\ \Longrightarrow\ &\frac{1}{2}\left(\frac{R_1R_2}{R_1+R_2}\right)^{-\frac{1}{2}} \frac{R_2^2}{(R_1+R_2)^2} \frac{2(R_1+R_2-d_{12})^{\frac{5}{2}}}{5D} + \left(\frac{R_1R_2}{R_1+R_2}\right)^{\frac{1}{2}} \frac{(R_1+R_2-d_{12})^{\frac{3}{2}}}{D} = 0, \end{aligned}$$

with a similar expression applying to cell 2. Again, $R_1 + R_2 = d_{12}$ is a solution, so two overlapping cells should shrink until they just touch. Figure 5.1B shows that our test simulation displayed the correct behaviour.

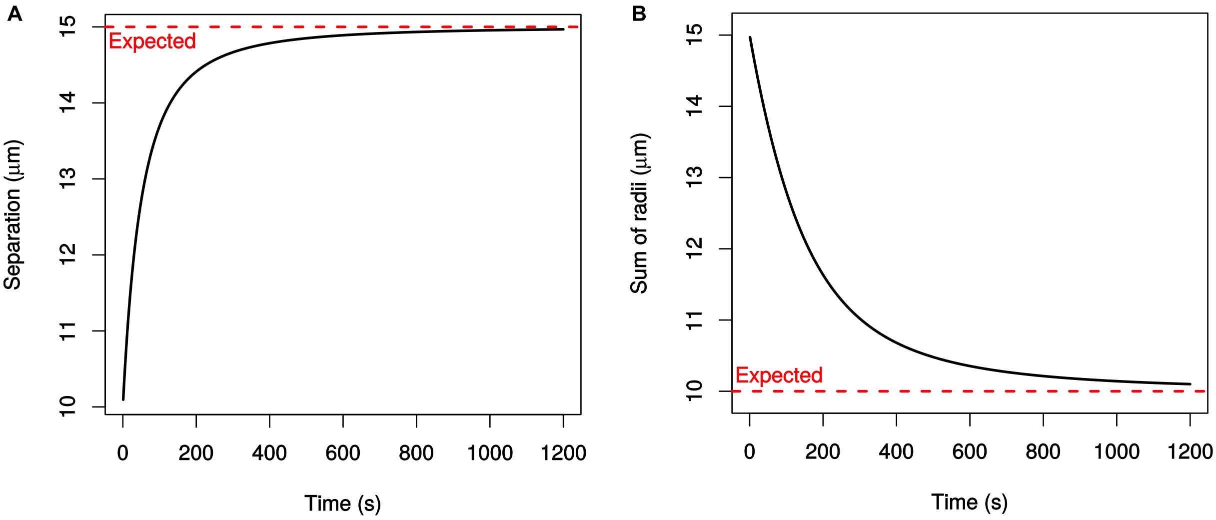


Figure 5.1: A Buske cell pair simulation with elastic interactions only. A) Cell separation over time when radial changes are disabled ($D = 1.33 \times 10^{-4} Pa^{-1}$). As expected, the two cells move apart until they no longer overlap. B) The change in $R_1 + R_2$ over time when cell movement is disabled ($D = 1.33 \times 10^{-4} Pa^{-1}$, $\eta_{VO} = 4$ Nsm^{-1}). Both cells shrink until their radial sum equals their separation, again eliminating the overlap.

Adhesive interactions only (Equation (5.2))

Next we allowed only adhesive cell-cell interactions (i.e. $W_{ij}^D = W_i^K = 0$). When cell radius changes are disabled, the equilibrium separation of a cell pair is given by

$$\begin{aligned}
&\frac{W_{12}^A}{\partial d_{12}} = 0, \\
&\quad \Longrightarrow -2\varepsilon\pi x_{12}\left(1 - \frac{x_{12}}{d_{12}}\right) = 0, \\
&\quad \Longrightarrow x_{12} = 0, \quad \text{or} \quad x_{12} = d_{12}, \\
&\quad \Longrightarrow d_{12} = \sqrt{R_2^2 - R_1^2}, \quad \text{or} \quad d_{12} = \sqrt{R_1^2 - R_2^2}.
\end{aligned}$$

Cell separation should therefore approach a limit, which for our two test cells with radii 5μm and 10μm is 8.66μm. Figure 5.2A shows that cell separation converged to the correct value in our test run.

When instead cells are fixed in place and only radial changes are permitted, an overlapping cell pair will seek to maximize their adhesion energy $W_{12}^A = \varepsilon\pi(R_1^2 - x_{12}^2)$. Setting $R_1 = R_2$, x_{12} becomes a constant (see Equation (5.4)). W_{12}^A can then be made arbitrarily large by letting R_1 and R_2 go to infinity at the same rate. The behaviour in our test simulation matched this description, with both cell radii converging then growing indefinitely (Figure 5.2B).

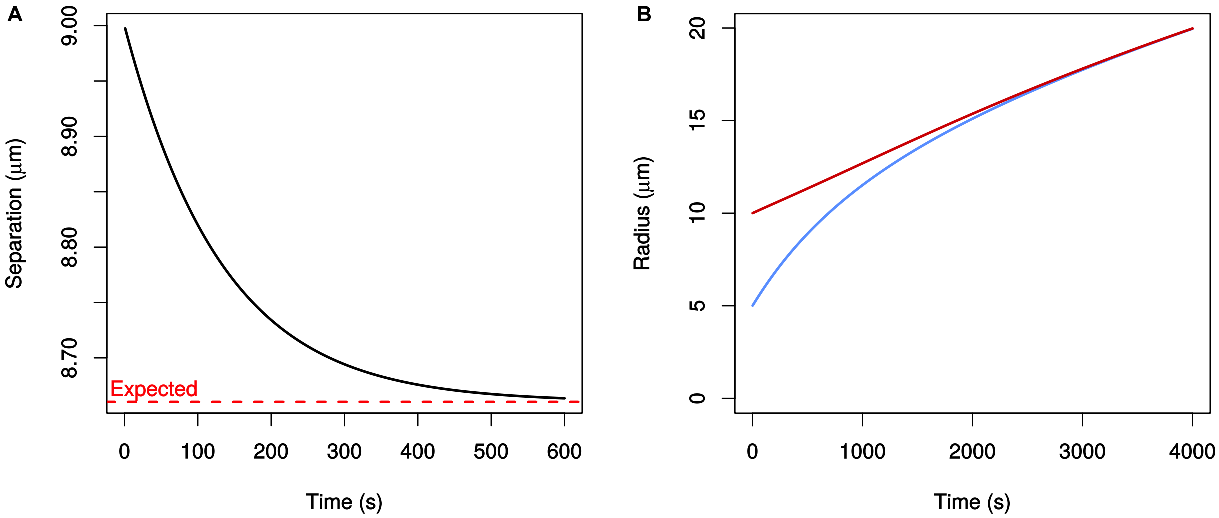


Figure 5.2: A Buske cell pair simulation with adhesive interactions only. A) Cell separation over time when radial changes are disabled ($\varepsilon = 2.0 \times 10^{-3}$ Nm^{-1}). The two overlapping cells moved closer together due to adhesion, with their separation tending to our calculated limit. B) Radial changes over time when cell movement is disabled ($\varepsilon = 4.0 \times 10^{-3}$ Nm^{-1}, $\eta_{VO} = 4$ Nsm^{-1}). The two colours indicate the two different test cells. In the absence of an energetic penalty for volume changes, both cells achieved an equal size, then grew without bound to maximize their adhesion energy.

Compression energy only (Equation (5.5))

Finally, we activated only the compression part of the force law (i.e. $W_{ij}^A = W_{ij}^D = 0$). When radial changes are disabled, a pair of overlapping cells can attain their rest volume by moving apart, eliminating any compression due to their overlap. This was observed in our test (Figure 5.3A). If instead the two cells are fixed in place but their radii are allowed to vary, they can still approach their target volumes by growing or shrinking. Figure 5.3B shows that the discrepancy between actual and target cell volumes tended to zero over the course of our test simulation.

Having confirmed that the Buske force law code works correctly, we next sought comparisons with the GLS force law in a simple test simulation.

5.1.3 Force law comparison for a cell pair

We used the following approach to compare the Buske and GLS force laws in a cell pair simulation. Initially, we sought to make the two force laws agree as closely as possible, by disabling several of the cell-cell interaction types in the Buske force law, and by choosing suitable values for the GLS parameters, μ (spring strength) and η (drag). The disabled components of the Buske force law were then reintroduced one by one, to observe the effect on the fit. By this incremental method, we sought to identify which biophysical details included by Buske *et al.* make a practical difference to cell behaviour.

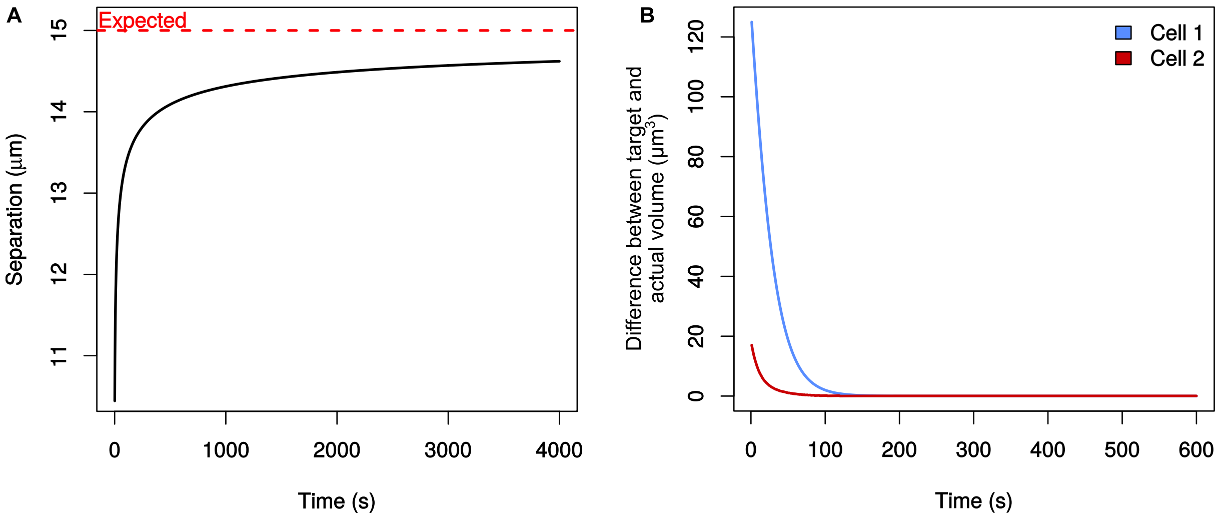


Figure 5.3: A Buske cell pair simulation with compression only. A) Cell separation over time when radial changes are disabled ($K = 10^5$ Pa). To eliminate compression by one another, the two test cells move apart until they no longer overlap. B) The difference between each cell's target volume and current volume over time, when cell movement is disabled ($K = 10^5$ Pa). Radial changes compensate for the cell-cell overlap, until both cells attain their relaxed volume.

Elastic interactions only (Equation (5.1))

We began with a Buske force law containing only elastic cell-cell interactions, without cell-cell friction and without radial updates. The Buske position update equation

$$\eta_{BM}\frac{d\mathbf{r}_1}{dt} = -\sqrt{\frac{R_1R_2}{R_1+R_2}}\frac{(R_1+R_2-d_{12})^{\frac{3}{2}}}{D}\hat{\mathbf{r}}_{12}, \tag{5.10}$$

is then directly comparable to the GLS update equation

$$\eta\frac{d\mathbf{r}_1}{dt} = \mu(R_1+R_2)\log\left(1-\frac{R_1+R_2-d_{12}}{R_1+R_2}\right)\hat{\mathbf{r}}_{12}. \tag{5.11}$$

Taking $R_1 = R_2 = 3\mu\text{m}$ (approximately germ cell size [15]), and letting $\eta = \eta_{BM} = 11520\text{kg}\,\text{h}^{-1}$ (Table 5.1), it is possible to determine the value of μ that produces the closest agreement between the GLS and Buske force laws under these conditions. The right hand sides of Equations (5.10) and (5.11) can be evaluated at a range of cell separations, and interval bisection can then be carried out on μ to minimize the difference between the two force magnitude vectors.

The best fit μ was found to be $14152\text{kg}\,\text{h}^{-2}$ (see Figure 5.4A). We then ran a test simulation in which two 3µm radius cells were placed 1µm apart and their separation over time was recorded. The best fit value of μ was used, with $\eta = \eta_{BM} = 11520\text{kg}\,\text{h}^{-1}$. Both force laws were found to give a very similar cell separation plot, as expected given the close force magnitude agreement

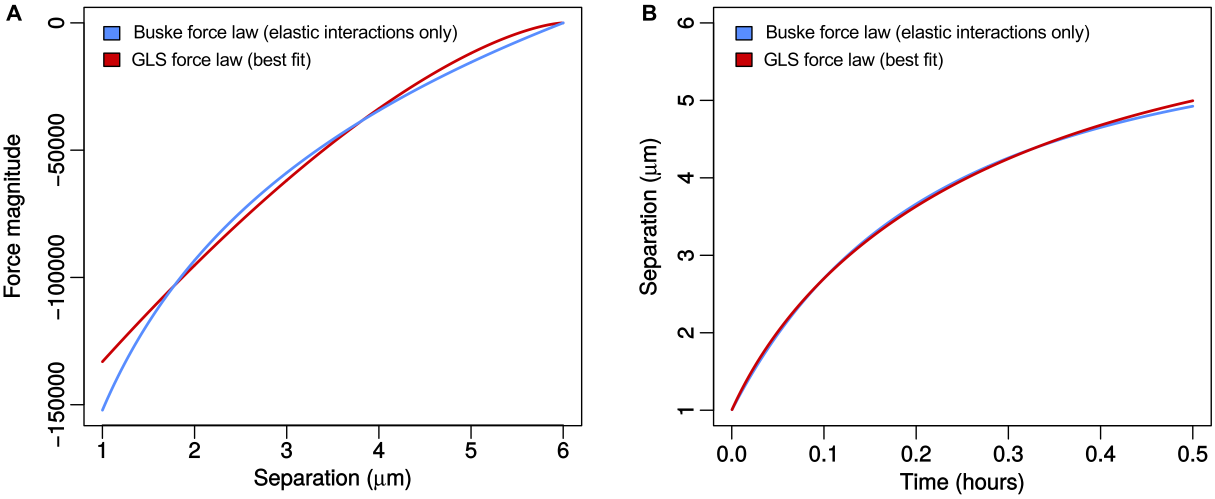


Figure 5.4: The closest fit possible between a Buske force law with elastic interactions only and a GLS force law. A) Taking $\mu = 14152\text{kg}\,\text{h}^{-2}$, the GLS force law generates very similar force magnitudes to the Buske force law with elastic interactions only. B) Taking $\mu = 14152\text{kg}\,\text{h}^{-2}$ and $\eta = \eta_{BM} = 11520\text{kg}\,\text{h}^{-1}$, the two force laws also produce almost identical cell separation plots in a test simulation. In panel B), the L2 norm of the difference between the two separation vectors is 1.29.

under this parameter set (Figure 5.4B). These parameter choices are very different from Chaste's default spring constant and drag coefficient (Table 3.1). However, the important quantity in the GLS force law is the ratio μ/η, so we can set η back to its default value provided μ is also scaled appropriately. This gives $\mu = 1.23\text{kg}\,\text{h}^{-2}$, much closer to the default spring strength of $15\text{kg}\,\text{h}^{-2}$

.

Elastic interactions and cell-cell friction

Next, we reintroduced cell-cell friction into the Buske update equation, by increasing the cell-cell friction coefficient η_C from $10^{-10}\text{kg}\,\mu\text{m}^{-2}\,\text{h}^{-1}$ back to its usual value of $180\text{kg}\,\mu\text{m}^{-2}\,\text{h}^{-1}$. The same cell pair simulation was then repeated, using a Buske model incorporating elastic interactions and cell-cell friction. Figure 5.5 shows the resulting cell separation over time, compared against the GLS force law run from Figure 5.4B. The close agreement between the two separation plots noted in Figure 5.4B persists almost unchanged in Figure 5.5. We therefore conclude that adding cell-cell friction had little practical effect, at least in this test.

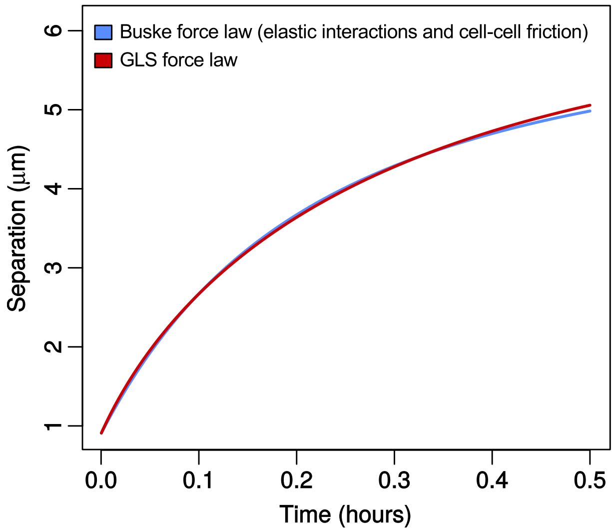


Figure 5.5: Reintroducing cell-cell friction into the Buske update equation. A Buske model incorporating elastic interactions and cell-cell friction continues to produce similar results to the GLS force law from Figure 5.4B. Introducing cell-cell friction therefore had little practical effect on movement in this test.

Elastic interactions and compression energy (Equations (5.1) and (5.5))

Building on the Buske model with elastic interactions and cell-cell friction, we then reintroduced the compression component of the force law given in Equation (5.5). Figure 5.6 shows the resulting cell separation plot for our test simulation, which is again compared with the GLS force law result from Figure 5.4B. Adding the compression energy term produced a noticeable effect, causing the cell pair to separate more rapidly and therefore changing behaviour relative to our best fit GLS run. With compression enabled, there is an energetic penalty associated with any volume lost in a cell-cell overlap. As a result, a larger force acts to push cells apart than in the purely elastic case.

Elastic interactions, compression and radial variation (Equations (5.1), (5.5) and (5.9))

Continuing to build up to the full Buske model, we next introduced the radial update equation (5.9). This had little impact on the rate of cell-cell separation in our test simulation (Figure 5.7A). Moreover, the cell radius changes that occurred using this Buske force law were small (Figure 5.7B). It appears that with the default parameter set, cells are much more likely to move apart than change in volume. However, it is interesting to note that in Figure 5.7B cell radii actually increased slightly under the Buske model before returning to their rest values. This can be thought of as cells responding to an overlap by bulging outward to maintain a constant volume. Since shape changes are not included in the Buske model, radial variations can appear counterintuitive, particularly when cells are visualised as spheres.

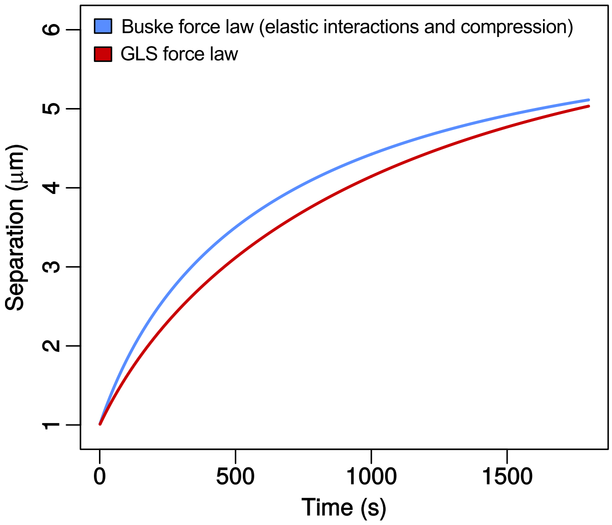


Figure 5.6: Cell separation under a Buske force law that includes elastic interactions and compression. When the compression energy term is introduced, cell separation becomes more rapid under the Buske model. This reduces the agreement with the GLS force law result from Figure 5.4B (the L2 norm of the separation difference is increased from 1.29 to 9.9).

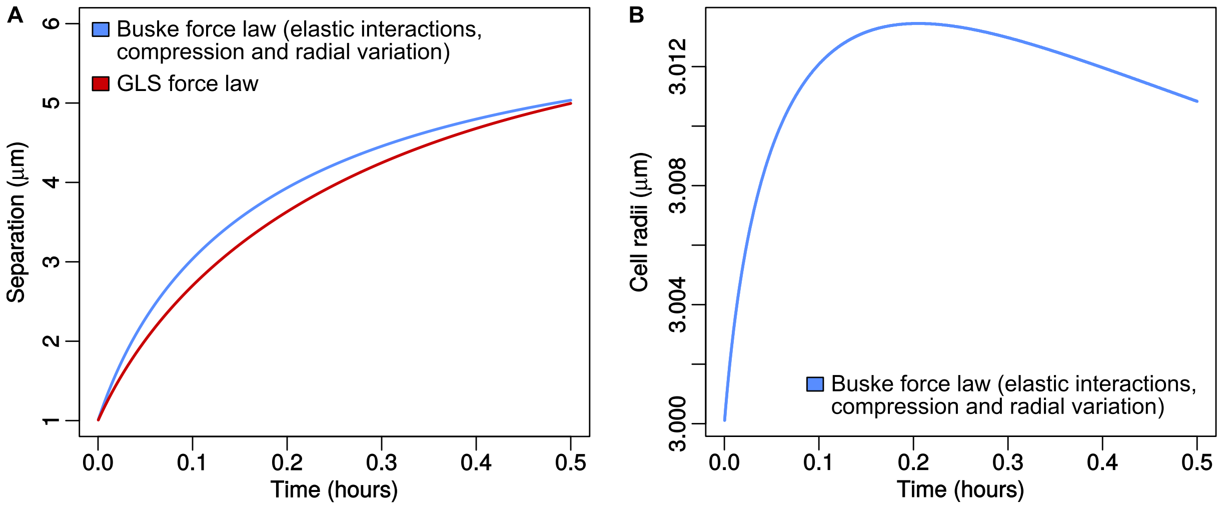


Figure 5.7: Cell behaviour under a Buske model including elasticity, compression and radial updates. A) Cell separation over time in our test simulation was barely affected by introducing radial updates. Relative to our best fit GLS force law, the L2 norm of the separation difference remained 9.90 to three significant figures. B) Using the parameter set in Table 5.1, cell radii did not change greatly in our test simulation. Under this Buske force law, both cells in the pair grew briefly to compensate for the volume lost in their overlap, before slowly returning to their rest sizes after separation.

The full Buske force law with adhesion (Equation (5.2))

Finally, we ran a cell pair simulation using the full Buske force law, including the adhesion energy term given in Equation (5.2). The corresponding cell separation results are shown in Figure 5.8. Introducing adhesion markedly changed the Buske simulation, causing the cells to separate more slowly and approach a steady state with a non-zero overlap. We tried varying the value of μ to see whether the GLS force law result could be brought back into agreement. However, the two separation curves now have different profiles, and varying μ did not achieve a close fit.

We conclude from this investigation that the most significant additions in the Buske force law are adhesion and compression, both of which can affect simulation results. The impact of cell-cell friction and radial updates, meanwhile, was negligible, although this may be due to the small scale nature of the test. This covers all possibilities for two overlapping cells. However, it is worth noting a final difference between these force laws for non-overlapping cells. The GLS force law will cause two nearby cells to approach each other until they touch, whereas under a Buske force law non-overlapping cells cannot interact. This disagreement can be eliminated by raising α, or by reverting to a pure repulsion force.

Having examined a simple case study, we now go on to run a simulation with a larger population of cells, more similar to our *C. elegans* germ line model. This requires the application of boundary

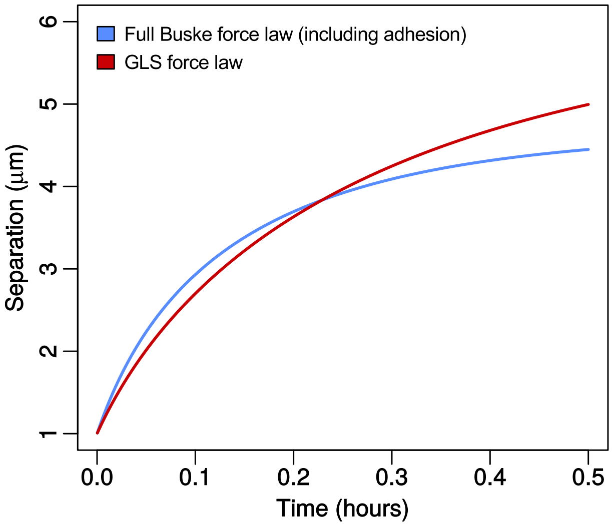


Figure 5.8: Cell separation under the full Buske force law. Reintroducing the adhesion term of the Buske force law changed the shape of the cell separation curve in our test simulation, causing the two cells to separate slowly and approach a steady state with a non-zero overlap. Relative to the GLS run from Figure 5.4B, the L2 norm of the separation difference was 11.4.

conditions to a cell population using Buske updates (Equations (5.7) and (5.9)). The next section describes how an appropriate type of boundary condition was chosen.

5.1.4 Imposing boundary conditions on a Buske simulation

Normally, Chaste uses a “position correction” type of boundary condition. This means boundaries are applied at the end of each time step, by detecting cells that have passed through a boundary and moving them back into the allowed region (Section 3.1.6). However, it is unclear whether this approach is appropriate for the Buske model. If cells are moved artificially at the end of each step, the boundary condition will not be taken into account in the energy calculation used to update positions and radii. As a result, cell radii may not behave correctly in response to confinement.

In the crypt model of Buske *et al.*, a different “knot-based” boundary condition was applied [33]. This involved adding extra nodes to the simulation called knots, which were placed throughout each boundary surface (Figure 5.9). Knots exert an additional force on nearby cells, which acts to maintain the boundary (see [33]). We tested this type of boundary condition, but found it to be poorly suited to simulations involving tight cell packing. When many cells are crowded into a box, large forces arise, and cells tend to squeeze out between the knots. Adding extra knots also adds more potential interactions, slowing the simulation down.

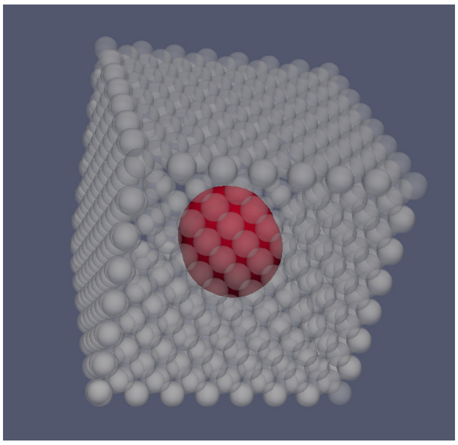


Figure 5.9: Visualisation of a knot-based boundary condition. Knots are represented by white spheres and cells by red ones.

A different possibility would be to use a "force-based" boundary; i.e. apply an extra repulsion force to cells, which grows very strong close to the boundary. For instance, to add a plane, the perpendicular distance d_i between the centre of cell i and the plane boundary is measured. A repulsion force is then added that grows exponentially as the cell approaches the plane

$$\mathbf{B}_i = \begin{cases} -S\exp(d_i)\hat{\mathbf{n}}, & \text{if } d_i > -R_i. \\ 0, & \text{if } d_i \leq -R_i. \end{cases} \tag{5.12}$$

Here, $\mathbf{p}$ is a point on the plane, while $\hat{\mathbf{n}}$ is the plane's unit normal, and $d_i = (\mathbf{r}_i - \mathbf{p}) \cdot \hat{\mathbf{n}}$. S is a dimensionless constant that is here set equal to 10^6; making the magnitude of the boundary force much greater than the typical magnitude of elastic repulsion, to facilitate containment.

We compared the position correction approach and the force-based approach to boundary conditions in a simple test simulation. A dividing cell population was confined inside a cube of side 6µm, formed from six planar boundaries. A simulation was run beginning with a single 3µm radius cell in the middle of the cube. Cells then divided synchronously every 8 hours, and the simulation was continued for 40 hours or until it terminated with an error (absolute movement threshold violation). The test was run once using the full Buske force law, and was also repeated with certain force law components disabled, for comparison.

Figure 5.10 shows our results. Surprisingly, the two boundary condition types produced much the same outcome. The only difference was that the force-based boundary sometimes struggled to contain tightly packed cells completely. We therefore concluded that, contrary to our expectations, standard Chaste boundary conditions are suitable for use in a Buske simulation. Interestingly, cells that were subject to elastic interactions only shrank in response to crowding after filling the box, whereas cells subject to the compression and adhesion terms of the Buske force law tended to

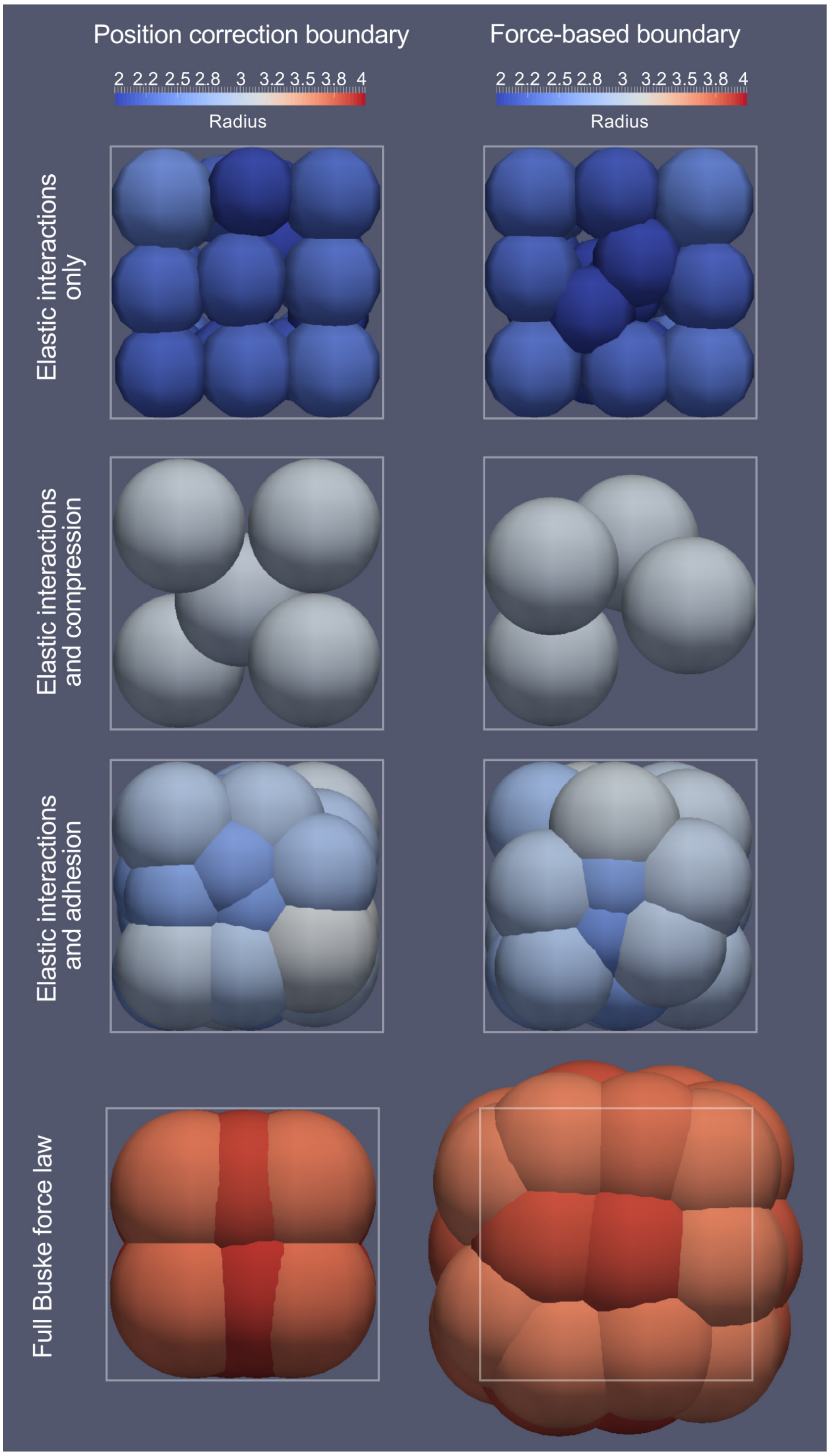


Figure 5.10: Confining a dividing Buske cell population to a box, using different types of boundary and different force law terms. The columns of this figure show snapshots of the same simulation run using two different types of boundary condition. Meanwhile, the rows correspond to different variations on the Buske force law (for example, the top row allows only elastic cell-cell interactions). The position correction boundary produced very similar results to the force-based boundary in terms of cell radii and positions. The only difference was that a force-based boundary sometimes failed to contain tightly packed cells completely (see the bottom right image where cells extend beyond the box).

grow. Mathematically speaking, this is a plausible way for cells to compensate for the volume they lose due to overlaps; it can be interpreted either as the cells bulging outward, or increasing their adhesion energy. However, it could also indicate that this model begins to break down for large cell-cell overlaps in a confined space.

Having established an appropriate way to apply boundary conditions, we went on to compare the Buske and GLS force laws in the context of a distal germ line simulation. The next section describes the simulation in more detail and presents the results.

5.1.5 Force law comparison for a distal germ line simulation

To compare the Buske and GLS force laws in a larger scale test, a simulation was written resembling the distal *C. elegans* germ line. Using a position correction boundary condition, cells were constrained to lie at the surface of a test tube shaped geometry with radius 11.3µm and length 248µm (all parameter values used in this model are based on the values for *C. elegans* reported in Sections 4.2.1 and 4.2.2). Since this simulation's only purpose was to compare force laws, certain biological details were greatly simplified. All cells were given a radius of 2.8µm, and cells less than 70µm from the distal tip were considered stem and permitted to cycle continuously with the phase lengths shown in Table 5.2. On travelling more than 70µm from the distal tip, cells immediately differentiated regardless of their current phase. Cells exiting the open end of the tube were removed from the simulation. Finally, the daughter cell separation immediately after division was set to $0.95 \times 2R_p$ (where R_p denotes the radius of the parent cell). This is necessary to prevent adhesion in the Buske force law dominating due to a large cell-cell overlap. At the beginning of each simulation, cells were seeded in rings of ten along the length of the tube, with a 5µm spacing between rings.

Phase	Length (hours)
G1	$0.16 + 0.016\,(U[0,1] - 0.5)$
S	4.56
G2	$3.12 + 0.312\,(U[0,1] - 0.5)$
M	0.16

Table 5.2: Cell cycle model for the distal germ line tests. The G1 and G2 phase lengths both include a stochastic contribution, where $U[0,1]$ denotes a uniformly distributed random variable between 0 and 1.

Buske force law simulations

We began by running a set of distal germ line simulations using a Buske force law with default parameters (Table 5.1). One run was performed using the full force law, while a number of other runs were performed with certain force law terms disabled. Figure 5.11 shows snapshots from a collection of Buske simulations in which radial updates were permitted. As in our box boundary test, when only elastic interactions were enabled (Equation (5.1)) cells shrank to reduce their overlap, reaching radii as small as 1.2µm. When the adhesion and compression terms were enabled (Equations (5.2) and (5.5)), cells compensated for their overlap by growing to increase their adhesion energy and overall volume. In this tight packing, large overlap scenario, the radial update equation therefore produces unreasonably large cell size changes.

We therefore ran a number of Buske force law simulations with radial updates *disabled* (see Equation (5.9)). Figure 5.12 shows some distal germ line snapshots that were generated using fixed cell radii. We see that, with this parameter set and a cell cycle length of 8 hours, cells accumulate in the proliferative part of the germ line. After approximately 35 hours, all simulations slowed down and ceased making reasonable progress due to the huge number of cells present. We conclude that applying a more detailed force law does not eliminate the problem of distal overcrowding discussed in Chapter 4. Buske *et al.*'s original paper included a contact inhibition assumption similar

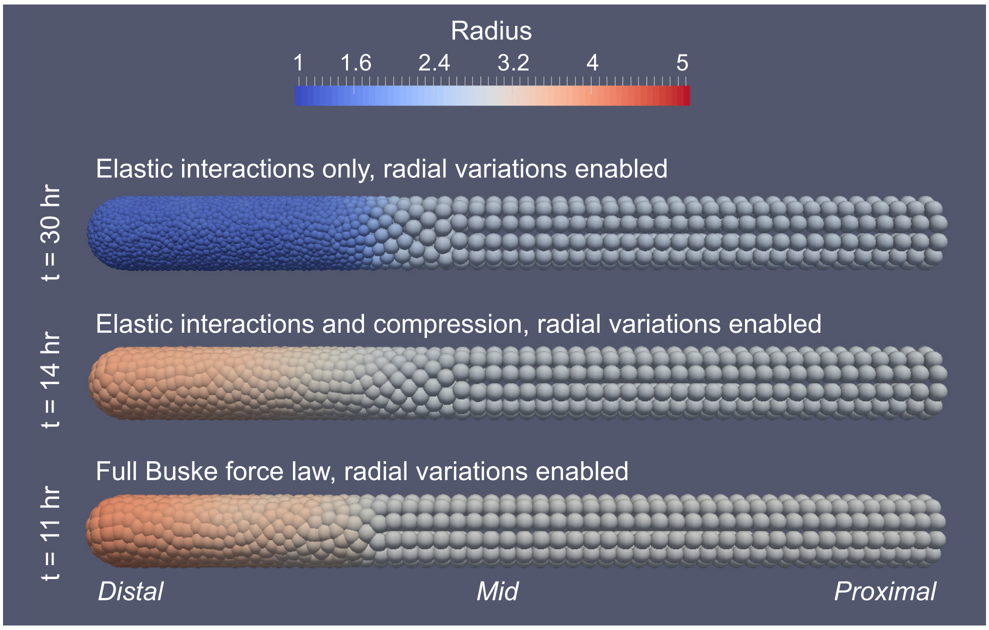


Figure 5.11: Distal germ line results using the Buske model with radial updates. The snapshots above show a set of distal germ line simulations run with variable radii. When only elastic cell-cell interactions were permitted, cells shrank rapidly to relieve overcrowding. When the compression and adhesion terms were enabled, cells instead grew; a response that may be more properly interpreted as a shape change (cells bulging outward).

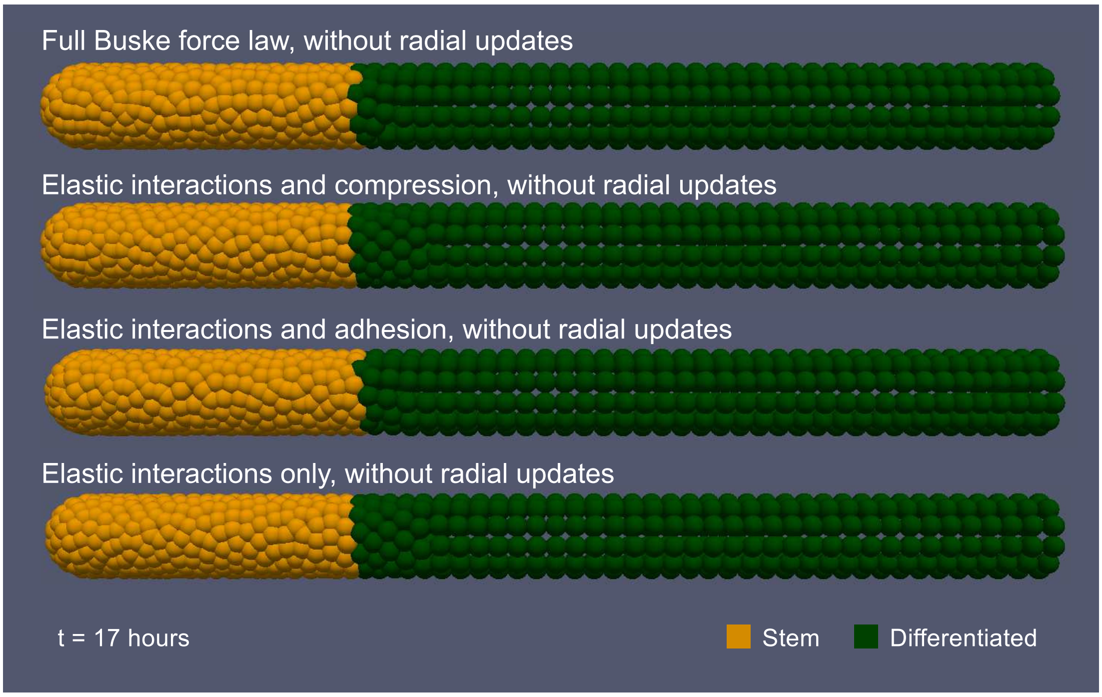


Figure 5.12: Distal germ line simulations using the Buske force law with radial updates disabled. Every variant of the Buske force law with default parameters produced similar looking results. A large number of cells accumulated in the distal third of the gonad, eventually making the simulation progress too slowly to continue.

to the one proposed in Chapter 4, which may have prevented cell accumulation [33].

Due to overcrowding, none of our Buske force law simulations were able to reach a steady state or progress to the point where cells began exiting the system. As such, the only measurement we can compare between runs is the cell count over time in each section of the gonad (Figure 5.13). It appears that disabling the compression and adhesion terms (Equations (5.5) and (5.2)) has little effect on the number of cells in the distal third of the gonad, where cell numbers are presumably dictated by the proliferation rate. However, cell counts in the middle third *are* influenced by which force law components are active, since this affects how rapidly the mass of compacted cells relaxes along the gonad. When adhesion is enabled, cells separate more gradually and take longer to enter the middle third. If resistance to compression is enabled, cells separate faster and the cell count in the middle third of the gonad rises more rapidly.

GLS force law simulations

For comparison, we ran some distal germ line simulations using the GLS force law. One set of simulations was performed using the default attraction decay rate ($\alpha = 5$), while a second set took $\alpha = 100$ — essentially a pure repulsion force. A range of values were tested for the ratio μ/η, and we sought to determine which parameter set most closely resembled our Buske results.

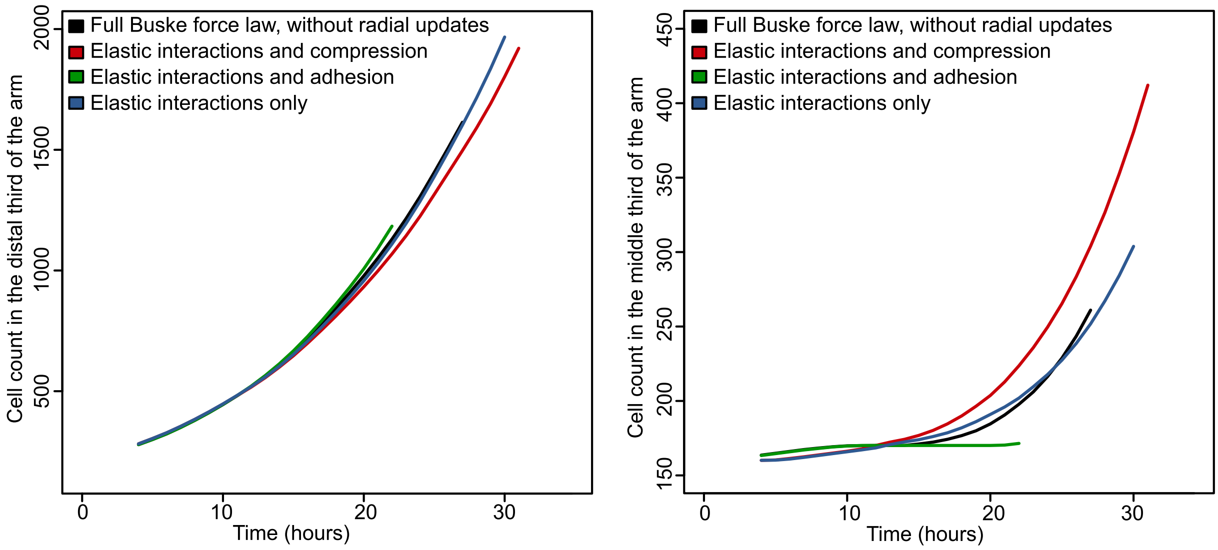


Figure 5.13: Cell counts over time for a number of distal germ line simulations, using a Buske force law without radial updates. The left hand panel shows cell counts in the distal third of the gonad, while the right hand panel shows counts in the middle third (counts in the final third remained constant). Changing which Buske force law terms were enabled had little effect on cell accumulation in the distal gonad. However, disabling adhesion caused cells to seperate more rapidly, leading to faster movement into the middle third.

Figure 5.14 presents cell count measurements for our GLS simulations. Each row of the figure corresponds to a particular region of the gonad, and the two columns show results for the two different values of α. A solid black line indicates the cell counts produced using the full Buske force law with fixed radii. With $\alpha = 5$, cells accumulated in the distal third of the gonad just as in the Buske case, with cell counts of 1000 or more recorded within 50 simulated hours. These runs eventually had to be abandoned because the simulations ceased to make progress. Increasing α to 100 led to lower cell counts for the same set of spring properties, and a steady state was achieved given a large enough value of μ. Of the parameter sets tested, $\alpha = 5$ with $\mu/\eta = 5$ gave the best fit to the Buske force law results.

For the GLS simulations, we also measured “compressed” cell volumes, using the same formula given in Section 4.1.3 and in [15]. These results underline our earlier comments about cell accumulation and packing density (see Figure 5.15). When $\alpha = 5$, cells overlap to such an extent that their effective volume is less than half of their relaxed volume. Given much larger choices of α and of μ/η, the amount of cell “compression” can be reduced to 15%.

Finally, Figure 5.16 shows the number of cells exiting the system over time in our GLS simulations (a possible measure of niche output). With $\alpha = 5$ and with $\mu/\eta < 70$, cells were not observed to exit the gonad during the first 50 hours. Even with $\mu/\eta = 234$, niche output was zero for the first 20 hours, since cells accumulated rather than ejecting their neighbours from the system. Once

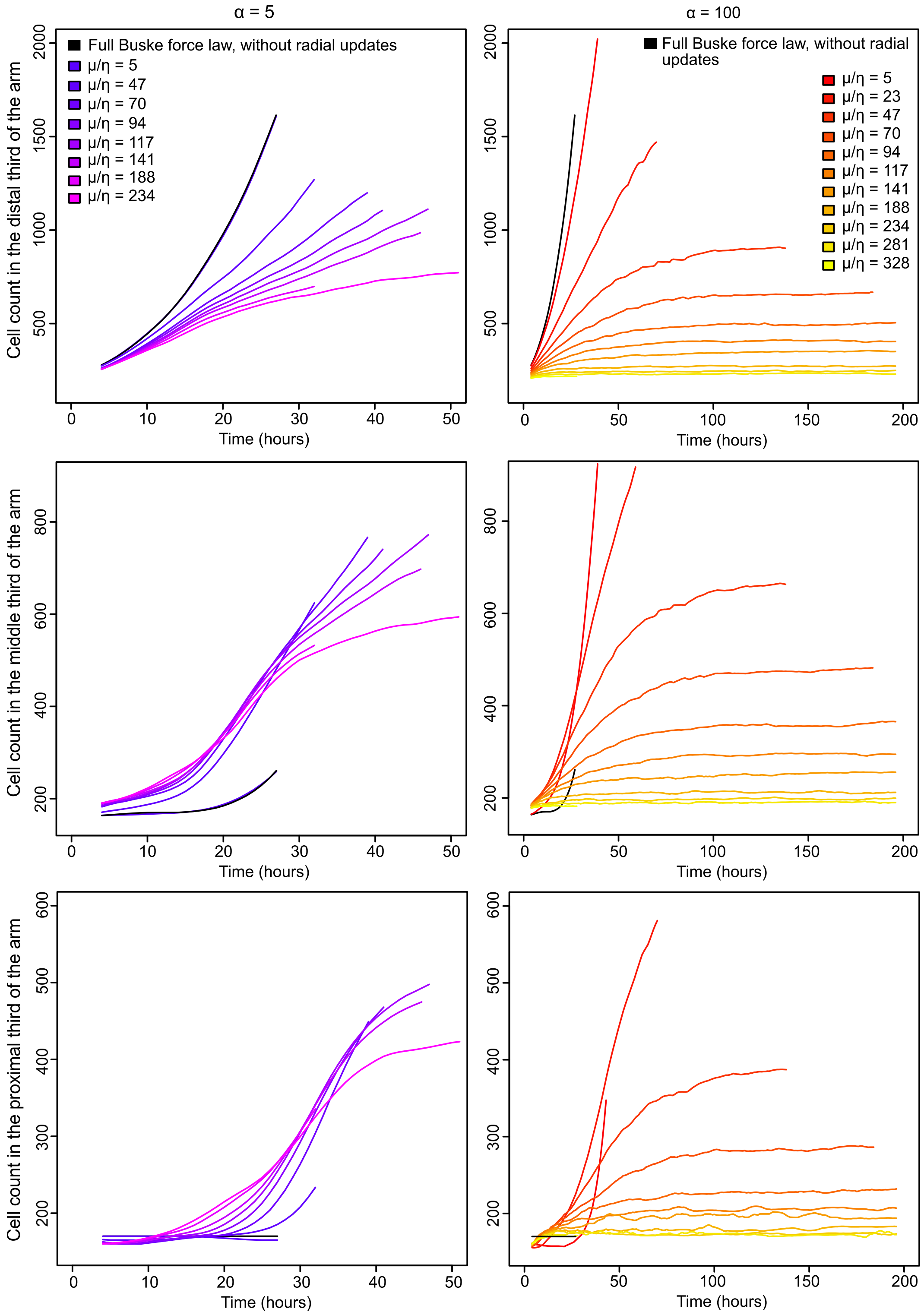


Figure 5.14: Cell counts over time for a set of GLS distal germ line simulations. Top to bottom, the three rows of this figure show cell counts in the distal, mid and proximal thirds of the gonad. Each column corresponds to a different choice of α: 5 on the left and 100 on the right. Some runs were terminated early due to large cell counts causing extremely slow progress. The best fit to the cell counts for a Buske simulation (black line) occurred at $\alpha = 5$, $\mu/\eta = 5$.

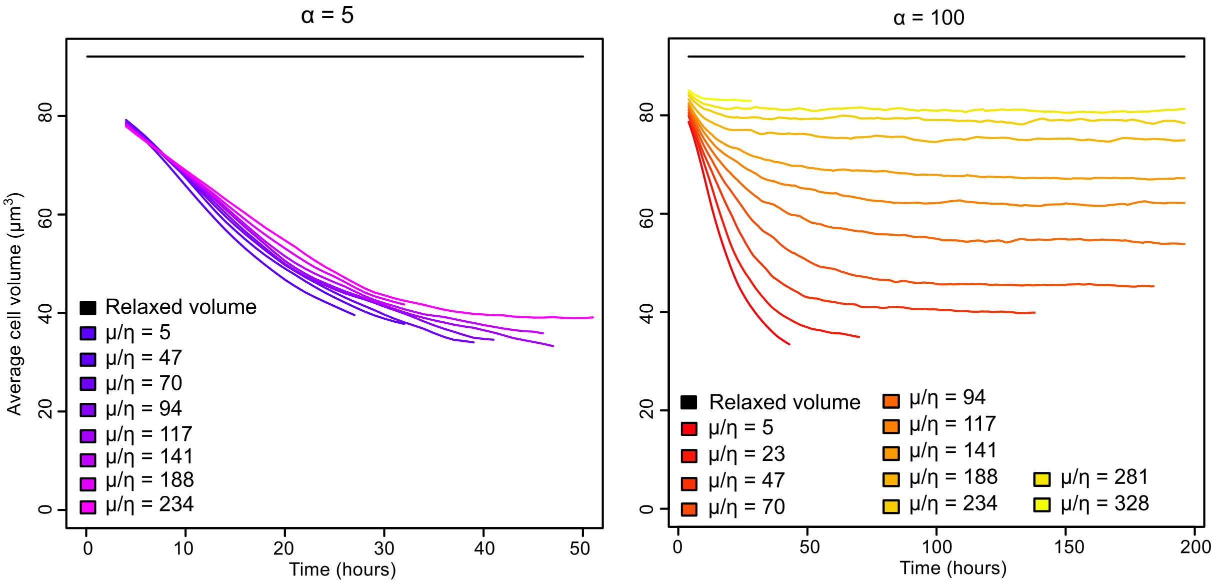


Figure 5.15: Mean "compressed" cell volume over time for a number of GLS simulations. When $\alpha = 5$ or μ/η is low, large cell-cell overlaps form and effective cell volumes drop far below their relaxed value. Lower levels of crowding can be achieved by increasing α and μ.

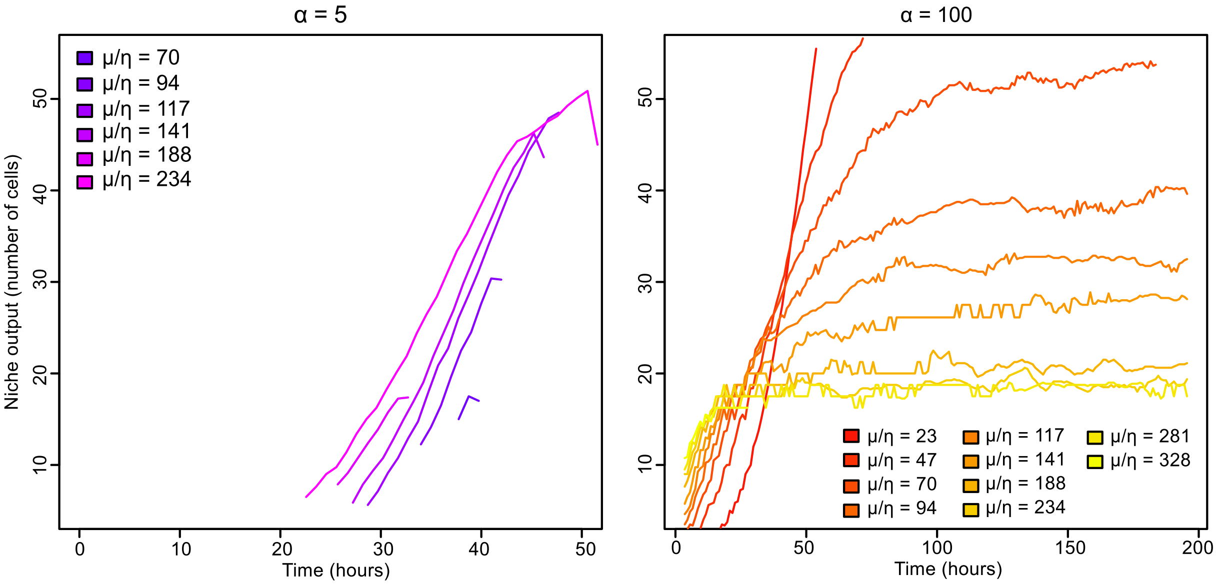


Figure 5.16: Cells exiting the gonad per hour for a number of GLS simulations. When $\alpha = 5$, niche output is initially zero, since cells accumulate rather than pushing apart. Once cells do begin to exit the gonad, they do so rapidly, and the output per hour becomes high. Taking $\alpha = 100$ and μ/η large, cell output can be stabilised at a more reasonable level for the *C. elegans* gonad.

cells did begin to leave the gonad, output quickly rose as high as 50 cells per hour. By contrast, with α set to 100 and with μ/η large, niche output stabilised rapidly at 10-20 cells per hour.

This concludes our comparison of the Buske and GLS force laws. We have seen how parameters might be chosen to minimize the difference between the two forces for a cell pair. We also found that adhesion is perhaps the most significant addition in the Buske force law, and we saw that cell radius changes can become counterintuitive in situations with tight cell packing. Finally, in a distal germ line simulation, we found that applying a Buske force law did not substantially change outcomes relative to a GLS force with weak springs.

We now go on to look at the second of the topics covered in this chapter, namely, whether an alternative numerical scheme might improve performance and accuracy in cell-based simulations.

5.2 Alternative numerical methods for cell-based simulation

In this section, we investigate the numerics of cell-based simulations in Chaste. The motivation is twofold. First, we want to understand how accurate Chaste simulations currently are. Second, we want to see whether changing the numerical method could improve performance.

With regard to accuracy, different users of the code will have different requirements. For instance, a scientist modelling tumour growth will be interested in large scale simulation summary statistics, such as the population's radial growth rate. By contrast, a developmental biologist is likely to be interested in the exact position of individual cells, in order to make comparisons with experimental tracking data, or confirm that correct neighbour interactions are established.

It is currently unclear whether simulations run using Chaste's default time step produce accurate cell positions, or whether they only produce accurate large scale properties. Even if a particular choice of time step generates converged positions in a short test, it may allow unacceptable errors to accumulate over the course of a longer run. Moreover, details such as the choice of cell-based model or choice of cell cycle could also affect how rapidly error accumulates. If Chaste is to be used for research in developmental biology, a better understanding of the numerical error in cell positions is required.

With regard to performance, 3D simulations can be slow to run [15, 62], and addressing this issue would allow larger problems to be tackled. Using an alternative numerical scheme in place of the forward Euler method would increase the computational cost of each time step. However, it should also produce a more accurate approximation to the solution of the position update equation. As a result, it might be possible to take larger steps while maintaining acceptable error levels. If the gains from taking larger time steps outweigh the additional cost per step, an alternative numerical scheme will improve run times. This trade-off has been investigated before for non-biological problems [98, 232], but cell-based simulations present special challenges that could lead to new behaviour, such as the abrupt increase in problem size after each cell division.

Here, we consider three alternative choices of numerical scheme, and compare their speed and accuracy across a range of test simulations. Initially, we focus on cell centre population models with synchronised divisions, before going on to look at other cell cycle choices. Finally, we investigate the potential benefits of adaptive time-stepping.

5.2.1 Current numerical scheme and alternative choices

Chaste supports a range of off-lattice cell population models, however, all of them evolve through time in the same way: by updating the positions of a set of nodes. A node is any point in space that helps define a cell, such as a vertex or cell centre. Recall that node positions are updated according to Newton's second law of motion with overdamping (Equation (3.2)). This is discretised using the forward Euler method to give Equation (3.16),

$$\mathbf{r}_{N+1} = \mathbf{r}_N + \frac{dt}{\eta}\mathbf{F}(\mathbf{r}_N),$$

where $\mathbf{r}_N$ is the vector of node positions at step N, $\mathbf{F}$ is a vector containing the net force on each node, η is a drag coefficient and dt is the time step.

It is recognised in Chaste that the forward Euler method may accumulate error over time or even become unstable. To address these problems, the user can specify an Absolute Movement Threshold (AMT). If any cell travels further than the AMT in one time step, the result is considered spurious, and the simulation halts with an error message. Users also typically repeat their simulations with a smaller time step, to ensure results are unchanged.

The global error of the current forward Euler method is $O(dt)$. A reasonable option would be to test a higher order method to try and achieve more accurate results. We have chosen the 4th order Runge-Kutta method (RK4), which is relatively simple to implement and widely used in scientific computing [196]. It evaluates the force at several intermediate times and positions, then takes a weighted average as follows

$$\mathbf{r}_{N+1} = \mathbf{r}_N + \frac{dt}{6\eta}(\mathbf{K}_1 + 2\mathbf{K}_2 + 2\mathbf{K}_3 + \mathbf{K}_4), \tag{5.13}$$

where

$$\mathbf{K}_1 = \mathbf{F}(\mathbf{r}_N), \tag{5.14a}$$

$$\mathbf{K}_2 = \mathbf{F}(\mathbf{r}_N + \frac{dt}{2\eta}\mathbf{K}_1), \tag{5.14b}$$

$$\mathbf{K}_3 = \mathbf{F}(\mathbf{r}_N + \frac{dt}{2\eta}\mathbf{K}_2), \tag{5.14c}$$

$$\mathbf{K}_4 = \mathbf{F}(\mathbf{r}_N + \frac{dt}{\eta}\mathbf{K}_3). \tag{5.14d}$$

The other methods we consider here are the backward Euler (BE) method and the 2nd order Adams-Moulton method (AM2), also known as the trapezoidal method. Whereas forward Euler and RK4 are both *explicit* methods, BE and AM2 are *implicit*, meaning $\mathbf{r}_{N+1}$ appears on both sides of the update equation. For BE we have

$$\mathbf{r}_{N+1} = \mathbf{r}_N + \frac{dt}{\eta}\mathbf{F}(\mathbf{r}_{N+1}), \tag{5.15}$$

and for AM2

$$\mathbf{r}_{N+1} = \mathbf{r}_N + \frac{dt}{2\eta}(\mathbf{F}(\mathbf{r}_{N+1}) + \mathbf{F}(\mathbf{r}_N)). \tag{5.16}$$

In this case, $\mathbf{F}(\mathbf{r})$ will be nonlinear, as the net force on a node depends on its distance from its neighbours. Consequently, Equations (5.15) and (5.16) cannot be solved by matrix inversion and must instead be treated as root finding problems. In the case of BE, we seek a vector $\mathbf{r}_{N+1}$ satisfying

$$\mathbf{G}(\mathbf{r}_{N+1}) = \mathbf{r}_{N+1} - \mathbf{r}_N - \frac{dt}{\eta}\mathbf{F}(\mathbf{r}_{N+1}) = 0. \tag{5.17}$$

In our implementation, Equation (5.17) is solved using the PETSc library's Newton solver with line search routine (SNESNEWTONLS) [20]. The Newton solver requires an initial guess, which we take to be the current vector of node positions plus one forward Euler step. The choice of initial guess does not affect the outcome of the step, but it must be close enough to the solution of Equation (5.17) for the solver to converge; hence a forward Euler step is necessary. The Newton solver also has a tolerance parameter, which determines how small $\mathbf{G}(\mathbf{r}_{N+1})$ must be for $\mathbf{r}_{N+1}$ to be accepted. The tolerance value is given alongside each set of implicit method results in this chapter, and tolerance choice is discussed further in Section 5.2.3.

Running a Newton solver every time step is costly. In particular, a Jacobian has to be calculated to capture how $\mathbf{G}(\mathbf{r}_{N+1})$ changes as node positions are varied. For optimal performance, an analytic

Jacobian should be provided by the user. However, that would require the addition of a Jacobian function to every force law class in Chaste, and calculating the Jacobian is not always straightforward. Moreover, matrix sparsity cannot be exploited in this application, because the sparsity pattern of the Jacobian changes as cells move. We therefore chose to use PETSc's SNESComputeJacobianDefault routine to compute a numerical Jacobian for the time being. While this is extremely inefficient, it still allows us to assess the step size each method needs to produce reasonably accurate results, putting an upper bound on efficiency.

The rationale for testing some implicit methods, despite their greater computational cost, is that they should cope better with stiff systems. A system is stiff if the time step must be suppressed beyond the size required for accuracy to maintain stability [120]. Stiffness arises in sets of equations involving two very different time scales, or in systems where a small change in some parameter produces a large change in the solution [53]. It is plausible that a cell-based simulation might be stiff, due to the abrupt changes that happen on division, and to force laws in which a small position change produces a large force change. Numerical methods that are *A-stable* handle stiffness better, where A-stability is defined as producing a solution for $dy/dt = ky$ that tends to 0 as $t \to \infty$, for all k with negative real part. Unfortunately, according to the second Dahlquist barrier theorem [54], explicit multistep methods (such as RK4 and forward Euler) cannot be A-stable, and implicit multistep methods can only be A-stable if they are of order 2 or below. Therefore, we test BE and AM2, methods that meet this criterion.

In the next section, we describe the test simulation that was used to compare the performance and accuracy of these methods when solving cell centre problems. We focussed on cell centre simulations initially because this type of population model is used throughout the rest of the thesis, and is particularly simple (only one node per cell).

5.2.2 A 3D cell centre test simulation

In an effort to create a typical cell centre simulation for testing, commonly used classes and default parameters were chosen wherever possible. As such, all cell radii were fixed at 0.5 units, and a GLS force law was applied with the default spring stiffness (see Section 3.1.6). Only three changes were made to Chaste's default parameter set. Firstly, the damping constant was increased from 1 to 1.1, as this provides a better test of whether damping is applied correctly in the new code. Secondly, the AMT was set to 10, allowing simulations to run uninterrupted. Since we were assessing error and convergence manually, there was no need for an AMT to provide warnings about spurious results. Finally, the cutoff distance for the spring force was extended from 1.5 to 3.1, to prevent any discontinuous jump in the force law that might affect the numerics. By the time cells are separated by 3.1 units of distance, the force between them has magnitude $<10^{-3}$.

Each cell centre test began with four non-overlapping cells positioned in a square (see Figure 5.17). A "fixed" cell cycle model was applied, in which cells divide synchronously exactly every 6 hours. In addition, the different values of the time step, dt, were chosen to divide one another perfectly, with an integer number of steps per hour. Both these measures help ensure that runs using different time steps are comparable, with data output and divisions occurring at identical times, and with minimal opportunity for rounding error.

The test simulation described above was repeated 20 times for each numerical method and time step, seeding the random number generator with the run number. Even when using a fixed cell cycle, the random number generator determines the direction of cell divisions, so reseeding produces a slightly different outcome each time.

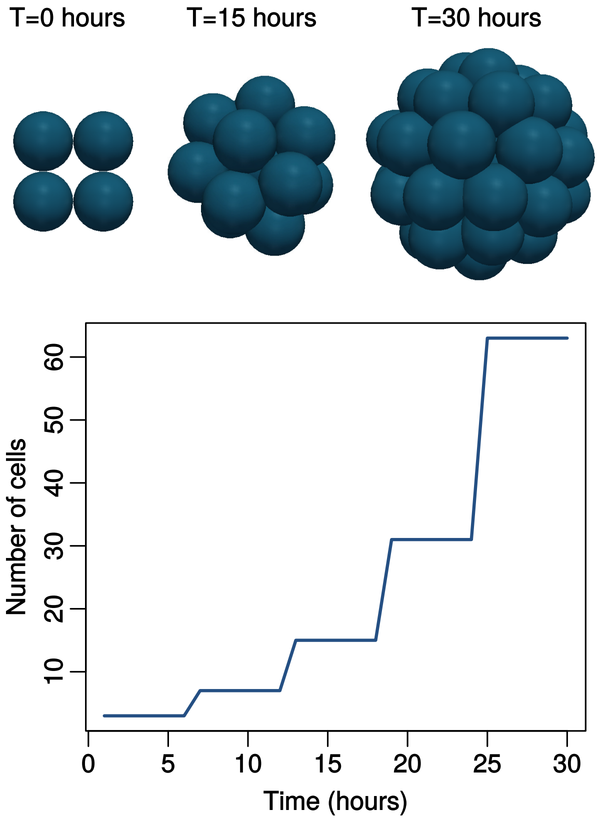


Figure 5.17: The 3D cell centre test simulation. Pictured is the state of a test simulation after 0, 15 and 30 hours. The graph shows the number of cells present over time, with each step up corresponding to a synchronised division event.

Two different types of data were recorded. Firstly, individual cell centre positions were saved hourly. The maximum position error in run A relative to a baseline run B can then be defined as

$$||\text{Position error}^A||_\infty = \max_i |\mathbf{r}_i^A - \mathbf{r}_i^B|, \tag{5.18}$$

where i ranges over all cells, $\mathbf{r}_i^A$ is the position of cell i in run A, and $\mathbf{r}_i^B$ is the position of cell i in the baseline run. Since we carry out replicates with different random seeds, the plots in the next section actually report the mean

$$\frac{1}{R}\sum_R ||\text{Position error}^R||_\infty, \tag{5.19}$$

where R ranges over all successful runs.

We also recorded the extent of the cell population hourly. Extent was calculated by averaging the maximum length of the population in the x, y and z directions at time t

$$\text{Extent} = \frac{1}{3}\left[(x_{\max} - x_{\min}) + (y_{\max} - y_{\min}) + (z_{\max} - z_{\min})\right]. \tag{5.20}$$

Again, the difference in extent can be calculated between two runs, and an average extent error can be determined for a collection of replicates. In terms of our aims in this chapter, population extent serves as an example of a large scale property, which may be simulated accurately even if individual cell positions are not.

5.2.3 Numerical method comparison for a cell centre population

Figure 5.18 summarises the position results we obtained from our cell centre tests. Each panel corresponds to a single numerical method, and we plot the maximum position error over all cell centres as a function of time. Position error is measured relative to an RK4 run with a very small time step ($dt = 2^{-14}$), which we take as a baseline.

Since the graphs in each panel approach zero as dt goes to zero, we conclude that all four methods converge on identical cell positions. This also suggests that cell positions are close to their true values in the baseline RK4 run. Another obvious feature of Figure 5.18 is the step-like profile of each curve. Every sudden increase in position error corresponds to a synchronised division event. It is unsurprising that numerical error mainly accumulates during cell divisions, since without any active migration the population is essentially static at other times.

Figure 5.18 can be used to determine the time step each method requires to generate reasonably accurate cell positions. Here, we take "reasonably accurate" to mean a maximum error of less than one cell radius - the most meaningful length scale for the problem. Given this definition, the forward Euler method requires a step size of around $dt = 2^{-12}$, BE requires $dt < 2^{-12}$, AM2 requires $dt = 2^{-9}$, and RK4 requires only $dt = 2^{-7}$. We therefore find that the RK4 and AM2

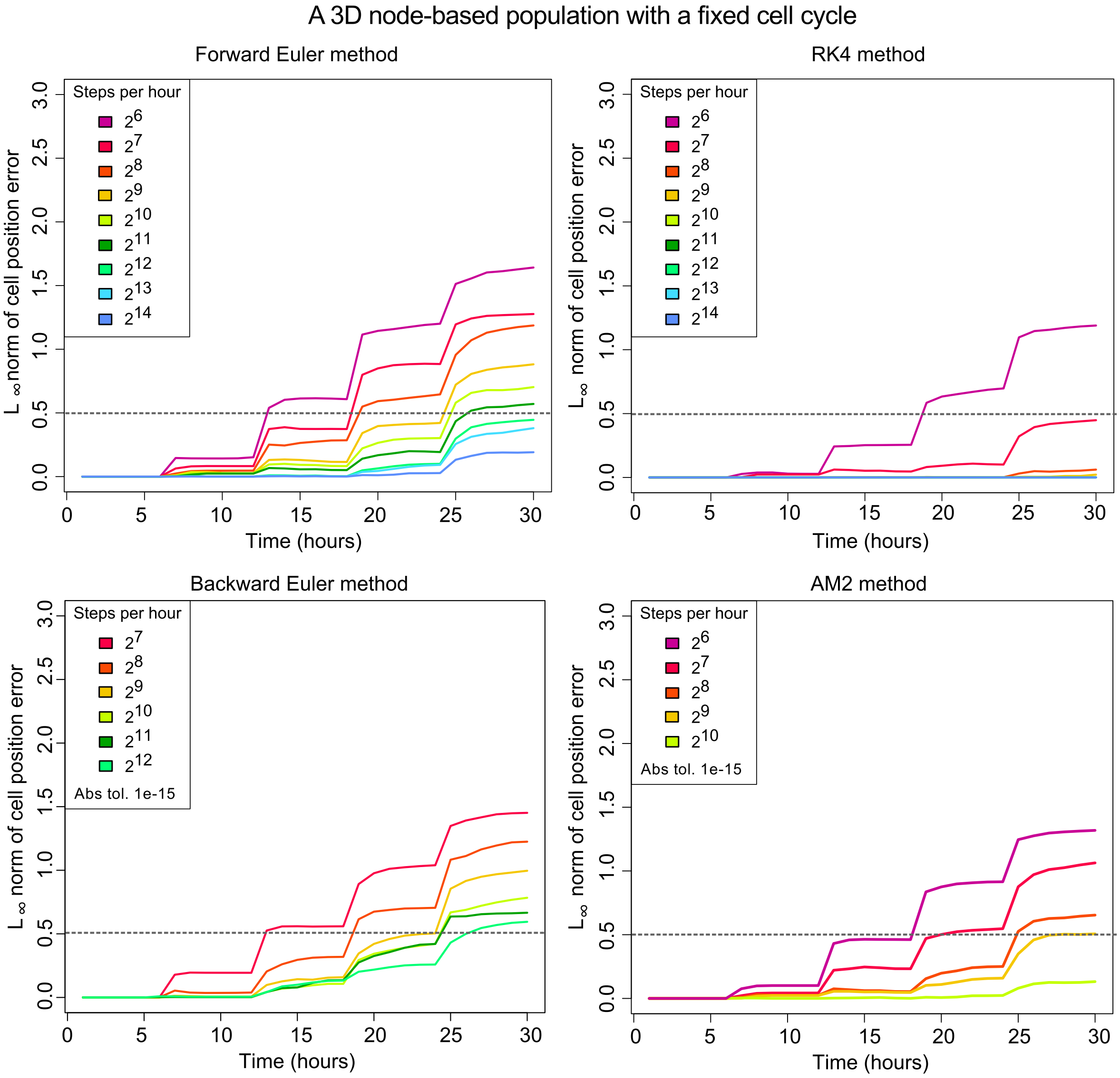


Figure 5.18: The numerical error in cell positions, using different methods and time steps. Plots in this figure show the maximum cell position error over time, measured relative to an RK4 baseline run with $dt = 2^{-14}$. A dashed grey line in each panel marks a position error of one cell radius - the most meaningful length scale for the problem. Traces are the mean of 20 replicates run with different random seeds, and tolerances were chosen for the two implicit methods by progressively lowering abs. tol. until a convergence trend emerged with decreasing dt. These graphs indicate that all four numerical methods converge on the same set of cell centre positions when dt is reduced. The RK4 and AM2 schemes can both produce accurate cell positions using a larger time step than forward Euler, with RK4 allowing the greatest improvement.

numerical schemes both allow a larger time step to be chosen while still preserving accuracy. Note that the default time step in the Chaste library is 120 steps per hour; approximately $dt = 2^{-7}$. Using the forward Euler method, this would *not* be small enough to give accurate cell positions after roughly four rounds of division.

Table 5.3 records the time taken to run each of the simulations in Figure 5.18. Highlighted in blue for each numerical method is the largest step size run that produced accurate positions. The results show that switching to the RK4 method allows a 32-fold increase in dt and a 10-fold increase in speed relative to forward Euler. However, both implicit methods ran far slower than forward Euler. While the implementation of the implicit methods could be improved, it seems unlikely that they will ever outperform RK4 for this problem. To offset their greater computational cost, implicit methods must allow a large increase in step size to give a performance benefit; on the order of 100-fold. Here, dt can only be increased 8-fold using the AM2 scheme. Moreover, the two implicit methods proved somewhat difficult to use, in the sense that if the tolerance parameter is set too low the Newton solver fails to converge, but if the tolerance is set too high the Newton solver returns partially converged positions with a high level of error. Figure 5.19 shows example output for the BE method where too large a tolerance has been applied. This extra complexity makes implicit methods a less attractive choice for use in Chaste, and it appears they do not offer sufficient benefit to be worthwhile.

Steps per hour	Method	Run time (minutes)
2^{14}	Forward Euler	6.47
	RK4	19.1
	Backward Euler	—
	Adams Moulton	—
2^{13}	Forward Euler	3.23
	RK4	9.55
	Backward Euler	—
	Adams Moulton	—
2^{12}	Forward Euler	1.62
	RK4	4.77
	Backward Euler	370
	Adams Moulton	—
2^{11}	Forward Euler	0.808
	RK4	2.38
	Backward Euler	216
	Adams Moulton	—
2^{10}	Forward Euler	0.403
	RK4	1.20
	Backward Euler	115
	Adams Moulton	472
2^{9}	Forward Euler	0.203
	RK4	0.597
	Backward Euler	56.5
	Adams Moulton	176
2^{8}	Forward Euler	0.102
	RK4	0.298
	Backward Euler	27.2
	Adams Moulton	66.5
2^{7}	Forward Euler	0.0508
	RK4	0.151
	Backward Euler	14.4
	Adams Moulton	28.2
2^{6}	Forward Euler	0.0267
	RK4	0.075
	Backward Euler	—
	Adams Moulton	13.9

Table 5.3: Run times for the cell centre test simulation, using different numerical methods and time steps. Values reported here are the mean of all successful replicates run using different random seeds. For each numerical method, the highlighted row indicates the largest dt simulation that gave accurate cell positions. RK4 outperforms forward Euler by a factor of 10, while the greater computational cost of the implicit methods makes them significantly slower.

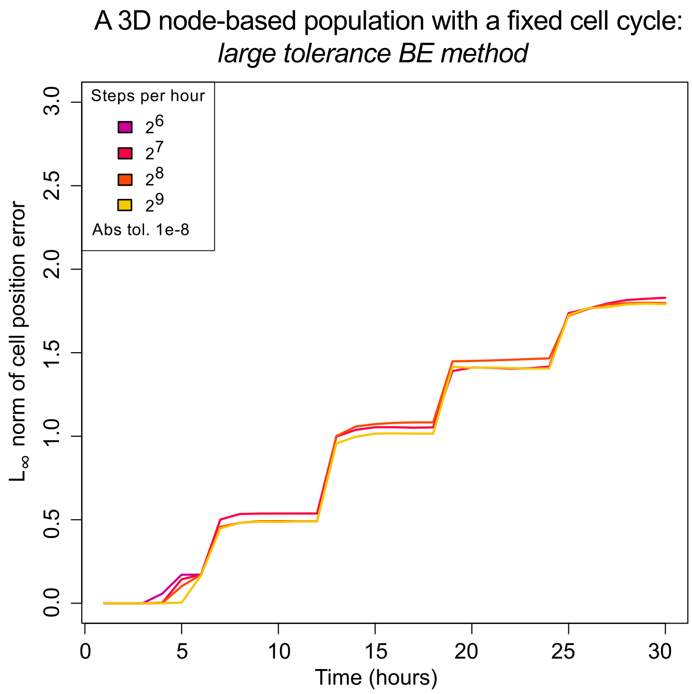


Figure 5.19: A set of BE simulations, run using a tolerance that is too large. This graph is similar to Figure 5.18, in that it shows maximum cell position error over time, measured relative to an RK4 baseline with $dt = 2^{-14}$. Here, however, the Newton solver tolerance has been set too high, causing partially converged results to be returned with a large amount of error. This is indicated by the fact that position error no longer tends to zero when dt is reduced.

Turning now to large scale simulation properties, Figure 5.20 summarises our population extent measurements. Again, each panel shows the error in extent over time, measured relative to an RK4 baseline run with $dt = 2^{-14}$. Our results indicate that all choices of numerical method and time step produce a similar estimate of population extent. Therefore, if we care only about large scale properties such as growth rates rather than individual cell positions, a much larger time step can be taken without significantly altering the results. In particular, the default Chaste value of 120 steps per hour is perfectly adequate for calculating whole population properties.

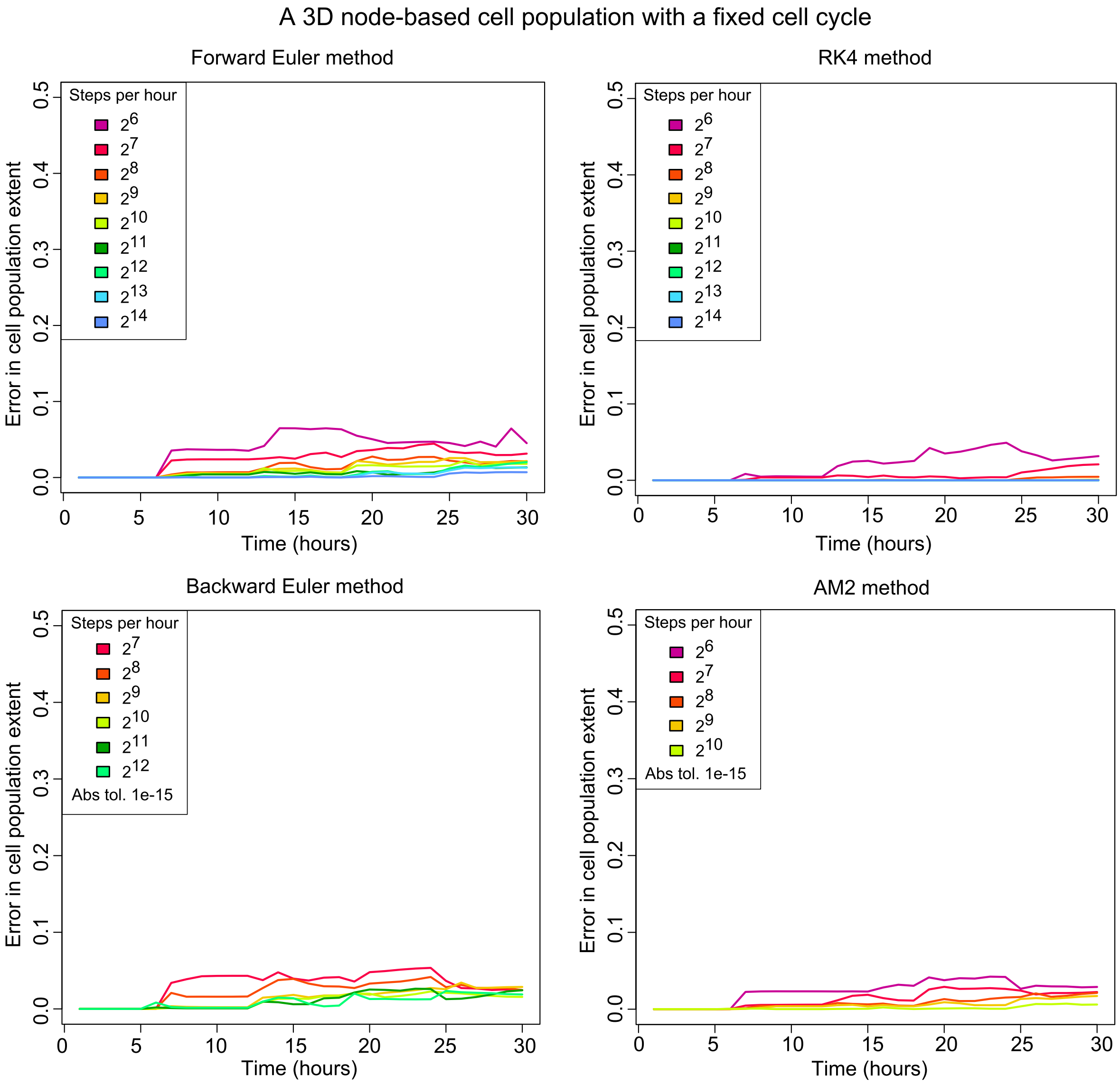


Figure 5.20: The numerical error in population extent, using different methods and time steps. In each panel above, an RK4 run with $dt = 2^{-14}$ is taken as a baseline. For all other simulations, we plot the difference in population extent relative to the baseline run. The maximum value on the y axis corresponds to one cell radius. All step sizes and numerical schemes tested were found to produce reasonably accurate extent results.

5.2.4 Changing the cell cycle model

Next, we investigated whether changing the cell cycle model in our test simulation would affect numerical method performance. In particular, we changed the cell cycle model from one with synchronous divisions every 6 hours to a stochastic model, in which the length of the G1 phase includes a random contribution $\sim U[0,2]$ hours. We also desynchronised the four cells present at the beginning of the simulation, assigning each one a random starting age. Stochastic cell division is more typical of Chaste simulations. However, it does introduce the possibility that runs using different time steps may disagree for reasons unrelated to numerics. In particular, because Chaste uses the same time step for the mechanics simulation, cell cycle model and data output, the same division event could occur at slightly different times in runs using different dt. For instance, if the intended cell division time is 1.05 hours, a simulation with $dt = 0.01$ will capture this exactly, while a run with $dt = 0.1$ will include a slightly early or late division. This effect will artificially inflate the amount of position error we observe.

Figure 5.21 shows the cell centre position results we obtained using a stochastic cell cycle. There are now no sudden jumps in position error, because divisions occur throughout the simulation

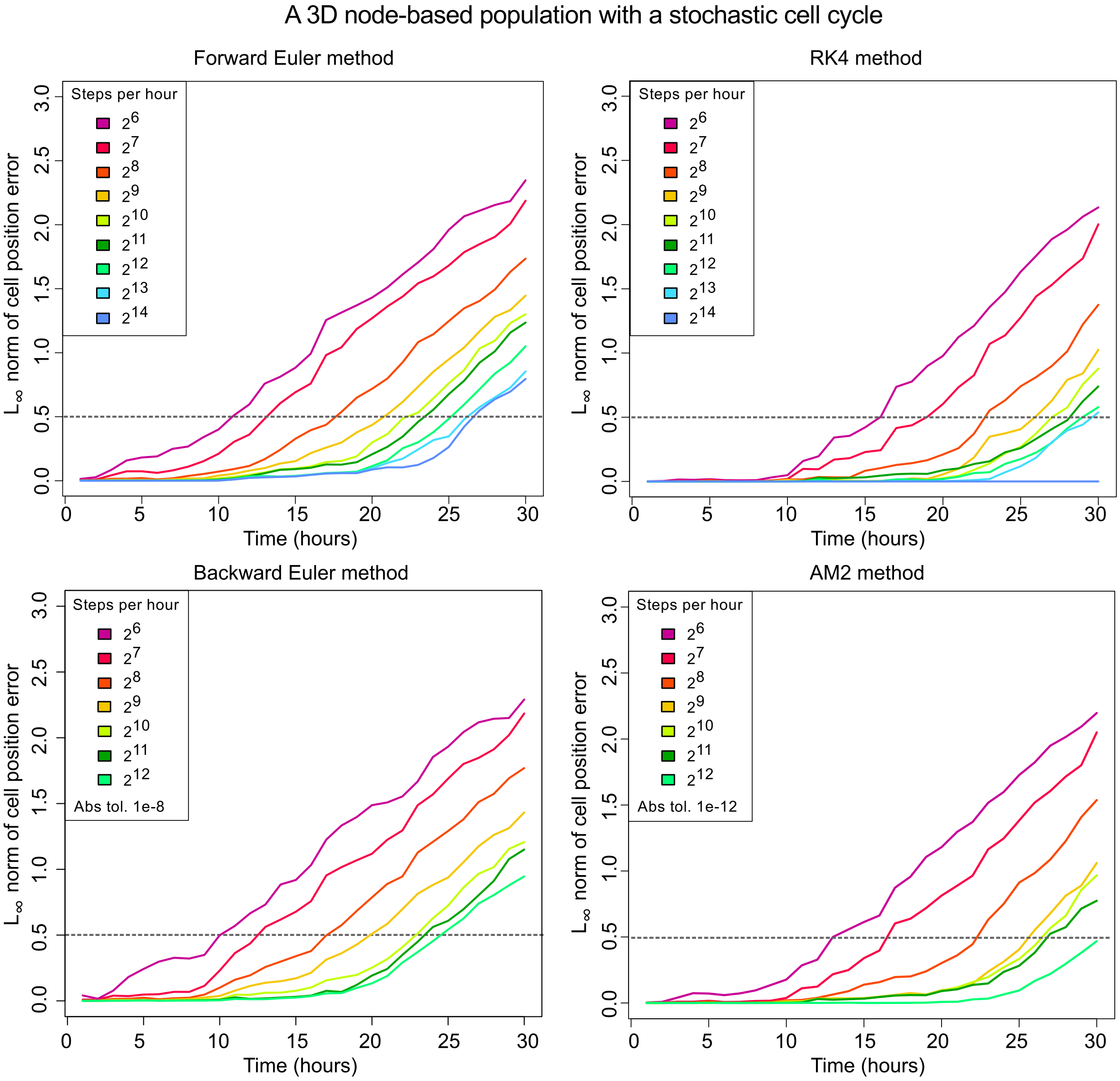


Figure 5.21: The numerical error in cell positions, using a stochastic cell cycle. A similar plot to Figure 5.18, but for a test simulation with stochastic cell divisions. Again, error is measured relative to an RK4 baseline with $dt = 2^{-14}$, and traces are the mean of 20 replicates with different random seeds. The two explicit methods now give a larger position error, possibly due to the effect of dt on division timing. The implicit methods, meanwhile, are able to converge on accurate cell positions using a much larger tolerance than before. However, they still do not allow a substantial enough increase in dt to see a performance benefit.

rather than in bursts. Again, position error decreases relative to the baseline run as dt tends to zero. However, convergence is now slower and generally larger errors appear. The forward Euler method no longer gives acceptably accurate results even with a step size of 2^{-14}, while the RK4 method now requires a time step of $\sim 2^{-13}$. The additional error may be ascribed to the division timing problem discussed above.

Interestingly, the two implicit methods actually worked better in some respects when using a stochastic cell cycle. We were able to apply a tolerance several orders of magnitude larger than in the synchronised divisions case and still observe convergence, which suggests the Newton solver is converging faster. Potentially, this may be because cell positions change less dramatically in a single time step when divisions are desynchronised. As a result, our initial guess positions are closer to the solution of the root finding problem (Equation (5.17)). The AM2 method now requires a time step of $dt = 2^{-12}$ to produce accurate cell positions. Meanwhile, the BE method had to be abandoned, since implicit runs with $dt<2^{-12}$ are infeasibly slow.

In terms of speed, the ranking of the four methods was similar to the synchronised divisions case. The time in minutes taken to produce accurate cell positions was 1190 for AM2, 445+ for BE, 6.5+ for FE, and 9.53 for RK4. We conclude that implicit methods are still unlikely to deliver a performance benefit when using a stochastic cell cycle model. The RK4 method again performed best, and gave more accurate cell positions for the same dt than forward Euler. The population extent results (not shown) were also similar to those for a fixed cell cycle model, with all choices of dt and all numerical schemes giving an accurate extent estimate.

5.2.5 Changing the cell population model

Having tested alternative numerical methods on a cell centre problem, we went on to consider other population types. Firstly, we wrote a test simulation using a mesh-based model, in which cell boundaries are defined by a Voronoi tessellation based on the underlying set of nodes. Figure 5.22 shows the progress of a typical mesh-based test. Initially, a honeycomb grid of cells is present, measuring two cells along each side. Two ghost node layers are added surrounding the actual population. These ghosts do not correspond to real cells, but instead prevent cells from extending to infinity in the tessellation (see Section 3.1.3). Our mesh-based test uses the same GLS force law and fixed cell cycle model described in Section 5.2.2.

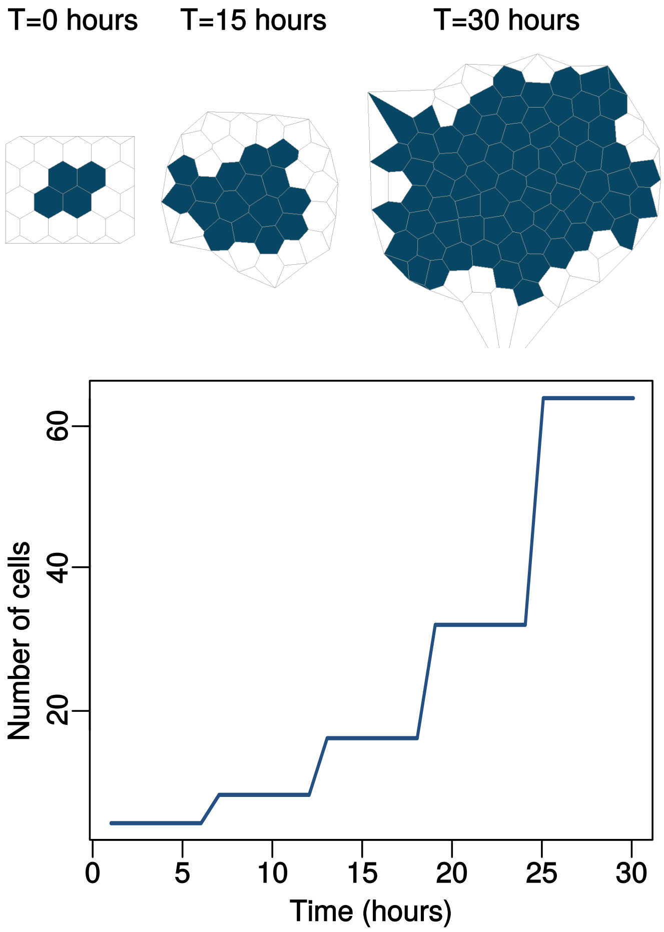


Figure 5.22: The mesh-based test simulation. Above are three snapshots from a typical mesh-based run, along with a graph showing the number of real cells present as a function of time.

Figure 5.23 presents the position error results we obtained for a mesh-based population. These results are based on real cell positions only, excluding ghost nodes. Here, we focus only on the performance of forward Euler relative to RK4, since the two implicit methods are costly to run and have so far failed to significantly raise the acceptable time step.

Figure 5.23 shows that both numerical methods converged on identical cell positions for sufficiently small dt. However, position error was generally larger than in the overlapping spheres case, particularly for the RK4 method. The time step required for acceptable accuracy was 2^{-12} for forward Euler and 2^{-10} for RK4; corresponding to a speed up from 41 to 23 seconds. This is a much smaller performance gain than in the cell centre test. Comparing simulation snapshots reveals that ghost node positions are the first thing to disagree between runs, and it seems possible that some aspect of their special treatment is affecting the result. For example, ghost nodes can move an arbitrary distance in one time step (no AMT is applied) and they experience a ghost node force rather than the GLS force. Although we calculated position error based on real cell locations only, ghost nodes can still influence the result via their effect on the Voronoi tessellation.

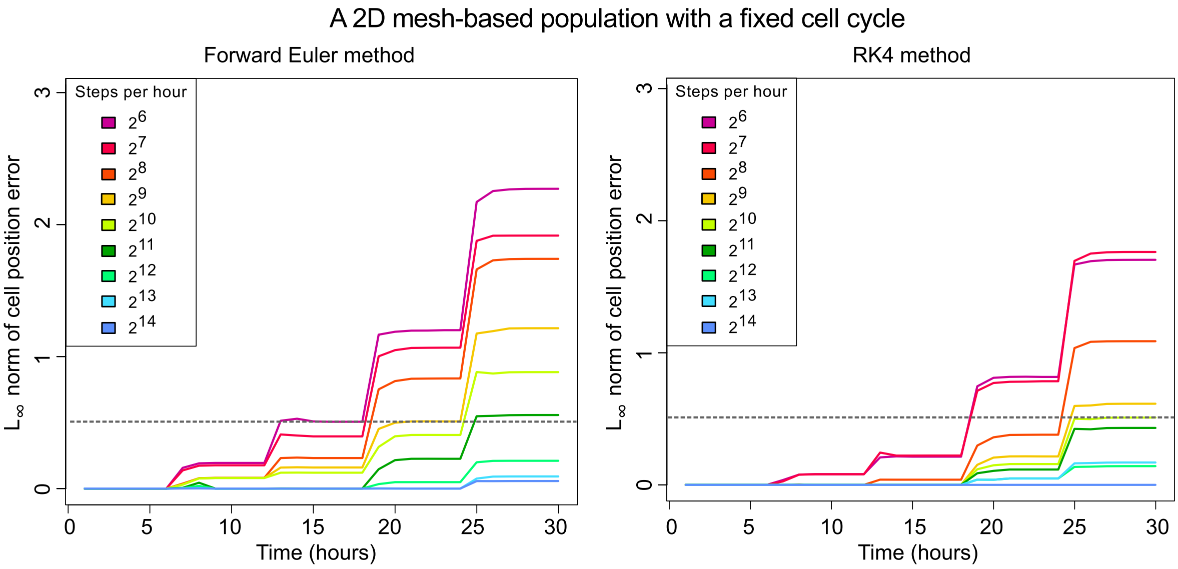


Figure 5.23: The numerical error in cell positions for a mesh-based simulation. In this figure, the position error calculation includes real cells only, and error is measured relative to an RK4 baseline with $dt = 2^{-14}$. Position error again tends to zero as dt decreases for both methods, but now RK4 outperforms forward Euler much more narrowly.

We also ran some simulations using a vertex-based population, in which cells are represented by polygons and each node corresponds to a boundary vertex (see Section 3.1.3). Figure 5.24 shows example output from a vertex-based test. Cells were again seeded in a 2×2 honeycomb arrangement, and, since the GLS force law is unsuitable for a vertex-based population, a Nagai-Honda force was applied instead [181]. The Nagai-Honda force law takes into account cell deformation energy, adhesion, and membrane surface tension, and vertices move so as to minimize the free energy of the cells that contain them. All Nagai-Honda force law parameters were left at their default values, as specified in the Chaste library.

Extra difficulties arise when varying dt in a vertex-based simulation. In Chaste, vertex-based populations have an additional constraint on node movement: a node cannot travel further than half the cell rearrangement threshold in one time step. This constraint ensures that cell intersections and inversions can be properly resolved. If the cell rearrangement threshold is violated, node movement becomes truncated, which changes the outcome of the time step and makes it difficult to observe trends. The default cell rearrangement threshold is 0.01 units of distance, which is highly restrictive. In order to observe the trend on varying dt, we ran our tests with a cell rearrangement threshold of 0.25. Raising the rearrangement threshold can cause other types of error that terminate the simulation (for example, cell intersection errors). As a result, we were only able to run vertex-based tests with a stochastic cell cycle (simultaneous divisions produced a large number of warnings), and then only with a limited range of dt.

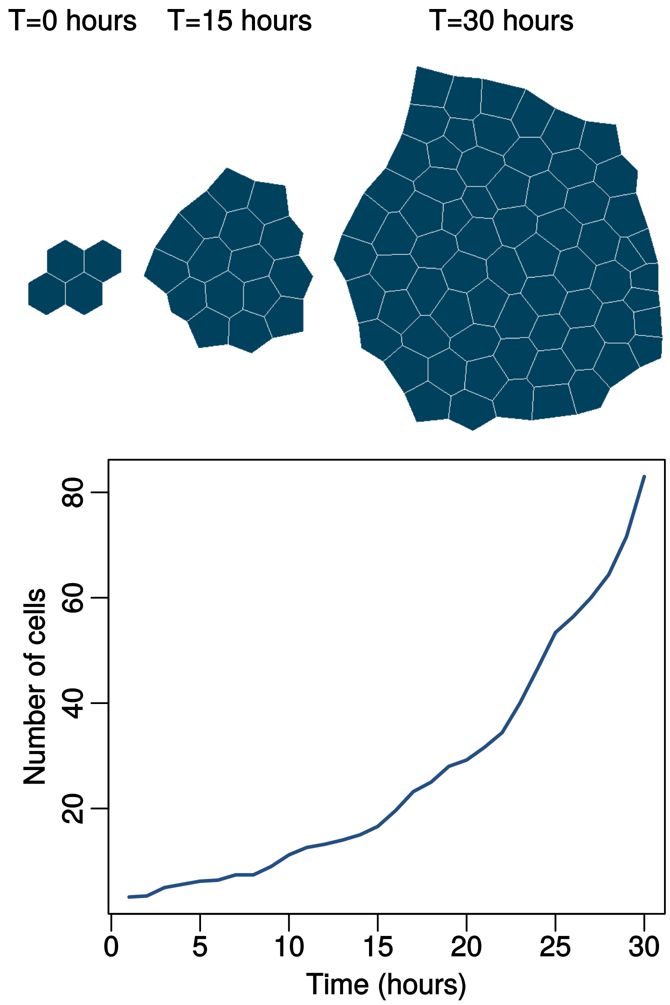


Figure 5.24: The vertex-based test simulation. The snapshots above show a typical vertex-based run, captured at 0, 15 and 30 hours. The graph gives the number of cells present over time, using a stochastic cell cycle model.

Figure 5.25 shows the position results we obtained for a vertex-based test. Overall, this simulation type produced the largest position errors, and although the RK4 method was more accurate for the same dt, it did not significantly lower the time step required to generate accurate positions. However, it is important to remember that these results do *not* reflect typical behaviour for a vertex-based simulation in Chaste. Normally the cell rearrangement threshold and other safeguards would warn the user to choose a smaller time step, preventing significant position errors from occurring. We disabled these errors and warnings in order to observe the trend on changing dt.

So far, all our tests have used a fixed time step throughout the simulation. However, since position error mainly accumulates during divisions and not while the system is static, it is reasonable to consider adaptive time-stepping. This will be the topic of the next section.

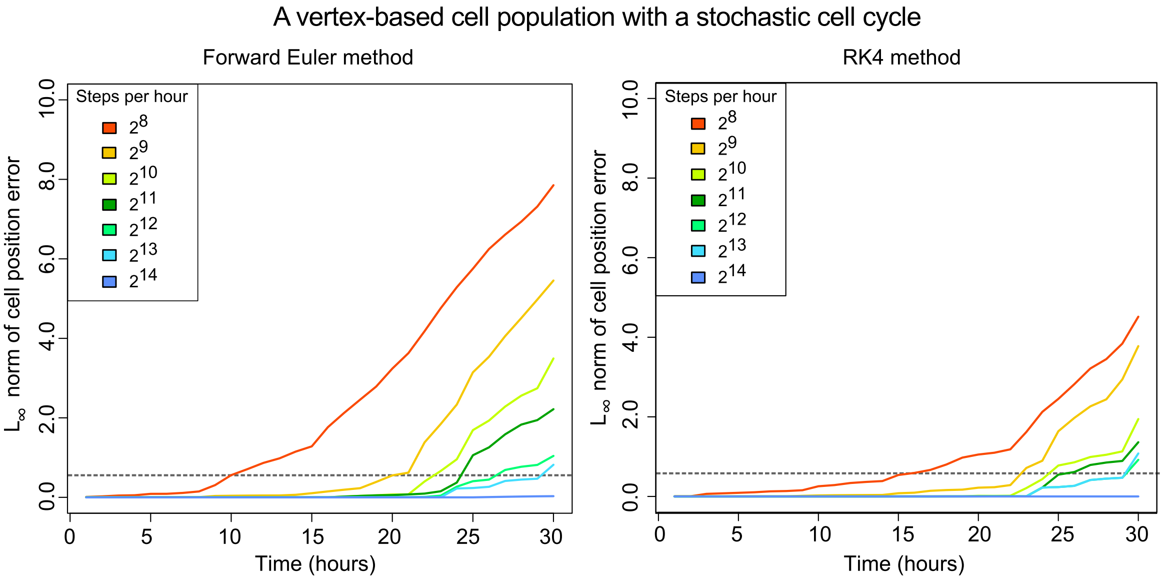


Figure 5.25: The numerical error in cell positions for a vertex-based simulation. Error is again measured relative to an RK4 baseline with $dt = 2^{-14}$. Although the RK4 method produces smaller errors for the same dt compared to forward Euler, it does not allow significantly larger time steps to be taken while maintaining acceptable accuracy.

5.2.6 Simple adaptive time-stepping

Using a fixed time step throughout a cell-based simulation is likely to be inefficient. By taking small time steps only when cells are dividing, it should be possible to improve run times while maintaining accuracy. Adaptive time-stepping could also allow a simulation to continue after the AMT has been exceeded. Rather than halting with an error message (and potentially losing useful data), the program could reduce dt automatically.

Initially, we implemented a very simple adaptive scheme, aimed primarily at handling AMT errors. Under this scheme, the user chooses a starting time step dt. Whenever a node exceeds the AMT, its displacement is recorded, and a new step size dt' is calculated that is small enough to avoid the error:

$$dt' = \frac{0.95\,\text{AMT}}{\text{displacement}}\,dt. \tag{5.21}$$

Nodes are then returned to their positions before the failed step, and the simulation continues with a step size of dt'. For every subsequent successful step, dt' is increased by 1%. Once the simulation time has advanced by dt, the original user provided step size is restored. If necessary, a remainder step is taken to ensure that the solution is evaluated at $N\,dt$ for integer N.

The scheme described above is not fully adaptive, because the time step can never be made larger than the user's original choice. However, steps that are too big can be automatically subdivided.

The benefit of this approach is that it cannot interfere with other parts of the code that may assume a fixed time step (for example, data output methods that save data every N steps). The substeps of size dt' are visible only to a single class; the rest of the program sees a fixed step of dt in use.

For optimal performance, the user should select a large starting dt and a reasonably stringent AMT. The only caveat is that Chaste uses the same time step to advance both the mechanics simulation and any logical model running inside cells. So, dt must be small enough for an ODE, PDE or statechart-based cell cycle model to update accurately. We judged that 2^6 time steps per hour is about the upper limit for dt, given this constraint.

Our new adaptive time-stepping scheme was tested by running the same cell centre simulation described in Section 5.2.2; initially with a fixed cell cycle model. A starting time step of $dt = 2^{-6}$ was used for all adaptive runs. By progressively lowering the AMT, it is then possible to force more and more of these large time steps to be subdivided. Figure 5.26 presents the position error results we obtained. As in previous sections, error is measured relative to a *fixed* time step RK4 baseline with $dt = 2^{-14}$.

Figure 5.26 demonstrates that position error tends to zero as the AMT is reduced, for two different underlying numerical schemes. We conclude that adaptive time-stepping can produce accurate cell positions for this problem, provided an appropriate AMT is chosen. Table 5.4 shows run times for

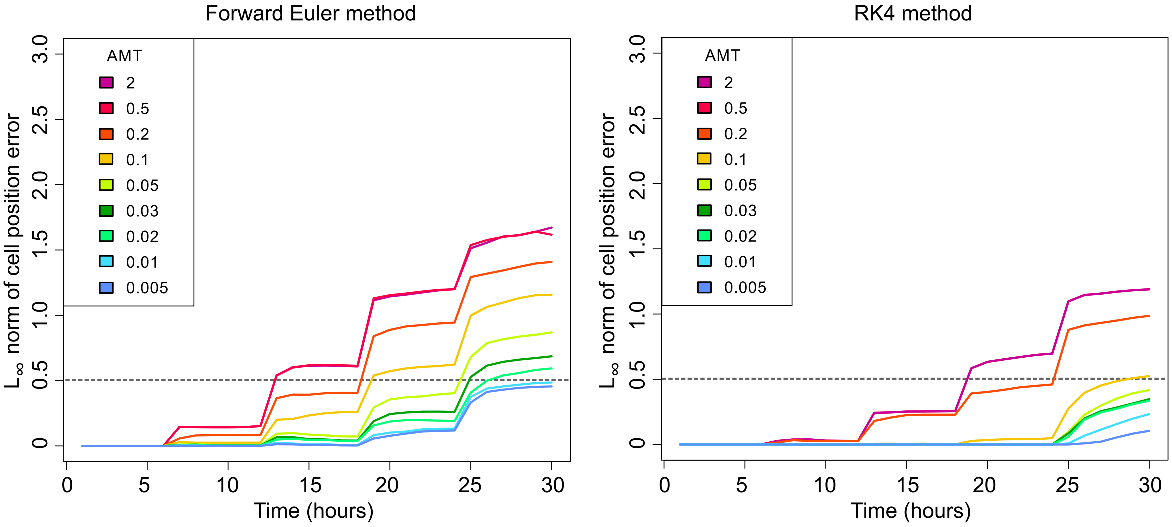


Figure 5.26: Numerical error in cell positions, using adaptive time-stepping and a fixed cell cycle. In both panels, a fixed time step RK4 run with $dt = 2^{-14}$ is taken as a baseline. Position error is then plotted relative to the baseline for a number of adaptive runs using different choices of AMT. For a sufficiently small AMT, the adaptive and fixed time step results agree. The RK4 method allowed a larger AMT to be chosen than forward Euler while still generating accurate results.

Method	Simulation parameters	Run time (s)	Acceptable accuracy?
Forward Euler adaptive	AMT = 0.1	2.35	N
Forward Euler adaptive	AMT = 0.03	2.55	N
Forward Euler adaptive	AMT = 0.01	2.95	Y
Forward Euler fixed step	$dt = 2^{-12}$	97	Y
RK4 adaptive	AMT = 0.5	4.55	N
RK4 adaptive	AMT = 0.2	4.75	N
RK4 adaptive	AMT = 0.05	4.95	Y
RK4 fixed step	$dt = 2^{-7}$	9	Y

Table 5.4: Run times using an adaptive time step. The table above lists run times for selected adaptive simulations, as well as for the two most efficient fixed time step runs, for comparison. Both numerical schemes (forward Euler and RK4) were found to perform better using adaptive time-stepping.

this test using adaptive time-stepping, as compared to the fixed time step approach from earlier in the chapter. Adaptivity was found to produce a large speed up in this case, particularly when using the forward Euler method. For forward Euler, the run time was cut from 97 to 2.95 seconds, while for RK4 the run time went from 9 to 4.95 seconds. Since cell divisions are synchronised in this test, long division-free periods occur during which large time steps can be taken. This situation is optimal for adaptive stepping, so we can expect the performance improvement observed here (around 30-fold for forward Euler) to represent an upper bound.

We then repeated our adaptive cell centre test, but using a stochastic cell cycle model. Figure 5.27 shows the position error results. This time, as the AMT was reduced, position error decreased up to a point but did not tend to zero. The problem appears to be the same one discussed in Section 5.2.4: Chaste uses the same time step to update cell positions, the cell cycle and data output. Subdividing large time steps in the mechanics simulation does improve the accuracy of cell movement, but it does not address inaccuracies in the timing of divisions. Therefore, it cannot change the fact that a run with $dt = 2^{-6}$ may place the same division event at a slightly different time compared to a run with $dt = 2^{-14}$. We conclude that adaptive time-stepping *can* substantially speed up cell-based simulations, particularly when divisions are highly synchronised. However, for the Chaste library to benefit from adaptivity, there would need to be a substantial rewrite, addressing the problem of dt affecting division timing. Interpolation could perhaps be used to better account for divisions that occur between time steps. Alternatively, dt could be made fully variable, in such a way that each division is guaranteed to fall at the end of a step.

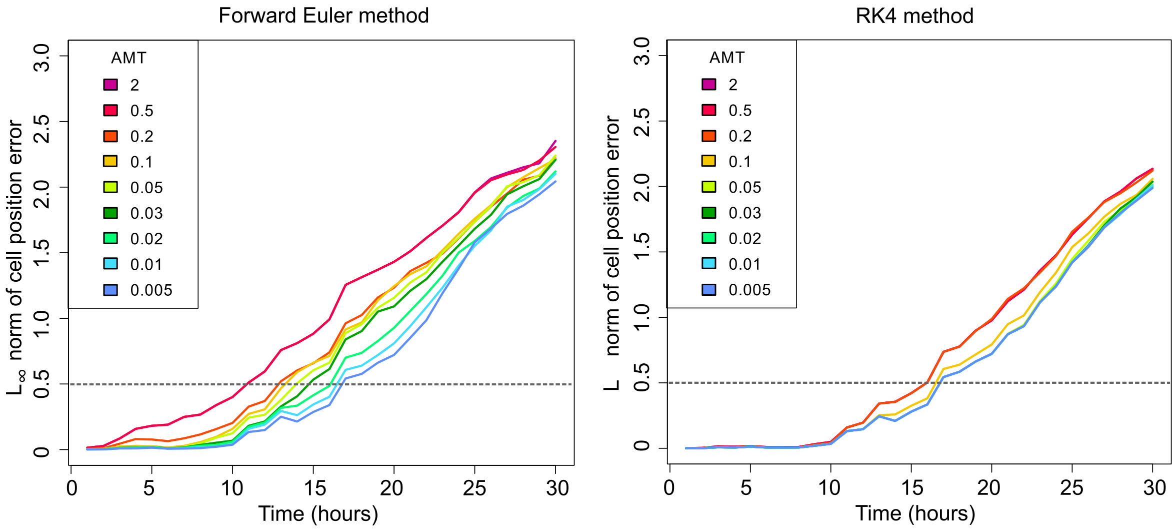


Figure 5.27: Numerical error in cell positions, using adaptive time-stepping and a stochastic cell cycle. A similar plot to Figure 5.26, but for a test simulation in which cells divide stochastically rather than at fixed intervals. Position error still decreases as the AMT is lowered, but it does not approach zero. Although subdividing time steps in the mechanics simulation can reduce the error in cell movement, it cannot address the error in division timing that arises from using a large starting dt of 2^{-6}.

5.2.7 Adaptive time-stepping with error estimation

The simple adaptive scheme described above already has the potential to improve performance. However, it still relies on a user-selected AMT to assess numerical error. Alternative methods exist that can estimate the error in a time step more rigorously. In particular, embedded Runge-Kutta schemes carry out each step using multiple different RK methods with different orders of convergence. The highest order method is used to produce the solution, while lower order methods are used to generate an error estimate and determine when dt should be reduced. By combining RK methods that use the same set of function evaluations but different weightings, an error estimate can be obtained at relatively little computational cost.

We hypothesised that an embedded RK scheme might allow longer simulations to be run while still controlling error. It would also eliminate the need for the user to guess a suitable AMT. The exact method we chose to test was the Dormand-Prince 853 scheme (DOP853), which has been shown to perform well in non-biological applications [98]. The full set of DOP853 equations is too long to reproduce here, since 12 intermediate calculations are involved using 144 non-integer constants. However, the details can be found in [98], and several example implementations are available online; including [161] on which we based the Chaste code. Essentially, an order 8 RK method is used to produce the solution, while order 3 and order 5 estimates of the error are calculated for each node position. A “MaxNormError” can then be computed as

$$\text{MaxNormError} = \frac{M_5^2}{\sqrt{M_3^2 + 0.01 M_5^2}}, \tag{5.22}$$

where M_5 and M_3 are maximum scaled error estimates, given by

$$M_3 = \max_i \frac{|\mathbf{E_3}_i|}{s_i}, \tag{5.23a}$$

$$M_5 = \max_i \frac{|\mathbf{E_5}_i|}{s_i}, \tag{5.23b}$$

$$s_i = \text{Tol}_\text{a} + \text{Tol}_\text{r} \max(|\mathbf{r}_i^{old}|, |\mathbf{r}_i^{new}|). \tag{5.23c}$$

Here, i ranges over all nodes in the population, with $\mathbf{r}_i^{old}$ denoting the node position before the step, and $\mathbf{r}_i^{new}$ denoting the position afterwards (according to an 8th order RK scheme). $\mathbf{E_3}$ and $\mathbf{E_5}$ are 3rd and 5th order estimates of the position error, while Tol_a and Tol_r are error tolerances set by the user. Whenever the MaxNormError exceeds 1, dt is halved. Once the simulation time has advanced by dt, the original step size is restored, as in our previous adaptive scheme.

Figure 5.28 shows cell position results for the test described in Section 5.2.2, using a fixed cell cycle of length 24 hours, and with the simulation extended to cover more division events. A DOP853 run with an initial dt of 2^{-6} is taken as a baseline, and position error is plotted for a number of other simulations using different time-stepping schemes. Our DOP853 simulation agreed with a fixed time step RK4 run using an extremely small dt. Looking at our simpler adaptive scheme, as the AMT was reduced the solution agreed for longer and longer with DOP853. This indicates that the DOP853 method is likely giving accurate cell positions over the full time course, while the various simple adaptive runs eventually diverge from the correct answer. We conclude that an embedded RK method can generate accurate cell positions over a long period, without requiring the user to guess a suitable AMT. With regard to run time, DOP853 took approximately 2 hours to complete the simulation, compared to more than 6 hours for the fixed time step run, and around 1 hour for the various simple adaptive schemes. Embedded RK methods are therefore slower than simpler adaptive approaches and will be worthwhile only when accurate positions are a priority.

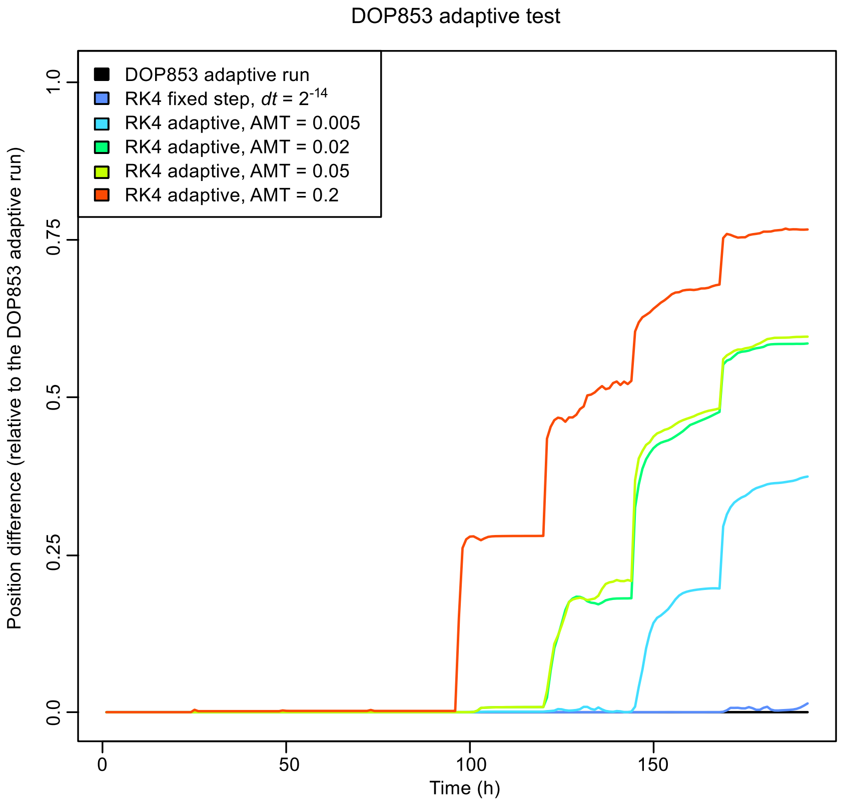


Figure 5.28: The cell position difference between a DOP853 run and simulations using other time-stepping schemes. Here, a single DOP853 run is taken as a baseline, and position difference is plotted for a number of other simulation types. The DOP853 result agrees with a fixed time step simulation using a very small choice of dt. When the AMT is lowered in our simple adaptive scheme, it too agrees with DOP853 for longer and longer. This suggests that the DOP853 method produces accurate cell positions over the full 200 hours.

5.3 An alternative intracellular modelling approach

In Chapter 4, we used a statechart to model individual germ cell behaviour. This approach works well, provided daughter cells simply inherit their parent's state. However, statecharts are arguably less suitable for representing more complex behaviour on division, such as the asymmetric inheritance that occurs when a stem cell produces a transit amplifying daughter. Since statecharts were not designed with biology in mind, the visual formalism lacks a place to display division rules. Moreover, in our particular implementation in Chaste, asymmetric inheritance is made harder to model by the order in which Chaste constructs a new cell. On division, a daughter cell cycle model is created *before* the cell that will contain it. This is not ideal, because statechart cell cycle models require access to their 'owning cell' from an early stage. We implemented a workaround for Chapter 4, but this only applies to state copying from parent to daughter.

There are other aspects of the statecharts approach that could also be improved upon. For example, while statecharts have a clear, readable organisation when represented graphically, their representation as C++ code is harder to interpret. It would be easier to make targeted changes to a model if the code controlling different cellular functions were kept separate, and if division/inheritance code had a well defined place. There are also occasions where having a default starting state for a cell is unnecessary, and where it would be better to be able to specify an initial state explicitly. These points motivated us to try to develop an alternative approach for representing cell decision-making.

5.3.1 Cell logic modules

To facilitate the investigations in Chapter 6, we designed and implemented a new type of intracellular model, an approach that we call *cell logic modules*. A logic module is analogous to one orthogonal region of a statechart. It controls a single intracellular function, and belongs to a category that describes that function. For instance, the module StochasticG1Duration would belong to the category CellCycle, while the module LinearCellGrowth would belong to the category Growth. Every module has a current state, which is represented as an enum (a named value) for readability. For example, modules in the Growth category support the states Growth::Growing and Growth::Static. A module's initial state is specified by the user when it is created, eliminating the concept of default transitions. Optionally, a module may also contain variables in addition to the state, for tracking other relevant quantities such as time delays.

All logic modules have two key methods: Update and Divide. Update is called at each time step and can make changes to the state and variable values. Meanwhile, the Divide method produces a daughter logic module, while also handling any division related changes to the state of the parent. Now, when a cell divides, an "empty" daughter cell is first created. Then Divide is called on each of the parent's logic modules, and the resulting daughter modules are assigned to the new cell. Since both the parent and the daughter cell are accessible during division, arbitrarily complex patterns of inheritance can easily be represented. Moreover, this approach increases modularity, making it easier to find the relevant piece of code controlling each part of the model, since different intracellular functions are kept separate and update logic is separated from division logic.

Listings A.2 and A.3 in Appendix A illustrate the difference between statecharts and cell logic modules in practice. Both pieces of code implement cell growth, but Listing A.2 does so as a statechart while Listing A.3 uses a logic module. In the former case three pages of code are required; in the latter only one. We therefore contend that the cell logic modules approach is more readable, more concise and more modular than Boost Statecharts.

5.3.2 Visualising cell logic modules

One advantage of statecharts is that they allow models to be represented graphically in an organised fashion. We have attempted to develop a similar visual formalism for representing cell logic modules. Figure 5.29 presents a simple example.

The category of the cell logic module is written in the top left hand corner. So, for example, the module shown in Figure 5.29 controls differentiation behaviour. The upper part of the diagram indicates how the state is updated at each time step. States are written in bold, transitions are indicated by arrows, and transition conditions are written alongside their arrows. Transition conditions may depend on signals external to the cell (such as IncomingSignal), or on information from other logic modules (such as CellCycle::G2, which specifies that a CellCycle module must exist and that it must be in the G2 state). If no other indication is given, states are assumed to be inherited unchanged on division. If more complex behaviour occurs, this is indicated by drawing a "division line" under a state, then listing possible outcomes. Here, a square indicates the state of the parent cell after division, while a circle indicates the state of the daughter. So, in Figure 5.29, transit amplifying (TA) cells divide into two TA cells if divCount<3, but two differentiating (D) cells otherwise. Finally, the bottom part of the diagram shows how any variables are updated and inherited, using similar conventions as for states. Here, there is a single variable divCount that does not change at each time step, but when a cell in state TA divides it increases by one.

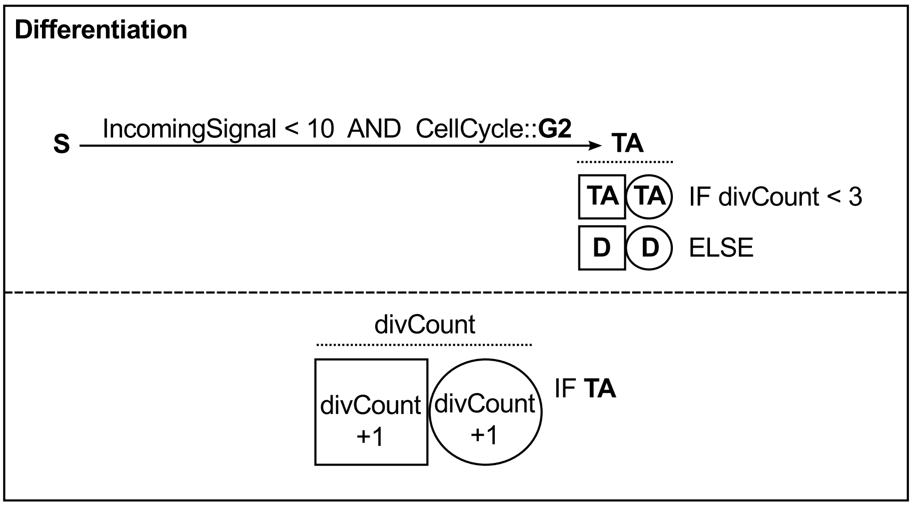


Figure 5.29: An example cell logic module controlling differentiation. For a detailed description of this diagram, see text.

5.4 Summary

In this chapter, we considered two main technical ways in which our cell-based simulations might be improved. First, we investigated replacing the GLS force with a more detailed, physically realistic force law proposed by Buske *et al.* [33]. We began by performing manageable, small scale tests on pairs of cells. These confirmed that our Buske force law was correctly implemented, and indicated that a reasonably close approximation to the Buske force can be obtained by applying a GLS force law with $\mu/\eta = 1.23$. The adhesion component of the Buske force law appears to be the most significant addition, producing a noticeable change in cell separation rates and in steady state separation.

A number of large scale tests were then carried out using a model of the distal germ line. We found that the Buske force law with its standard parameter set results in unrealistically large radial variations when tight cell packing is present. With fixed cell radii, and with a cell cycle length of 8 hours, every variant of the Buske force law caused cells to accumulate in the distal gonad, eventually reaching such high numbers that the simulation could not continue. We conclude that replacing the GLS force law with a more detailed Buske force does not eliminate the overcrowding problem described in Chapter 4. The only means we found to prevent cell accumulation was to take the GLS force law and increase the spring strength dramatically. It is arguable whether such an increase is physically realistic; answering that question would require more experimental data on the mechanical properties of germ cells.

In the second half of this chapter, we looked at whether changing the numerical method used by Chaste could boost performance and accuracy. We found that Chaste's default time step and numerical method are already sufficient to produce accurate large scale simulation properties. However, producing accurate cell positions requires a much smaller time step. In a cell centre test, alternative numerical methods allow the step size to be increased without sacrificing accuracy. In the case of RK4 this led to a 10-fold improvement in run times. Unfortunately, implicit methods did not provide sufficient benefit to make up for their greater computational cost.

Adaptive time-stepping was also tested and was found to substantially speed up simulations with synchronised, deterministic divisions, while also eliminating AMT errors. In particular, embedded RK methods such as DOP853 could remove the need for an AMT altogether, replacing it with a more rigorous error estimator. However, to benefit from adaptive time-stepping the Chaste library would need to be altered to ensure precise division timings. Based on this investigation, we conclude that moving to an RK4 numerical scheme could provide immediate benefit, while the DOP853 method could be adopted in the longer term for applications where precise cell positions are critical. This would allow larger problems to be tackled using Chaste, while still generating the accurate solutions necessary for developmental biology research.

Finally, we introduced an alternative approach to modelling intracellular decision-making, using cell logic modules. This approach promotes a clear separation between different intracellular functions, and between update and division logic, leading to more manageable code. It also simplifies the task of implementing complex inheritance rules within Chaste. As such, we will use cell logic modules extensively in the next chapter, which involves examples of asymmetric inheritance.

Chapter 6

Evaluating mechanisms of proliferative zone maintenance

The details of proliferative zone maintenance are a contentious issue in *C. elegans* germline research. Are all germ cells functionally equivalent, or are specialised stem cells involved? The term *germline stem cell* is used extensively throughout the literature [13, 47, 114, 260], and, in 2010, an experiment performed by Cinquin *et al.* provided evidence for the existence of immature distal cells [47] (see Section 2.8). Recently, however, Fox and Schedl repeated the same experiment using a different labelling technique, and found no evidence of a distinct stem cell population [81]. Rather, their results suggest that all germ cells are equivalent and that proliferative capacity is conferred through interactions with the DTC. In such instances of conflicting experimental evidence, modelling can be a useful tool to compare hypotheses.

This chapter goes beyond the simple model of the proliferative zone in Chapter 4, by using *in silico* modelling to evaluate several different potential maintenance mechanisms; some of which include distinct stem cells, and others equivalent germ cells. For each mechanism, we generate data on a wide range of germline properties, then compare the simulated results with experimental measurements. We also simulate the response of each mechanism to withdrawal of DTC signalling (as performed by Cinquin *et al.* and Fox and Schedl [47, 81]). The aim is to clarify the predicted behaviour of the proliferative zone under each mechanism, and to see whether any hypothesis gives a clearly superior fit. Finally, we consider whether any maintenance mechanism can explain the decline in proliferative cell count now known to occur with age [197].

6.1 Hypothesised mechanisms

We begin by introducing the various proliferative zone maintenance mechanisms under consideration. Three different hypotheses are modelled, each of which can also be simulated with or without germ cell contact inhibition.

6.1.1 Mechanism 1: Equivalent germ cells

The first maintenance mechanism we consider is similar to that proposed by Fox and Schedl [81], and also to the model of DTC signalling in Chapter 4. All germ cells are assumed to be functionally equivalent, and those within Xµm of the DTC receive a signal that confers proliferative capacity (Figure 6.1). Once a germ cell moves out of range of the DTC signal, it enters meiosis from the G1 phase. Note that there are no transit amplifying cells in this model: germ cells more than Xµm from the DTC complete their current cell cycle but cannot divide again.

From a biological perspective, the Xµm threshold distance can be interpreted as the region within which germ cells make close contact with the DTC membrane, allowing GLP-1 on germ cells to bind the membrane-bound ligands LAG-2/APX-1. Alternatively, the threshold distance could be interpreted as the region within which the concentration of some diffusible signal exceeds a critical level.

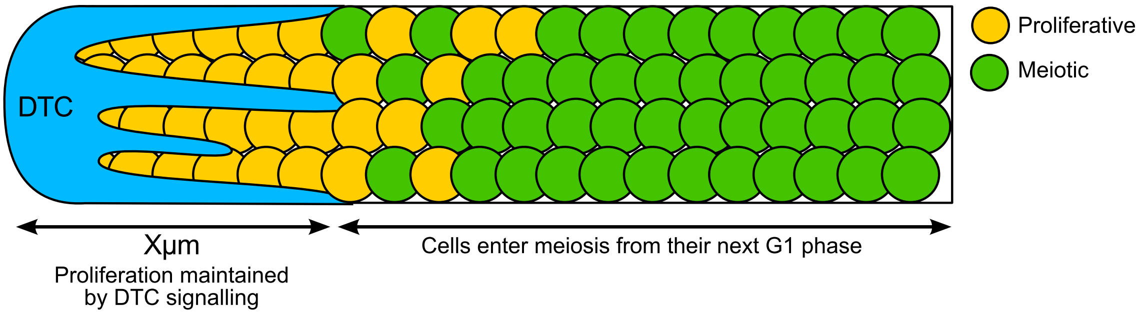


Figure 6.1: The equivalent germ cells mechanism. Proliferative capacity is conferred on cells within Xµm of the DTC. On moving out of range of the DTC signal, germ cells complete their current cell cycle then enter meiosis from G1.

6.1.2 Mechanism 2: Stem and transit amplifying cells

The second maintenance mechanism is similar to that proposed by Cinquin *et al.* in [47]. Germline stem cells reside in a niche within the first Xµm of the DTC (Figure 6.2). Stem cells in this region produce two stem daughters on division. Once out of range of the DTC signal, stem cells produce two transit amplifying daughters the next time they divide. Transit amplifying cells then undergo a further n_{Div} divisions, before entering meiosis from the G1 phase.

While there is arguably little difference between germ cells that proliferate due to a DTC signal, and stem cells that are maintained by a DTC signal, this mechanism does differ from the previous one in two important ways. Firstly, cells close to the DTC are less mature; even if $n_{Div} = 0$ they still require an extra division to become transit amplifying before they can enter meiosis. Secondly, transit amplifying cells can divide multiple times in this model, rather than just completing their current cell cycle.

6.1.3 Mechanism 3: Asymmetric divisions

The final mechanism we consider involves germline stem cells that can divide asymmetrically with some probability. Initially, in the young adult, all germ cells within Xµm of the DTC are stem (Figure 6.3). Whenever a stem cell divides in this region, it can produce either a stem daughter (symmetric division, probability p), or a transit amplifying daughter (asymmetric division, probability $1 - p$). In both cases the parent cell remains stem. Stem cells further than Xµm from the distal tip lie outside the DTC's signalling range and always produce two transit amplifying cells on division. Transit amplifying cells undergo n_{Div} symmetric divisions before entering meiosis from G1.

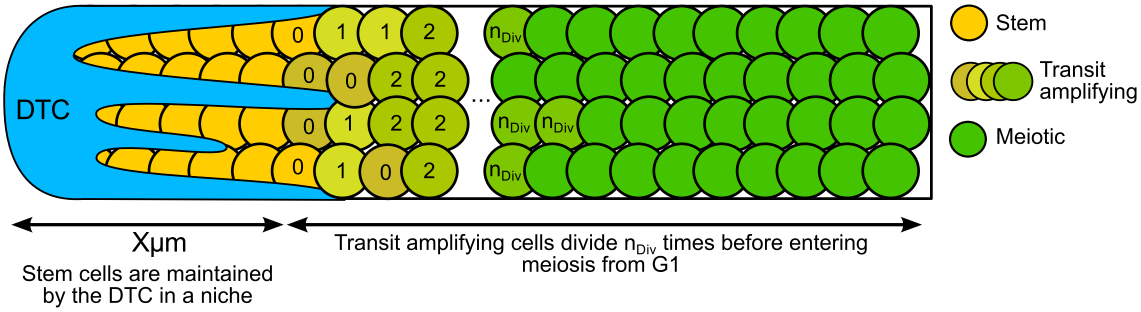


Figure 6.2: The stem and transit amplifying cells mechanism. Cells within the first Xµm of the DTC are stem, due to a supporting signal from the niche. On leaving this region, germ cells become transit amplifying at their next division. Transit amplifying cells can then divide n_{Div} times before entering meiosis from G1.

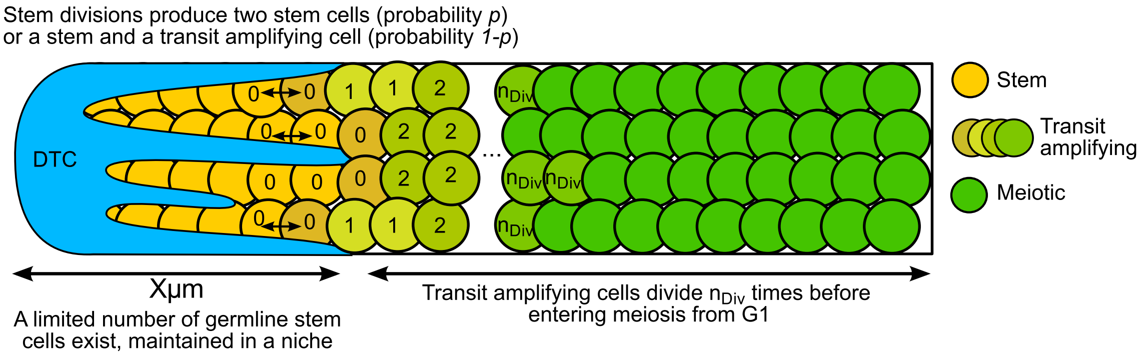


Figure 6.3: The asymmetric divisions mechanism. All germ cells within Xμm of the DTC are initially stem. However, stem cells can divide asymmetrically within this region, producing transit amplifying cells that displace stem cells from the niche. On moving more than Xμm from the DTC, stem cells become transit amplifying at their next division. Transit amplifying cells then undergo n_{Div} divisions before entering meiosis from G1.

The major difference between this model and the previous one is that transit amplifying cells can now be produced within Xμm of the DTC. In the absence of adhesion and oriented divisions, this will inevitably lead to the gradual displacement of stem cells from the niche, which are lost permanently. Recent experiments have shown that proliferative germ cell count declines with age in *C. elegans* [197]; therefore it is worth considering a mechanism that incorporates stem cell loss. Indeed, by varying the value of the parameter p it should be possible to quantify the probability of symmetric division required in the niche to account for the observed proliferative cell count decline.

6.1.4 Presence or absence of contact inhibition

In Chapter 4, we hypothesised that adult germ cells may be subject to contact inhibition, which would act to prevent overproliferation and maintain a stable number of cells in the germline. In this investigation, contact inhibition is an additional factor that can be turned on or off in each of the three maintenance mechanisms. It is modelled in the same way as in Chapter 4, i.e. by halting progress through the G2 cell cycle phase for as long as a cell's "compressed" volume is less than 70% of its rest volume.

6.2 Modelling the distal germ line

Each of the mechanisms described in Section 6.1 was implemented as a cell-based model. Since this investigation focusses entirely on the proliferative zone and distal gonad, the decision was made to simplify our *in silico* model, rather than using the whole germ line system from Chapter 4. In this section, we describe our simplified distal germ line model. In particular, we explain how the intracellular program was modified to allow non-trivial behaviour on division to be expressed clearly.

6.2.1 Cell mechanics and the distal gonad boundary condition

In many respects, the cell mechanics simulation used in this chapter is identical to that described in Chapter 4. We use Chaste, with germ cells represented in an overlapping spheres, off-lattice fashion. A pure repulsion force law with cell size correction is applied between cells (see Sections 3.1.6 and 4.1.1), and the same force law parameter values are used as listed in Section 4.2.2.

The model in this chapter does, however, implement a different boundary condition. Here, we model only the section of the adult germ line lying between the distal tip and the turn. The boundary condition is therefore static, and forces all cells to lie a distance $r - R_i$ from the organ midline, where R_i is the radius of cell i and r is the radius of the adult gonad. The resulting boundary is cylindrical, with an excluded volume representing the rachis in the centre, and with a hemispherical cap closing the distal end (see Figure 6.4). Based on the *C. elegans* gonad measurements reported in Section 4.2.1, the length of the distal gonad is taken to be 248µm, while the adult gonad radius is set to 11.3µm. A cell killer is added to the proximal end of the tube that removes all germ cells on reaching the turn. This is the most notable difference relative to the whole germ line model, since germ cells will not experience any resistance to forward motion due to cells in the

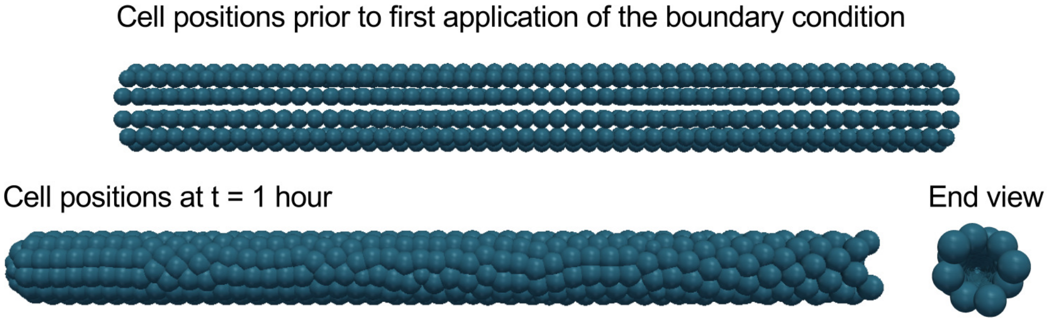


Figure 6.4: Cell positions at the beginning of a distal germ line simulation. Germ cells are restricted to lie at the surface of a cylindrical gonad boundary, with a hemispherical cap closing the distal end. A space representing the rachis is left empty in the centre of the organ. Initially, cells are placed in rows along the length of the gonad, with 10 cells per circular row and 50 rows in total.

proximal gonad. At time 0, germ cells are seeded in circular rows along the length of the organ, with 10 cells per row and 50 rows in total (Figure 6.4).

6.2.2 Intracellular program

The hypotheses tested in this chapter include examples of asymmetric fate outcomes on division. To make it easier to model this behaviour in Chaste, we adopted the cell logic modules approach proposed in Section 5.3. Here, we describe the intracellular model used throughout this chapter, making use of diagrams of the kind introduced in Section 5.3.2. Our model comprises three cell logic modules, governing growth, the cell cycle and differentiation.

Germ cell growth

Figure 6.5 shows the logic module governing cell growth, which is common to all three proliferative zone maintenance mechanisms modelled here. On update, if a cell's current radius is less than the maximum radius the growing state is entered, otherwise the cell becomes static. On division, both the parent and the daughter cell always enter the growing state.

The remainder of the diagram shows how variables are updated and inherited. Germ cell radius increases at a growth rate of 1μm h^{-1} for cells in the growing state. The maximum cell radius

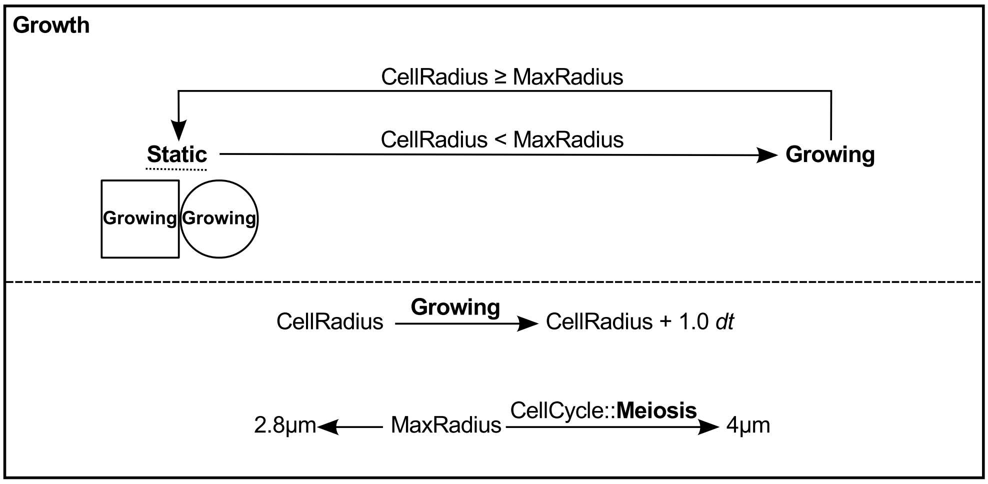


Figure 6.5: The germ cell growth model. Cells can be in one of two growth states: growing or static. On update, cells become growing if their radius is less than the maximum value and become static otherwise. On division, both resulting cells enter the growing state. There are also two variables in this model: CellRadius and MaxRadius. CellRadius increases by *dt* each time step for growing cells, while MaxRadius is set to 2.8μm for proliferative cells and 4μm for meiotic cells.

is 2.8μm initially and switches to 4μm when the cell cycle module enters meiosis. This reflects the difference in size between proliferative and meiotic germ cells, using the same parameters as in Sections 4.2.1 and 4.2.2. Here, we choose for daughter cells to inherit their parent's radius unchanged, consistent with earlier simulations.

The germ cell cycle

Figure 6.6 shows the logic module governing the cell cycle across all three maintenance mechanisms. Proliferative cells cycle through the states G1, S, G2 and M, and divide on transition into M phase. Germ cells in G1 can enter the meiosis pathway if their differentiation module is in state D. They then undergo a fixed length G1 and meiotic S before entering meiosis proper. Most states have a fixed length, with a transition occurring once sufficient time has elapsed. These phase lengths are based on an adult cell cycle of 8 hours, with the same phase proportions as in Section 4.2.2. The lengths of the G1 and G2 phases, meanwhile, are sampled from a normal distribution on entry into those states (G1 length $\sim N(0.16, 0.016)$h, G2 length $\sim N(3.16, 0.316)$h). Cells inherit

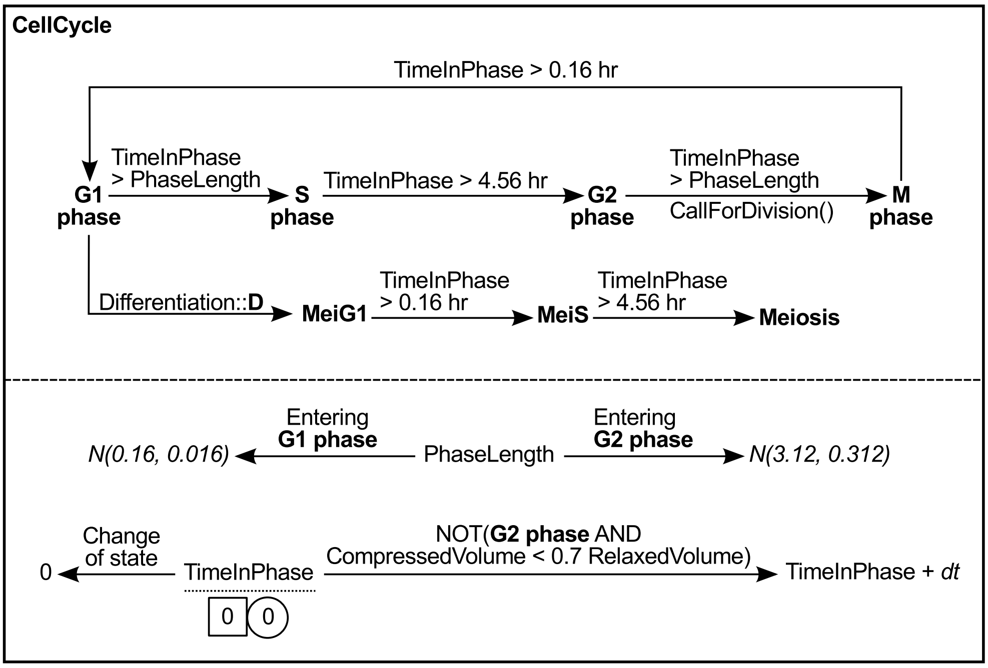


Figure 6.6: The germ cell cycle model. The upper part of this diagram shows how the current state is updated. Mitotic cells cycle through the states G1, S, G2 and M, with each one having a predetermined length. The pathway to meiosis is accessible only from G1, and is enabled when the differentiation logic module enters state D. The lower part of the diagram relates to variable updates and inheritance. PhaseLength, a variable that pertains only to G1 and G2, is set on entry into those states by sampling from a normal distribution. Meanwhile, the variable TimeInPhase increases by *dt* each time step, except for cells that are contact inhibited. It resets to zero on division and at the start of a new cell cycle phase.

their parent's cell cycle state on division, which in practice is always M.

As for variables, TimeInPhase normally increases by *dt* each time step, resetting to 0 on division or whenever the cell cycle phase changes. However, if contact inhibition is enabled, TimeInPhase does *not* increase for cells in G2 that are subject to extensive crowding, leading to a temporary cell cycle arrest. "Extensive crowding" is defined using the same volume based procedure as in the whole germ line model (see Section 4.1.3).

Germ cell differentiation

Finally, Figure 6.7 shows three differentiation logic modules. Since this module controls whether a cell is stem, transit amplifying, or differentiating, it is the only part of the model that differs between our three proliferative zone maintenance mechanisms.

Figure 6.7A represents the equivalent germ cells hypothesis. Cells within a certain distance of the DTC are in a self-renewing state (S) by virtue of their position. When a germ cell receives a signal indicating that it has moved out of the stem region, it immediately becomes differentiating (D), and will enter meiosis during its next G1.

Figure 6.7B represents the stem and transit amplifying cells hypothesis. Again, cells sufficiently close to the DTC are stem, but once they move out of the niche, stem cells become transit amplifying at their next division. Transit amplifying cells become differentiating once they exceed a maximum number of divisions, n_{Div}.

Finally, Figure 6.7C represents the asymmetric divisions hypothesis. It is largely identical to 6.7B, except that when a stem cell close to the DTC divides, it can now produce a stem daughter with probability p, or a transit amplifying daughter with probability $1 - p$. Transit amplifying cells can therefore gradually displace stem cells from the niche.

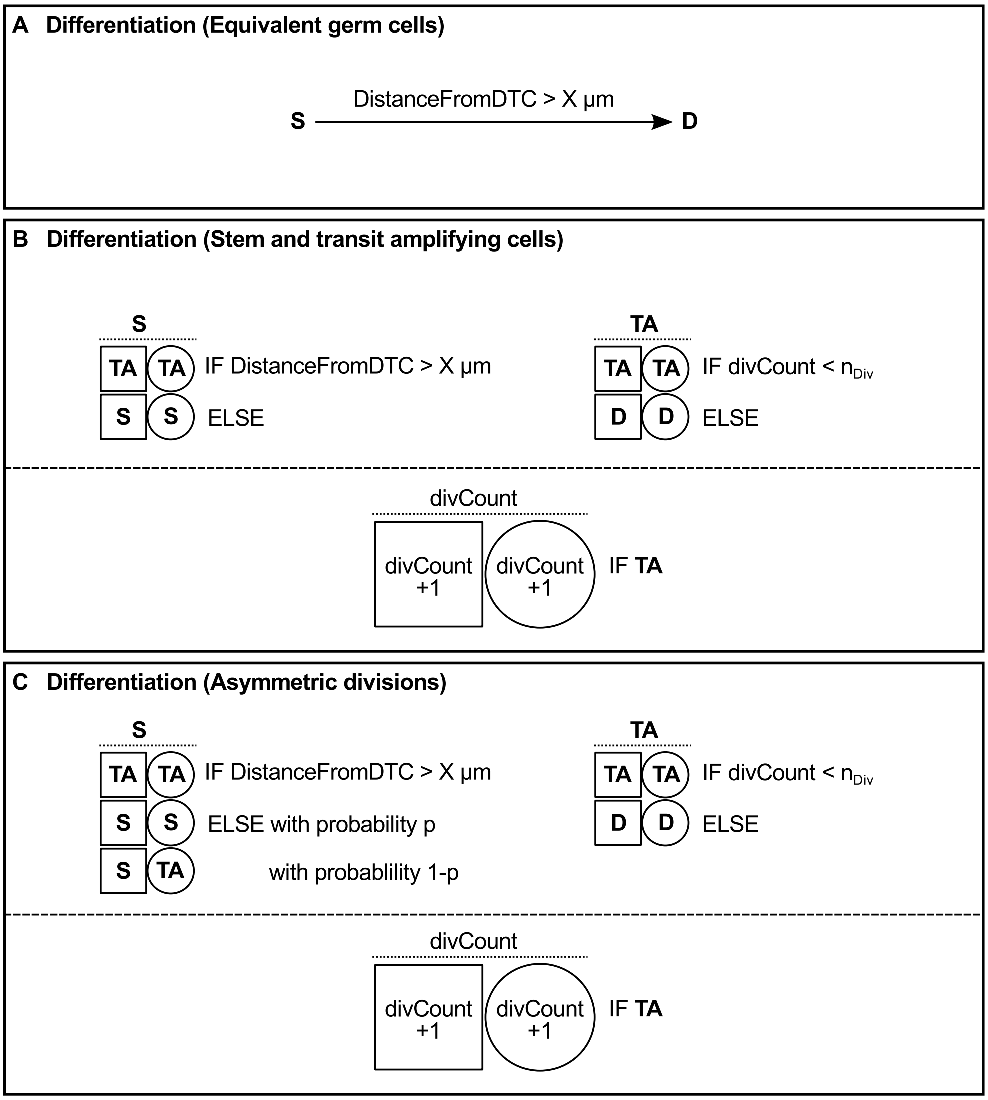


Figure 6.7: Three different models of germ cell differentiation. Each part of this figure shows a logic module corresponding to one of the maintenance mechanisms introduced in Section 6.1. A) Equivalent germ cells. Cells within a certain distance of the DTC are self-renewing, and they become differentiating on leaving this region. B) Stem and transit amplifying cells. Cells within a certain distance of the DTC are stem, and on leaving this region they produce two transit amplifying cells at their next division. Transit amplifying cells undergo n_{Div} divisions before becoming differentiating. C) Asymmetric divisions. This model is identical to B), except that transit amplifying cells can be produced inside the niche with some probability. Transit amplifying cells can therefore displace stem cells.

6.3 Comparison with experimental measurements

Having described our model of the distal germ line, the next section presents simulation results for each potential maintenance mechanism. These results are then compared with experimental observations. Throughout this process, we consider a range of germline properties including proliferative cell count, mitotic index profile, cell cycle phase breakdown, and hourly meiotic output.

6.3.1 Distal germ line appearance

We begin by presenting snapshots of each maintenance model, captured after 100 hours of simulation. We then comment on whether the resulting distal germ line has a realistic appearance. This gives us the opportunity to display typical model output and highlight the effect of varying parameters.

Equivalent germ cells

Firstly, Figure 6.8 shows snapshots of the equivalent germ cells model, run using three different

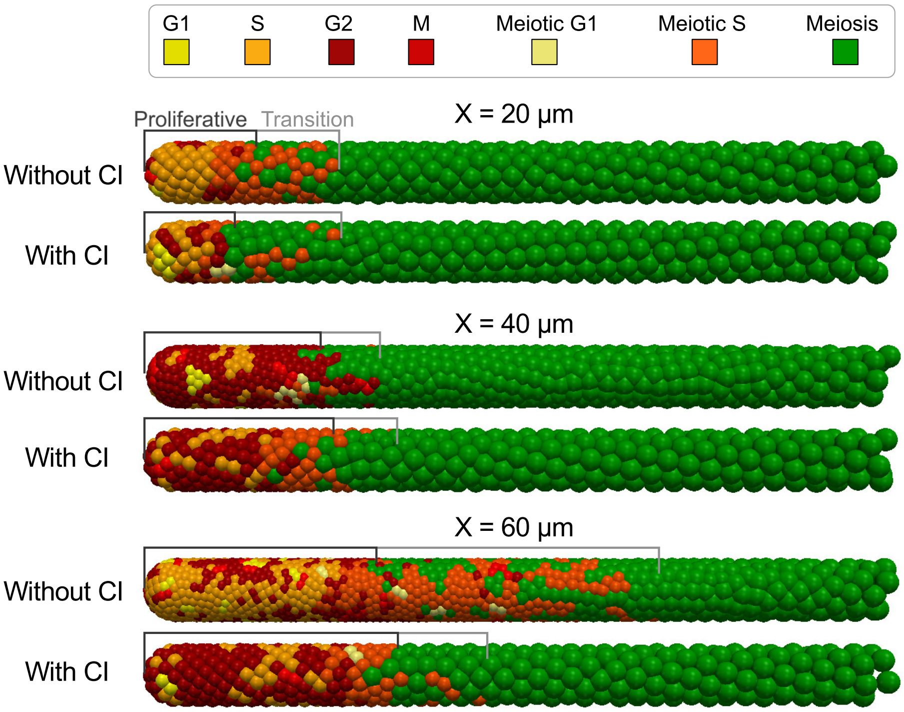


Figure 6.8: Snapshots of the equivalent germ cells model, run using different values of X, either with or without contact inhibition (CI). The images above show the distal germ line after 100 hours, with germ cells coloured according to their cell cycle state. Increasing the range of the DTC signal increases proliferative zone length, while adding contact inhibition reduces it. Setting $X = 40\mu m$ gives the most realistic appearance of the parameter sets tested (see text).

DTC signalling ranges X, both with and without contact inhibition. A distinct proliferative zone is visible in all snapshots, and, as expected, we find that the length of the proliferative zone increases with X and decreases with contact inhibition. The experimentally observed length of the proliferative zone is $20-30$ cell rows [114], and based on manual row counts $X = 40-60\mu$m most closely reproduces this value. However, the $X = 60\mu$m simulation without contact inhibition appears overcrowded, and has a transition zone that is far too long (31 cell rows compared to an expected value of 5 [114]). We therefore conclude that, of the parameters sets tested for this model, $X = 40\mu$m or $X = 60\mu$m with contact inhibition give the most realistic appearance.

Stem and transit amplifying cells

Figure 6.9 presents similar snapshots for the stem and transit amplifying cells model. There are too many potential parameter sets to show every combination tested, so we include a sample of representative images with different values of X, n_{Div} and contact inhibition status. We find that allowing multiple transit amplifying divisions greatly increases the length of the proliferative zone. In particular, when $X = 20\mu$m and $n_{Div} = 3$ all germ cells become proliferative (Figure 6.9). To achieve a reasonable germ line appearance, a larger value of n_{Div} must be paired with a shorter DTC signalling range. However, $X \leq 5\mu$m is too restrictive, since 5μm is less than one cell

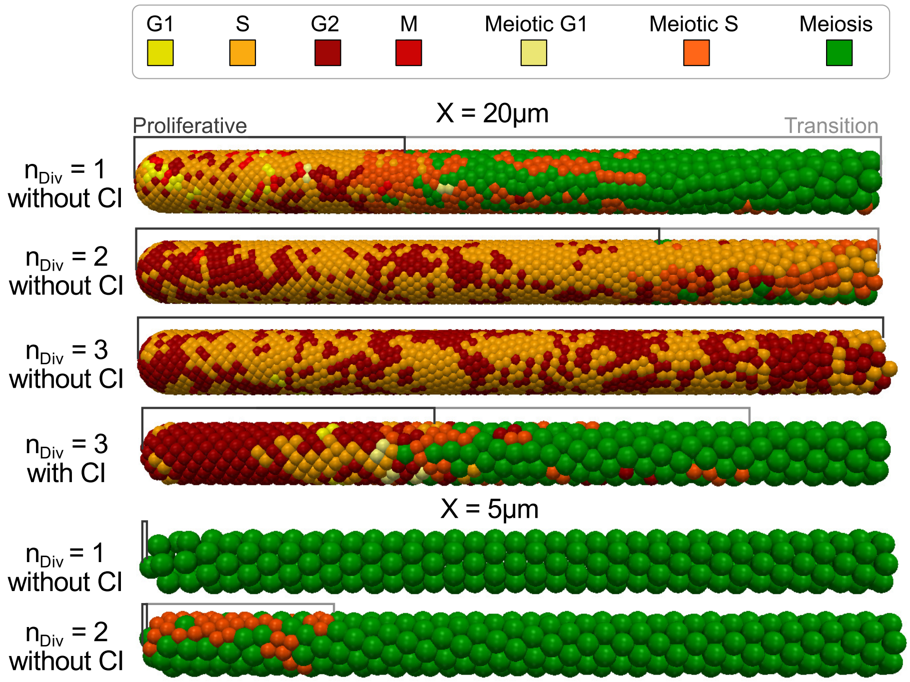


Figure 6.9: Snapshots of the stem and transit amplifying cells model, run using a range of parameter values. Introducing multiple transit amplifying divisions ($n_{Div} = 2, 3$) tends to make the proliferative zone unrealistically long. This effect must be countered to achieve a realistic appearance, either by introducing contact inhibition or decreasing the DTC signalling range, X. However, setting $X \leq 5\mu$m causes the opposite problem: the proliferative zone is frequently lost.

diameter, and the proliferative zone is easily lost in the course of random cell movement. Adding contact inhibition can also compensate for multiple transit amplifying divisions, recovering a more reasonable proliferative zone length. However, the transition zone still tends to be unrealistically long. Of the parameters we tested, the best fitting appearance is achieved with $n_{Div} = 1$ and $X = 10 - 20\mu m$.

Asymmetric divisions

Finally, Figure 6.10 shows selected snapshots of the asymmetric divisions model. As expected, the larger the stem-stem division probability, p, the longer the proliferative zone after 100 hours, as fewer stem cells are displaced from the niche over time. More interestingly, transit amplifying cells that are born close to the distal tip in this model do not always move away before entering meiosis. This leads to meiotic cells (green in Figure 6.10) appearing throughout the proliferative zone, eventually turning it into a mixed transition region. While isolated meiotic cells do sometimes appear inside the proliferative zone *in vivo*, the extensive fragmentation seen in Figure 6.10 is unrealistic, and in terms of appearance the asymmetric divisions model produces the weakest fit.

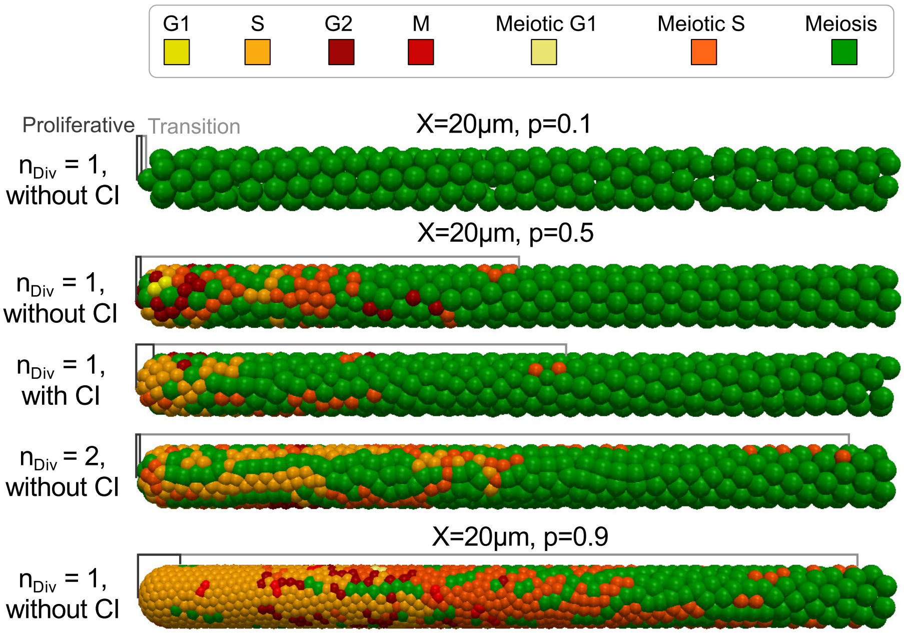


Figure 6.10: Snapshots of the asymmetric divisions model, run with a range of parameter values. Increasing the stem-stem division probability, p, results in a longer proliferative zone after 100 hours, since fewer stem cells are displaced from the niche. We also observe that transit amplifying cells born near the distal tip do not always move away before entering meiosis. This leads to meiotic (green) cells appearing unrealistically often throughout the proliferative zone.

6.3.2 Proliferative cell count

Next, we looked at the total number of proliferative cells under each maintenance model. According to the literature, the expected proliferative cell count for a young adult germ line is 200-250 cells [81, 82, 114]. This figure includes germ cells in meiotic S phase, which are experimentally indistinguishable from cells in mitotic S.

Equivalent germ cells

Figure 6.11 shows simulated proliferative cell counts for the equivalent germ cells model. Two trends are visible: as the DTC signalling range, X, increases, the total number of proliferative cells increases and the steady state takes longer to reach. Adding contact inhibition drastically reduces the proliferative cell count, and the steady state is attained relatively quickly. Comparing these results to the observed value of ~ 225 proliferative cells, it appears that either $X = 30$µm without contact inhibition or $X = 40$µm with contact inhibition would give the best fit.

Stem and transit amplifying cells

Figure 6.12 presents similar counts for the stem and transit amplifying cells model. As expected, increasing the number of transit amplifying divisions, n_{Div}, produces a corresponding increase in

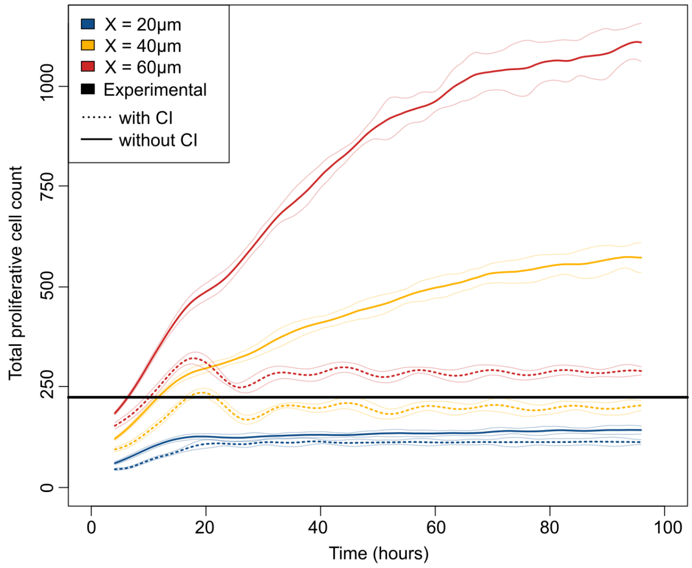


Figure 6.11: Proliferative cell count over time for the equivalent germ cells model. The three different colours above correspond to different values of the DTC signalling range, X (see legend). Solid lines indicate results without contact inhibition, while dotted lines indicate results with contact inhibition (mean of 10 simulations per condition, sliding window averaged with an 8 hour window to remove phasing). Thin lines mark a region ±2 standard deviations either side of the mean for each parameter set. Finally, the experimentally observed value is indicated by a solid black line. Taking $20 < X < 40$µm without contact inhibition, or $X = 40$µm with contact inhibition gave the best fitting proliferative cell count.

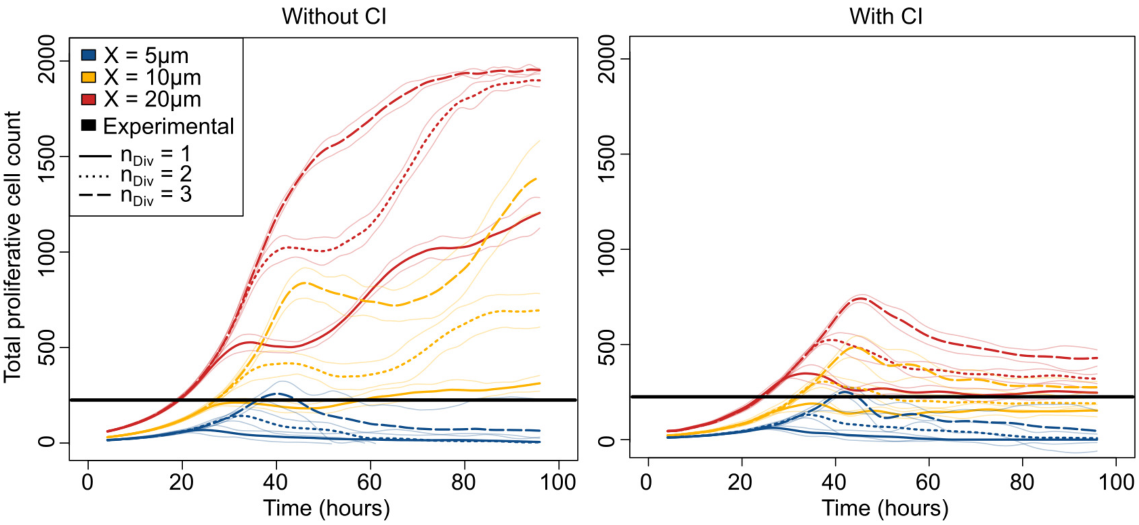


Figure 6.12: Proliferative cell count over time for the stem and transit amplifying cells model. Results were processed as described in the legend of Figure 6.11. Here, line type indicates the value of the parameter n_{Div}, and runs with and without contact inhibition are shown in separate panels. In general, setting $X = 5\mu m$ tends to lead to proliferative zone loss, while setting $n_{Div} = 3$ tends to produce proliferative cell counts that are too high. A good fit can be achieved without contact inhibition by taking $X = 10\mu m$ and $n_{Div} = 1$. Introducing contact inhibition increases the number of acceptable parameter sets.

the number of proliferative cells. Indeed, prior to the first transit amplifying cells entering meiosis, there is a period where the proliferative cell count grows exponentially, the length of which is determined by n_{Div}. The effect of varying X and introducing contact inhibition, meanwhile, is similar to the previous model. A good fit can be obtained without contact inhibition by taking $X = 10\mu m$ and $n_{Div} = 1$, or with a range of parameter sets when contact inhibition is included.

Asymmetric divisions

Finally, Figure 6.13 shows proliferative cell counts for the asymmetric divisions model. Here, and again in Figure 6.23, we do not present data for simulations with contact inhibition enabled. Since the effects of contact inhibition are similar to in previous models, and since these plots already contain a lot of information, we instead focus on the behaviour as p and n_{Div} are varied. The overall appearance of these graphs is similar to the previous model, except that when the stem-stem division probability is low, the number of proliferative cells eventually enters a sharp decline. Interestingly, although we chose to study this model for its potential to explain proliferative cell loss with age, it can also generate a relatively stable proliferative cell count. For example, when $X = 10\mu m$, $n_{Div} = 1$ and $p = 0.9$, the gradual displacement of stem cells from the niche is counterbalanced by the accumulation of proliferative cells that this parameter set would otherwise produce.

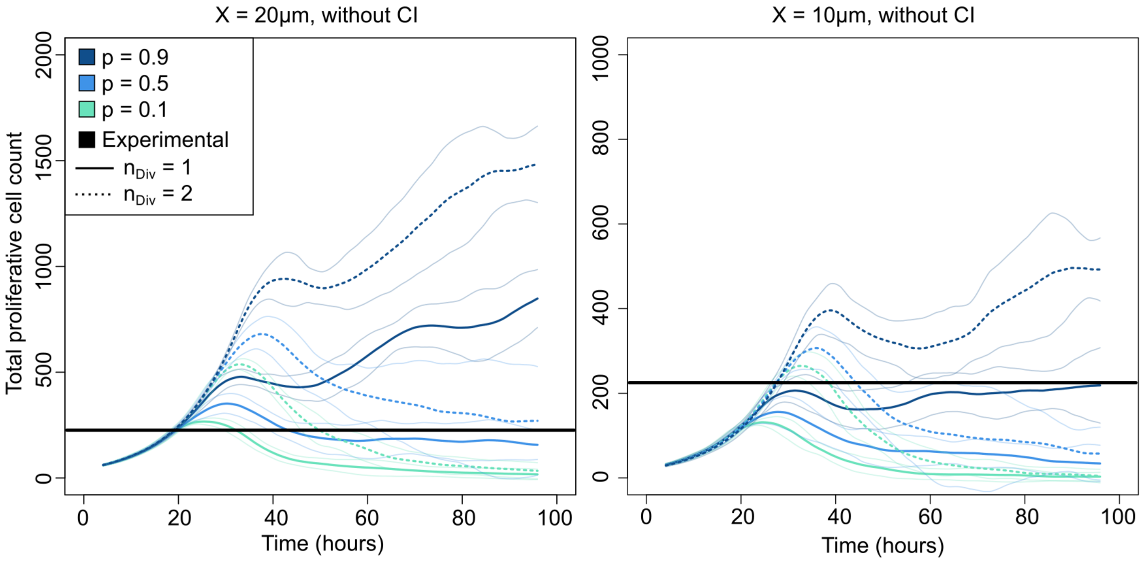


Figure 6.13: Proliferative cell count over time for the asymmetric divisions model. Again, results were processed as described in the Figure 6.11 legend. We find that selecting a low stem-stem division probability, p, produces an eventual decline in the number of proliferative cells. However, a relatively stable proliferative cell count can also be achieved, by increasing the likelihood of symmetric vs. asymmetric renewal (see $X = 10\mu\text{m}$ with $n_{Div} = 1$ and $p = 0.9$).

6.3.3 Mitotic index profile

Another germline property to consider is the proliferative zone mitotic index (see Sections 2.4 and 4.3.2). Equation (2.1) defined the mitotic index, which can also be expressed in percentage form as:

$$\text{Mitotic Index (\%)} = 100 \times \frac{\text{Number of M phase cells}}{\text{Number of cells}}.$$

Most literature sources report a single mitotic index for the entire proliferative zone. However, spatial mitotic index profiles also exist, and we use one here, as it provides a trend to compare our simulations against rather than a single value. The mitotic index data we use comes from Maciejowski *et al.* (2006) [154]. It was obtained by taking 587 gonad arms from worms of a similar age and capturing a confocal z-stack image of each one. These z-stack images were processed to identify all germ cells in M phase. A separate mitotic index was then calculated for each cell row of the organ, based on aggregate counts across all biological replicates.

Generating mitotic index profiles *in silico*

To generate comparable mitotic index profiles *in silico*, each maintenance mechanism and parameter set was simulated 30 times, using a different random seed for each replicate. At regular intervals during a simulation, we captured a “snapshot” of mitotic activity. This involved dividing the long axis of the gonad up into 5μm wide bins, then recording the number of M phase cells and the total cells in each bin. The mitotic index of bin i at snapshot time T is given by

$$\text{Mitotic index of bin } i \text{ (T)} = 100 \times \frac{\sum_r \text{Number of M phase cells in bin } i \text{ (T)}}{\sum_r \text{Number of cells in bin } i \text{ (T)}},$$

where r ranges over all replicates. This calculation produces a mitotic index profile for time T, where the number of bins determines the spatial resolution.

Ideally, we would like to use just one snapshot per replicate simulation, always captured at the same time point, so as to exactly reproduce the experimental procedure from [154]. Unfortunately, the number of replicates we can carry out is limited by time constraints, since simulations take on the order of hours and we want to investigate a range of parameter sets. Therefore, we decided to pool multiple snapshots from each replicate, in order to gather enough data to generate a smooth profile. We took into account a set of snapshots from each simulation captured between 93 and 100 hours at half hour intervals. Since M phase in our model only lasts 10 minutes, spacing the snapshots half an hour apart prevents double counting of the same division event. Meanwhile, taking snapshots over a 7 hour period avoids time correlated divisions, since the germ cell cycle length is 8 hours. Figure 6.14 illustrates the impact of combining multiple snapshots per replicate simulation, and shows that the result is reasonably representative of the set of single snapshot

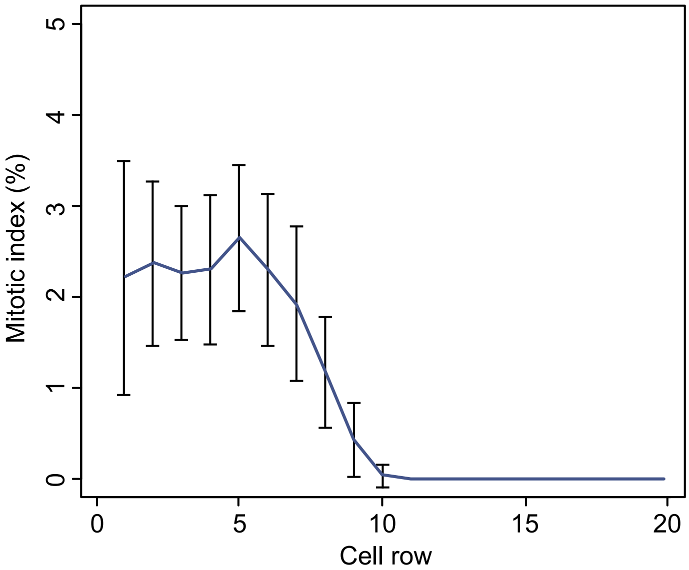


Figure 6.14: An assessment of our method for calculating mitotic index profiles. Shown in blue is a mitotic index profile for the equivalent germ cells model, with $X = 20\mu\text{m}$ and without contact inhibition. It was computed by pooling snapshots taken at multiple timepoints as described above. Meanwhile, the error bars are based on a set of profiles, each calculated using a single timepoint per replicate simulation. Each error bar shows one standard deviation either side of the mean value for this profile set. While there is a lot of variation, the shape of the pooled profile is generally representative of the set of single snapshot results.

profiles. We therefore argue that pooling multiple snapshots produces a valid mitotic index profile for the system.

Mitotic index profile results

Equivalent germ cells

Figure 6.15A shows the mitotic index profiles we obtained for the equivalent germ cells model. To make the general shape of the profiles clearer, Figure 6.15B presents the same data, but with a sliding window average of width 20μm applied. We find that without contact inhibition, this model results in a largely flat mitotic index profile, followed by a sharp fall-off in mitotic activity. As X increases, the fall-off begins at a greater distance from the DTC, since germ cells remain in the signalling region for longer and travel further during their final cell cycle. When contact inhibition is enabled, it depresses the mitotic index most near the distal tip, resulting in a spatial profile closer to the experimentally observed result. However, the fall-off in mitotic activity is unchanged and is sharper than expected. Overall, taking $X = 40\mu m$ with contact inhibition produces the best fit.

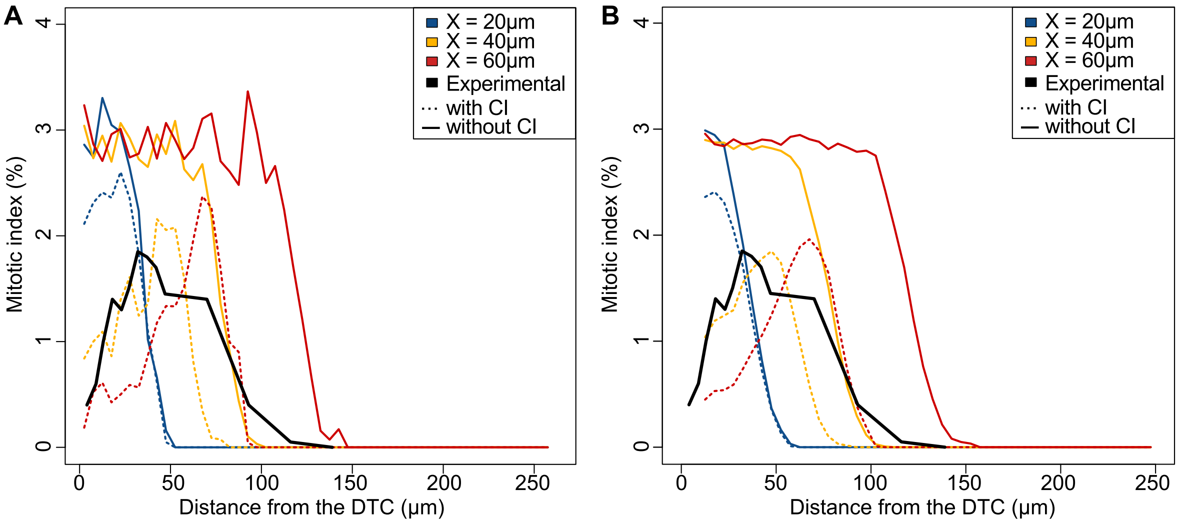


Figure 6.15: Mitotic index profiles for the equivalent germ cells model, with a range of parameters. A) Simulated mitotic index profiles calculated as described in the text. Experimental data is taken from [154]. Without contact inhibition, this model produces a flat mitotic index profile followed by a sharp fall-off. The length of the flat region increases with X. Introducing contact inhibition suppresses the mitotic index near the distal tip, resulting in a profile closer to the experimental result. B) The same profiles as A) but with a sliding window average applied to clarify the overall shape.

Stem and transit amplifying cells

Figure 6.16 shows similar smoothed mitotic index profiles for the stem and transit amplifying cells model. The general shape of these profiles resembles those in Figure 6.15B: flat initially, with mitotic activity falling off further from the DTC. However, introducing transit amplifying divisions results in a longer region of mitotic activity, and a more gradual fall-off that better matches the experimental profile. Again, enabling contact inhibition suppresses the mitotic index near the distal tip, leading to a reasonably good fit to the data from [154]; particularly when $X = 10$µm and $n_{Div} = 1$ or 2.

Asymmetric divisions

Lastly, Figure 6.17 shows smoothed mitotic index profiles for the asymmetric divisions model. A notable feature here is that the mitotic index profile no longer has a flat portion. Instead, mitotic activity simply declines with distance from the distal tip. This is likely due to the fact that transit amplifying cells can now be produced inside the niche region. These transit amplifying cells have a reasonable probability of moving in a proximal direction before they enter meiosis. Therefore, stem cell loss/displacement is more prevalent at the proximal end of the niche than near the distal tip, resulting in a spatial trend in mitotic activity. As expected, when the stem-stem division probability is low, mitotic activity is reduced after ~ 100 hours, due to extensive stem cell

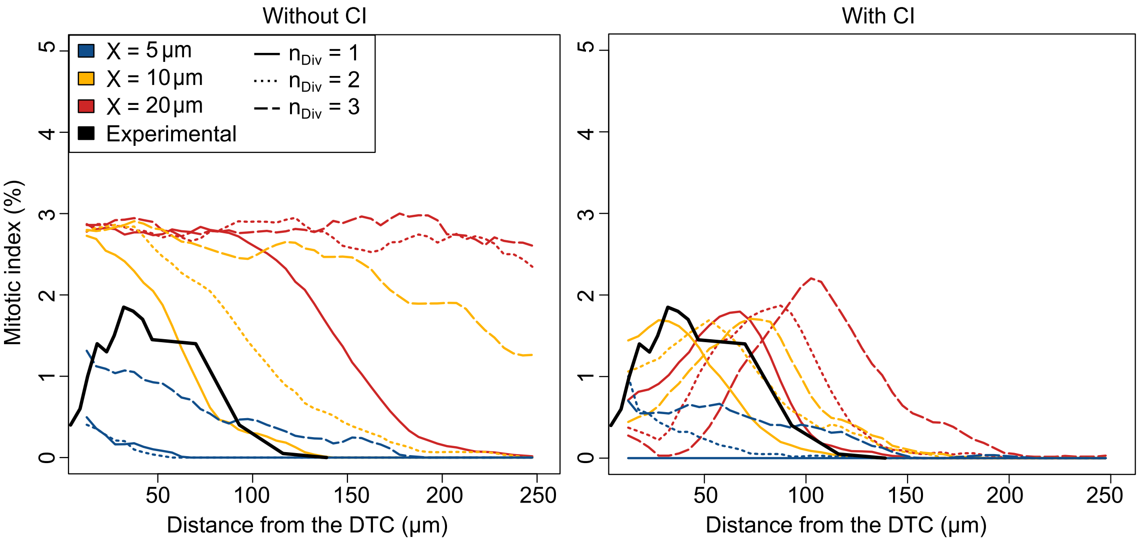


Figure 6.16: Smoothed mitotic index profiles for the stem and transit amplifying cells model. The left hand panel presents results without contact inhibition enabled. This model again generates a flat mitotic index profile, but the subsequent fall-off in mitotic activity is more gradual than for the equivalent germ cells model (Figure 6.15). When contact inhibition is enabled (right panel), the mitotic index near the distal tip falls, resulting in a better fit to the experimental profile (see $X = 10$µm with $n_{Div} = 1$ or 2).

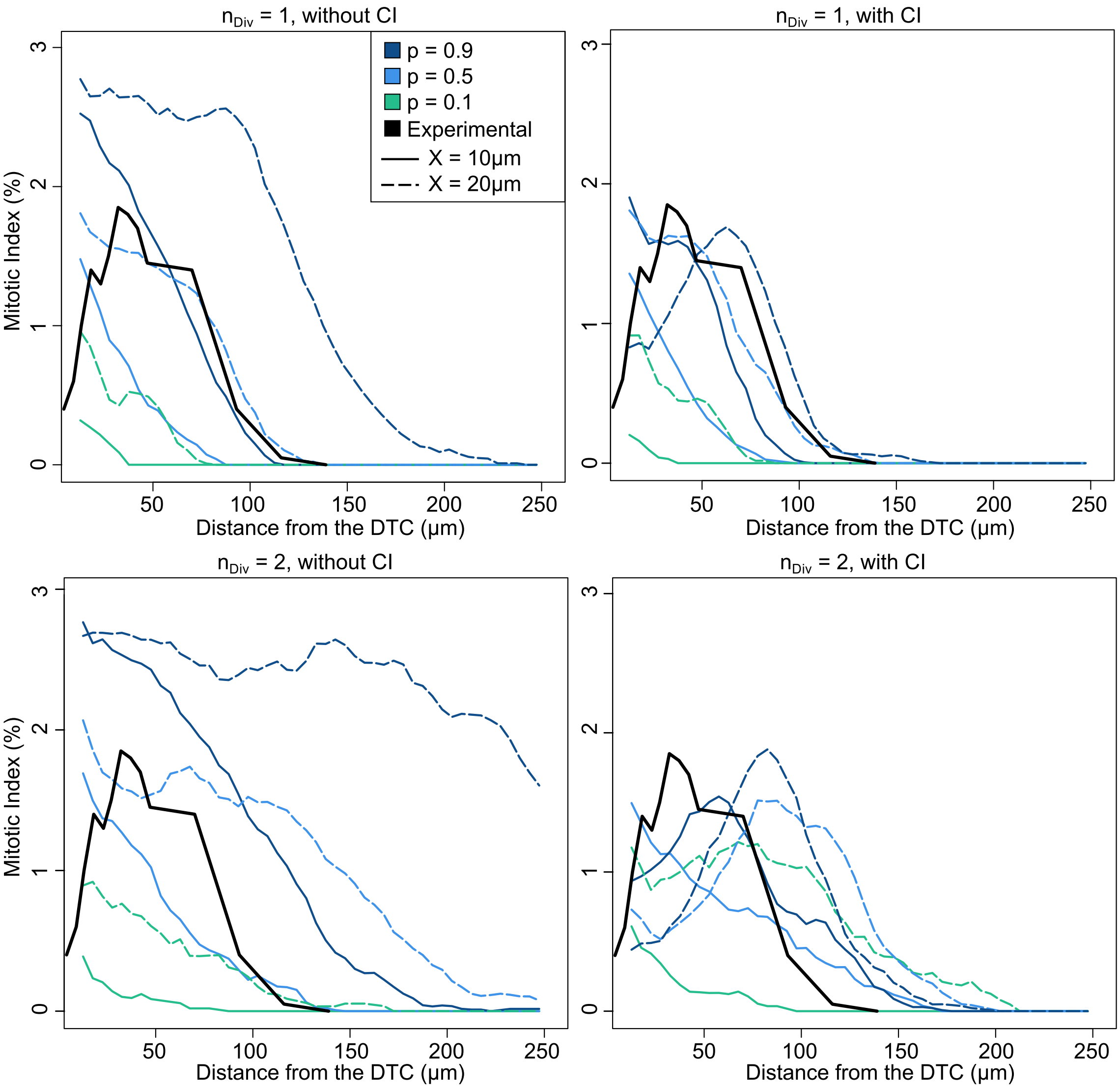


Figure 6.17: Smoothed mitotic index profiles for the asymmetric divisions model. When the stem-stem division probability, p, is low, mitotic activity after 100 hours is also low, due to stem cell loss/displacement. Adding contact inhibition still suppresses the mitotic index near the distal tip, but only in scenarios where overcrowding is likely (high values of X, n_{Div} and p). Of the parameter sets tested, the best fit is achieved by taking $X = 20\mu\text{m}$, $n_{Div} = 1$ and $0.5 \leq p \leq 0.9$ with contact inhibition enabled.

loss. Introducing contact inhibition still suppresses mitotic activity near the distal tip, but only for parameter sets likely to produce overcrowding (high values of X, n_{Div} and p). None of the parameter sets we tested for this model produced a very good fit to the experimental profile, although $X = 20\mu\text{m}$, $n_{Div} = 1$, $p = 0.5$ with contact inhibition fits the tail of the distribution well.

6.3.4 Cell cycle phase percentages

Next, we explore the percentage of proliferative cells found in each cell cycle phase over the course of a simulation. Experimental phase percentages are available for comparison, although these values cannot be measured directly, and must instead be inferred from time course experiments using multiple labelling techniques. Here, we compare our simulated cell cycle phase percentages against estimates reported by Fox *et al.* in (2011) and (2015) [81, 82]. Table 6.1 gives their experimental results, along with an indication of how each value was obtained.

Equivalent germ cells

Figure 6.18 shows simulated cell cycle phase percentages for the equivalent germ cells model. Across all parameter sets, there is a good fit for the percentage of cells in G1 and M phase. This is unsurprising, given that we hard-coded reasonable phase lengths into the cell cycle model. More notable is the fact that, when contact inhibition is applied, the G2 phase percentage rises well above expected levels. This reinforces our earlier point from Chapter 4, that a model reliant on many germ cells being in G2 arrest will have a distorted cell cycle phase distribution that may become unrealistic. The effect is more pronounced when X is large, since more proliferative cells means greater cell crowding, causing contact inhibition to trigger more frequently. A corresponding drop occurs in the percentage of S phase and meiotic S germ cells when contact inhibition is present, but this is a consequence of the rise in G2 occupancy, combined with the fact that we report percentages not counts. The best fitting parameter set for this mechanism is $X = 20\mu\text{m}$ with contact inhibition. Using these parameters, the short DTC signalling range moderates the effects of contact inhibition.

Phase	% of proliferative cells	Experimental details
G1	2%	Remainder after accounting for other phases.
S	57%	Percentage of REC-8 positive cells labelled by a short EdU pulse.
G2	39%	Continuous EdU labelling was applied. The G2 duration was estimated based on the time taken for the first wave of EdU positive cells to label positive for PH3.
M	2%	Percentage of REC-8 positive cells that are also PH3 positive.
Meiotic S	30%	Based on balancing proliferation and meiotic output in a mathematical model in [81].
Mitotic S	27%	Correcting for the number of cells in meiotic S.

Table 6.1: The percentage of proliferative cells in each cell cycle phase, according to Fox *et al.* [81, 82]. This table gives an estimated occupancy percentage for each phase, as well as a brief description of the experiment that produced the result. A number of cell labelling techniques are referenced, including: EdU labelling, which permanently marks cells passing through S phase; PH3 labelling, which marks cells in M phase; and REC-8 labelling, which marks proliferative cells (either in mitosis or meiotic S).

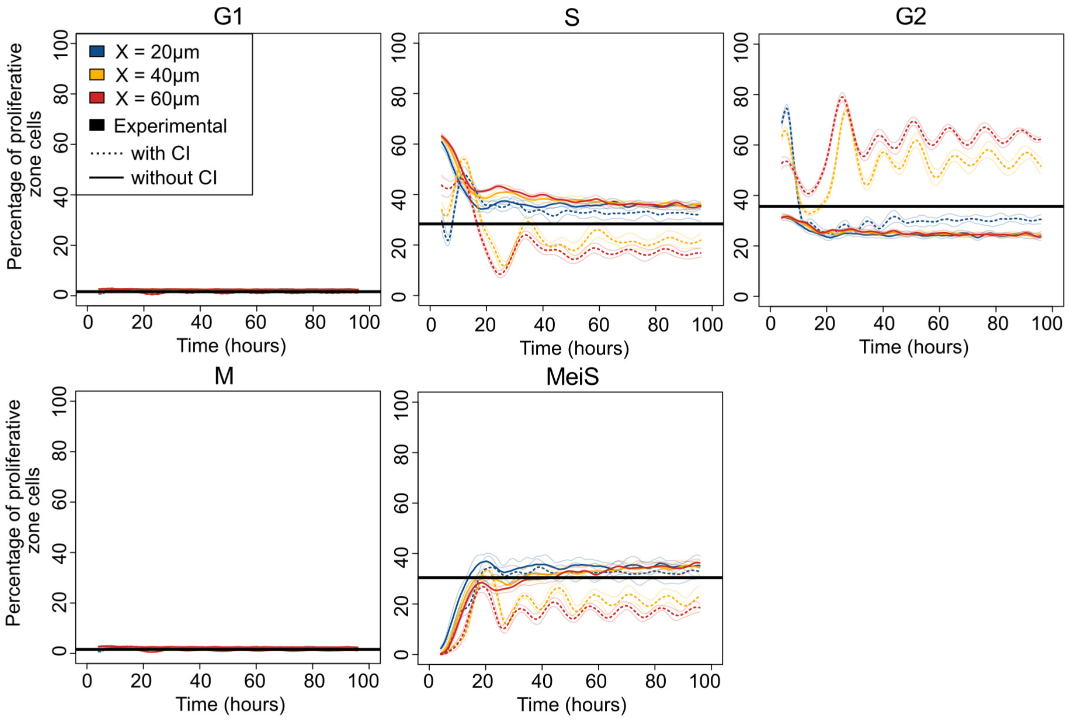


Figure 6.18: Cell cycle phase percentages for the equivalent germ cells model. Each panel corresponds to one cell cycle phase, and a solid black line shows the expected percentage of proliferative cells found in that phase. The other traces represent simulated results, with colour indicating the size of the parameter X and line type indicating whether contact inhibition is enabled. Results are the mean of 10 replicates with a sliding window average of length 8 hours applied. A region ±2 standard deviations is shown either side of the mean. We find that introducing contact inhibition causes a noticeable increase in the G2 percentage, particularly when X is large. This has a corresponding effect on the percentage of cells in S phase and meiotic S.

In Figure 6.18 and in several other figures in this section, oscillations can be seen in the percentage of G2 phase cells over time. Since this oscillation is more pronounced in simulations with contact inhibition, we suggest that germ cells are repeatedly becoming tightly packed, arresting in G2, relaxing, dividing, then becoming tightly packed again; leading to some amount of cell cycle synchronisation. Presumably the small amount of stochasticity introduced into phase lengths in our model is insufficient to offset this effect.

Stem and transit amplifying cells

Figure 6.19 shows the corresponding percentages for the stem and transit amplifying cells model. Looking at the results without contact inhibition first, some new trends appear. The introduction of transit amplifying divisions means that germ cells now take longer to enter meiotic S. Therefore, we see a transient period at the beginning of each simulation where the meiotic S percentage is zero and the mitotic S percentage is high. For parameter sets where all germ cells eventually become proliferative (e.g. $X = 20\mu\text{m}$ with $n_{Div} = 2$), this situation reoccurs at the end of each simulation. Once again, introducing contact inhibition increases the percentage of G2 cells, with the size of the effect depending on X and n_{Div}. Proliferative zone loss when $X = 5$µm can also be seen clearly in these plots (the meiotic S percentage tends to 100). A moderately good fit can be achieved under this model by taking $X = 10$µm and $n_{Div} = 1$; consistent with several other germline properties.

Asymmetric divisions

Finally, Figure 6.20 shows cell cycle phase percentages for the asymmetric divisions model. Many of the trends are similar to those in Figure 6.19; particularly the low initial percentage of meiotic S cells, which only increases once germ cells have had time to exhaust their transit amplifying divisions. When p is low, the percentage of meiotic S cells tends to be higher, since more transit amplifying cells are produced within the niche, which go on to become meiotic. Introducing contact inhibition still causes the percentage of G2 cells to rise in general, but the effect is much more pronounced when p and X are both large, as this makes overcrowding and arrests more likely. The best fitting parameter set of those tested is $X = 10$µm, $n_{Div} = 1$, $p = 0.9$ with contact inhibition enabled.

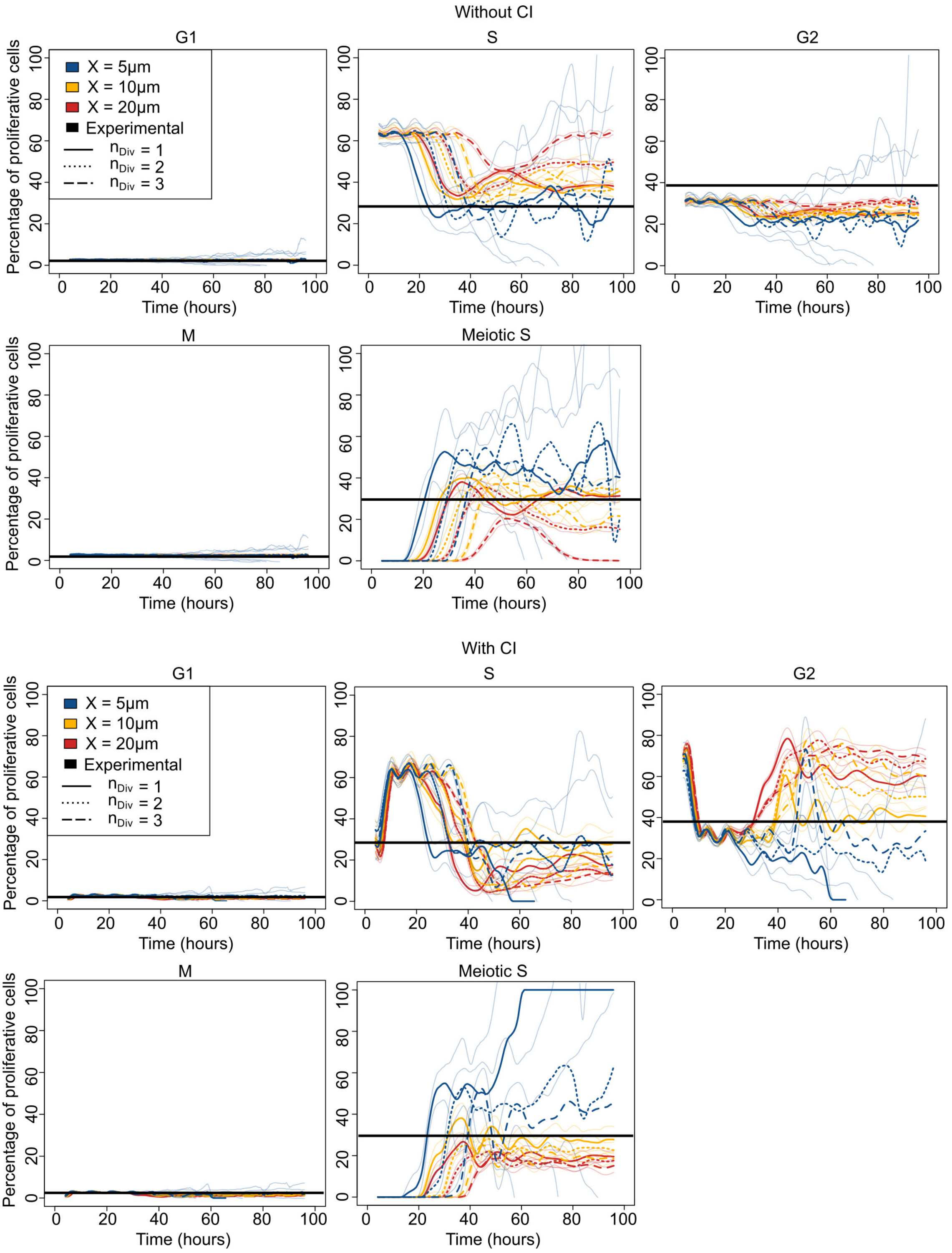


Figure 6.19: Cell cycle phase percentages for the stem and transit amplifying cells model. Results were processed as described in the legend of Figure 6.18. Line type now indicates the value of the parameter n_{Div}, while runs with and without contact inhibition are displayed separately. Under this model, we observe a transient period at the beginning of each simulation where transit amplifying cells have yet to enter meiotic S. Introducing contact inhibition still causes an increase in the G2 percentage, but the magnitude of the effect depends on X and n_{Div}.

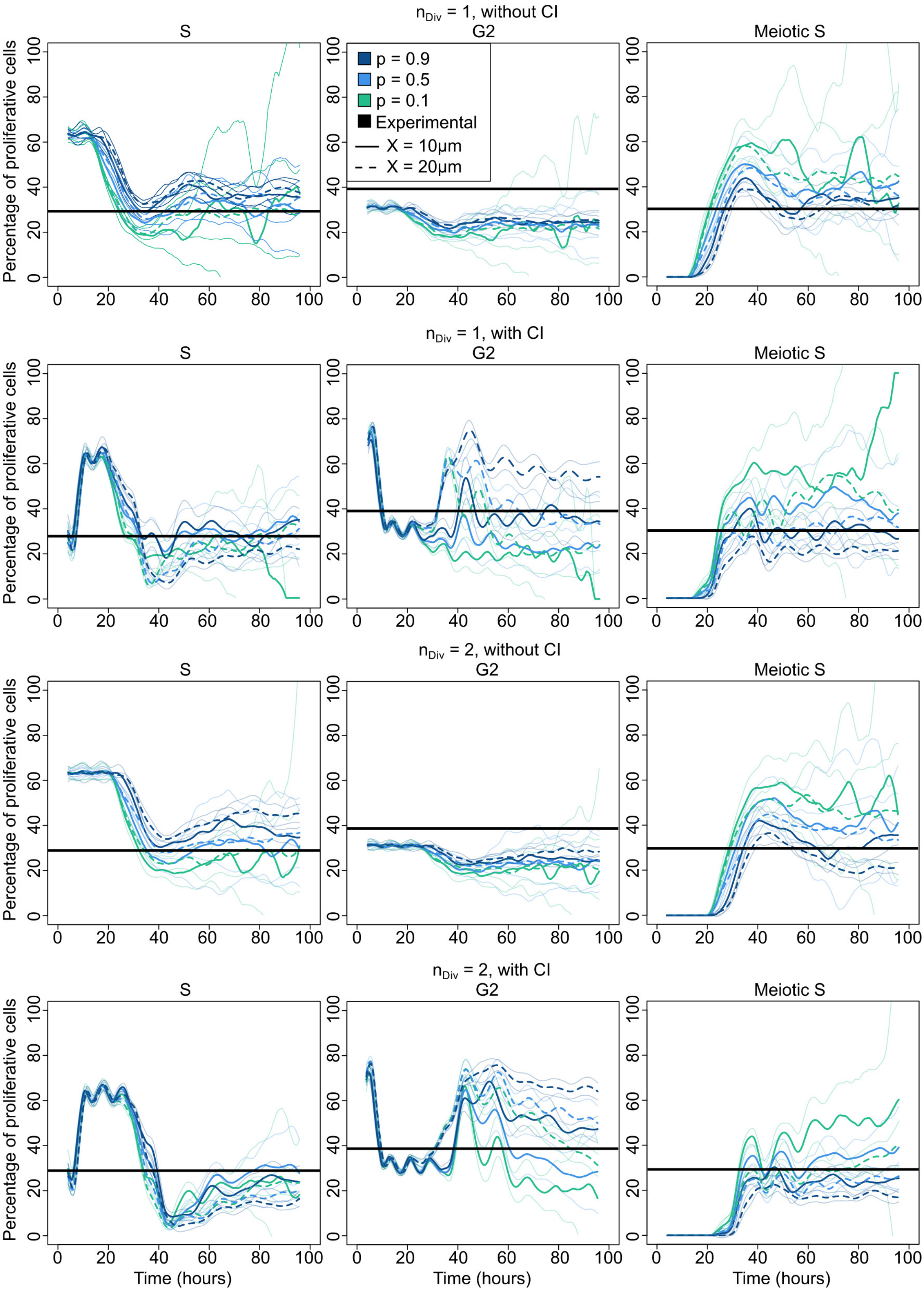


Figure 6.20: Cell cycle phase percentages for the asymmetric divisions model. Results were processed as described in the legend of Figure 6.18. Here we show only the S, G2 and meiotic S phase percentages, since the G1 and M phase plots are similar across all models and parameter sets. Line colour indicates the value of p, while line type indicates the value of X. Most trends are similar to the stem and transit amplifying cells model; for instance, an initial shift occurs in the percentage of mitotic vs. meiotic S phase cells. At late times, simulations with a low probability of stem-stem renewal contain a high percentage of meiotic S cells. Adding contact inhibition still increases the G2 phase percentage, but now the size of the effect also depends on p.

6.3.5 Hourly meiotic output

The final germline property we consider is hourly meiotic output, i.e. the number of germ cells entering meiosis per hour. Fox *et al.* have measured this quantity experimentally, by continuously labelling cells passing through S phase using the nucleotide analogue EdU [81]. After an N hour chase period, an antibody against HIM-3 was applied to mark cells in meiosis. The hourly meiotic output can then be calculated as

$$\text{Hourly meiotic output} = \frac{1}{N}(\text{Number of EdU and HIM-3 positive cells}).$$

This method estimates the output of the adult germ line to be 20 meioses per hour.

Equivalent germ cells

Figure 6.21 shows simulated meiotic output data for the equivalent germ cells model. We see that runs with a long DTC signalling range ($X = 40 - 60$μm) and no contact inhibition produce an unrealistically high meiotic output. The presence of a longer proliferative zone likely increases mitotic pressure in the simulated gonad, forcing germ cells to move faster and therefore enter meiosis at a higher rate. By contrast, runs with contact inhibition enabled, or with $X = 20$μm,

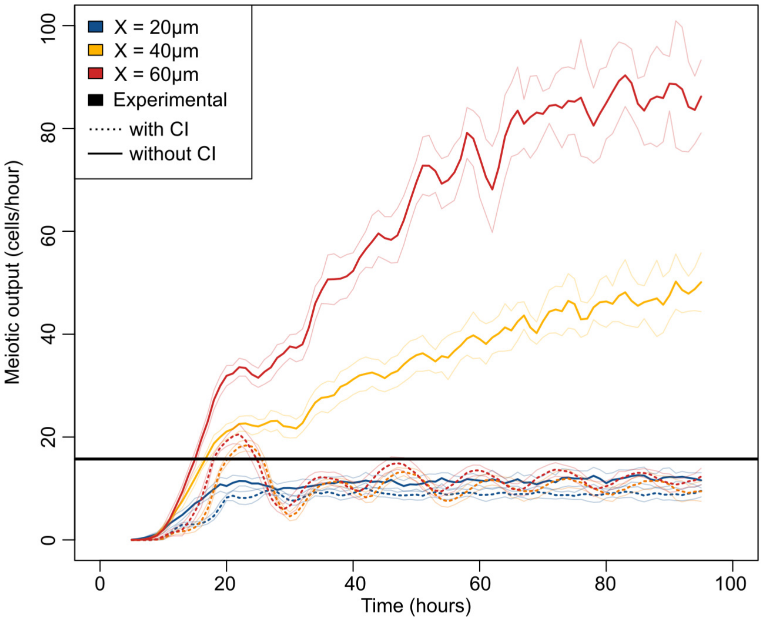


Figure 6.21: Hourly meiotic output for the equivalent germ cells model. Line colour indicates the value of X, line type indicates whether contact inhibition is applied, and thin lines show a region ±2 standard deviations either side of the mean (10 replicates per condition, with an 8 hour sliding window average). We find that a DTC signalling range ≥ 40μm produces an unrealistically high meiotic output, while runs with $X = 20$μm or with contact inhibition enabled give an output that is too low.

produce lower than expected output. While none of the parameter sets we tested fits the data well, we predict that a DTC signalling range of $X = 30\mu m$ without contact inhibition would provide the closest match for this model.

Stem and transit amplifying cells

Figure 6.22 shows the corresponding results for the stem and transit amplifying cells model. Here, certain extreme parameter sets produce undesirable behaviour. Meiotic output tends to zero in two cases: when $X = 5\mu m$ (because the proliferative zone is lost); and when $X = 20\mu m$ with $n_{Div} = 3$ (because all germ cells have become proliferative). Most parameter sets show quite variable meiotic output, depending on how many transit amplifying divisions have been completed, and on the balance between proliferative and meiotic cells at each point in the simulation. Adding contact inhibition significantly suppresses meiotic output, making it unrealistically low in most cases. The best fitting parameters are $X = 10\mu m$ with $n_{Div} = 1$, as for previous germline properties.

Asymmetric divisions

Finally, Figure 6.23 shows meiotic output data for the asymmetric divisions model. As expected, when the stem-stem division probability is low, meiotic output eventually tends to zero as stem cells are progressively lost. However, choosing a larger symmetric renewal probability can produce a surprisingly stable meiotic output over time. This is true for $X = 10\mu m$ with $n_{Div} = 1$ and

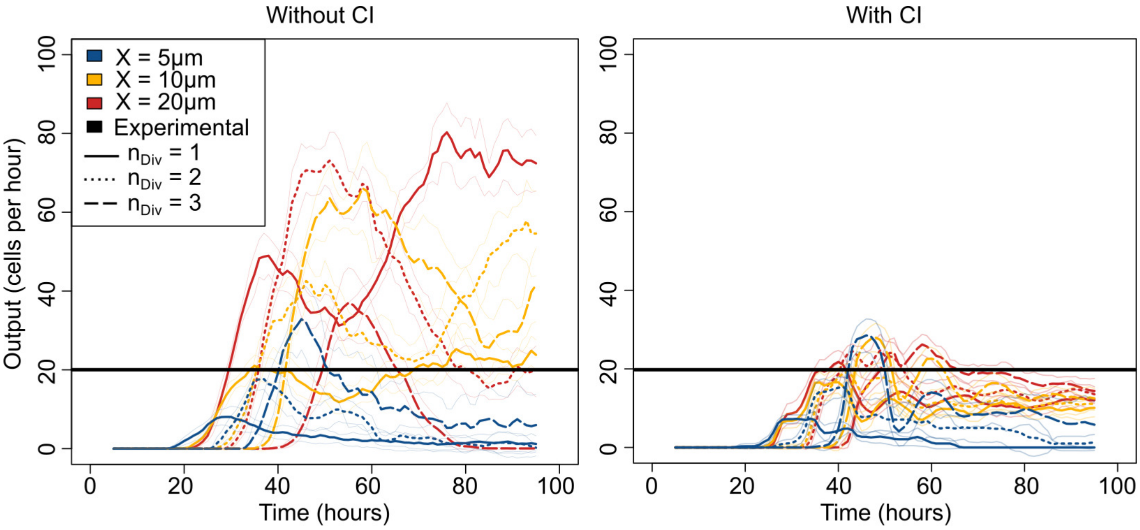


Figure 6.22: Hourly meiotic output for the stem and transit amplifying cells model. Results were processed as described in the Figure 6.21 legend. Under this model without contact inhibition, meiotic output tends to vary over the course of a simulation, depending on the number of transit amplifying divisions that have been completed and on the balance between proliferative and meiotic cells. When contact inhibition is introduced, it lowers the meiotic output to below expected levels across all parameter sets tested. The best fit is achieved with $X = 10\mu m$ and $n_{Div} = 1$.

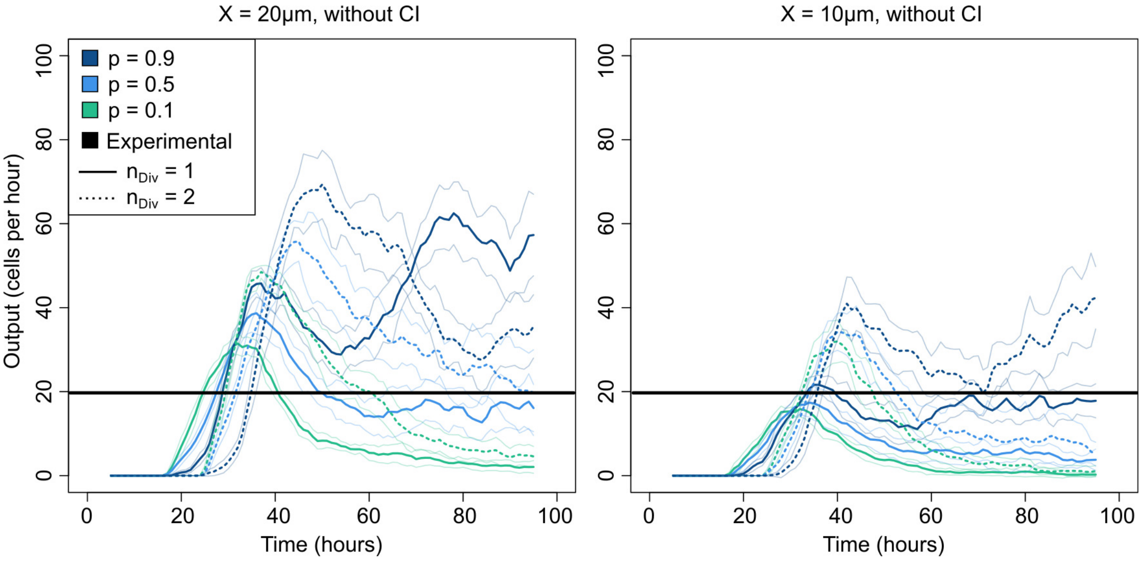


Figure 6.23: Hourly meiotic output for the asymmetric divisions model. Results were processed as described in the Figure 6.21 legend. Under this model, meiotic output tends to zero whenever a low value of p is chosen. However, when stem-stem divisions are made more likely, a surprisingly stable meiotic output can be achieved.

$p = 0.9$, which is also the best fitting parameter set of the values we tested.

6.3.6 Comparison summary

Table 6.2 summarises how well each maintenance model is able to fit the experimental data. Certain germline properties proved more useful than others for distinguishing between models and parameter sets. For example, proliferative cell count and meiotic output can generally be fit well by any of our three models, provided appropriate parameters are chosen. Cell cycle phase percentages and the spatial mitotic index profile proved somewhat more useful. Although this data did not allow us to definitively rule out any model, it did reveal general trends in which parameter sets are permissible; particularly with regard to contact inhibition status. For example, parameter sets that trigger contact inhibition very frequently are unacceptable, since they lead to an unrealistic percentage of germ cells in G2 arrest. However, contact inhibition combined with milder overcrowding *can* produce well-fitting cell cycle phase percentages. Regarding the mitotic index profile, the long tail of the experimentally observed distribution is best fit by allowing a single transit amplifying division. Meanwhile, the reduced mitotic activity observed near the distal tip can be replicated in our models by introducing contact inhibition within an appropriate parameter set.

Property		**Model**	
	Equivalent germ cells	*Stem and transit amplifying cells*	*Asymmetric divisions*
Appearance	Good fit ($X = 40$μm; $X = 60$μm + CI)	Moderate fit, long transition zone ($X =$ 10-20μm, $n_{Div} = 1$)	Poor fit, meiotic cells too close to the DTC
Proliferative cell count	Good fit ($X = 30$μm; $X = 40$μm + CI)	Good fit ($X = 10$μm, $n_{Div} = 1$)	Good fit ($X = 10$μm, $n_{Div} = 1$, $p = 0.9$)
Mitotic index profile	Moderate fit ($X = 40$μm + CI)	Moderate fit ($X = 10$μm, $n_{Div} = 2$ + CI)	Moderate fit ($X = 20$μm, $n_{Div} = 1$, $p = 0.5$ + CI)
Cell cycle phase percentages	Moderate fit ($X = 20$μm + CI)	Moderate fit ($X = 10$μm, $n_{Div} = 1$)	Good fit ($X = 10$μm, $n_{Div} = 1$, $p = 0.9$ + CI)
Hourly meiotic output	Poor fit ($X = 30$μm would be better)	Good fit ($X = 10$μm, $n_{Div} = 1$)	Good fit ($X = 10$μm, $n_{Div} = 1$, $p = 0.9$)

Table 6.2: A summary of the fit obtained using different maintenance models. For each model and germline property, we give the level of fit to the data that was achieved, along with the best fitting parameter set. In the process, a reasonable choice of baseline parameters emerges for each model.

These results do not allow us to select a clearly superior model of the three presented. However, the weakest hypothesis at this stage is the asymmetric divisions model. The fact that transit amplifying cells born near the DTC often fail to migrate proximally before entering meiosis means that meiotic cells appear unrealistically close to the distal tip. The proliferative zone then breaks up into a long transition region, something that is not reported in ageing *C. elegans in vivo* [197]. This problem can only be partially mitigated by changing the model parameters, since it arises from the probabilistic nature of transit amplifying cell production, together with the rate of germ cell movement. If we wish to modify the asymmetric divisions hypothesis to give a more realistic appearance, the best option would be to make transit amplifying cell production much more likely at the proximal end of the niche. Introducing directed divisions would not help, since it would not prevent transit amplifying cells being born near the DTC, nor would it rule out slow or backward germ cell movement. Meanwhile, introducing cell adhesion would negate the purpose of considering a model where stem cells are gradually displaced. It remains to be seen whether the current asymmetric divisions model can at least explain the loss of proliferative cells that occurs with age (a question we explore in Section 6.5).

While these results do not allow us to differentiate between the equivalent germ cells model and the stem and transit amplifying cells model, they *do* help to determine reasonable parameters in each case. This lets us focus on well-fitting parameter sets in subsequent investigations.

We have now completed our study of the baseline behaviour of each maintenance model. In the next section, we look at how the same three models respond to a perturbation; namely, the sudden withdrawal of DTC signalling.

6.4 Disrupting the DTC signal

Having established reasonable baseline parameters for each maintenance model, we next study what happens when DTC signalling is suddenly withdrawn. As described in Section 2.8, two experimental studies have been carried out in which DTC signalling was disrupted in adult worms [47, 81]. These studies disagreed on the manner in which germ cells enter meiosis, once the DTC signal is removed. Cinquin *et al.* (2010) reported seeing a proximal to distal wave of differentiation, with the final five rows of germ cells entering meiosis near-simultaneously. By contrast, Fox and Schedl (2015) reported that the last cells to enter meiosis are distributed throughout the proliferative zone. We would like to clarify the expected behaviour under each of our models, and to make comparisons with these two sets of observations.

In all of our distal germ line models, the DTC influences cell behaviour by sending a position signal to the differentiation logic module. This signal normally communicates how far away a cell is from the distal tip, and a response is triggered when that distance exceeds Xµm (see Figures 6.7A-C). In this investigation, we ran all models as normal for the first 100 simulated hours. The position signal was then set to a large positive value for all germ cells. As a result, cells behave as if they are out of range of the DTC's influence from 100 hours onward, regardless of their actual location.

Equivalent germ cells

We begin by examining the results obtained for the equivalent germ cells model. Figure 6.24 summarises when and where germ cells entered meiosis after the DTC signal was disabled. The various histograms show number of cells entering meiosis as a function of position along the gonad. Time runs down the figure, with each row corresponding to a 5 hour interval following signal withdrawal. Meanwhile, the two columns present data generated using different well-fitting parameter sets: $X = 30$µm without contact inhibition (left), and $X = 40$µm with contact inhibition (right).

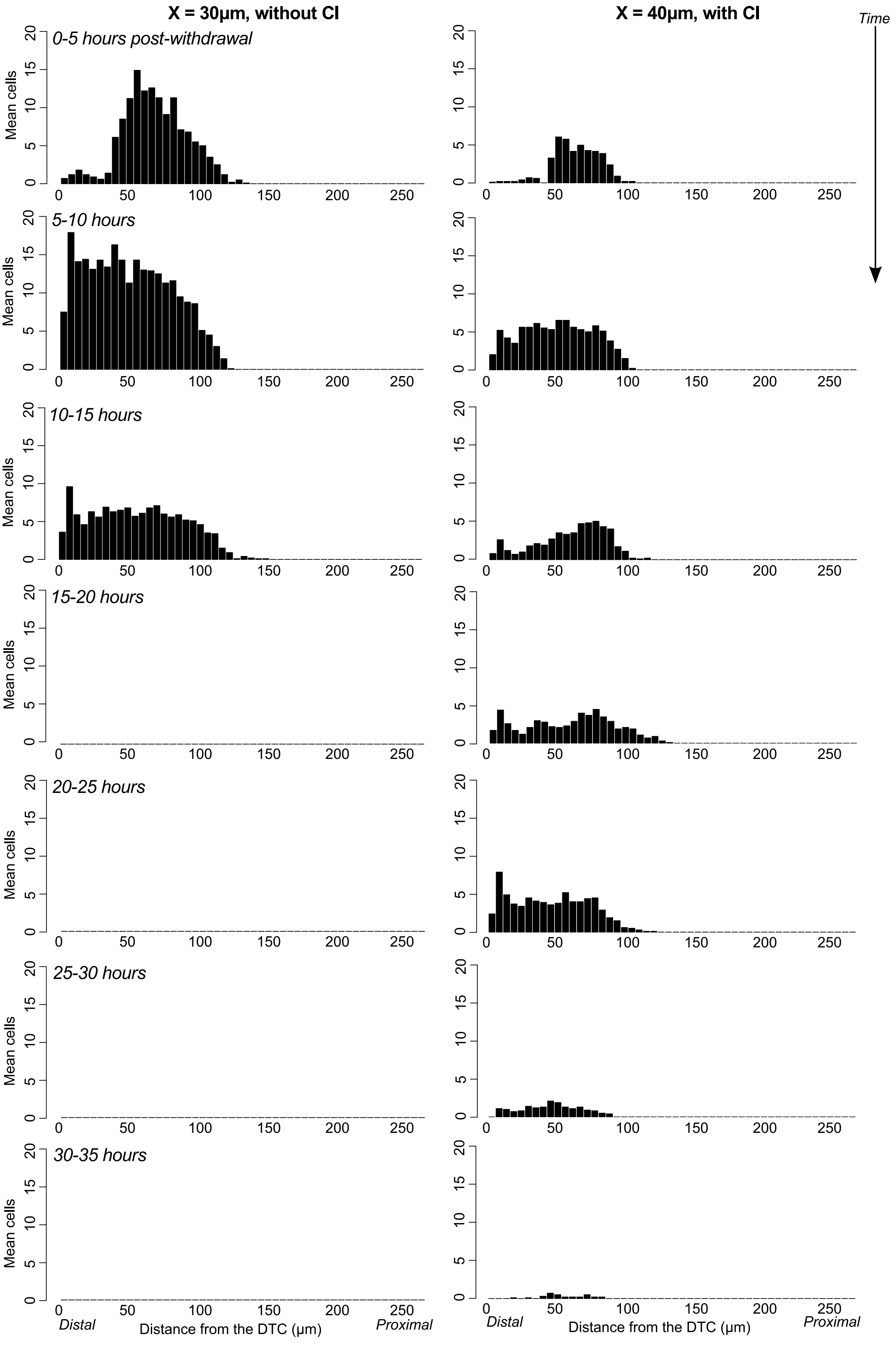


Figure 6.24: The location and timing of meiotic entry events following DTC signal withdrawal - equivalent germ cells model. The two columns of this figure correspond to two different parameter sets. In particular, the left column does not apply contact inhibition, while the right column does. Each histogram shows the number of cells entering meiosis at different positions along the distal gonad (mean of 10 simulations). The top row includes only meiosis events occurring within 5 hours of signal withdrawal; each subsequent row then corresponds to the next 5 hour interval. We find that introducing contact inhibition significantly lengthens the time taken for the entire germ line to enter meiosis. Under both parameter sets, few distal germ cells enter meiosis within 5 hours of DTC signal withdrawal. Only at later times does the meiosis distribution become more spatially uniform.

We find that the two proposed parameter sets lead to a very different delay before the entire germ line becomes meiotic. Without contact inhibition, all germ cells enter meiosis within 15 hours of DTC signal withdrawal. By contrast, when contact inhibition is enabled, it takes 35 hours for the whole germ line to enter meiosis, since overcrowding in the proliferative zone takes time to dissipate, and arrested cells take time to progress from G2 phase through to G1. This increase in the time taken for complete meiotic entry is only associated with the introduction of contact inhibition; simply increasing X from 30μm to 40μm does not produce the same effect. In this respect, the model without contact inhibition matches both sets of experimental observations much more closely. Fox *et al.* report complete entry into meiosis within 10 hours, while Cinquin *et al.* report 6 hours.

Another notable feature of Figure 6.24 is that the meiotic entry distribution in the first 5 hours is spatially non-uniform, despite all germ cells being functionally equivalent. Initially, most meioses take place beyond the normal DTC signalling range of Xμm. It takes several hours for germ cells closer to the DTC to reach G1 phase, pass through meiotic S, and enter meiosis. This pattern is consistent with remarks made by Fox *et al.* in [81]. However, it is easy to see how Cinquin *et al.* could have interpreted the same spatial trend as a proximal to distal wave of differentiation. At later times, meiosis events occur throughout the proliferative zone under this model, and there is no synchronised meiotic entry of the distal-most germ cell rows.

Stem and transit amplifying cells

Figure 6.25 shows the corresponding results for the stem and transit amplifying cells model. The well-fitting parameter set used here takes $n_{Div} = 1$. Therefore the main change relative to the previous model is that germ cells must now undergo a transit amplifying division before entering meiosis. We find that introducing a single amplifying division greatly increases the time taken for the entire germ line to become meiotic. Under the equivalent germ cells model, complete meiotic entry was achieved 15 hours after signal withdrawal; under the stem and transit amplifying cells model the process takes 30 hours. Given that both Fox *et al.* and Cinquin *et al.* report full meiotic entry in under 10 hours, this provides evidence against the existence of transit amplifying divisions in the germ line; at least, as we have modelled them here. It remains possible that amplifying divisions could occur, but they would need to be contingent on the DTC signal and terminate once signalling is disrupted. This conflicts with the idea that germ cells can divide multiple times after leaving the niche under normal, unperturbed circumstances.

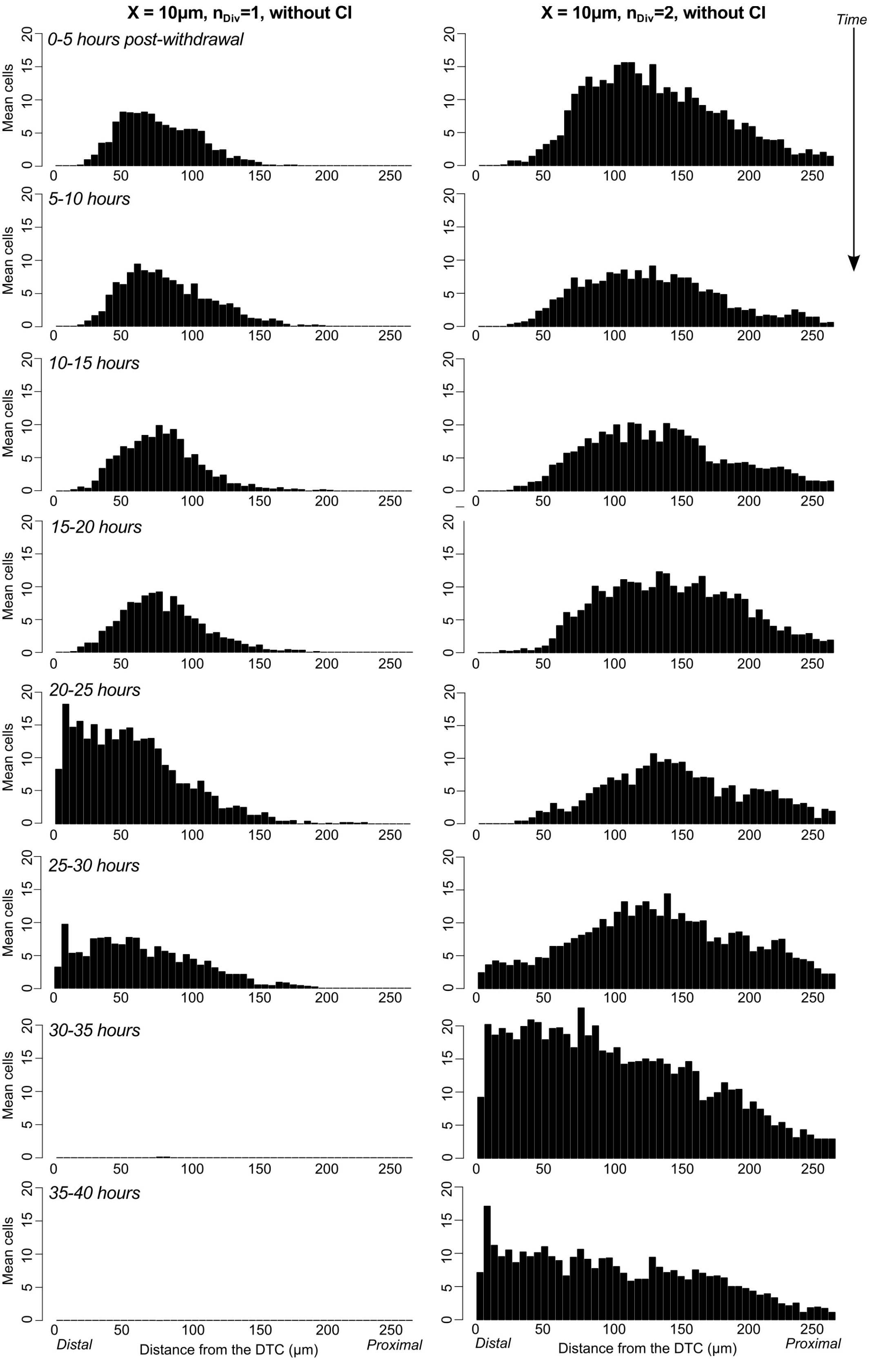


Figure 6.25: The location and timing of meiotic entry events after DTC signal withdrawal - stem and transit amplifying cells model. Results are presented as described in the Figure 6.24 legend. Here, we focus on a single well-fitting parameter set, shown in the left hand column ($X = 10\mu\text{m}$, $n_{Div} = 1$, without contact inhibition). The right hand column presents results using a larger value of n_{Div}, for comparison. Introducing transit amplifying divisions significantly increases the time taken for the whole germ line to enter meiosis. Again, the spatial distribution of meiosis events is non-uniform, with distal germ cells entering meiosis later in time. However, there is no evidence that distal cells enter meiosis simultaneously, for any value of n_{Div}.

Turning to the spatial distribution of meiosis events in Figure 6.25, distal germ cells again do not enter meiosis in the first few hours after signal withdrawal, but rather at later times, leading to a non-uniform distribution. This spatial pattern could easily be interpreted as a proximal to distal wave. However, in our simulations, the distal-most germ cell rows do *not* enter meiosis simultaneously over the course of a single hour, as described experimentally by Cinquin *et al.* [47]. While it is true that the distal germ cells in this model are measurably less mature (having undergone zero transit amplifying divisions) that does not make them a homogeneous population. Distal cells still differ in terms of their position in the cell cycle. Since germ cells must wait until they reach G1 phase before entering the pathway to meiosis, these cell cycle differences prevent synchronised meiotic entry. The existence of an immature stem cell population therefore need not imply synchronised meiosis of that population. To produce the degree of synchronisation reported by Cinquin *et al.*, the meiosis decision would either have to be decoupled from the cell cycle, or distal germ cell divisions would have to be highly synchronised, neither of which appears likely given current evidence [114, 123, 154].

Asymmetric divisions

Finally, Figure 6.26 shows similar meiosis data for the asymmetric divisions model, using two different, well-fitting parameter sets. We do not discuss these results in detail, as they exhibit similar trends to the stem and transit amplifying cells model (except that fewer meioses occur overall, since a smaller number of proliferative cells remain prior to signal withdrawal when $p < 1$).

In summary, this investigation provides evidence *against* the existence of either contact inhibition or transit amplifying divisions in the germ line. Both of these features would produce a delay in complete meiotic entry after DTC signal withdrawal; contradicting two independent sets of experimental observations [47, 81]. In Chapter 7, we will discuss in greater detail the evidence for and against these hypotheses, focussing particularly on contact inhibition. Based on these perturbation simulations though, the straightforward equivalent germ cells mechanism emerges as the best fitting model. Even when all germ cells are taken to be functionally equivalent, there remains a spatial pattern in meiotic entry events over time, which may account for the proximal to distal wave of differentiation described by Cinquin *et al.* [47]. However, none of our three models produces tightly synchronised meiotic entry of the distal-most germ cell rows. This would require decoupling the meiosis decision from the cell cycle, or assuming highly synchronised divisions. Since neither of these amendments seems plausible, and since synchronised distal meiosis was

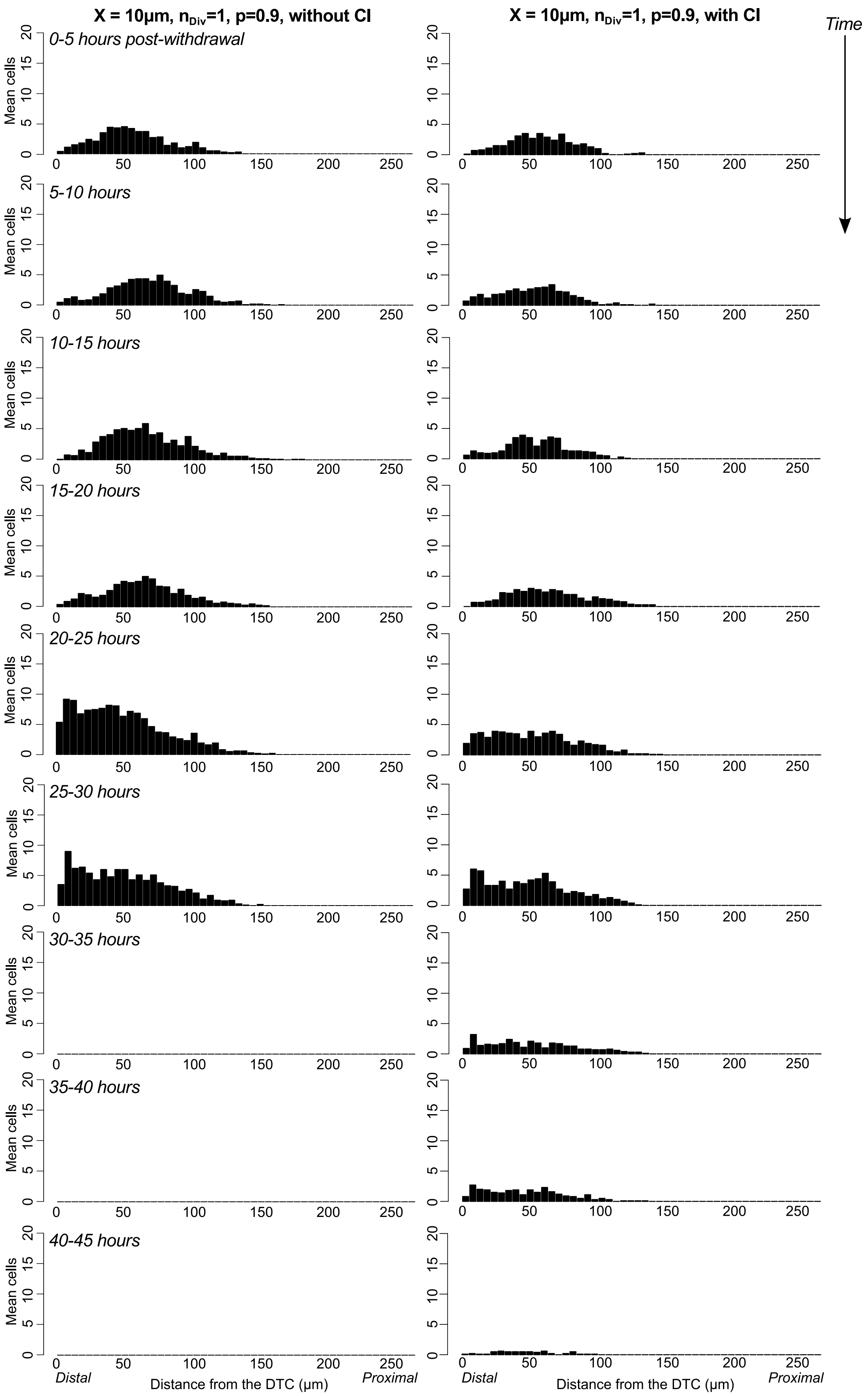


Figure 6.26: The location and timing of meiotic entry events after DTC signal withdrawal - asymmetric divisions model. Results are presented as described in the Figure 6.24 legend. Here we show data for two well-fitting parameter sets: $X = 10\mu\text{m}$, $n_{Div} = 1$, $p = 0.9$ without contact inhibition (left column); and the same parameter set with contact inhibition (right column). Similar trends are observed to the stem and transit amplifying cells model, namely: that adding transit amplifying divisions delays full meiotic entry; that the spatial distribution is initially non-uniform; and that synchronous meiosis of the distal-most cells does not occur.

not observed by Fox *et al.*, we are inclined to believe this feature was an artefact arising from the labelling technique used in [47].

6.5 The effect of ageing on proliferative cell count

A final germline behaviour that we will discuss briefly is proliferative zone ageing. Recent work by Qin and Hubbard has shown that the number of proliferative cells in the *C. elegans* germ line declines with age [197] (see Section 2.8). Intriguingly, Qin and Hubbard also found that slowing the rate of germ cell flux through the gonad protects against proliferative cell loss. For example, mutant animals without sperm lack the chemical signals to stimulate ovulation. As a result, germ cells exit the proximal gonad more slowly, and this is associated with better long term maintenance of proliferative cells. It would be satisfying if a proliferative zone maintenance mechanism could explain this relationship between age, germ cell flux and proliferative cell count.

Of our three proposed mechanisms, the asymmetric divisions model stands out as providing a potential explanation for the observed ageing effect. Proliferative cells in this model are lost over time, because each stem cell division can produce a transit amplifying daughter with some probability. Transit amplifying cells have a limited self-renewal potential, and they displace stem cells from the niche. Moreover, if we accept that a high germ cell flux is usually associated with frequent germ cell divisions, then this model could also account for the flux relationship. Since stem cell loss occurs with some probability per division, the more often cells divide, the faster proliferative cell count should decline.

To investigate whether the asymmetric divisions model could explain the data in [197], we ran a set of ageing simulations. The distal germ line was simulated as normal for 100 hours, using a parameter set that maintains the proliferative cell count ($X = 10\mu\text{m}$, $n_{Div} = 1$, $p = 0.9$, without contact inhibition). After 100 hours, p was reduced, and we observed the effect of more frequent stem-transit amplifying divisions on the germ line. In particular, the proliferative cell count was tracked for 100 hours after the shift in p, to record any resulting decline.

Figure 6.27A presents our results for a range of values of p. The solid black line on the same figure shows the experimentally observed proliferative cell count decline with age [197]. From these relatively short simulations, we can see that a stem-stem renewal probability of between 0.7 and 0.8 comes closest to replicating the observed decline. Having narrowed down the appropriate

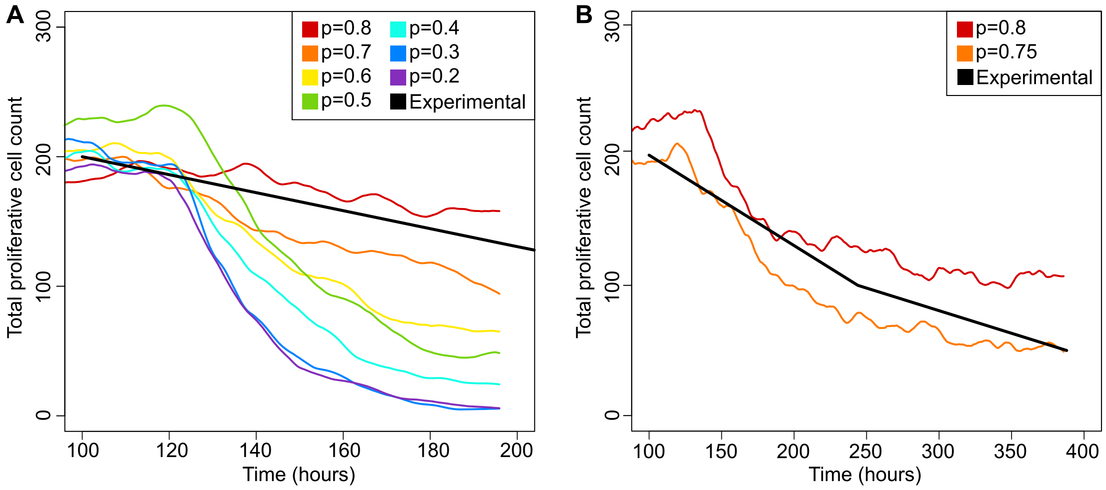


Figure 6.27: The effect on proliferative cell count of altering the stem-stem division probability. All simulations shown above used the asymmetric divisions model. The distal germ line was simulated for 100 hours using a parameter set that maintains the number of proliferative cells (see text). p was then reduced, and we show the effect of this change on the proliferative cell count (mean of 10 simulations). A) Short simulations covering a wide range of p. A value of p between 0.8 and 0.7 comes closest to reproducing the experimentally observed decline [197]. B) Longer simulations using selected values of p. The overall magnitude of the decline can be reproduced using this model, but a more detailed assessment of the fit awaits further experimental data.

range of p, a number of longer simulations were carried out, following the germ line for 100 hours + 12 days. This is the same period over which Qin and Hubbard monitored the wild-type germ line. Figure 6.27B presents these longer term results. It appears that the asymmetric divisions model may produce too sharp a decline in cell count initially. However, the simulated proliferative cell counts with $p = 0.75$ are similar to the observed counts overall. To fully assess the fit of this model would require data at more time points than are reported in [197]; experimental data that should be feasible to collect in principle.

While these results are very preliminary, they do suggest one potential mechanism to explain the observations in [197], and they quantify the imbalance between stem and transit amplifying cell production that would be required. A more detailed analysis awaits the availability of more quantitative experimental data on this phenomenon. Qin and Hubbard themselves propose a biochemical explanation for their results, involving an unknown signal produced in the proximal gonad that communicates cell flux levels to the proliferative zone. In support of a biochemical mechanism, they report that the germ cell flux effect is partially mediated by the transcription factor DAF-16/FOXO [197], and that it is reduced in *daf-16(-)* mutants. However, this accounts for only half of the observed effect, leaving open the possibility of an additional population-level mechanism of proliferative cell loss.

6.6 Summary

In this chapter, we evaluated three possible mechanisms of proliferative zone maintenance: functionally equivalent germ cells; stem and transit amplifying cells; and asymmetric divisions. These mechanisms were implemented within cell-based models, using our new cell logic modules approach to specify the germ cell program. Each model was then simulated using a range of parameter values, and the output was compared with experimentally observed germline properties. Runs were carried out both with and without germ cell contact inhibition applied. We were unable to identify a clearly superior model on the basis of these results. However, we were able to narrow down the acceptable range of parameter values. Key findings were that: 1) the asymmetric divisions model produces an unrealistic proliferative zone appearance; 2) including contact inhibition can help reproduce the observed shape of the mitotic index profile; and 3) overreliance on contact inhibition results in an unrealistic percentage of cells in G2 arrest.

We also studied the response of these three models to withdrawal of DTC signalling. In particular, we recorded the timing and position of meiotic entry events, after the DTC signal was removed. Our main finding was that introducing contact inhibition or transit amplifying divisions significantly delays full meiotic entry; a process that should be completed in less than 10 hours [47, 81]. We will discuss the implications for a contact inhibition hypothesis in Chapter 7. Briefly though, this suggests that if contact inhibition occurs *in vivo*, it must be a relatively small effect causing only transient arrests. In all of our DTC signal withdrawal simulations, the meiosis distribution was spatially non-uniform at first, producing an effect similar to the proximal to distal wave of differentiation described by Cinquin *et al.* [47]. However, we did not see the synchronised meiosis of the distal-most germ cell rows reported in the same paper; even when using a model containing an immature stem cell population.

Finally, we made a preliminary study of whether the asymmetric divisions model could explain declining proliferative cell counts with age [197]. We found that a stem-stem renewal probability of ~ 0.75 produces a decrease in proliferative cell numbers comparable to that reported in [197]. In light of our earlier results, however, the asymmetric divisions model would need to be modified to give a realistic distal germ line appearance before it would be a reasonable hypothesis; for example, by restricting transit amplifying cell production to the proximal part of the niche.

One caveat regarding the results in this Chapter is that they apply to a truncated domain. To simplify the problem and reduce run times, we modeled only the distal gonad and excluded the turn and proximal arm. As a result, germ cells experience less resistance to forward motion than in the full germ line system. It is our expectation that certain effects would be amplified in the full system. For instance, greater resistance to forward motion should cause cells to pack together more tightly in the distal arm. This would trigger more frequent contact inhibition and would amplify the associated effects: a high percentage of cells in G2 phase, and a suppressed mitotic index near the distal tip. Cell counts would be generally higher, and meiotic output would likely be lower, due to slower movement of germ cells out of the DTC signaling region. All of this would affect the precise values in our results, though it need not necessarily alter the general trends.

Chapter 7

Discussion

Cell proliferation and differentiation are fundamental processes in biology, underlying metazoan development, tissue maintenance and repair. These processes must be properly coordinated across populations of cells to give rise to normal organ development and function. Dysregulation has serious consequences; for example, uncontrolled proliferation leads to cancer. Since regulatory mechanisms can be conserved over evolutionary time scales, it has become common to research the control of proliferation and differentiation using experimentally tractable model organisms. In this thesis, we focussed on the *C. elegans* gonad and germ line, a system used to study organogenesis, tissue maintenance and ageing effects.

This chapter summarises the main conclusions of our work. We begin by briefly reviewing the contents of this thesis. We then discuss our results in two separate sections. The biological conclusions section looks at what was learned about the *C. elegans* germ line specifically. It seeks to address some of the open questions posed in Chapter 2. Meanwhile, the technical contributions section discusses work relevant to cell-based modelling more generally. It includes comments on our model and code development. Finally, we propose some ways in which our work could be extended in the near future.

7.1 Thesis summary

The first two chapters of this thesis review pre-existing work. In particular, Chapter 2 summarises key aspects of *C. elegans* germ line biology; how the germ line develops in the larva, the processes that maintain the adult tissue, and how normal germ line function can be disrupted. We also introduce several open questions in the field, including: how are germ line expansion and gonadogenesis interrelated?; how is a stable cell count maintained in young adults?; and does a population of specialised germline stem cells exist? We assert that these questions can be studied using *in silico* models, particularly if those models incorporate both germ cell movement and decision-making.

Chapter 3 reviews available tools and techniques for biological modelling. We compare different approaches to model an interacting population of cells, and make the case that an overlapping spheres model implemented using the Chaste library would be appropriate to represent the germ line. We then discuss techniques for modelling intracellular decision-making, highlighting statecharts as an appropriate formalism for representing germ cell behaviour. Our conclusions about the suitability of each technique are motivated by a number of considerations, including an appropriate level of biological detail, computational tractability, and the availability of open source libraries. Finally, we review the few existing *C. elegans* germ line models, indicating where further improvements might be made.

Chapter 4 presents our 3D *C. elegans* germ line model, which extends previous work in the field in two respects. Firstly, it integrates a statechart representation of cell behaviour into a wider mechanics simulation, allowing the interplay between cell movement and behaviour to be studied. Secondly, it covers the L3 and L4 larval stages as well as adult germ line maintenance. We discuss how a wild-type parameter set was chosen, using a combination of new experimental data, the literature, and manual searches. We also assess the model's fit to experimental data, and display the effect of small parameter variations. Out of the model development process, three main findings emerged. First, we found that proximal gonad growth during late L4 is a significant component of gonadogenesis, which does not appear to be driven by DTC migration. This is relevant to the first of the biological questions introduced at the beginning of the thesis, since proximal gonad growth could be caused by pressure generated by the developing germ line. Second, we found it necessary to introduce negative feedback from crowding on the germ cell cycle to prevent overproliferation. We proposed a regulatory mechanism akin to contact inhibition, and our collaborators have found evidence consistent with such a mechanism *in vivo*. This contributes toward answering our second biological question - how is a stable cell count maintained in young adult animals? Finally, Chapter 4 presents germ cell tracking and labelling predictions for the late larval development period, showing that germ cells largely maintain their relative positions.

Chapter 5 addresses two important technical questions arising from our modelling work. Firstly, we consider whether using a more physically realistic force law changes the outcome of simulations. We implement a force law proposed by Buske *et al.* [33], which has previously been used to model the colonic crypt (a system with similar geometry to the distal germ line). We then compare it with the simpler spring-like force law used up to this point. Our results show that, for most practical purposes, the Buske force law with its default parameters behaves like a weak spring force. Some extreme radius changes occur when cells are tightly packed, a behaviour that is perhaps best interpreted as a cell shape change. Otherwise, distal germ line simulations are not greatly affected by the change of force law. Secondly, we assess the numerical accuracy of our cell-based simulations, and seek to determine whether an alternative numerical scheme could boost accuracy or performance. We find that the forward Euler method with a moderately-sized time step is sufficient for estimating whole population properties. However, generating accurate individual cell positions requires a much suppressed time step. Switching to an RK4 or adaptive method can improve accuracy and performance, but with the caveat that care must be taken to capture accurate division timings. Finally, we propose a new *cell logic modules* approach to representing intracellular decision-making, which facilitates modelling complex behaviour on division.

Lastly, Chapter 6 returns to the third biological question raised at the beginning of this thesis: how is the proliferative zone maintained on a cellular level, and does a distinct stem cell population exist? We simulate and compare three different maintenance mechanisms; an equivalent germ cells model, a stem and transit amplifying cells model, and an asymmetric divisions model. To allow more simulations to be run in a reasonable time, we focus on the distal germ line only. We find that the asymmetric divisions model leads to an unrealistic distal germ line appearance, although it does show some potential to explain ageing effects. All three models were able to reproduce a range of wild-type germline properties, given an appropriate choice of parameters. However, in a simulation involving a perturbation (disruption of the DTC signal), the equivalent germ cells model provided a superior fit.

7.2 Biological conclusions

The first biological finding to emerge from this work is the importance of proximal arm growth as a component of gonadogenesis. While late L4 proximal arm growth has been noted before [115], the measurements made as part of our modelling effort quantify its extent and highlight its significance. During model development, we found it impossible to produce a realistically sized adult organ without incorporating centrifugal movement of the gonad turn. The question therefore is, what mechanism drives this proximal growth? Our initial hypothesis was that germ cell mitotic pressure might cause the gonad membrane to stretch. That has since been ruled out by our collaborators: preventing germ cell proliferation during late L4 does not lead to a significantly shorter adult gonad. It remains possible that oocyte growth might provide the pressure to stretch the gonad membrane. Alternatively, there may be a separate genetic program responsible for the observed growth. Regardless of the exact mechanism, this finding challenges the idea that gonadogenesis is driven solely by leader cell migration.

A major topic discussed in this thesis is whether adult germ cells could be subject to a feedback mechanism that limits their proliferation. This idea first arose because of the difficulty we had stabilising cell numbers in our *in silico* model. We therefore proposed a negative feedback mechanism from crowding on proliferation, akin to contact inhibition. This is a simple and direct hypothesis as to how overproliferation might be prevented, and is a reasonably common assumption in cell-based modelling [33, 86]. Now, having investigated the idea more thoroughly, we can weigh up the evidence for and against contact inhibition among adult germ cells.

In support of contact inhibition, we have the fact that it robustly stabilises cell numbers in our germ line model. Other changes, such as varying the cell death rate, were not able to robustly produce homeostasis. Contact inhibition could also explain why *ced-3* mutants, in which cells are not removed from the gonad by apoptosis, do not show any signs of overcrowding or an abnormal proliferative cell count. In fact, our collaborators found evidence consistent with a contact inhibition mechanism in *ced-3* animals: germ cells *are* more tightly packed, and the mitotic index *is* generally reduced, relative to wild type. Finally, when we simulated a range of proliferative zone maintenance mechanisms, we found that introducing contact inhibition gave a good fit to properties such as phase percentages and proliferative cell count, provided cell cycle arrests were not too frequent. Contact inhibition also helped reproduce the slightly lower mitotic index of the distal-most cell rows, although other plausible explanations for that effect exist, such as an inhibitory signal from the DTC.

Our main result that contradicts a contact inhibition hypothesis concerns how the germ line enters meiosis after DTC signalling is disrupted. Experimental studies show that the entire germ line becomes meiotic within 10 hours of DTC signal withdrawal [47, 81]. Meanwhile, our simulations indicate that full meiotic entry takes much longer than that when contact inhibition is present. This is because inhibited germ cells take time to come out of arrest and progress through the cell cycle to the point where they can undergo meiosis. In addition, we find that frequent germ cell arrests lead to an unrealistic number of cells in the G2 phase *in silico*. Experimental objections to contact inhibition include a labelling study by Fox *et al.*, which shows that the entire germ line can be labelled with the S phase marker EdU within an 8 hour time frame [82]. While this does not completely rule out contact inhibition, it does rule out long periods of arrest for individual cells. The existence of germline tumours also appears to be evidence against contact inhibition, although loss of responsiveness to normal anti-growth signals is recognised as a key step in cancer development [101].

It may be possible to further refine the contact inhibition hypothesis presented in this thesis. For example, instead of applying a uniform volume threshold, germ cells could enter arrest at a range of different volume thresholds, or based on their rate of volume change. At present, though, these results suggest that if adult *C. elegans* germ cells *are* subject to contact inhibition, it must cause only transient, infrequent arrests, and is likely to be only part of a wider mechanism maintaining homeostasis. These results also caution against adding contact inhibition to a cell-based model purely for the sake of achieving stable cell counts, without also exploring the wider effects of that assumption on the system.

Turning to the work in Chapter 6, we evaluated several possible mechanisms of proliferative zone maintenance, to see which provides the best fit to known germline properties. Our work on this topic is somewhat inconclusive, as all three models tested were able to fit most properties, given an appropriate choice of parameters. However, we did find that a straightforward model involving stem cell displacement produces an unrealistic germ line appearance. While gradual stem cell loss could account for some of the changes that occur during ageing, further constraints are needed to make this hypothesis plausible. For example, asymmetric divisions could be limited to the proximal end of the proliferative zone by some signalling gradient. We also find a mixed picture regarding transit amplifying divisions. On the one hand, allowing amplifying divisions means that the germ line takes longer to enter meiosis in a simulation where the DTC signal is withdrawn. This does not match the experimental evidence, as we have already discussed. However, introducing one transit amplifying division does better fit the long tail of the observed mitotic index profile. On balance, our results seem to point to an equivalent germ cells maintenance mechanism, in which some cells undergo two divisions rather than one after leaving the DTC signalling region. Indeed, Fox *et al.* do report seeing rare germ cells divide twice following DTC signal disruption [81].

Finally, we present some preliminary results on *in silico* cell tracking and clonal labelling. These results suggest that the ordering of germ cells is largely preserved during larval development, with distal germ cells at L3 contributing to the proliferative zone and proximal germ cells at L3 developing into self-sperm. There could be interesting implications for germ cell specialisation; if germ cell fate is already predetermined by position at L3, then potentially some cell specialisation could take place at an early stage in development. While 3D *in silico* cell tracking is a promising technique, our work in Chapter 5 sounds a note of caution by showing it is difficult to obtain accurate individual cell positions over the course of a long simulation. For this reason, we confine ourselves to discussing general movement trends.

7.3 Technical contributions

The main technical contribution of this thesis is the code for our 3D germ line model. While other 3D models of the germ line do now exist [99], our program remains the only one freely available under an open source license. Code and usage instructions can be be found at https://github.com/Katwell/ElegansGermline. Other researchers therefore have all the necessary details to reproduce our results and build on this work.

In the process of developing our model, we worked on several new features that may prove helpful to other users of the Chaste library studying related biological questions. For instance, we implemented a growing boundary condition that produces organ morphology based on the path of a leader cell. Since the leader cell program can be easily changed, the same basic code could be used to study DTC pathfinding mutations or gonadogenesis in the *C. elegans* male. We also introduced more realistic movement for cells of different sizes, by scaling the drag coefficient according to cell radius. Some code developed as part of this project is now incorporated into the main Chaste library; for example, an alternative way to store node neighbours that substantially improves performance, and a restructuring of the code to make it easier to experiment with different numerical schemes.

Most importantly, this work demonstrates how complex cell decision-making that goes beyond just cell cycle control can be incorporated into a Chaste simulation. Initially, we used the statecharts formalism to model germ cell decision-making. This has the advantage of being a well-established method, excellent at representing complex behaviour in a structured, hierarchical fashion. While a statecharts approach worked well initially, these lack a place to describe non-trivial behaviour on division, and the particular library we chose led to C++ code that was verbose and difficult to interpret. This prompted us to propose an alternative, *cell logic modules* approach in Chapter 5. Cell logic modules could represent an extensible way to add complex logic to multicellular simulations in future. Particular advantages are the ability to specify the starting state of a cell easily, and improved readability through representing states as enums. Cell logic modules also obey the software design principle of separation of concerns; code for different cellular functions is found in separate modules, and division logic is separated from update logic. It is possible that at some point, the statecharts formalism could be extended to include the representation of division behaviour found in cell logic modules, making it easier to illustrate patterns of inheritance in a statechart.

Finally, the other technical work in Chapter 5 may be of interest to cell-based model developers more generally. There, we carry out an in-depth comparison of two force laws, and find that adding extra physical detail to the force does not necessarily alter the outcome of simulations. We also highlight the difficulty of interpreting cell radius changes in the model proposed by Buske *et al.* [33]. Under tight packing conditions, cells can actually grow in radius to compensate for overlaps or gain adhesion energy. The second half of Chapter 5 then explores different choices of numerical scheme for cell centre simulations. While different numerical methods have been compared many times before in a non-biological context, this work is unusual in focussing on simulating a growing cell population specifically. We conclude that accurate estimates of large scale properties can be obtained quite easily by any method, whereas accurate individual cell positions are difficult to maintain beyond three or four divisions. Use of the RK4 method instead of forward Euler seems to mitigate this problem, and can even provide a moderate performance benefit. The best long term solution to maintain accurate positions seems to be adaptive time-stepping coupled with error estimation, which can also generate a speed up of up to 30 fold (for highly synchronised cells). However, this procedure requires extra care to be taken to ensure that step size changes do not affect the precise timing of division events.

7.4 Future work

Some of the future work suggested by this thesis is experimental in nature. We have proposed a number of biological hypotheses that can now be followed up. For instance, we theorise that pressure inside the gonad membrane may cause a form of stretching growth, contributing to L4 gonadogenesis. It should be possible to study mutant animals with impaired oocyte growth, to see whether they have an abnormally short proximal gonad as a result. We also propose that negative feedback might exist from cell crowding on germ cell proliferation. Future experimental work would focus on looking for a potential molecular mechanism for this. For example, the Hippo pathway is thought to be responsible for contact inhibition behaviour in mammalian cells, and certain elements of it are conserved in *C. elegans* [270]. It may be possible to disrupt this pathway, then observe the effects on the germ line. Finally, our work could motivate further quantitative studies of germ line ageing. The experimental techniques already exist to measure proliferative cell count at many more time points than in Qin and Hubbard [197], building up a better picture of age related decline. Now that it is possible to compare that trend with model predictions, there is arguably all the more reason for experimentalists to gather such data.

As for the model itself, an interesting future development would be to make the gonad membrane deformable, allowing the feedback between germ cell proliferation and gonadogenesis to be captured more faithfully. This would not be a trivial extension of the model for two reasons. Firstly, biological tissues do not usually conform to simple mathematical assumptions such as linear elasticity. Rather, tissues behave elasto-plastically, deforming and growing over time. How to model this behaviour is an area of active research. Secondly, there is a lack of data on the mechanical properties of the gonad membrane. Plausible parameter ranges could of course be estimated by comparison with other materials. Experimentally, it may also be possible to get an idea of gonad membrane elasticity by measuring its contraction when the germ line is eliminated; for instance, following ARD.

Other alterations to the model boundary condition could be attempted. Perhaps the simplest change would be to include radial variation along the length of the gonad. We assert that the overall effect of this change would be small, and indeed, most current germline models do not take radial variation into account [25, 99]. However, the necessary data is readily available, and it could allow very precise cell count predictions to be made for different sections of the organ. A more interesting challenge would be to investigate alternative ways of representing the rachis, apart from using an excluded volume. In reality, the cytoplasmic connections between germ cells and the rachis could potentially influence their movement. However, cell detachments and rearrangements are known to occur, suggesting that germ cells are not fixed in place like leaves on a stem as they travel down the organ. To compare the two extreme interpretations, one could use a 2D model, or even a cellular automaton, to represent the scenario where rachis connections keep germ cells in a strict monolayer and in fixed relative positions. A 3D model with a deformable or weakly-enforced rachis boundary could then be used for comparison.

A different potential modelling direction would be to consider the male germ line. Our model and code should be relatively easy to adapt for this purpose, since it uses a leader cell based boundary condition. The different morphology of the male gonad could therefore be reproduced simply by changing the DTC movement program. *C. elegans* males display subtly different germline properties from hermaphrodites, including a faster division rate and quicker progress through meiotic prophase [123, 176]. One motivation for modelling the male might be to try to understand the reasons behind these differences. *In silico* models have already been used to explore the evolutionary pressures that shaped hermaphrodite germline biology [42], and this is certainly a topic that is easier to investigate through modelling than in the lab.

Turning to the intracellular part of the model, the hypotheses tested in this thesis involve a fairly high-level view of germ cell behaviour. Germ cells respond to cues such as distance from the DTC and number of transit amplifying divisions, and we investigate the effect of these behaviours on whole organ properties. This was an appropriate level of detail to focus on initially, as many high-level questions about germ cell behaviour remain unanswered. In reality, though, germ cells cannot respond directly to factors such as distance from the DTC. Rather, these inputs must be sensed indirectly through molecular mechanisms. We demonstrate that our model can handle some level of molecular detail by including LAG-2/GLP-1 signalling in our intracellular statechart in Chapter 4. Moving forward, it should be possible to add further molecular detail governing the meiotic entry and sex determination decisions. That would enable our model to reflect the latest biological knowledge, and would allow testable predictions to be made about the effects of mutations/RNAi.

Finally, from a tools perspective, further work is needed to realise the potential performance and accuracy benefits of adaptive time-stepping. Simulations using a deterministic cell division model show that varying the step size can help to compute cell positions accurately and efficiently. However, when a more typical stochastic cell cycle model is used, different time steps currently lead to slightly different effective division times, which can prevent cell positions converging. To really benefit from adaptivity, ideally the next cell division time should be precomputed and then step sizes chosen so that the division coincides exactly with the end of a step. This should be possible, but would require a significant development effort. Equally, the cell logic modules code used in Chapter 6 introduces a possible design for implementing complex cell programs within Chaste, but it is not yet a complete solution. Further work is needed on archiving and data output. Future developers are now in a good position to decide whether to embark on this work.

7.5 Summary

The *C. elegans* germ line is a useful experimental system to study the regulation of proliferation and differentiation. In this thesis, we presented a number of new *in silico* models of the germ line system. Novel aspects included working in three dimensions, coverage of the L3 and L4 larval stages, and the integration of a logical model of germ cell behaviour into a wider cell mechanics simulation. Our work suggested a number of new biological hypotheses, and enabled potential mechanisms of proliferative zone maintenance to be compared. We also produced results relevant to cell-based modelling more generally, relating to choice of force law and choice of numerical method.

In this chapter, we put forward a number of ways in which our research could be extended in future. Modelling work on the *C. elegans* germ line is still in its early stages. However, particularly as quantitative studies become more common, models will be helpful in interpreting complex data and comparing subtly different hypotheses. Modelling therefore has the potential to make a strong contribution to this field, complementing experimental approaches.

References

[1] N. Abdulkader and J. Bruin. “Induction, detection and isolation of temperature-sensitive lethal and/or sterile mutants in nematodes”. In: *Journal of Nematology* 1 (1978), pp. 27–37.

[2] M. Abercrombie. “Contact inhibition and malignancy”. In: *Nature* 281.5729 (1979), pp. 259–262.

[3] M. Abercrombie and J. Heaysman. “Observations on the social behaviour of cells in tissue culture: II. monolayering of fibroblasts”. In: *Experimental Cell Research* 6.2 (1954), pp. 293–306.

[4] J. Ahrens, B. Geveci, and C. Law. “Paraview: an end-user tool for large-data visualization”. In: *The Visualization Handbook*. Ed. by C. Hansen and C. Johnson. Elsevier, 2005. Chap. 36, pp. 717–731.

[5] T. Alarcón, H. Byrne, and P. Maini. “A mathematical model of the effects of hypoxia on the cell-cycle of normal and cancer cells”. In: *Journal of Theoretical Biology* 229.3 (2004), pp. 395–411.

[6] R. Albert and H. G. Othmer. “The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster”. In: *Journal of Theoretical Biology* 223.1 (2003), pp. 1–18.

[7] D. Albertson, A. Rose, and A. Villeneuve. “Chapter 3: Chromosome organization, mitosis, and meiosis”. In: *C. elegans II 2nd edition*. Ed. by D. Riddle, T. Blumenthal, B. Meyer, and J. R. Priess. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1997.

[8] G. An, Q. Mi, J. Dutta-Moscato, and Y. Vodovotz. “Agent-based models in translational systems biology”. In: *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 1.2 (2009), pp. 159–171.

[9] *Anatomy of an adult hermaphrodite*. http://www.wormatlas.org/hermaphrodite/introduction/ IMAGES/introfig1lr.jpg. Accessed 19 Feb 2017.

[10] V. Andasari, R. Roper, M. Swat, and M. Chaplain. “Integrating intracellular dynamics using CompuCell3D and Bionetsolver: applications to multiscale modelling of cancer cell growth and invasion”. In: *PLoS One* 7.3 (2012), e33726.

[11] S. Andux and R. E. Ellis. “Apoptosis maintains oocyte quality in ageing Caenorhabditis elegans females”. In: *PLoS Genetics* 4.12 (2008), e1000295.

[12] G. Angelo and M. R. Van Gilst. “Starvation protects germline stem cells and extends reproductive longevity in C. elegans”. In: *Science* 326.5955 (2009), pp. 954–958.

[13] N. Arantes-Oliveira, J. Apfeld, A. Dillin, and C. Kenyon. “Regulation of life-span by germ-line stem cells in Caenorhabditis elegans”. In: *Science* 295.5554 (2002), pp. 502–505.

[14] S. Athanassopoulos, C. Kaklamanis, G. Kalfoutzos, and E. Papaioannou. “Cellular automata: simulations using matlab”. In: *The Sixth International Conference on Digital Society*. 2012.

[15] K. Atwell, Z. Qin, D. Gavaghan, H. Kugler, E. Hubbard, and J. Osborne. “Mechano-logical model of C. elegans germ line suggests feedback on the cell cycle”. In: *Development* 142.22 (2015), pp. 3902–3911.

[16] J. Austin and J. Kimble. “Glp-1 is required in the germ line for regulation of the decision between mitosis and meiosis in C. elegans”. In: *Cell* 51.4 (1987), pp. 589–599.

[17] W. de Back, Y. Yamaguchi, K. Yoshimura, M. Herrero, and T. Miura. “Dynamics of VEGF matrix-retention in vascular network patterning”. In: *Physical Biology* 10.6 (2013), p. 066007.

[18] W. de Back, J. Zhou, and L. Brusch. “On the role of lateral stabilization during early patterning in the pancreas”. In: *Journal of The Royal Society Interface* 10.79 (2013), p. 20120766.

[19] A. Bailly and A. Gartner. “Chapter 9. Germ cell apoptosis and DNA damage responses”. In: *Germ Cell Development in C. elegans*. Ed. by T. Schedl. New York: Springer, 2013.

[20] S. Balay, S. Abhyankar, M. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. Gropp, D. Kaushik, M. Knepley, L. McInnes, K. Rupp, B. Smith, S. Zampini, and H. Zhang. *PETSc Users Manual*. Tech. rep. ANL-95/11 - Revision 3.6. Argonne National Laboratory, 2015. URL: http://www.mcs.anl.gov/petsc.

[21] M. Barton and J. Kimble. “fog-1, a regulatory gene required for specification of spermatogenesis in the germ line of Caenorhabditis elegans”. In: *Genetics* 125.1 (1990), pp. 29–39.

[22] G. Batchelor. *An introduction to fluid dynamics*. Cambridge University Press, 2000. Chap. 4, pp. 229–240.

[23] D. Benque, S. Bourton, C. Cockerton, B. Cook, J. Fisher, S. Ishtiaq, N. Piterman, A. Taylor, and M. Vardi. “BMA: visual tool for modeling and analyzing biological networks”. In: *Computer Aided Verification*. Springer. 2012, pp. 686–692.

[24] A. Beyer, R. Eberhard, N. Piterman, M. O. Hengartner, A. Hajnal, and J. Fisher. “A dynamic physical model of cell migration, differentiation and apoptosis in Caenorhabditis elegans”. In: *Advances in Systems Biology*. Springer, 2012, pp. 211–233.

[25] A. Beyer, R. Eberhard, N. Piterman, M. O. Hengartner, A. Hajnal, and J. Fisher. “Predictive modelling of stem cell differentiation and apoptosis in C. elegans”. In: *Information Processing in Cells and Tissues*. Springer, 2012, pp. 99–104.

[26] A. Beyer, P. Thomason, X. Li, J. Scott, and J. Fisher. “Mechanistic insights into metabolic disturbance during type-2 diabetes and obesity using qualitative networks”. In: *Transactions on Computational Systems Biology XII*. Springer, 2010, pp. 146–162.

[27] D. Beysens, G. Forgacs, and J. Glazier. “Cell sorting is analogous to phase ordering in fluids”. In: *Proceedings of the National Academy of Sciences* 97.17 (2000), pp. 9467–9471.

[28] B. Biedermann, J. Wright, M. Senften, I. Kalchhauser, G. Sarathy, M.-H. Lee, and R. Ciosk. “Translational repression of cyclin E prevents precocious mitosis and embryonic gene activation during C. elegans meiosis”. In: *Developmental Cell* 17.3 (2009), pp. 355–364.

[29] R. Blelloch and J. Kimble. “Control of organ shape by a secreted metalloprotease in the nematode Caenorhabditis elegans”. In: *Nature* 399.6736 (1999), pp. 586–590.

[30] S. Boas and R. Merks. "Synergy of cell-cell repulsion and vacuolation in a computational model of lumen formation". In: *Journal of The Royal Society Interface* 11.92 (2014), p. 20131049.

[31] V. Borde, A. Goldman, and M. Lichten. "Direct coupling between meiotic DNA replication and recombination initiation". In: *Science* 290.5492 (2000), pp. 806–809.

[32] S. Brenner. "The genetics of Caenorhabditis elegans". In: *Genetics* 77.1 (1974), pp. 71–94.

[33] P. Buske, J. Galle, N. Barker, G. Aust, H. Clevers, and M. Loeffler. "A comprehensive model of the spatio-temporal stem cell and tissue organisation in the intestinal crypt". In: *PLoS Computational Biology* 7 (2011), e1001045.

[34] P. Buske, J. Przybilla, M. Loeffler, N. Sachs, T. Sato, H. Clevers, and J. Galle. "On the biomechanics of stem cell niche formation in the gut - modelling growing organoids". In: *FEBS Journal* 279.18 (2012), pp. 3475–3487.

[35] D. Byrd, K. Knobel, K. Affeldt, S. Crittenden, and J. Kimble. "A DTC niche plexus surrounds the germline stem cell pool in Caenorhabditis elegans". In: *PLoS One* 9.2 (2014), e88372.

[36] I. Carmi, J. Kopczynski, and B. Meyer. "The nuclear hormone receptor SEX-1 is an X-chromosome signal that determines nematode sex". In: *Nature* 396.6707 (1998), pp. 168–173.

[37] R. Cassada and R. Russell. "The dauer larva, a post-embryonic developmental variant of the nematode Caenorhabditis elegans". In: *Developmental Biology* 46.2 (1975), pp. 326–342.

[38] M. Chalfie. "GFP: lighting up life (Nobel Lecture)". In: *Angewandte Chemie International Edition* 48.31 (2009), pp. 5603–5611.

[39] C. Chaouiya. "Petri net modelling of biological networks". In: *Briefings in Bioinformatics* 8.4 (2007), pp. 210–219.

[40] P. Chen and R. Ellis. "TRA-1A regulates transcription of fog-3, which controls germ cell fate in C. elegans". In: *Development* 127.14 (2000), pp. 3119–3129.

[41] X. Chen and G. Brodland. "Multi-scale finite element modeling allows the mechanics of amphibian neurulation to be elucidated". In: *Physical Biology* 5.1 (2008), p. 015003.

[42] M. Chiang, A. Cinquin, A. Paz, E. Meeds, C. Price, M. Welling, and O. Cinquin. "Control of Caenorhabditis elegans germ-line stem-cell cycling speed meets requirements of design to minimize mutation accumulation". In: *BMC Biology* 13.1 (2015), p. 1.

[43] D. Chihara and J. Nance. "An E-cadherin-mediated hitchhiking mechanism for C. elegans germ cell internalization during gastrulation". In: *Development* 139.14 (2012), pp. 2547–2556.

[44] S. Christensen, V. Kodoyianni, M. Bosenberg, L. Friedman, and J. Kimble. "Lag-1, a gene required for lin-12 and glp-1 signaling in Caenorhabditis elegans, is homologous to human CBF1 and Drosophila Su(H)". In: *Development* 122.5 (1996), pp. 1373–1383.

[45] H. Christophe. *Boost Meta State Machine (MSM)*. http://www.boost.org/doc/libs/1_60_0/libs/msm/doc/HTML/index.html. [Online; accessed 17-February-2016].

[46] O. Cinquin. "Purpose and regulation of stem cells: a systems biology view from the Caenorhabditis elegans germ line". In: *The Journal of Pathology* 217.2 (2009), pp. 186–198.

[47] O. Cinquin, S. Crittenden, D. Morgan, and J. Kimble. "Progression from a stem cell-like state to early differentiation in the C. elegans germ line". In: *Proceedings of the National Academy of Sciences* 107.5 (2010), pp. 2048–2053.

[48] R. Clifford, M.-H. Lee, S. Nayak, M. Ohmachi, F. Giorgini, and T. Schedl. "FOG-2, a novel F-box containing protein, associates with the GLD-1 RNA binding protein and directs male sex determination in the C. elegans hermaphrodite germline". In: *Development* 127.24 (2000), pp. 5265–5276.

[49] S. Crittenden, D. Bernstein, J. Bachorik, B. Thompson, M. Gallegos, A. Petcherski, G. Moulder, R. Barstead, M. Wickens, and J. Kimble. "A conserved RNA-binding protein controls germline stem cells in Caenorhabditis elegans". In: *Nature* 417.6889 (2002), pp. 660–663.

[50] S. Crittenden, K. Leonhard, D. Byrd, and J. Kimble. "Cellular analyses of the mitotic region in the Caenorhabditis elegans adult germ line". In: *Molecular Biology of the Cell* 17.7 (2006), pp. 3051–3061.

[51] S. Crittenden, E. Troemel, T. Evans, and J. Kimble. "GLP-1 is localized to the mitotic region of the C. elegans germ line". In: *Development* 120.10 (1994), pp. 2901–2911.

[52] A. Csikász-Nagy, D. Battogtokh, K. Chen, B. Novák, and J. Tyson. "Analysis of a generic model of eukaryotic cell-cycle regulation". In: *Biophysical Journal* 90.12 (2006), pp. 4361–4379.

[53] C. Curtiss and J. Hirschfelder. "Integration of stiff equations". In: *Proceedings of the National Academy of Sciences* 38 (1952), pp. 235–243.

[54] G. Dahlquist. "A special stability problem for linear multistep methods". In: *BIT Numerical Mathematics* 3.1 (1963), pp. 27–43.

[55] D. Dalfó, D. Michaelson, and E. Hubbard. "Sensory regulation of the C. elegans germline through TGF-β-dependent signaling in the niche". In: *Current Biology* 22.8 (2012), pp. 712–719.

[56] W. Damm and D. Harel. *LSCs: breathing life into Message Sequence Charts*. Springer, 1999.

[57] A. Dernberg and J. Sedat. *Methods in cell biology volume 53: Nuclear structure and function*. Academic Press, 1998. Chap. 10.

[58] T. Doniach. "Activity of the sex-determining gene tra-2 is modulated to allow spermatogenesis in the C. elegans hermaphrodite". In: *Genetics* 114.1 (1986), pp. 53–76.

[59] A. Dönni. *The Boost Statechart Library*. http://www.boost.org/doc/libs/1_60_0/libs/statechart/doc/index.html. [Online; accessed 17-February-2016].

[60] D. Doupé, A. Klein, B. Simons, and P. Jones. "The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate". In: *Developmental Cell* 18.2 (2010), pp. 317–323.

[61] T. Doyle, C. Wen, and I. Greenwald. "SEL-8, a nuclear protein required for LIN-12 and GLP-1 signaling in Caenorhabditis elegans". In: *Proceedings of the National Academy of Sciences* 97.14 (2000), pp. 7877–7881.

[62] S.-J. Dunn. “Towards a computational model of the colonic crypt with a realistic, deformable geometry”. PhD thesis. University of Oxford, 2011.

[63] S.-J. Dunn, P. Appleton, S. Nelson, I. Näthke, D. Gavaghan, and J. Osborne. “A two-dimensional model of the colonic crypt accounting for the role of the basement membrane and pericryptal fibroblast sheath”. In: *PLoS Computational Biology* 8.5 (2012), e1002515.

[64] S.-J. Dunn, S. Emmott, C. Gravill, H. Kugler, C. Wintersteiger, and B. Yordanov. *Reasoning Engine for Interaction Networks (RE:IN).* http://research.microsoft.com/en-us/projects/rein/. [Online; accessed 18-February-2016].

[65] S.-J. Dunn, A. Fletcher, S. Chapman, D. Gavaghan, and J. Osborne. “Modelling the role of the basement membrane beneath a growing epithelial monolayer”. In: *Journal of Theoretical Biology* 298 (2012), pp. 82–91.

[66] S.-J. Dunn, G. Martello, B. Yordanov, S. Emmott, and A. Smith. “Defining an essential transcription factor program for naïve pluripotency”. In: *Science* 344.6188 (2014), pp. 1156–1160.

[67] S.-J. Dunn, I. Näthke, and J. Osborne. “Computational models reveal a passive mechanism for cell migration in the crypt”. In: *PLoS One* 8 (2013), e80516.

[68] S.-J. Dunn, J. Osborne, P. Appleton, and I. Näthke. “Combined changes in Wnt signaling response and contact inhibition induce altered proliferation in radiation-treated intestinal crypts”. In: *Molecular Biology of the Cell* 27.11 (2016), pp. 1863–1874.

[69] C. Eckmann, S. Crittenden, N. Suh, and J. Kimble. “GLD-3 and control of the mitosis/meiosis decision in the germline of Caenorhabditis elegans”. In: *Genetics* 168.1 (2004), pp. 147–160.

[70] B. Edgar. “From cell structure to transcription: Hippo forges a new path”. In: *Cell* 124.2 (2006), pp. 267–273.

[71] S. Efroni, D. Harel, and I. Cohen. “Toward rigorous comprehension of biological complexity: modeling, execution, and visualization of thymic T-cell maturation”. In: *Genome Research* 13.11 (2003), pp. 2485–2497.

[72] R. Ellis and J. Kimble. “The fog-3 gene and regulation of cell fate in the germ line of Caenorhabditis elegans”. In: *Genetics* 139.2 (1995), pp. 561–577.

[73] R. Farhadifar, J.-C. Röper, B. Aigouy, S. Eaton, and F. Jülicher. “The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing”. In: *Current Biology* 17.24 (2007), pp. 2095–2104.

[74] G. Figueredo, T. Joshi, J. Osborne, H. Byrne, and M. Owen. “On-lattice agent-based simulation of populations of cells within the open-source Chaste framework”. In: *Interface Focus* 3.2 (2013), p. 20120081.

[75] A. Fire, S. Xu, M. Montgomery, S. Kostas, S. Driver, and C. Mello. “Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans”. In: *Nature* 391.6669 (1998), pp. 806–811.

[76] J. Fisher and T. Henzinger. “Executable cell biology”. In: *Nature Biotechnology* 25.11 (2007), pp. 1239–1249.

[77] J. Fisher, N. Piterman, A. Hajnal, and T. A. Henzinger. “Predictive modeling of signaling crosstalk during C. elegans vulval development”. In: *PLoS Computational Biology* 3.5 (2007), e92.

[78] J. Fisher, N. Piterman, E. Hubbard, M. Stern, and D. Harel. "Computational insights into Caenorhabditis elegans vulval development". In: *Proceedings of the National Academy of Sciences* 102.6 (2005), pp. 1951–1956.

[79] A. Fletcher, J. Osborne, P. Maini, and D. Gavaghan. "Implementing vertex dynamics models of cell populations in biology within a consistent computational framework". In: *Progress in Biophysics and Molecular Biology* 113.2 (2013), pp. 299–326.

[80] A. Fletcher, M. Osterfield, R. Baker, and S. Shvartsman. "Vertex models of epithelial morphogenesis". In: *Biophysical Journal* 106.11 (2014), pp. 2291–2304.

[81] P. Fox and T. Schedl. "Analysis of germline stem cell differentiation following loss of GLP-1 notch activity in Caenorhabditis elegans". In: *Genetics* 201.1 (2015), pp. 167–184.

[82] P. Fox, V. Vought, M. Hanazawa, M.-H. Lee, E. Maine, and T. Schedl. "Cyclin E and CDK-2 regulate proliferative cell fate and cell cycle progression in the C. elegans germline". In: *Development* 138.11 (2011), pp. 2223–2234.

[83] R. Francis, M. Barton, J. Kimble, and T. Schedl. "Gld-1, a tumor suppressor gene required for oocyte development in Caenorhabditis elegans". In: *Genetics* 139.2 (1995), p. 579.

[84] T. Frisch and O. Thoumine. "Predicting the kinetics of cell spreading". In: *Journal of Biomechanics* 35.8 (2002), pp. 1137–1141.

[85] M. Fukuyama, A. Rougvie, and J. Rothman. "C. elegans DAF-18/PTEN mediates nutrient-dependent arrest of cell cycle and growth in the germline". In: *Current Biology* 16.8 (2006), pp. 773–779.

[86] J. Galle, M. Loeffler, and D. Drasdo. "Modeling the effect of deregulated proliferation and apoptosis on the growth dynamics of epithelial cell populations in vitro". In: *Biophysical Journal* 88.1 (2005), pp. 62–75.

[87] R. Gao, W. Hu, and T.-J. Tarn. "The application of finite state machine in modeling and control of gene mutation process". In: *IEEE Transactions on NanoBioscience* 12.4 (2013), pp. 265–274.

[88] D. Garigan, A.-L. Hsu, A. Fraser, R. Kamath, J. Ahringer, and C. Kenyon. "Genetic analysis of tissue aging in Caenorhabditis elegans: a role for heat-shock factor and bacterial proliferation". In: *Genetics* 161.3 (2002), pp. 1101–1112.

[89] A. Gartner, S. Milstein, S. Ahmed, J. Hodgkin, and M. Hengartner. "A conserved checkpoint pathway mediates DNA damage-induced apoptosis and cell cycle arrest in C. elegans". In: *Molecular Cell* 5.3 (2000), pp. 435–443.

[90] A. Gerhold, J. Ryan, J.-N. Vallée-Trudeau, J. Dorn, J.-C. Labbé, and P. Maddox. "Investigating the regulation of stem and progenitor cell mitotic progression by in situ imaging". In: *Current Biology* 25.9 (2015), pp. 1123–1134.

[91] A. Gill. *Introduction to the theory of finite-state machines*. McGraw-Hill, 1962.

[92] J. Grabowski, P. Graubmann, and E. Rudolph. "The standardization of Message Sequence Charts". In: *Software Engineering Standards Symposium, Proceedings 1993*. IEEE. 1993, pp. 48–63.

[93] F. Graner and J. Glazier. "Simulation of biological cell sorting using a two-dimensional extended Potts model". In: *Physical Review Letters* 69.13 (1992), p. 2013.

[94] J. Guck, R. Ananthakrishnan, H. Mahmood, T. Moon, C. Cunningham, and J. Käs. "The optical stretcher: a novel laser tool to micromanipulate cells". In: *Biophysical Journal* 81.2 (2001), pp. 767–784.

[95] T. Gumienny, E. Lambie, E. Hartwieg, H. R. Horvitz, and M. Hengartner. "Genetic control of programmed cell death in the Caenorhabditis elegans hermaphrodite germ line". In: *Development* 126.5 (1999), pp. 1011–1022.

[96] W. van Gunsteren and H. Berendsen. "Algorithms for Brownian dynamics". In: *Molecular Physics* 45.3 (1982), pp. 637–647.

[97] K. Hackl and M. Goodarzi. *An introduction to linear continuum mechanics*. 2010.

[98] E. Haier, S. Norsett, and G. Wanner. *Solving ordinary differential equations I, nonstiff problems*. Springer, 1993. Chap. II.5.

[99] B. Hall, N. Piterman, A. Hajnal, and J. Fisher. "Emergent stem cell homeostasis in the C. elegans germline is revealed by hybrid modeling". In: *Biophysical Journal* 109.2 (2015), pp. 428–438.

[100] B. Hamaoka, C. Dann, B. Geisbrecht, and D. Leahy. "Crystal structure of Caenorhabditis elegans HER-1 and characterization of the interaction between HER-1 and TRA-2A". In: *Proceedings of the National Academy of Sciences* 101.32 (2004), pp. 11673–11678.

[101] D. Hanahan and R. Weinberg. "The hallmarks of cancer". In: *Cell* 100.1 (2000), pp. 57–70.

[102] D. Hansen, E. Hubbard, and T. Schedl. "Multi-pathway control of the proliferation versus meiotic development decision in the Caenorhabditis elegans germline". In: *Developmental Biology* 268.2 (2004), pp. 342–357.

[103] D. Hansen and T. Schedl. "Chapter 4. Stem cell proliferation versus meiotic fate decision in C. elegans". In: *Germ Cell Development in C. elegans*. Ed. by T. Schedl. New York: Springer, 2013.

[104] D. Hansen, L. Wilson-Berry, T. Dang, and T. Schedl. "Control of the proliferation versus meiotic development decision in the C. elegans germline through regulation of GLD-1 protein accumulation". In: *Development* 131.1 (2004), pp. 93–104.

[105] D. Harel. "Statecharts: A visual formalism for complex systems". In: *Science of Computer Programming* 8.3 (1987), pp. 231–274.

[106] D. Harel and H. Kugler. "The rhapsody semantics of statecharts (or, on the executable core of the UML)". In: *Integration of Software Specification Techniques for Applications in Engineering*. Springer, 2004, pp. 325–354.

[107] D. Harel, H. Kugler, and A. Pnueli. "Synthesis revisited: generating statechart models from scenario-based requirements". In: *Formal Methods in Software and Systems Modeling*. Springer, 2005, pp. 309–324.

[108] S. Henderson, D. Gao, E. Lambie, and J. Kimble. "Lag-2 may encode a signalling ligand for the GLP-1 and LIN-12 receptors of C. elegans". In: *Development* 120.10 (1994), pp. 2913–2924.

[109] D. Hirsh, D. Oppenheim, and M. Klass. "Development of the reproductive system of Caenorhabditis elegans". In: *Developmental Biology* 49.1 (1976), pp. 200–219.

[110] A. Hodgkin and A. Huxley. "A quantitative description of membrane current and its application to conduction and excitation in nerve". In: *The Journal of Physiology* 117.4 (1952), pp. 500–544.

[111] J. Hodgkin. “Sex determination in the nematode C. elegans: analysis of tra-3 suppressors and characterization of fem genes”. In: *Genetics* 114.1 (1986), pp. 15–52.

[112] J. Hodgkin, H. Horvitz, and S. Brenner. “Nondisjunction mutants of the nematode Caenorhabditis elegans”. In: *Genetics* 91.1 (1979), pp. 67–94.

[113] J. Hodgkin and S. Brenner. “Mutations causing transformation of sexual phenotype in the nematode Caenorhabditis elegans”. In: *Genetics* 86.2 (1977), pp. 275–287.

[114] E. Hubbard. “Caenorhabditis elegans germ line: a model for stem cell biology”. In: *Developmental Dynamics* 236.12 (2007), pp. 3343–3357.

[115] E. Hubbard and R. Stupay. *GermMOVIE 1: Animation of hermaphrodite gonadogenesis*. http://www.wormatlas.org/hermaphrodite/germ%20line/Germframeset.html. [Online; accessed 10-March-2016].

[116] M. Hucka, A. Finney, H. Sauro, H. Bolouri, J. Doyle, H. Kitano, A. Arkin, B. Bornstein, D. Bray, A. Cornish-Bowden, A. Cuellar, S. Dronov, E. Gilles, M. Ginkel, V. Gor, I. Goryanin, W. Hedley, T. Hodgman, J. Hofmeyr, P. Hunter, N. Juty, J. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. Loew, D. Lucio, P. Mendes, E. Minch, E. Mjolsness, Y. Nakayama, M. Nelson, P. Nielsen, T. Sakurada, J. Schaff, B. Shapiro, T. Shimizu, H. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang. “The Systems Biology Markup Language (SBML): a medium for representation and exchange of biochemical network models”. In: *Bioinformatics* 19.4 (2003), pp. 524–531.

[117] S. Hughes, K. Evason, C. Xiong, and K. Kornfeld. “Genetic and pharmacological factors that influence reproductive aging in nematodes”. In: *PLoS Genetics* 3.2 (2007), e25.

[118] IBM. *Rational Rhapsody business case studies*. http://www-03.ibm.com/software/businesscasestudies/us/en/corp?search=true& text=&sw=U299676P31954H43&dr=. [Online; accessed 17-February-2016].

[119] IBM. *The Rational Rhapsody product family*. http://www-03.ibm.com/software/products/en/ratirhapfami. [Online; accessed 17-February-2016].

[120] A. Iserles. *Numerical Analysis lecture notes*. http://www.damtp.cam.ac.uk/user/na/ PartIB/Lect10.pdf. [Online; accessed 3-December-2015].

[121] J. J Roa, G. Oncins, J. Diaz, F. Sanz, and M. Segarra. “Calculation of Young’s modulus value by means of AFM”. In: *Recent Patents on Nanotechnology* 5.1 (2011), pp. 27–36.

[122] E. Jan, C. Motzny, L. Graves, and E. Goodwin. “The STAR protein, GLD-1, is a translational regulator of sexual identity in Caenorhabditis elegans”. In: *The EMBO Journal* 18.1 (1999), pp. 258–269.

[123] A. Jaramillo-Lambert, M. Ellefson, A. Villeneuve, and J. Engebrecht. “Differential timing of S phases, X chromosome replication, and meiotic prophase in the C. elegans germ line”. In: *Developmental Biology* 308.1 (2007), pp. 206–221.

[124] A. Jaramillo-Lambert, Y. Harigaya, J. Vitt, A. Villeneuve, and J. Engebrecht. “Meiotic errors activate checkpoints that improve gamete quality without triggering apoptosis in male germ cells”. In: *Current Biology* 20.23 (2010), pp. 2078–2089.

[125] J. Jeong, J. Verheyden, and J. Kimble. “Cyclin E and CDK-2 control GLD-1, the mitosis/meiosis decision, and germline stem cells in Caenorhabditis elegans”. In: *PLoS Genetics* 7.3 (2011), e1001348–e1001348.

[126] J. Jeong, J. Snay, and M. Ataai. “A mathematical model for examining growth and sporulation processes of Bacillus subtilis”. In: *Biotechnology and Bioengineering* 35.2 (1990), pp. 160–184.

[127] S.-W. Jin, N. Arno, A. Cohen, A. Shah, Q. Xu, N. Chen, and R. Ellis. “In Caenorhabditis elegans, the RNA-binding domains of the cytoplasmic polyadenylation element binding protein FOG-1 are needed to regulate germ cell fates”. In: *Genetics* 159.4 (2001), pp. 1617–1630.

[128] K. Johnson, K. Kendall, and A. Roberts. “Surface energy and the contact of elastic solids”. In: *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*. Vol. 324. 1558. The Royal Society. 1971, pp. 301–313.

[129] T. Johnson, D. Mitchell, S. Kline, R. Kemal, and J. Foy. “Arresting development arrests ageing in the nematode Caenorhabditis elegans”. In: *Mechanisms of Ageing and Development* 28.1 (1984), pp. 23–40.

[130] A.-C. Jungkamp, M. Stoeckius, D. Mecenas, D. Grün, G. Mastrobuoni, S. Kempa, and N. Rajewsky. “In vivo and transcriptome-wide identification of RNA binding protein target sites”. In: *Molecular Cell* 44.5 (2011), pp. 828–840.

[131] L. Kadyk and J. Kimble. “Genetic regulation of entry into meiosis in Caenorhabditis elegans”. In: *Development* 125.10 (1998), pp. 1803–1813.

[132] N. Kam, D. Harel, H. Kugler, R. Marelly, A. Pnueli, E. Hubbard, and M. Stern. “Formal modelling of C. elegans development, a scenario-based approach”. In: *Modelling in Molecular Biology*. Springer, 2004, pp. 151–173.

[133] A. Kamkin and I. Kiseleva. *Mechanosensitivity in cells and tissues*. Springer, 2005.

[134] S. Kauffman. “Metabolic stability and epigenesis in randomly constructed genetic nets”. In: *Journal of Theoretical Biology* 22.3 (1969), pp. 437–467.

[135] T. Kawano, H. Zheng, D. Merz, Y. Kohara, K. Tamai, K. Nishiwaki, and J. Culotti. “C. elegans mig-6 encodes papilin isoforms that affect distinct aspects of DTC migration, and interacts genetically with mig-17 and collagen IV”. In: *Development* 136.9 (2009), pp. 1433–1442.

[136] E. Keller and L. Segel. “Traveling bands of chemotactic bacteria: a theoretical analysis”. In: *Journal of Theoretical Biology* 30.2 (1971), pp. 235–248.

[137] C. Kenyon, J. Chang, E. Gensch, A. Rudner, and R. Tabtiang. “A C. elegans mutant that lives twice as long as wild type”. In: *Nature* 366.6454 (1993), pp. 461–464.

[138] A. Kershner and J. Kimble. “Genome-wide analysis of mRNA targets for Caenorhabditis elegans FBF, a conserved stem cell regulator”. In: *Proceedings of the National Academy of Sciences* 107.8 (2010), pp. 3936–3941.

[139] D. Killian and E. Hubbard. “Caenorhabditis elegans germline patterning requires coordinated development of the somatic gonadal sheath and the germ line”. In: *Developmental Biology* 279.2 (2005), pp. 322–335.

[140] Y. Kim, M. Stolarska, and H. Othmer. “A hybrid model for tumor spheroid growth in vitro I: theoretical development and early results”. In: *Mathematical Models and Methods in Applied Sciences* 17.supp01 (2007), pp. 1773–1798.

[141] J. Kimble and D. Hirsh. “The postembryonic cell lineages of the hermaphrodite and male gonads in Caenorhabditis elegans”. In: *Developmental Biology* 70.2 (1979), pp. 396–417.

[142] J. Kimble and W. Sharrock. “Tissue-specific synthesis of yolk proteins in Caenorhabditis elegans”. In: *Developmental Biology* 96.1 (1983), pp. 189–196.

[143] J. Kimble and S. Ward. “Germ-line development and fertilization”. In: *Cold Spring Harbor Monograph Archive* 17 (1988), pp. 191–213.

[144] J. Kimble and J. White. “On the control of germ cell development in Caenorhabditis elegans”. In: *Developmental Biology* 81.2 (1981), pp. 208–219.

[145] E. Kipreos, L. Lander, J. Wing, W. He, and E. Hedgecock. “cul-1 is required for cell cycle exit in C. elegans and identifies a novel gene family”. In: *Cell* 85.6 (1996), pp. 829–839.

[146] M. Klass. “Aging in the nematode Caenorhabditis elegans: major biological and environmental factors influencing life span”. In: *Mechanisms of Ageing and Development* 6 (1977), pp. 413–429.

[147] D. Korta, S. Tuck, and E. Hubbard. “S6K links cell fate, cell cycle and nutrient response in C. elegans germline stem/progenitor cells”. In: *Development* 139.5 (2012), pp. 859–870.

[148] H. Kugler, R. Marelly, L. Appleby, J. Fisher, A. Pnueli, D. Harel, M. Stern, and E. Hubbard. “A scenario-based approach to modeling development: a prototype model of C. elegans vulval fate specification”. In: *Developmental Biology* 323.1 (2008), pp. 1–5.

[149] J. Lewis and J. Fleming. “Basic culture methods”. In: *Caenorhabditis elegans: modern biological analysis of an organism* 48 (1995), pp. 25–26.

[150] C. Li, M. Nagasaki, C. H. Koh, and S. Miyano. “Online model checking approach based parameter estimation to a neuronal fate decision simulation model in Caenorhabditis elegans with hybrid functional Petri net with extension”. In: *Molecular Biosystems* 7.5 (2011), pp. 1576–1592.

[151] D. Y. Lui and M. P. Colaiácovo. “Chapter 6. Meiotic development in Caenorhabditis elegans”. In: *Germ Cell Development in C. elegans*. Ed. by T. Schedl. New York: Springer, 2013.

[152] S. Luo, G. Kleemann, J. Ashraf, W. Shaw, and C. Murphy. “TGF-β and insulin signaling regulate reproductive aging via oocyte and germline quality maintenance”. In: *Cell* 143.2 (2010), pp. 299–312.

[153] D. Machado, R. Costa, M. Rocha, E. Ferreira, B. Tidor, and I. Rocha. “Modeling formalisms in systems biology”. In: *AMB Express* 1.1 (2011), pp. 1–14.

[154] J. Maciejowski, N. Ugel, B. Mishra, M. Isopi, and E. Hubbard. “Quantitative analysis of germline mitosis in adult C. elegans”. In: *Developmental Biology* 292.1 (2006), pp. 142–151.

[155] A. Maddox, B. Habermann, A. Desai, and K. Oegema. “Distinct roles for two C. elegans anillins in the gonad and early embryo”. In: *Development* 132.12 (2005), pp. 2837–2848.

[156] R. Mahaffy, C. Shih, F. MacKintosh, and J. Käs. "Scanning probe-based frequency-dependent microrheology of polymer gels and biological cells". In: *Physical Review Letters* 85.4 (2000), p. 880.

[157] T. Maiwald and J. Timmer. "Dynamical modeling and multi-experiment fitting with PottersWheel". In: *Bioinformatics* 24.18 (2008), pp. 2037–2043.

[158] A. Maniotis, C. Chen, and D. Ingber. "Demonstration of mechanical connections between integrins, cytoskeletal filaments, and nucleoplasm that stabilize nuclear structure". In: *Proceedings of the National Academy of Sciences* 94.3 (1997), pp. 849–854.

[159] H. Matsuno, A. Doi, M. Nagasaki, and S. Miyano. "Hybrid Petri net representation of gene regulatory network". In: *Pacific Symposium on Biocomputing 2000*. World Scientific. 2000, pp. 341–352.

[160] H. Mattingly, J. Chen, S. Arur, and S. Shvartsman. "A transport model for estimating the time course of ERK activation in the C. elegans germline". In: *Biophysical Journal* 109.11 (2015), pp. 2436–2445.

[161] H. Matuschek. *Dopri853Stepper; a part of the code for the Intrinsic Noise Analyzer project*. https://code.google.com/p/intrinsic-noise-analyzer/source/browse/lib/ode/dopri853.hh. [Online; accessed 29-December-2015].

[162] J. McCarter, B. Bartlett, T. Dang, and T. Schedl. "On the control of oocyte meiotic maturation and ovulation in Caenorhabditis elegans". In: *Developmental Biology* 205.1 (1999), pp. 111–128.

[163] J. McCarter, B. Bartlett, T. Dang, and T. Schedl. "Soma-germ cell interactions in Caenorhabditis elegans: multiple events of hermaphrodite germline development require the somatic sheath and spermathecal lineages". In: *Developmental Biology* 181.2 (1997), pp. 121–143.

[164] M. McGovern, R. Voutev, J. Maciejowski, A. K. Corsi, and E. Hubbard. "A latent niche mechanism for tumor initiation". In: *Proceedings of the National Academy of Sciences* 106.28 (2009), pp. 11617–11622.

[165] A. Mehra, J. Gaudet, L. Heck, P. Kuwabara, and A. Spence. "Negative regulation of male development in Caenorhabditis elegans by a protein-protein interaction between TRA-2A and FEM-3". In: *Genes & Development* 13.11 (1999), pp. 1453–1463.

[166] F. Meineke, C. Potten, and M. Loeffler. "Cell migration and organization in the intestinal crypt using a lattice-free model". In: *Cell Proliferation* 34.4 (2001), pp. 253–266.

[167] R. Merks, M. Guravage, D. Inzé, and G. Beemster. "VirtualLeaf: an open-source framework for cell-based modeling of plant tissue growth and development". In: *Plant Physiology* 155.2 (2011), pp. 656–666.

[168] D. Michaelson, D. Korta, Y. Capua, and E. Hubbard. "Insulin signaling promotes germline proliferation in C. elegans". In: *Development* 137.4 (2010), pp. 671–680.

[169] M. Miller, V. Nguyen, M.-H. Lee, M. Kosinski, T. Schedl, R. Caprioli, and D. Greenstein. "A sperm cytoskeletal protein that signals oocyte meiotic maturation and ovulation". In: *Science* 291.5511 (2001), pp. 2144–2147.

[170] M. Miller, P. Ruest, M. Kosinski, S. Hanks, and D. Greenstein. "An Eph receptor sperm sensing control mechanism for oocyte meiotic maturation in Caenorhabditis elegans". In: *Genes & Development* 17.2 (2003), pp. 187–200.

[171] G. Mirams, C. Arthurs, M. Bernabeu, R. Bordas, J. Cooper, A. Corrias, Y. Davit, S.-J. Dunn, A. Fletcher, D. Harvey, M. Marsh, J. Osborne, P. Pathmanathan, J. Pitt-Francis, J. Southern, N. Zemzemi, and D. Gavaghan. "Chaste: an open source C++ library for computational physiology and biology". In: *PLoS Computational Biology* 9 (2013), e1002970.

[172] S. Miyano. "Boundary formation by Notch signalling in Drosophila multicellular systems: experimental observations and gene network modelling by Genomic Object Net". In: *Pacific Symposium on Biocomputing 2003*. World Scientific. 2002, p. 152.

[173] Modeliosoft. *Modelio: the open source modeling environment*. https://www.modelio.org/. [Online; accessed 17-February-2016].

[174] D. Morel, R. Marcelpoil, and G. Brugal. "A proliferation control network model: the simulation of two-dimensional epithelial homeostasis". In: *Acta Biotheoretica* 49.4 (2001), pp. 219–234.

[175] C. Morgan, M.-H. Lee, and J. Kimble. "Chemical reprogramming of Caenorhabditis elegans germ cell fate". In: *Nature Chemical Biology* 6.2 (2010), pp. 102–104.

[176] D. Morgan, S. Crittenden, and J. Kimble. "The C. elegans adult male germline: stem cells and sexual dimorphism". In: *Developmental Biology* 346.2 (2010), pp. 204–214.

[177] S. van Mourik, K. Kaufmann, A. van Dijk, G. Angenent, R. Merks, and J. Molenaar. "Simulation of organ patterning on the floral meristem using a polar auxin transport model". In: *PLoS One* 7.1 (2012), e28762.

[178] T. Murata. "Petri nets: properties, analysis and applications". In: *Proceedings of the IEEE* 77.4 (1989), pp. 541–580.

[179] C. Müssel, M. Hopfensitz, and H. Kestler. "BoolNet - an R package for generation, reconstruction and analysis of Boolean networks". In: *Bioinformatics* 26.10 (2010), pp. 1378–1380.

[180] S. Nadarajan, J. Govindan, M. McGovern, E. Hubbard, and D. Greenstein. "MSP and GLP-1/Notch signaling coordinately regulate actomyosin-dependent cytoplasmic streaming and oocyte growth in C. elegans". In: *Development* 136.13 (2009), pp. 2223–2234.

[181] T. Nagai and H. Honda. "A dynamic cell model for the formation of epithelial tissues". In: *Philosophical Magazine B* 81.7 (2001), pp. 699–719.

[182] M. Nagasaki, A. Doi, H. Matsuno, and S. Miyano. "Genomic Object Net: I. A platform for modelling and simulating biopathways". In: *Applied Bioinformatics* 2.3 (2002), pp. 181–184.

[183] P. Narbonne and R. Roy. "Inhibition of germline proliferation during C. elegans dauer development requires PTEN, LKB1 and AMPK signalling". In: *Development* 133.4 (2006), pp. 611–619.

[184] M. Nicoll, C. C. Akerib, and B. J. Meyer. "X chromosome-counting mechanisms that determine nematode sex". In: *Nature* 388.6638 (1997), pp. 200–204.

[185] S. Nusser-Stein, A. Beyer, I. Rimann, M. Adamczyk, N. Piterman, A. Hajnal, and J. Fisher. “Cell-cycle regulation of NOTCH signaling during C. elegans vulval development”. In: *Molecular Systems Biology* 8.1 (2012), p. 618.

[186] Obeo. *UML Designer: graphical tooling to edit and visualize UML models*. http://www.umldesigner.org/. [Online; accessed 17-February-2016].

[187] J. Osborne. “Numerical and computational methods for simulating multiphase models of tissue growth”. PhD thesis. University of Oxford, 2009.

[188] H. Othmer, S. Dunbar, and W. Alt. “Models of dispersal in biological systems”. In: *Journal of Mathematical Biology* 26.3 (1988), pp. 263–298.

[189] M. Otori, T. Karashima, and M. Yamamoto. “The Caenorhabditis elegans homologue of deleted in azoospermia is involved in the sperm/oocyte switch”. In: *Molecular Biology of the Cell* 17.7 (2006), pp. 3147–3155.

[190] N. Paoletti, B. Yordanov, Y. Hamadi, C. Wintersteiger, and H. Kugler. “Analyzing and synthesizing genomic logic functions”. In: *Computer Aided Verification*. Springer. 2014, pp. 343–357.

[191] P. Pasierbek, M. Jantsch, M. Melcher, A. Schleiffer, D. Schweizer, and J. Loidl. “A Caenorhabditis elegans cohesion protein with functions in meiotic chromosome pairing and disjunction”. In: *Genes & Development* 15.11 (2001), pp. 1349–1360.

[192] P. Pathmanathan, J. Cooper, A. Fletcher, G. Mirams, P. Murray, J. Osborne, J. Pitt-Francis, A. Walter, and S. Chapman. “A computational study of discrete mechanical tissue models”. In: *Physical Biology* 6.3 (2009), p. 036001.

[193] I. Peter, E. Faure, and E. Davidson. “Predictive computation of genomic logic processing functions in embryonic development”. In: *Proceedings of the National Academy of Sciences* 109.41 (2012), pp. 16434–16442.

[194] V. Popov. *Contact mechanics and friction*. Springer-Verlag Berlin, 2010. Chap. 5.

[195] R. Potts. “Some generalized order-disorder transformations”. In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 48. 01. Cambridge University Press. 1952, pp. 106–109.

[196] W. Press, A. Saul, W. Vetterling, and B. Flannery. *Numerical Recipes in Fortran 77: The Art of Scientific Computing*. Cambridge University Press, 1992. Chap. 16.

[197] Z. Qin and E. Hubbard. “Non-autonomous DAF-16/FOXO activity antagonizes age-related loss of C. elegans germline stem/progenitor cells”. In: *Nature Communications* 6 (2015).

[198] A. Rajeev and A. Thomas. “Reactive modules”. In: *Eleventh Annual IEEE Symposium on Logic in Computer Science*. IEEE. 1996, pp. 207–218.

[199] M. Rauzi, P. Verant, T. Lecuit, and P.-F. Lenne. “Nature and anisotropy of cortical forces orienting Drosophila tissue morphogenesis”. In: *Nature Cell Biology* 10.12 (2008), pp. 1401–1410.

[200] V. Reddy, M. Liebman, and M. Mavrovouniotis. “Qualitative analysis of biochemical reaction systems”. In: *Computers in Biology and Medicine* 26.1 (1996), pp. 9–24.

[201] M. Robertson-Tessi, R. Gillies, R. Gatenby, and A. Anderson. “Impact of metabolic heterogeneity on tumor growth, invasion, and treatment outcomes”. In: *Cancer Research* 75.8 (2015), pp. 1567–1579.

[202] T. Roose, S. Chapman, and P. Maini. “Mathematical models of avascular tumor growth”. In: *SIAM Review* 49.2 (2007), pp. 179–208.

[203] A. Rose and D. Baillie. “The effect of temperature and parental age on recombination and nondisjunction in Caenorhabditis elegans”. In: *Genetics* 92.2 (1979), pp. 409–418.

[204] T. Rudge, P. Steiner, A. Phillips, and J. Haseloff. “Computational modeling of synthetic microbial biofilms”. In: *ACS Synthetic Biology* 1.8 (2012), pp. 345–352.

[205] A. Saadatpour, R.-S. Wang, A. Liao, X. Liu, T. Loughran, I. Albert, and R. Albert. “Dynamical and structural analysis of a T cell survival network identifies novel candidate therapeutic targets for large granular lymphocyte leukemia”. In: *PLoS Computational Biology* 7.11 (2011), e1002267.

[206] J. Saez-Rodriguez, L. Alexopoulos, J. Epperlein, R. Samaga, D. Lauffenburger, S. Klamt, and P. Sorger. “Discrete logic modelling as a means to link protein signalling networks with functional analysis of mammalian signal transduction”. In: *Molecular Systems Biology* 5.1 (2009), p. 331.

[207] B. Schäling. *The Boost C++ libraries*. Boris Schäling, 2011.

[208] G. Schaller and M. Meyer-Hermann. “Multicellular tumor spheroid in an off-lattice Voronoi-Delaunay cell model”. In: *Physical Review E* 71.5 (2005), p. 051910.

[209] M. Schaub, T. Henzinger, and J. Fisher. “Qualitative networks: a symbolic approach to analyze biological signaling networks”. In: *BMC Systems Biology* 1.1 (2007), p. 4.

[210] P. Schedin, P. Jonas, and W. Wood. “Function of the her-1 gene is required for maintenance of male differentiation in adult tissues of C. elegans”. In: *Developmental Genetics* 15.3 (1994), pp. 231–239.

[211] T. Schedl. “Chapter 10: developmental genetics of the germ line”. In: *C. elegans II 2nd edition*. Ed. by D. Riddle, T. Blumenthal, B. Meyer, and J. Priess. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press, 1997.

[212] T. olde Scheper, D. Klinkenberg, C. Pennartz, and J. Van Pelt. “A mathematical model for the intracellular circadian rhythm generator”. In: *The Journal of Neuroscience* 19.1 (1999), pp. 40–47.

[213] R. Schlatter, K. Schmich, I. Avalos Vizcarra, P. Scheurich, T. Sauter, C. Borner, M. Ederer, I. Merfort, and O. Sawodny. “ON/OFF and beyond–a Boolean model of apoptosis”. In: *PLoS Computational Biology* 5.12 (2009), e1000595.

[214] C. Schneider, W. Rasband, and K. Eliceiri. “NIH Image to ImageJ: 25 years of image analysis”. In: *Nature Methods* 9.7 (2012), pp. 671–675.

[215] B. Schumacher, M. Hanazawa, M.-H. Lee, S. Nayak, K. Volkmann, R. Hofmann, M. Hengartner, T. Schedl, and A. Gartner. “Translational repression of C. elegans p53 by GLD-1 regulates DNA damage-induced apoptosis”. In: *Cell* 120.3 (2005), pp. 357–368.

[216] B. Schumacher, C. Schertel, N. Wittenburg, S. Tuck, S. Mitani, A. Gartner, B. Conradt, and S. Shaham. “C. elegans ced-13 can promote apoptosis and is induced in response to DNA damage”. In: *Cell Death And Differentiation* 12.2 (2005), pp. 153–161.

[217] M. Scianna, L. Preziosi, and K. Wolf. “A cellular potts model simulating cell migration on and in matrix environments”. In: *Mathematical Biosciences and Engineering* 10.1 (2013), pp. 235–261.

[218] S. Segal, L. Graves, J. Verheyden, and E. Goodwin. “RNA-regulated TRA-1 nuclear export controls sexual fate”. In: *Developmental Cell* 1.4 (2001), pp. 539–551.

[219] H. Seidel and J. Kimble. “The oogenic germline starvation response in C. elegans”. In: *PLoS One* 6.12 (2011), e28074–e28074.

[220] Y. Setty, I. Cohen, Y. Dor, and D. Harel. “Four-dimensional realistic modeling of pancreatic organogenesis”. In: *Proceedings of the National Academy of Sciences* 105.51 (2008), pp. 20374–20379.

[221] Y. Setty, D. Dalfó, D. Korta, E. Hubbard, and H. Kugler. “A model of stem cell population dynamics: in silico analysis and in vivo validation”. In: *Development* 139.1 (2012), pp. 47–56.

[222] W. Sha, J. Moore, K. Chen, A. Lassaletta, C.-S. Yi, J. Tyson, and J. Sible. “Hysteresis drives cell-cycle transitions in Xenopus laevis egg extracts”. In: *Proceedings of the National Academy of Sciences* 100.3 (2003), pp. 975–980.

[223] S. Shaham, P. Reddien, B. Davies, and H. R. Horvitz. “Mutational analysis of the Caenorhabditis elegans cell-death gene ced-3”. In: *Genetics* 153.4 (1999), pp. 1655–1671.

[224] D. Shakes, B. Neva, H. Huynh, J. Chaudhuri, and A. Pires-da-Silva. “Asymmetric spermatocyte division as a mechanism for controlling sex ratios”. In: *Nature Communications* 2 (2011), p. 157.

[225] J. Sherratt and J. Murray. “Epidermal wound healing: a theoretical approach”. In: *Comments in Theoretical Biology* 2.5 (1992), pp. 315–333.

[226] A. Shirinifard, J. Gens, B. Zaitlen, N. Popławski, M. Swat, and J. Glazier. “3D multi-cell simulation of tumor growth and angiogenesis”. In: *PLoS One* 4.10 (2009), e7190.

[227] A. Silva, A. Anderson, and R. Gatenby. “A multiscale model of the bone marrow and hematopoiesis”. In: *Mathematical Biosciences and Engineering* 8.2 (2011), p. 643.

[228] H. Snippert, L. van der Flier, T. Sato, J. van Es, M. van den Born, C. Kroon-Veenboer, N. Barker, A. Klein, J. van Rheenen, B. Simons, and H. Clevers. “Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells”. In: *Cell* 143.1 (2010), pp. 134–144.

[229] T. Starich, D. Hall, and D. Greenstein. “Two classes of gap junction channels mediate soma-germline interactions essential for germline proliferation and gametogenesis in Caenorhabditis elegans”. In: *Genetics* 198.3 (2014), pp. 1127–1153.

[230] N. Starostina, J. Lim, M. Schvarzstein, L. Wells, A. Spence, and E. T. Kipreos. “A CUL-2 ubiquitin ligase containing three FEM proteins degrades TRA-1 to regulate C. elegans sex determination”. In: *Developmental Cell* 13.1 (2007), pp. 127–139.

[231] J. Starruß, W. de Back, L. Brusch, and A. Deutsch. “Morpheus: a user-friendly modeling environment for multiscale and multicellular systems biology”. In: *Bioinformatics* 30 (2014), pp. 1331–1332.

[232] J. Stoer and R. Bulirsch. *Introduction to numerical analysis*. Vol. 12. Springer Science & Business Media, 2013. Chap. 7.

[233] M. Su, D. Merz, M. Killeen, Y. Zhou, H. Zheng, J. Kramer, E. Hedgecock, and J. Culotti. “Regulation of the UNC-5 netrin receptor initiates the first reorientation of migrating distal tip cells in Caenorhabditis elegans”. In: *Development* 127.3 (2000), pp. 585–594.

[234] J. Sulston and H. Horvitz. “Post-embryonic cell lineages of the nematode, Caenorhabditis elegans”. In: *Developmental Biology* 56.1 (1977), pp. 110–156.

[235] *Summary of spermatogenesis*. http://www.wormbook.org/chapters/www_spermatogenesis/spermatogenesis.html. Accessed 11th June 2015.

[236] X. Sun and P. Hong. “Computational modeling of Caenorhabditis elegans vulval induction”. In: *Bioinformatics* 23.13 (2007), pp. 499–507.

[237] M. Swat, G. Thomas, J. Belmonte, A. Shirinifard, D. Hmeljak, and J. Glazier. “Multi-scale modeling of tissues using CompuCell3D”. In: *Methods in Cell Biology* 110 (2012), p. 325.

[238] H. Tabara, M. Sarkissian, W. Kelly, J. Fleenor, A. Grishok, L. Timmons, A. Fire, and C. Mello. “The rde-1 gene, RNA interference, and transposon silencing in C. elegans”. In: *Cell* 99.2 (1999), pp. 123–132.

[239] F. Tax, J. Yeargers, and J. Thomas. “Sequence of C. elegans lag-2 reveals a cell-signalling domain shared with Delta and Serrate of Drosophila”. In: *Nature* (1994).

[240] The MathWorks Inc. *SimBiology: model, simulate, and analyze biological systems*. http://uk.mathworks.com/products/simbiology/. [Online; accessed 18-February-2016].

[241] The MathWorks Inc. *Simulink: simulation and model-based design*. http://uk.mathworks.com/help/simulink/index.html. [Online; accessed 10-March-2016].

[242] The Object Management Group. *What is UML?* http://www.omg.org/gettingstarted/what_is_uml.htm. [Online; accessed 17-February-2016].

[243] The Papyrus project group. *Papyrus modeling environment*. https://eclipse.org/papyrus/. [Online; accessed 17-February-2016].

[244] The Papyrus project group. *Papyrus use case stories*. https://eclipse.org/papyrus/testimonials.html. [Online; accessed 17-February-2016].

[245] The Umbrello project group. *Umbrello the UML modeller*. https://umbrello.kde.org/. [Online; accessed 17-February-2016].

[246] R. Thomas. “Boolean formalization of genetic control circuits”. In: *Journal of Theoretical Biology* 42.3 (1973), pp. 563–585.

[247] B. Thompson, D. Bernstein, J. Bachorik, A. Petcherski, M. Wickens, and J. Kimble. “Dose-dependent control of proliferation and sperm specification by FOG-1/CPEB”. In: *Development* 132.15 (2005), pp. 3471–3481.

[248] R. Thompson, C. Yates, and R. Baker. “Modelling cell migration and adhesion during development”. In: *Bulletin of Mathematical Biology* 74.12 (2012), pp. 2793–2809.

[249] A. Trevorrow and T. Rokicki. *Golly: an open source, cross-platform cellular automaton application*. http://golly.sourceforge.net/. [Online; accessed 9-February-2016].

[250] S. Turner and J. Sherratt. “Intercellular adhesion and cancer invasion: a discrete simulation using the extended Potts model”. In: *Journal of Theoretical Biology* 216.1 (2002), pp. 85–100.

[251] J. Tyson, K. Chen, and B. Novak. “Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell”. In: *Current Opinion in Cell Biology* 15.2 (2003), pp. 221–231.

[252] University of Tokyo, Cell Illustrator team. *Cell Illustrator 5.0*. http://www.cellillustrator.com/home. [Online; accessed 18-February-2016].

[253] B. Vasiev, A. Balter, M. Chaplain, J. Glazier, and C. Weijer. “Modeling gastrulation in the chick embryo: formation of the primitive streak”. In: *PLoS One* 5.5 (2010), e10571.

[254] L. Verlet. “Computer “experiments” on classical fluids I. Thermodynamical properties of Lennard-Jones molecules”. In: *Physical Review* 159.1 (1967), p. 98.

[255] J. Vogel, D. Michaelson, A. Santella, E. Hubbard, and Z. Bao. “Irises: a practical tool for image-based analysis of cellular DNA content”. In: *Worm*. Vol. 3. 2. Taylor & Francis. 2014, e29041.

[256] J. Von Neumann and A. Burks. *Theory of self-reproducing automata*. University of Illinois Press, 1966.

[257] D. de Vos, K. Vissenberg, J. Broeckhove, and G. Beemster. “Putting theory to the test: which regulatory mechanisms can drive realistic growth of a root?” In: *PLoS Computational Biology* 10.10 (2014), e1003910.

[258] W. Wadsworth. “Moving around in a worm: netrin UNC-6 and circumferential axon guidance in C. elegans”. In: *Trends in Neurosciences* 25.8 (2002), pp. 423–429.

[259] G. Walker, K. Houthoofd, J. Vanfleteren, and D. Gems. “Dietary restriction in C. elegans: from rate-of-living effects to nutrient sensing pathways”. In: *Mechanisms of Ageing and Development* 126.9 (2005), pp. 929–937.

[260] M. Wang, E. O’Rourke, and G. Ruvkun. “Fat metabolism links germline stem cells and longevity in C. elegans”. In: *Science* 322.5903 (2008), pp. 957–960.

[261] R.-S. Wang and R. Albert. “Elementary signaling modes predict the essentiality of signal transduction network components”. In: *BMC Systems Biology* 5.1 (2011), p. 44.

[262] R.-S. Wang, A. Saadatpour, and R. Albert. “Boolean modeling in systems biology: an overview of methodology and applications”. In: *Physical Biology* 9.5 (2012), p. 055001.

[263] S. Ward, Y. Argon, and G. Nelson. “Sperm morphogenesis in wild-type and fertilization-defective mutants of Caenorhabditis elegans”. In: *The Journal of Cell Biology* 91.1 (1981), pp. 26–44.

[264] M. Weliky and G. Oster. “The mechanical basis of cell rearrangement”. In: *Development* 109.2 (1990), pp. 373–386.

[265] J. White, E. Southgate, J. Thomson, and S. Brenner. “The structure of the nervous system of the nematode Caenorhabditis elegans: the mind of a worm”. In: *Philosophical Transactions of the Royal Society B* 314 (1986), pp. 1–340.

[266] Wolfram Research. *Cellular Automata; Wolfram Language Tutorial*. https://reference.wolfram.com/language/tutorial/CellularAutomata.html. [Online; accessed 9-February-2016].

[267] U. Wolke, E. Jezuit, and J. Priess. “Actin-dependent cytoplasmic streaming in C. elegans oogenesis”. In: *Development* 134.12 (2007), pp. 2227–2236.

[268] M. Wong and J. Schwarzbauer. “Gonad morphogenesis and distal tip cell migration in the Caenorhabditis elegans hermaphrodite”. In: *Wiley Interdisciplinary Reviews: Developmental Biology* 1.4 (2012), pp. 519–531.

[269] W. Wood, R. Hecht, S. Carr, R. Vanderslice, N. Wolf, and D. Hirsh. “Parental effects and phenotypic characterization of mutations that affect early development in Caenorhabditis elegans”. In: *Developmental Biology* 74.2 (1980), pp. 446–469.

[270] Z. Yang and Y. Hata. “What is the Hippo pathway? Is the Hippo pathway conserved in Caenorhabditis elegans?” In: *Journal of Biochemistry* 154.3 (2013), pp. 207–209.

[271] J. Yochem and I. Greenwald. “Glp-1 and lin-12, genes implicated in distinct cell-cell interactions in C. elegans, encode similar transmembrane proteins”. In: *Cell* 58.3 (1989), pp. 553–563.

[272] B. Yordanov, S.-J. Dunn, H. Kugler, A. Smith, G. Martello, and S. Emmott. “A method to identify and analyze biological programs through automated reasoning”. In: *NPJ Systems Biology and Applications* 2 (2016), p. 16010.

[273] B. Yordanov, C. Wintersteiger, Y. Hamadi, and H. Kugler. “Z34Bio: An SMT-based framework for analyzing biological computation”. In: *Proceedings of SMT* 13 (2013).

[274] S. Zanetti and A. Puoti. “Chapter 3. Sex determination in the Caenorhabditis elegans germ line”. In: *Germ Cell Development in C. elegans*. Ed. by T. Schedl. New York: Springer, 2013.

[275] E. Zeiser, C. Frøkjær-Jensen, E. Jorgensen, and J. Ahringer. “MosSCI and gateway compatible plasmid toolkit for constitutive and inducible expression of transgenes in the C. elegans germline”. In: *PLoS One* 6.5 (2011), e20082.

[276] Q. Zeng and W. Hong. “The emerging role of the hippo pathway in cell contact inhibition, organ size control, and cancer development in mammals”. In: *Cancer Cell* 13.3 (2008), pp. 188–192.

[277] M. Zetka, I. Kawasaki, S. Strome, and F. Müller. “Synapsis and chiasma formation in Caenorhabditis elegans require HIM-3, a meiotic chromosome core component that functions in chromosome segregation”. In: *Genes And Development* 13.17 (1999), pp. 2258–2270.

[278] B. Zhang, M. Gallegos, A. Puoti, E. Durkin, S. Fields, J. Kimble, and M. Wickens. “A conserved RNA binding protein that regulates sexual fates in the C. elegans hermaphrodite germ line”. In: *Nature* 390.6659 (1997), pp. 477–484.

[279] J. Ziel and D. Sherwood. “Roles for netrin signalling outside of axon guidance: a view from the worm”. In: *Developmental Dynamics* 239.5 (2010), pp. 1296–1305.

Appendix A

Code samples

This appendix contains three example pieces of simulation code. First, we include a typical Chaste test file defining an overlapping spheres simulation. The test file demonstrates how force laws, boundary conditions, and cell killers can be added to a simulation in a modular fashion. Second, we include some example statechart code controlling cell growth. Third, we include similar cell growth code, but written in the form of a cell logic module. This allows the statechart and cell logic module approaches to be compared side by side.

Listing A.1: An example Chaste test file.

```
//--------------------- CREATE A CELL POPULATION ---------------------

// Specify 2 cell positions and package them into a mesh.
std::vector<Node<3>*> nodes;
nodes.push_back(new Node<3>(0, false,  1, 0, 0));
nodes.push_back(new Node<3>(1, false, -1, 0, 0));
NodesOnlyMesh<3> mesh;
mesh.ConstructNodesWithoutMesh(nodes, 1.5);

// Make an overlapping spheres population of stem cells with a
// stochastic cell cycle.
std::vector<CellPtr> cells;
MAKE_PTR(StemCellProliferativeType, p_stemType);
CellsGenerator< StochasticDurationCellCycleModel, 3 > cellsGenerator;
cellsGenerator.GenerateBasic(cells, mesh.GetNumNodes(), p_stemType);
NodeBasedCellPopulation<3> nodeBasedCellPopulation(mesh, cells);

//-------------------- CREATE A SIMULATION OBJECT ---------------------

OffLatticeSimulation<3> simulation(cellPopulation);
simulation.SetOutputDirectory("SimpleCellCentreSimulation");
simulation.SetEndTime(100.0);

//--------------------- ADD SIMULATION FEATURES -----------------------

// Apply a Generalised Linear Spring force law.
MAKE_PTR(GeneralisedLinearSpringForce<3>, p_force);
p_force->SetCutOffLength(1.5);
simulation.AddForce(p_force);

// Add a plane boundary condition.
c_vector<double,3> pointOnPlane = zero_vector<double>(3);
pointOnPlane(0) = -2;
c_vector<double,3> normalToPlane = zero_vector<double>(3);
normalToPlane(0) = -1;
MAKE_PTR_ARGS(PlaneBoundaryCondition<3>, p_boundary,
             (&nodeBasedCellPopulation, pointOnPlane, normalToPlane));
simulation.AddCellPopulationBoundaryCondition(p_boundary);

// Add a cell killer that removes cells above a certain plane.
c_vector<double,3> point = zero_vector<double>(3);
pointOnPlane(0) = 2;
c_vector<double,3> normal = zero_vector<double>(3);
normalToPlane(0) = 1;
MAKE_PTR_ARGS(PlaneBasedCellKiller<3>, p_killer,
             (&nodeBasedCellPopulation, point, normal));
simulation.AddCellKiller(p_killer);

// Add a modifier to track cell volumes.
MAKE_PTR(VolumeTrackingModifier<3>, p_modifier);
simulator.AddSimulationModifier(p_modifier);

//------------------------------ RUN ------------------------------

simulation.Solve();
```

Listing A.2: Implementing cell growth as a statechart.

```
//----------- IN THE TEST FILE:-------------

CellsGenerator<StatechartCellCycleModel<GermCellBehaviourModel>, 3>
                                                     cellsGenerator;
cellsGenerator.GenerateBasicRandom(cells,
                                   p_mesh->GetNumNodes(),
                                   pStemType);

//----- IN THE STATECHART HEADER FILE:-------

// Declare all states:
struct Running;
struct Growth;
struct Growing;
struct Static;

// Declare all update and goto events:
struct EvGrowthUpdate : sc::event<EvGrowthUpdate> {};
struct EvGoToGrowing : sc::event<EvGoToGrowing> {};
struct EvGoToStatic : sc::event<EvGoToStatic> {};

// Fixed model parameters:
double maxRadius = 4;
double growthRate = 0.5;

// Define the statechart object:
struct GermCellBehaviourModel: sc::state_machine<GermCellBehaviourModel,
                                                 Running>
{
  GermCellBehaviourModel();

  CellPtr ownerCell;
  void SetCell(CellPtr newCell);

  // Required utility functions for copying and saving:
  boost::shared_ptr<GermCellBehaviourModel> CopyIntoChart(
   boost::shared_ptr<GermCellBehaviourModel> myNewStatechart);
  std::bitset<MAX_STATES> GetState();
  void SetState(std::bitset<MAX_STATES> state);
};


// Define the chart's structure and the events each state responds to:
struct Running: sc::simple_state<Running,GermCellBehaviourModel,Growth>
{
  typedef sc::custom_reaction<EvCheckCellData> reactions;
  sc::result react( const EvCheckCellData & );
};

struct Growth: sc::state<Growth, Running, Growing>
{
  Growth(my_context ctx);

  typedef mpl::list<
    sc::custom_reaction<EvGoToGrowing>,
    sc::custom_reaction<EvGoToStatic>,
  > reactions;
```

```
  sc::result react( const EvGoToGrowing & )
  {
     return transit<EvGoToGrowing>();
  };
  sc::result react( const EvGoToStatic & ){
     return transit<EvGoToStatic>();
  };
};

struct Growing: sc::state<Growing, Growth>
{
  Growing(my_context ctx);

  typedef mpl::list< sc::custom_reaction<EvGrowthUpdate> > reactions;
  sc::result react( const EvGrowthUpdate & );
};

struct Static: sc::state<Static, Growth>
{
  Static(my_context ctx);

  typedef mpl::list< sc::custom_reaction<EvGrowthUpdate> > reactions;
  sc::result react( const EvGrowthUpdate & );
};

//----- IN THE STATECHART SOURCE FILE:-------

// Standard functions required for every statechart model:
GermCellBehaviourModel::GermCellBehaviourModel(){
  ownerCell = boost::shared_ptr<Cell>();
};

void GermCellBehaviourModel::SetCell(CellPtr newCell){
  assert(newCell!=NULL);
  ownerCell=newCell;
};

std::bitset< MAX_STATES > GermCellBehaviourModel::GetState(){
  ...
}

void GermCellBehaviourModel::SetState(std::bitset<MAX_STATES>state){
  ...
}

sc::result Running::react( const EvCheckCellData & ){
  post_event(EvGrowthUpdate());
  return discard_event();
};

// Simple state inheritance:
boost::shared_ptr<GermCellBehaviourModel>GermCellBehaviourModel::
 CopyInto(boost::shared_ptr<GermCellBehaviourModel> myNewStatechart)
{
  myNewStatechart->initiate();

  if(state_cast<const Growing*>()!=0){
      myNewStatechart->process_event(EvGoToGrowing());
  }
  if(state_cast<const Static*>()!=0){
```

```
      myNewStatechart->process_event(EvGoToStatic());
  }
  return myNewStatechart;
};

// The actual model - defines behaviour in each state:

// The growth region
Growth::Growth( my_context ctx ):my_base( ctx ){};

// The growing state
Growing::Growing( my_context ctx ):my_base( ctx ){
        CellPtr myCell = context<GermCellBehaviourModel>().ownerCell;
        myCell->GetCellData->SetItem("Growing", 1.0);
};

sc::result Growing::react( const EvGrowthUpdate & )
{
  CellPtr myCell = context<GermCellBehaviourModel>().ownerCell;
  double currentRadius = myCell->GetCellData()->GetItem("Radius");

  if( currentRadius >= maxRadius ){
    return transit<Static>();
  }

  myCell->GetCellData->SetItem("Radius",
                          currentRadius + growthRate *
                          SimulationTime::Instance()->GetTimeStep());
  return discard_event();
};

// The static state
Static::Static( my_context ctx ):my_base( ctx ){
        CellPtr myCell = context<GermCellBehaviourModel>().ownerCell;
        myCell->GetCellData->SetItem("Growing", 0.0);
};

sc::result Static::react( const EvGrowthUpdate & )
{
  CellPtr myCell = context<GermCellBehaviourModel>().ownerCell;
  double currentRadius = myCell->GetCellData()->GetItem("Radius");

  if(currentRadius < maxRadius){
    return transit<Growing>();
  }

  return discard_event();
};
```

Listing A.3: Implementing cell growth as a logic module.

```
//----------- IN THE TEST FILE:-------------

// Make an empty logic cell:
MAKE_PTR(DummyMutationState, pDummyMut);
MAKE_PTR_ARGS(LogicCell, pNewCell, (pDummyMut, 0));

// Add cell growth logic module:
double startingRadius = 2;
double growthRate = 0.5;
double maxRadius = 4;
CellGrowth* myGrowthLogic = new CellGrowth(pNewCell.get(),
                                           Growth::GROWING,
                                           startingRadius,
                                           growthRate,
                                           maxRadius);
pNewCell->SetLogic<Growth>(myGrowthLogic);

//------- IN THE LOGIC MODULE SOURCE FILE:---------

CellGrowth::CellGrowth(LogicCell* ownerCell, int startingState,
                       double startingRadius, double growthRate,
                       double maxRadius):
  AbstractCellLogic (ownerCell, startingState),
  cellRadius (initialRadius),
  radialGrowthPerHour (growthRate),
  maximumRadius (maxRadius)
{}

// Update at each time step:
void CellGrowth::Update(){

   if(cellRadius >= maximumRadius) {
      state = Growth::STATIC;
   }else{
      state = Growth::GROWING;
   }

   if(state == Growth::GROWING) {
      cellRadius += radialGrowthPerHour
                  * SimulationTime::Instance()->GetTimeStep();
   }

   OutputVisualisationInfo();
}

// Create a daughter logic module on division:
AbstractCellLogic* CellGrowth::Divide(LogicCell* daughterCell){

   // Most information is inherited, but the daughter state is GROWING
   // and the radius of both cells is scaled to halve their volume
   AbstractCellLogic* daughterLogic = new CellGrowth(daughterCell,
                                           Growth::GROWING,
                                           cellRadius/pow(2,0.33),
                                           radialGrowthPerHour,
                                           maximumRadius);
   cellRadius = cellRadius/pow(2,0.33);
   return daughterLogic;
}
```

```
// Data output
void CellGrowth::OutputVisualisationInfo(){

   owningCell->GetCellData()->SetItem("Radius", cellRadius);
   AbstractCellLogic::OutputVisualisationInfo();
}
```

Appendix B

Computational details

This thesis contains substantial simulations of a variety of types.

- Whole *C. elegans* germ line simulations (Chapter 4).
- Force law comparison simulations using a test tube geometry (Chapter 5).
- Numerical method comparison simulations (Chapter 5).
- Distal germ line simulations using different proliferative zone maintenance mechanisms (Chapter 6).

Here we give details of the machines used to run these simulations, and an idea of typical run times.

Whole *C. elegans* germ line simulations

The majority of this work was performed using an Apple laptop, with simulations running on an Intel Core i5 processor. The simulations used to generate the results in Section 4.3 take between 10 and 12 hours each to complete. Local parameter variation work was carried out using ARCUS-B (an HPC cluster composed of Dual Haswell CPU nodes, provided by Oxford's Advanced Research Computing team). Each simulation was run on a single core, with a typical run time of 8 hours.

Force law comparison simulations using a test tube geometry

These simulations were carried out using computational resources provided by Microsoft Research Cambridge. Ten remote Linux machines were provided, and each simulation ran on a single Intel Xeon CPU. Typical run times for the tests described in Section 5.1.5 were $\sim$20 hours using the GLS force law and 48+ hours using the Buske force law (note that some simulations accumulated a large number of cells and had to be terminated after a day or more of slow progress).

Numerical method comparison simulations

The majority of these simulations were performed using the Microsoft Research machines described above. The exception is the DOP853 test (Section 5.2.7), which was run on a separate Linux laptop with an Intel Core i5 processor. Typical run times for the numerical method tests are reported in Chapter 5.

Distal germ line simulations using different proliferative zone maintenance mechanisms

All of these simulations were carried out using the Microsoft Research machines described above. The typical run time for a test used to produce the data in Section 6.2.2 was 4 hours.