Published in final edited form as:

Mol Reprod Dev. 2016 November ; 83(11): 944–957. doi:10.1002/mrd.22735.

How computational models contribute to our understanding of the germ line

Kathryn Atwell^{1,2}, Sara-Jane Dunn², James M. Osborne³, Hillel Kugler^{2,4}, and E. Jane Albert Hubbard⁵

¹Computational Biology Group, Department of Computer Science, University of Oxford, Oxford, UK

²Biological Computation, Microsoft Research, Cambridge, UK

³School of Mathematics and Statistics, University of Melbourne, Melbourne 3010, Australia

⁴Faculty of Engineering, Bar-Ilan University, Ramat Gan 5290002, Israel

⁵Skirball Institute of Biomolecular Medicine, Department of Cell Biology, and Kimmel Center for Stem Cell Biology, New York University School of Medicine, New York, USA

SUMMARY

Computational models are an invaluable tool in modern biology. They provide a framework within which to summarize existing knowledge, enable competing hypotheses to be compared qualitatively and quantitatively, and to facilitate the interpretation of complex data. Moreover, models allow questions to be investigated that are difficult to approach experimentally. Theories can be tested in context, identifying the gaps in our understanding and potentially leading to new hypotheses. Models can be developed on a variety of scales and with different levels of mechanistic detail, depending on the available data, the biological questions of interest, and the available mathematical and computational tools. The goal of this review is to provide a broad picture of how modeling has been applied to reproductive biology. Specifically, we look at four uses of modeling: i) comparing hypotheses, ii) interpreting data, iii) exploring experimentally challenging questions, and iv) hypothesis evaluation and generation. We present examples of each of these applications in reproductive biology, drawing from a range of organisms – including *Drosophila*, *Caenorhabditis elegans*, mouse, and humans. We aim to describe the data and techniques used to construct each model, and to highlight the benefits of modeling to the field, as complementary to experimental work.

Keywords

stem cell; modeling; germ cell; *in silico*

INTRODUCTION

The use of computational models is growing rapidly in many fields of biology, from protein structure prediction, to cell cycle dynamics, cancer research, and developmental biology (Tyson and Novak, 2001; Byrne et al., 2006; Tomlin and Axelrod, 2007; Peter et al., 2010; Dill and MacCallum, 2012). The term “computational model” generally applies to any formal, quantitative description of a biological system that can be used to explore the behavior arising from a set of input parameters. This broad definition encompasses everything from mathematical equations to genetic networks to simulated populations of interacting cells. Models can therefore be constructed at various levels of abstraction, ranging from largely phenomenological (the ideal level for hypothesis generation in systems where few mechanistic details are known), to highly detailed, multi-scale models parametrised by experimental data.

In principle, a verbal or diagrammatic model can be transformed into a computational model by selecting an appropriate computational framework and making concrete choices about the model representation, implementation, parameter values, and analysis. A wide range of mathematical formalisms can be used in model development: representations can be either continuous, discrete, or hybrid, and can lead to system-level behavior that is either deterministic (fixed in outcome), stochastic (random in outcome), or some combination of the two. Numerous techniques also exist for model analysis: some result in exact mathematical solutions, some provide an exhaustive description of all possible model behavior, while others make use of simulation to study behavior *in silico*. Regardless of the exact approach, the process of developing and “running” a formal model requires that the assumptions underlying specific hypotheses are made explicit, and exposes potential gaps in current knowledge.

Computational models can be used to study the effect of different parameter values and perturbations on a system, and to predict the phenotypes obtained. If these predictions fail to match experimental observations, it can reveal incorrect assumptions, and the model must be revised to resolve the discrepancy. This can lead to the formulation of entirely new hypotheses for testing, in a “virtuous cycle” between modeling and experiment (Fig. 1). This way of working complements traditional experimental approaches, in which mechanistic behavior is inferred by observing the phenotype that results from mutations, anatomical manipulations, or varied environmental conditions (Harel and Kugler, 2010).

Crucially, models also allow us to extract more information from existing experimental data. Given a small set of starting assumptions, a model can be constructed and its parameters tuned to fit a complex dataset. The model parameters may then reveal quantitative information about the biological system that was not immediately apparent from the data, such as predicted rates of cell proliferation or the relative frequency of symmetric versus asymmetric divisions. Finally, models enable us to explore biological questions that are difficult to approach experimentally, due to limitations such as cost, feasibility, or ethics.

This review highlights modeling work in the field of reproductive biology. It is divided into four sections, each addressing a different application of computational modeling. We begin

by illustrating how models can be used to distinguish between competing hypotheses. We then examine how modeling enables complex experimental data to be interpreted and relevant information extracted. Next, we describe studies that used models to explore questions that are difficult to approach experimentally. Finally, we show how modeling can be used to test ideas and to generate new hypotheses. Throughout, we draw on examples from a range of biological systems, including spermatogenesis in mouse, humans, and *Drosophila*, as well as the development and homeostasis of the *Caenorhabditis elegans* germ line. Many outstanding review articles provide details and primary references for these systems (see Lehmann, 2012; Schedl, 2013; Yoshida, 2016), so we describe only the key papers and results as needed in the course of the discussion.

COMPARING HYPOTHESES QUANTITATIVELY

We begin by demonstrating how computational models permit rigorous and quantitative hypothesis comparison, enabling the more cogent explanation for the data to be identified. We look at two examples of this approach: The first concerns a comparison of potential mechanisms by which mammalian spermatogenic stem cells might be maintained, while the second compares two possible explanations for the unusual prevalence of a specific genetic mutation in humans.

Comparing potential mechanisms of spermatogenic stem cell maintenance

Mammalian spermatogenesis occurs in the seminiferous tubule, which consists of four layers of tissue surrounding a central lumen (Fig. 2A). Stem cells are located in the basal layer of the tubule, while differentiating cells move toward the lumen. The prevailing view was that long-lived stem cells maintain this tissue by dividing, mostly asymmetrically, to produce one stem daughter and one differentiating daughter, thereby ensuring a constant stem cell count. Yet, stem cells *can* divide symmetrically and still maintain their numbers, provided the probability of producing two stem daughters equals the probability of producing two differentiating daughters. The extent to which a given stem cell population self-renews in this way remains an area of intensive research that is amenable to both experimental and modeling approaches (see also Rulands and Simons, 2016, and references therein).

Klein et al. (2010) combined modeling with experiment to compare different stem cell-maintenance hypotheses in the context of mouse spermatogenesis. First, a clonal labeling experiment was performed in which a small percentage of basal layer spermatogonia expressed lacZ. The labeling density was sufficiently low that each clone corresponded to a single founder cell. Clone development was followed for over a year, and the data were combined with the results of earlier labeling studies (Nakagawa et al., 2007). Some clear trends emerged: The number of distinct clones in the testis decreased continuously, with many clones becoming extinct while others expanded. Meanwhile, the distribution of clone lengths, kept the same form over time, but with a scaling transformation applied. Finally, the average clone length at time t was proportional to t (Fig. 2B).

Based on these data, the long-lived stem cells hypothesis with strict asymmetric division was immediately rejected since clones derived from the hypothesized long-lived stem cells should have persisted indefinitely; instead, a continual decline in clone number was

observed. A variety of other mechanisms could explain these data. For example, cells might differ in their proliferative capacity, with the expanding clones containing all the true stem cells and the unlabeled areas consisting of differentiating cells that are progressively lost. Alternatively, all basal-layer spermatogonia might have equal proliferative potential, with stem cell loss and replacement occurring – i.e. stem cells could divide symmetrically in response to the loss of a neighbor, or individual cells could choose a fate autonomously and independently of their surroundings. Since these hypotheses are difficult to distinguish using experiments alone, mathematical modeling was used to derive detailed predictions about expected clone lengths under each mechanism, which were then compared to observations.

Here we explain the modeling process used by Klein et al. to demonstrate this approach, taking as an example the hypothesis that labeled and unlabeled cells are equivalent and that symmetric division is coupled to stem cell loss. Consider a pair of basal-layer stem cells at the boundary of a clone, one labeled and one unlabeled. There are two possible outcomes when these cells are involved in a loss-and-replacement event: i) the unlabeled cell could be lost from the basal layer and the labeled cell could divide to replace it, thereby increasing the number of stem cells in the clone by one, or ii) the labeled cell could be lost and the unlabeled cell could divide to replace it, decreasing the number of stem cells in the clone by one. Assuming equivalent labeled and unlabeled cells, both outcomes are equally likely.

Let $P_n(t)$ stand for the probability that a clone contains n stem cells at time t . From the argument above, we can construct an equation describing how $P_n(t)$ changes over time (Fig.

2C). The left side of the equation is the rate of change of $P_n(t)$, written $\frac{dP_n(t)}{dt}$, where τ^{-1} is the average rate of loss and replacement. The first two terms on the right side of the equation define how cells can be added to or removed from the clone. The first term corresponds to adding a cell ($n \rightarrow n+1$) while the second corresponds to losing a cell ($n \rightarrow n-1$). The

third term is included to ensure that $\frac{dP_n(t)}{dt} = 0$ when $n = 0$, so that once a clone has become extinct, it stays extinct permanently. Here, the symbol $\delta_{a,b}$ is used to represent a constant that is 1 if $a = b$, otherwise it is 0. Finally, the fourth term ensures that clones contain exactly one cell at $t = 0$, in accordance with conditions at the beginning of the labeling experiment. In this case, $\delta(t)$ is a generalized function with two special properties: $\delta(t) = 0$ except when $t = 0$, and the area under the graph of $\delta(t)$ is 1.

Mathematical techniques exist that can solve the above equation (Fig. 2C) to obtain a formula for $P_n(t)$. The details are given in the supplementary information of Klein et al. (2010), and involve transforming the problem into a related equation for which the solution is already known. The result (Fig. 2D) is expressed in terms of a set of special functions, $I_n(x)$, which stands for the n^{th} modified Bessel function of the first kind. The value of $I_n(x)$ cannot be written in terms of simpler mathematical functions, but its behavior for selected values of n can be illustrated graphically (Fig. 2D). In practice, we are only interested in the properties of the clones that survive at time t , since extinct clones are not measured experimentally. The probability of a *surviving* clone hosting n stem cells at time t (Fig. 2E) and the average number of stem cells in a clone (Fig. 2F) can both be calculated. By using well-known properties of Bessel functions (such as their behavior when x is large), we can

determine that the average number of stem cells per clone grows proportional to t at late times. This proportionality remains when the expression defining n stem cells (Fig. 2F) is converted into a formula for average clone length.

A similar procedure can be followed for the other hypothesized mechanisms by changing the model's starting assumptions. Above, we assumed that a labeled cell is just as likely to replace an unlabeled one and vice versa; altering that assumption can produce a similarly structured model in which labeled cells have greater proliferative capacity. Equally possible would be modeling cell autonomous fate decisions by adding the possibility that a symmetric division occurs without stem cell loss, or that stem cell loss occurs without replacement. By solving the resulting model equations in an analogous manner to the presented example, however, these proposed mechanisms did not reproduce the observed features of clone size distribution. In particular, these hypotheses predict that, at late times, the average clone length will grow in proportion to t rather than $\sqrt{t}$. This type of analysis is made possible by formalizing a hypothesis as a mathematical model, which, in contrast to verbal reasoning, provides an analytical expression for the predicted average clone length that can then be used to distinguish hypotheses based on statistical data.

Klein et al. (2010) also studied other aspects of spermatogenic stem cell behavior by conducting simulations of clone development, subject to stem cell loss and replacement. Values calculated from this model – such as the probability of clone fragmentation and the rate of stem cell replacement – matched experimental results well. Overall, the close quantitative agreement between model predictions and experimental data provided strong evidence in favor of the hypothesis that stochastic, coupled loss and replacement occurs among mouse spermatogenic stem cells. Similar investigations have been conducted into the maintenance of the epidermis and intestinal crypt (Clayton et al., 2007; Doupé et al., 2010; Lopez-Garcia et al., 2010), revealing that stochastic loss and replacement of equipotent stem cells is a common feature of many self-renewing tissues (Simons and Clevers, 2011). These publications demonstrate how formalizing a biological hypothesis as a mathematical model that can be solved analytically allows for the comparison and exclusion of competing theories of cell behavior.

Comparing two possible explanations for the prevalence of Apert's syndrome in humans

The previous example highlighted the potential of a simple mathematical model to distinguish between competing hypotheses. In some cases, a more detailed simulation is required, particularly when the available experimental data relates to the distribution of cells in three dimensions.

A good example can be found in the work of Qin et al. (2007), concerning a genetic mutation causative for Apert's syndrome, a disorder characterized by the premature fusion of the skull bones. The authors found that the “Apert's mutation” is far more prevalent in the human population than might be expected, given estimates of the average human mutation rate. The simplest explanation would be that this particular locus mutates unusually frequently, an idea that we term the “high-mutation-rate” hypothesis. The authors, however, speculated that spermatogenic stem cells carrying the mutant gene might have a competitive advantage in the testis, producing larger clones and therefore more affected sperm (see also

Goriely et al., 2003); we refer to this alternate proposal as the “mutant-advantage” hypothesis.

As with Klein et al. (2010), a combination of experimental and modeling work was used to distinguish between the two hypotheses. First, several sections of donor testes were analysed using a PCR-based technique to determine the prevalence of the Apert's mutation in different regions of the organ. The results revealed strong spatial clustering of mutant cells, with larger clusters appearing in older individuals – which seemed to support the mutant-advantage hypothesis.

To provide a more rigorous and quantitative assessment, a computational model was developed to simulate each hypothesis and to compare the spatial arrangement of mutant cells produced *in silico* to experimental data. The testis was modeled as a three-dimensional grid, with simulated spermatogenic stem cells occupying the boxes therein (Fig. 3). The simulation was divided into two phases: an initial growth phase and a subsequent adult phase. Stem cells divided symmetrically during the growth phase versus asymmetrically during the adult phase. Apert's mutation was assumed to occur with some probability at each division. As a result, mutations arising during the growth phase were passed on to a large number of stem cell daughters, whereas mutations arising during the adult phase were only passed on to differentiating cells (see Fig. 3).

Within the mutant-advantage hypothesis model, mutant stem cells were permitted to divide symmetrically into a stem-stem pair with some small probability, allowing mutant clones to expand gradually even during the adult phase. The model was run thousands of times, both with and without “mutant advantage”. A range of values were also tested for two important model parameters: the probability of mutation and the probability of symmetric division in the adult. The parameter values giving the best fit to experimental observations were identified. Finally, statistical tests were performed to compare the best-fit high-mutation-rate model and the best-fit mutant-advantage model to the spatial clustering data, resulting in a rejection of the high-mutation-rate hypothesis; conversely, the concept that mutant clones gradually out-compete their neighbors was consistent with the empirical data. This study was therefore able to make a rigorous assessment of two competing hypotheses by exploiting the ability of simulations to be repeated thousands of times, generating statistics, and identifying best-fit parameters.

Similar investigations have since yielded the same result for a second Apert's syndrome mutation, and for mutations causing Nooan syndrome (Choi et al., 2008; Yoon et al., 2013). Taken together, this body of work suggests that competition between spermatagonial stem cell clones may be an important factor influencing the prevalence of some genetic diseases.

INTERPRETING COMPLEX DATA

Computational models also play an important role in interpreting experimental data. By fitting a simple model, more quantitative information about a biological system may be inferred than is immediately apparent. We present two examples of this, both in the context of *Drosophila* spermatogenesis. The first study sought to determine the relative frequency of

symmetric versus asymmetric divisions among germ-line stem cells, while the second study quantified the proliferative advantage enjoyed by mutant cyst stem cells over their wild-type counterparts.

Determining the frequency of symmetric divisions in *Drosophila* spermatogenesis

Drosophila spermatogenesis begins around the “hub”, a group of cells to which germ-line stem cells and somatically-derived cyst stem cells attach (see Fig. 4A). Germ-line stem cell usually divide asymmetrically and perpendicular to the hub, producing a gonialblast daughter. The cyst cell daughters produced by cyst stem cell then surround the gonialblast, and eventually give rise to mature sperm – after four transit-amplifying divisions, meiosis, and spermiogenesis. Although *Drosophila* spermatogenesis is often cited as a prime example in which oriented asymmetric divisions maintain a tissue, the picture is less clear than it first appears since gonialblasts can reattach to the hub and regain stem status. Evidence for these symmetric stem-stem divisions (e.g. Sheng and Matunis, 2011) has prompted studies aimed at determining the frequency and significance of symmetric self-renewal.

One such investigation was carried out by Salzmann et al. (2013), who began by performing a clonal labeling experiment that induced GFP expression in a small percentage of *Drosophila* germ-line stem cells. The testis was then re-examined after a 24-hour chase period, revealing a large number of isolated labeled cells, as well as a small number of labeled doublets and still fewer labeled triplets. The authors were not immediately able to deduce the frequency of symmetric division from the number of singlets, doublets, and triplets observed. While it is tempting to assume that each doublet corresponds to a symmetric germ-line stem cell division and each singlet corresponds to purely asymmetric divisions, this need not be the case as many sequences of events can result in a given number of labeled cells. For example, a singlet could also arise from one symmetric germ-line stem cell division, followed by one symmetric differentiation (i.e. germ-line stem cell loss).

The authors constructed a simple mathematical model to help interpret their labeling data. Their model consisted of a set of equations specifying the current number of singlets, doublets, and triplets in terms of the number present at the previous cell cycle (see Fig. 4B). These equations took into account the many different ways in which singlets, doublets, and triplets can be produced, and also incorporated three parameters, representing the frequency of asymmetric division, the frequency of symmetric division, and the frequency of symmetric differentiation. By simulating the model using a range of parameter values, the authors were able to determine which set of division frequencies most closely reproduced their labeling observations. The conclusion was that symmetric self-renewal accounts for around 1.5% of *Drosophila* germ-line stem cell divisions.

The available clonal-labeling data therefore proved sufficient to infer a symmetric germ-line stem cell division frequency. Without a computational model, however, relating these experimental measurements to the quantity of interest would have been extremely difficult. This study therefore exemplifies how a computational model can help to extract relevant metrics, thereby getting more information out of a complex dataset.

Quantifying mutant stem cell advantage

The work of Amoyel et al. (2014) provides another good example of how modeling can extract quantitative information from experimental data, here in the context of cyst stem cell maintenance in the *Drosophila* testis. The authors began by investigating how wild-type cyst stem cells self-renew. One hypothesis is that cyst stem cells are long-lived and divide strictly asymmetrically to produce cyst cells. Alternatively, cyst stem cells could be subject to stochastic loss and replacement. To distinguish between these hypotheses, the authors used a very similar modeling approach to Klein et al. (2010) (since we have already discussed this type of modeling procedure, we do not give full details here): An equation was formulated for $P_n(t)$ (the probability of finding a clone of size n at time t), assuming that all cyst stem cells are functionally equivalent and are subject to coupled loss and replacement. The equation was solved analytically, and then the behavior of the solution was compared with experimental clonal labeling data. The results strongly supported a coupled-loss-and-replacement hypothesis. Amoyel et al. (2014) also described some of the difficulties peculiar to this system, particularly regarding the identification of cyst stem cells and determination of the initial proportion of labeled cells.

Amoyel et al. (2014) went on to investigate certain cyst stem cell mutants and whether or not they have a competitive advantage over their wild-type neighbours. Specifically, they studied cyst stem cell clones mutant for the gene *patched* (*ptc*). Cyst stem cells homozygous mutant for *ptc* experience constitutive activation of the Hedgehog pathway, leading to faster-than-normal proliferation. In an experimental study, *ptc*-mutant clones were found to expand rapidly, replacing non-mutant cyst stem cells and eventually displacing even germ-line stem cells from the hub. An updated version of the computational model was developed in order to quantify the competitive advantage enjoyed by mutant cyst stem cells. A bias parameter was added to the equation for $P_n(t)$, representing the amount by which mutant clones are more likely to expand than shrink. The bias takes on a value between 0 and 1, where 0 corresponds to clone growth and shrinkage being equally likely, and 1 corresponds to constant clone expansion. By comparing the solution of the model equation with the observed labeling data, it was possible to identify the best-fit bias value. The conclusion was that *ptc* mutant clones show a moderate bias toward expansion of 0.15.

While experimental work alone was able to identify that *ptc* mutants have a competitive advantage, computational modeling made it possible to quantify that advantage – and thus for comparing the advantages conferred by different genetic mutations. Indeed, as part of their argument that faster proliferation is the key factor underlying mutant advantage, Amoyel et al. (2014) fit their mathematical model to data on *hippo*-mutant clones, commenting that the value of the bias parameter is similar to that obtained for *ptc*. Ultimately, these results may be relevant and translatable to how tumor-initiating cells become established in a self-renewing tissue, in addition to highlighting the effects of mutations in one stem cell population on another population sharing the same niche.

INVESTIGATING EXPERIMENTALLY CHALLENGING QUESTIONS

Computational modeling can be used to explore questions that are difficult to address experimentally, due to feasibility, cost, or ethical issues. We present two examples in this

section, both relating to the *C. elegans* germ line. Despite its optical transparency, a limitation of working with this species is that their germ cells cannot be tracked using long-term live-imaging due to their extreme sensitivity to stress. The first computational model we discuss sought to overcome this constraint by enabling germ cell movement information to be reconstructed from still images. The second example explored the evolutionary pressures that may have shaped the germ line – a challenging topic to study experimentally due to the long time scales of evolutionary change.

Reconstructing *C. elegans* germ cell movement

The *C. elegans* hermaphrodite germ line is a powerful model system used to study cell proliferation, differentiation, and tissue maintenance. The germ line is contained within two U-shaped gonad arms that meet at a central uterus (Fig. 5A). Key to this system is the distal tip cell, positioned at the end of the arm, that acts as a leader cell during gonad development as well as an important signaling cell in the adult. Signaling from the distal tip cell activates GLP-1, a member of the Notch family of transmembrane receptors, on nearby germ cells, keeping them in an undifferentiated mitotic state and thereby establishing a proliferative zone. As germ cells move out of this zone, they enter meiosis. During larval stages, meiotic cells differentiate into sperm that are subsequently stored in the spermatheca. By contrast, many meiotic cells undergo apoptosis in the adult; surviving germ cells mature into oocytes. As a result, the hermaphrodite contains gametes of both sexes and is self-fertile.

Despite offering many advantages for the experimentalist, the *C. elegans* germ line still presents certain challenges. One limitation is that techniques have yet to be developed to allow long-term live-imaging of the gonad without interfering with germ cell behavior. This is a significant obstacle for researchers wishing to study germ cell movement and the timing of germ-line events. A recent paper by Mattingly et al. (2015) circumvented this problem using a mathematical model to reconstruct information about germ cell movement from still images. The authors began by making the reasonable assumption that germ cells can not accumulate anywhere in the gonad. It follows that there must be a relationship between germ cell production and germ cell flux, in which an increase in cell production is associated with faster movement, preventing accumulation. This relationship is captured by a differential equation that contains variables representing cell production, $S(x)$, apoptosis, $R(x)$, and cell flux, $J(x)$ (Fig. 5B). All of these quantities are functions of x , the distance from the distal tip cell, measured in microns.

An expression for cell velocity, $v(x)$, can be derived from the differential equation (Fig. 5B) by writing the cell flux as $Nf(x)v(x)$, where N is the total number of distal germ cells and $f(x)$ is the probability density function of x (Fig. 5C). A few more logical steps produces an equation for $t(x)$, the time a germ cell takes to travel from the distal tip to position x (Fig. 5D). At this stage, the expression for t contains a number of unknown functions and parameters. $f(x)$ can be determined based on z-stack images, which leaves only the cell birth rate, apoptosis rate, and total number of distal cells as undetermined. Experimental measurements, however, provide a feasible range for each of these parameters. The authors therefore tested a large number of plausible parameter sets in the model, and for each one calculated the predicted time a cell would take to travel from the distal tip to the turn. A

parameter set was chosen for which this prediction was also consistent with experimental estimates.

The derived equation (Fig. 5D) can then be used to calculate, for any germ cell in a still image, approximately how long ago that cell lost contact with the distal tip cell. The authors were thus able to plot the activation of MPK-1/ERK (a regulator of meiosis and oogenesis) as a function of time since leaving the distal tip. Based on pooled data from several gonad arms, they concluded that ERK seems to be temporally regulated, with activation beginning 30 hrs after leaving the distal tip and peaking at 37 hrs. This trend was not obvious from plots of ERK activity against position. The work of Mattingly et al. (2015) is therefore an excellent example of how modeling can be used to circumvent current experimental limitations, making a wider range of investigations possible.

Exploring evolutionary pressures on the *C. elegans* germ line

Germ-line stem cell behavior may be subject to a number of contradictory evolutionary pressures. On the one hand, if stem cells divide slowly, this reduces the length of the chain of divisions leading to each differentiated daughter. Since each cell division carries a risk of introducing a mutation during DNA duplication, fewer divisions should correspond to gametes with fewer mutations. On the other hand, if germ-line stem cells divide quickly, they can produce the same number of differentiated progeny in a shorter time, resulting in faster development and earlier fertility. It is difficult to investigate experimentally which of these factors dominates.

Recently, Chiang et al. (2015) explored whether the *C. elegans* germ line is optimised for mutation resistance or developmental speed. The authors devised a computational model of a single *C. elegans* gonad arm in which the space inside the gonad was represented by a rectangular grid with one closed end, and each square within could contain exactly one germ cell (Fig. 5E). The grid was further subdivided into two zones, such that cells in the distal zone were proliferative while the remaining cells were meiotic. Following each division, a daughter cell was placed in a neighboring empty square, if available; neighboring cells were otherwise pushed aside, causing germ cells to move in a distal-to-proximal direction through the simulated organ. Finally, the model incorporated an optional cell-cycle-length gradient in the proliferative zone. When active, this gradient caused putative germ-line stem cells near the distal tip to divide at a slower rate.

The full model included six parameters: the length and width of the proliferative zone, and the germ-cell-cycle length during four different developmental stages. The authors began by searching for a set of parameter values that minimizes the number of divisions leading to each differentiated cell, and thus minimizes mutation rate. A large number of simulations were carried out, using Markov Chain Monte Carlo parameter space exploration to home in on an optimal set of values (Gilks et al., 1996) by stochastically generating and then testing new parameter sets. If a new parameter set performs better than the current choice, it is likely to be accepted; if not, it is likely to be rejected. Over time, the sequence of accepted parameters gradually approaches optimality. The authors found a lower mutation rate was achieved when the model incorporated a cell cycle gradient in the proliferative zone. In principle, slow cycling of germ-line stem cells *can* reduce mutation rates in this system. Yet

the optimal proliferative zone size predicted by the model was approximately 1.5 times larger than the proliferative zone *in vivo*. This poor fit suggests that the actual *C. elegans* germ line is not simply optimised for low mutation rates, but may also be shaped by other evolutionary pressures.

The authors then searched for a parameter set that would not substantially delay development, while still keeping the mutation rate as low as possible. This time, the optimal parameter set included more realistic dimensions for the proliferative zone. Moreover, the model predicted that a twofold difference in rate between the fastest and the slowest cycling germ cells would be ideal for minimizing mutations. Spatial differences had previously been reported in the cell-cycle rate of germ cells along the *C. elegans* gonad (Maciejowski et al., 2006), and Chiang et al. (2015) carried out some additional measurements. The new results supported the existence of a gradient of roughly the magnitude predicted by the model. The authors therefore suggested that the slower cycling of distal germ cells in *C. elegans* may be an evolutionary adaptation to minimize the germ-line mutation rate. This adaptation appears to have evolved alongside the important constraint that reproductive development must not be substantially delayed. While it will always be difficult to address questions about the priorities and trade-offs driving evolution, this study facilitated a consideration of the relative effects of possible evolutionary pressures.

TESTING HYPOTHESES IN CONTEXT, AND GENERATING NEW PREDICTIONS

So far, we have discussed models of one part of an organ or tissue, developed with a very specific biological question in mind. In such models, sufficient detail was included to investigate the question at hand, while extraneous details were ignored. Yet, many circumstances exist in which a more complex computational model is necessary. For example, if a large number of processes significantly affect system behavior, it may not be possible to neglect any one of them, particularly if the intent is to make quantitative predictions. Detailed, whole-organ models allow a range of hypotheses to be tested in context, highlighting problems and potentially generating new predictions. We illustrate this point using examples drawn from *C. elegans* germ line modeling. While these models are, as a rule, multi-purpose, the first section emphasizes investigations of proliferative germ cell behavior while the second section focusses more on germ-line apoptosis and meiotic progression.

Exploring proliferative germ cell behavior in *C. elegans*

In 2012, Setty et al. produced one of the first computational models of the whole *C. elegans* germ line, and used it to investigate germ cell behavior in the context of the gonad. The model represented the space inside each *C. elegans* gonad arm by a U-shaped grid (Fig. 6A). By changing the grid shape over the course of a simulation, the model was able to capture larval gonad growth and development. Germ cells occupied squares within the grid, and were permitted to move to a neighboring empty square at regular intervals. The direction of germ cell movement was largely random, but with a slight encoded bias for movement toward the uterus during the adult stages to maintain cell density.

Individual germ cell behavior was modeled using a statechart – a visual formalism for representing the behavior of complex systems (Harel, 1987). Statecharts contain a number of boxes corresponding to different states and arrows corresponding to the transitions between them (Fig. 6B). The main advantage of statecharts over a traditional programming language is its visual, intuitive representation, making the assumptions in the model explicit and concise. Intricate system behavior was encoded by requiring certain transitions to be conditional. Meanwhile, by nesting states and by dividing the chart into regions that run in parallel, the diagram of system behavior can be kept simple and readable.

The Setty et al. (2012) statechart governing germ cell behavior included regions dealing with i) the state of the extracellular GLP-1/Notch receptor; ii) the state of the cell cycle; and iii) meiotic progression (Fig. 6B). This model enabled the authors to investigate a wide range of hypotheses and perturbations in a whole-organ context. First, they varied the range over which the distal tip cell was able to signal to germ cells. This change had a significant effect in the model, wherein a short signaling range frequently led to proliferative zone loss while a longer signaling range maintained proliferation. Interestingly, in some runs, only one of the two gonad arms lost its proliferative zone – a phenotype that is also observed in some mutants. The authors therefore suggested that factors like stochastic cell movement might contribute to the incomplete penetrance of this phenotype *in vivo*.

Next, the larval cell division rate was perturbed *in silico*. A low rate of larval division led to a greater-than-expected drop in the number of adult proliferative cells, a prediction that was subsequently confirmed experimentally. The same computational model was then used to study the pattern of germ cell differentiation that occurs when distal tip cell signaling is interrupted in the adult. The authors were able to simulate a specific pattern of differentiation described in a previous publication (Cinquin et al., 2010), offering an alternate interpretation for how the observed behavior can be produced by functionally equivalent germ cells. Finally, individual germ cells were tracked *in silico*, measuring the time spent in the distal niche-contact region. In general, Setty et al. (2012) demonstrated how a single, detailed model allows a variety of biological questions to be addressed and compared with experimental data, particularly when that model simultaneously incorporates aspects of cell movement, division, and differentiation.

More recently, Atwell et al. (2015) proposed a three-dimensional model of the germ line, building on the earlier work of Setty et al (2012). A similar statechart representation of germ cell behavior was implemented in the more recent work, but the germ cell movement model was updated to address concerns associated with two-dimensional, grid-based movement. The first change was to move away from the grid-based model, in which cells need not obey any particular physical laws and which require assumptions about the dominant direction of cell motion. Grid-based models also keep maximal cell density constant, making it impossible to investigate density and packing over time. Instead, Atwell et al. (2015) took a mechanical approach, in which germ cells were represented by spheres that exerted a repulsive force on their overlapping neighbors (Fig. 6C). Cells moved according to Newton's law of motion, depending on the net force they experienced, with drag assumed to dominate over acceleration. The mechanics code made use of a C++ library called Chaste, software that had previously been used to simulate tissue renewal in the intestinal crypt

(Mirams et al., 2013; Dunn et al., 2013). Changing the shape of the gonad boundary over time allowed the study of both larval development and adulthood.

The output produced by the Atwell et al. (2015) model successfully reproduces key features of the germ line during development and adulthood, and was thus used to investigate several aspects of germ-line development and homeostasis. For example, simulated germ cell tracking and clonal labeling were conducted over the course of gonad development. The model predicted that germ cells roughly maintain their relative position in the gonad during the larval stages, with cells close to the distal tip cell in the larva giving rise to the proliferative zone while cells close to the developing uterus in the larva generate sperm. This is an example of an investigation that cannot as yet be performed experimentally, but for which results can nevertheless be predicted from a detailed whole-organ model.

The process of developing a three-dimensional model also highlighted previously overlooked aspects of germ-line development and revealed possible gaps in our knowledge. For instance, when larval gonad development was simulated by simply forming a tube behind the migrating distal tip cell, germ cells often failed to fill the available space inside the gonad, causing a gap to form behind the distal tip cell and eventually leading to proliferative zone loss. This modeling result prompted new experimental measurements that pointed to significant lengthening of the proximal gonad during the late larval stages. Incorporating this “stretching” into the model prevented unrealistic outcomes. In another example, maintaining a constant number of proliferative cells proved difficult in the model. Even assuming a high cell death rate, the proliferative cell count tended to increase unrealistically over time. The authors therefore hypothesized that a form of mechanical feedback might exist, preventing germ cells from dividing as rapidly when they are densely packed. This “contact inhibition”-like mechanism was incorporated into the model, after which homeostasis was achieved. Subsequently, new experiments using mutants that blocked germ-line progression and caused distal crowding showed behavior consistent with a mechanical feedback mechanism in vivo (see details in Atwell et al., 2015). These examples illustrate how a comprehensive model can lead to new, testable hypotheses and contribute to the cyclical process of model refinement.

Exploring the regulation of germ-line apoptosis and meiotic progression

Other whole-organ modeling work on the *C. elegans* germ line has produced interesting insights into the control of germ-line apoptosis, meiosis, and oogenesis. An example is the work of Beyer et al., whose two-dimensional model of the germ line focused on the adult developmental stage (Beyer et al., 2012a). Rather than confining germ cells to a grid, however, a more physically realistic “molecular dynamics” approach was taken, with germ cells represented by circles that repel each other when they overlap (similar to the model of Atwell et al. (2015)). Cell positions and velocities were updated at each simulation time step according to Newton's law of motion ($F = ma$). Beyer et al. also modeled individual germ cell behavior by running a program inside each cell that updated a collection of variables representing the cell state. These variables included differentiation status, growth rate, and the levels of certain signaling factors. The simulated gonad was divided into regions, and each region provided a different signaling environment to germ cells. For example, cells

close to the distal tip cell accumulated Notch signaling activity while cells further from the distal tip cell accumulated GLD-1 activity, and cells still further down the gonad accumulated Ras (a protein involved in meiotic progression) activity. Behaviors like meiotic entry were driven by the relative levels of these signals.

Beyer et al. demonstrated that their initial computational model could reproduce realistic cell division and cell death rates; their subsequent model (Fig. 7A) made further improvements (Beyer et al., 2012b). The two-dimensional gonad arm was wrapped to form a cylinder that better reflects the geometric arrangement of germ cells around a central cytoplasm reservoir (see Fig. 5A). The improved model was used to investigate germ-line apoptosis, which occurs with high frequency even among normal oogenic germ cells in *C. elegans*. The authors simulated various mechanisms to select which germ cells undergo apoptosis, ranging from random cell death to density-dependent apoptosis. They concluded that apoptosis of the largest germ cells with the longest developmental time yield the most realistic-looking simulations, with stable ovulation and death rates. Computational modeling thus enabled the simulation of hypotheses about a specific regulatory mechanism, linking these hypotheses with their predicted whole-organ effects.

A more recent publication by Hall et al. (2015) substantially extended this work to a three-dimensional model of the adult germ line (Fig. 7B). Hall et al. used an updated version of the Beyer et al. (2012) molecular dynamics approach to control germ cell movement, adding random Brownian motion of cells and determining germ cell behavior by a new qualitative network model (Fig. 7C). A qualitative network consists of a number of components (genes or proteins) linked by positive or negative interactions. Each component has an activity in some finite range, and activities are updated using target functions. A target function takes in the activities of all the components connected to a particular node, then computes the node's new activity from them (Schaub et al., 2007).

Hall et al. initially used their model to study germ cell lineages and clonal groupings in the simulated adult gonad. They found that germ cells in the proliferative zone tended to move around the gonad circumference more than meiotic cells, which were simply pushed directly toward the uterus. The authors suggested that this germ cell mixing might be biologically important, in that it reduces the probability that progeny from a single genetically damaged cell would dominate the stem-cell pool. Various mutant germ lines were also simulated, including a mutant with faster cell division and a mutant lacking apoptosis. The model germ line was robust to changes in proliferation rate: when cell division was increased, a backlog of cells developed in the apoptotic zone, leading to a compensatory high cell death rate. Loss of apoptosis, however, caused gradual morphological changes. Interestingly, to properly model a loss-of-apoptosis mutant required a feedback mechanism to prevent over-proliferation: germ cells under high pressure were not allowed to divide, which is a concept similar to the contact-inhibition mechanism proposed and tested by Atwell et al. (2015).

Finally, Hall et al. used computational verification techniques to perform an exhaustive analysis of a simplified version of their model. By considering only the changes in a single cell's qualitative network as that cell progressed through a linear sequence of signaling environments, the authors determined all possible sequences of fates that a cell might adopt.

They concluded that the gonad structure helps ensure correct fate progression by exposing cells to a strict sequence of environments. Moreover, a number of hypotheses were tested for how Ras might be activated in germ cells at the beginning of the diplotene stage of meiosis. The authors concluded that persistent activation of Ras, after a certain amount of time spent in the pachytene stage, is more likely than Ras activation by a localized ligand since timed activation produced more robust meiotic progression in simulations – a conclusion similar to that reached by Mattingly et al. (2015). Exhaustive fate analysis of this kind is currently practically feasible only for relatively simple models of cell behavior; however, complex and/or hybrid models may soon become feasible with additional progress in the computational methods and more advanced algorithms.

CONCLUDING REMARKS

Computational modeling has already contributed greatly to our understanding of reproductive biology in a number of organisms and areas. In the case of spermatogenesis in mammals and *Drosophila*, computational models have distinguished between hypothesized stem cell maintenance mechanisms, resulting in the re-assessment of the hypothesized long-lived, asymmetrically dividing stem cells in favor of more flexible stochastic loss and replacement. Modeling has also been used to extract more quantitative information from the available data, allowing values like the frequency of stem-stem division or the size of a mutant cell's competitive advantage to be inferred from experiments. In *C. elegans* research, computational modeling has helped address problems that are difficult to approach experimentally, such as questions concerning germ cell movement or evolutionary pressures that shape the germ line. Finally, increasingly detailed whole-organ models enabled the study of precise perturbations to the *C. elegans* germ line. The modeling process has also highlighted gaps in our knowledge of this system, leading to new testable hypotheses.

This review has touched on just a few of the techniques used in computational modeling, from mathematical equations that describe some value as a function of time to tissue models composed of grids of cells that divide and move according to a set of rules to cell population models that incorporate mechanical forces. A crucial point is that models do not need to include every detail of a biological system to be useful. In fact, many valuable insights have come from simple models that reduce the system to a small number of variables and equations, and do not aim to produce visually realistic simulations. The key is that a model should contain sufficient detail to address the questions at hand, while remaining tractable. It is our expectation that computational modeling methods will be increasingly used to drive hypothesis testing, interpretation, and prediction. Moreover we envision that computational modeling will become a key component of the scientific method.

Acknowledgments

Supporting grants:

EP/G037280/1 to the SABS Industrial Doctorate Centre; NIH R01GM61706, NIH R01GM102254 and R03HD084660 to EJA.H.

References

- Amoyel M, Simons BD, Bach EA. Neutral competition of stem cells is skewed by proliferative changes downstream of Hh and Hpo. *EMBO J.* 2014; 33:2295–2313. [PubMed: 25092766]
- Atwell K, Qin Z, Gavaghan D, Kugler H, Hubbard EJA, Osborne JM. Mechano-logical model of *C. elegans* germ line suggests feedback on the cell cycle. *Development.* 2015; 142:3902–3911. [PubMed: 26428008]
- Beyer A, Eberhard R, Piterman N, Hengartner MO, Hajnal A, Fisher J. A dynamic physical model of cell migration, differentiation and apoptosis in *Caenorhabditis elegans*. *Adv Exp Med Biol.* 2012a; 736:211–33. [PubMed: 22161331]
- Beyer, A., Eberhard, R., Piterman, N., Hengartner, MO., Hajnal, A., Fisher, J. Predictive modelling of stem cell differentiation and apoptosis in *C. elegans*. In: Lones, MA.Smith, SL.Teichmann, S.Naef, F.Jonsson, A., Trefzer, MA., editors. *Information Processing in Cells and Tissues (Proceedings)*. Berlin-Heidelberg: Springer; 2012b. p. 99-104.
- Byrne HM, Alarcon T, Owen MR, Webb SD, Maini PK. Modelling aspects of cancer dynamics: a review. *Phil Trans R Soc A.* 2006; 364:1563–1578. [PubMed: 16766361]
- Chiang M, Cinquin A, Paz A, Meeds E, Price CA, Welling M, Cinquin O. Control of *Caenorhabditis elegans* germ-line stem-cell cycling speed meets requirements of design to minimize mutation accumulation. *BMC Biol.* 2015; 13:51. [PubMed: 26187634]
- Choi S-K, Yoon S-R, Calabrese P, Arnheim N. A germ-line-selective advantage rather than an increased mutation rate can explain some unexpectedly common human disease mutations. *PNAS.* 2008; 105:10143–10148. [PubMed: 18632557]
- Cinquin O, Crittenden SL, Morgan DE, Kimble J. Progression from a stem cell-like state to early differentiation in the *C. elegans* germ line. *PNAS.* 2010; 107:2048–2053. [PubMed: 20080700]
- Clayton E, Doupe DP, Klein AM, Winton DJ, Simons BD, Jones PH. A single type of progenitor cell maintains normal epidermis. *Nature.* 2007; 446:185–189. [PubMed: 17330052]
- Dill KA, MacCallum JL. The protein-folding problem, 50 years on. *Science.* 2012; 338:1042–1046. [PubMed: 23180855]
- Doupe DP, Klein AM, Simons BD, Jones PH. The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate. *Dev Cell.* 2010; 18:317–323. [PubMed: 20159601]
- Dunn S-J, Näthke IS, Osborne JM. Computational models reveal a passive mechanism for cell migration in the crypt. *PLoS ONE.* 2013; 8:e80516. [PubMed: 24260407]
- Gilks, W., Richardson, S., Spiegelhalter, D. *Markov Chain Monte Carlo in practice*. London: Chapman and Hall; 1996.
- Goriely A, McVean GA, Röjmyr M, Ingemarsson B, Wilkie AO. Evidence for selective advantage of pathogenic FGFR2 mutations in the male germ line. *Science.* 2003; 301:643–6. [PubMed: 12893942]
- Hall BA, Piterman N, Hajnal A, Fisher J. Emergent stem cell homeostasis in the *C. elegans* germline is revealed by hybrid modeling. *Biophys J.* 2015; 109:428–438. [PubMed: 26200879]
- Harel D. Statecharts: a visual formalism for complex systems. *Sci Comput Program.* 1987; 8:231–274.
- Harel, D., Kugler, H. Some thoughts on the semantics of Biocharts. In: Manna, Z., Peled, DA., editors. *Time for Verification*. Springer; 2010. p. 185-194.
- Klein AM, Nakagawa T, Ichikawa R, Yoshida S, Simons BD. Mouse germ line stem cells undergo rapid and stochastic turnover. *Cell Stem Cell.* 2010; 7:214–224. [PubMed: 20682447]
- Lehmann R. Germline Stem Cells: Origin and Destiny. *Cell Stem Cell.* 2012; 10:729–739. [PubMed: 22704513]
- Lopez-Garcia C, Klein AM, Simons BD, Winton DJ. Intestinal stem cell replacement follows a pattern of neutral drift. *Science.* 2010; 330:822–825. [PubMed: 20929733]
- Maciejowski J, Ugel N, Mishra B, Isopi M, Hubbard EJA. Quantitative analysis of germline mitosis in adult *C. elegans*. *Dev Biol.* 2006; 292:142–151. [PubMed: 16480707]

- Mattingly HH, Chen JJ, Arur S, Shvartsman SY. A transport model for estimating the time course of ERK activation in the *C. elegans* germline. *Biophys J*. 2015; 109:2436–2445. [PubMed: 26636953]
- Mirams GR, Arthurs CJ, Bernabeu MO, Bordas R, Cooper J, Corrias A, Davit Y, Dunn S-J, Fletcher AG, Harvey DG, Marsh ME, Osborne JM, Pathmanathan P, Pitt-Francis J, Southern J, Zemzemi N, Gavaghan DJ. Chaste: an open source C++ library for computational physiology and biology. *PLoS Comput Biol*. 2013; 9:e1002970. [PubMed: 23516352]
- Nakagawa T, Nabeshima Y, Yoshida S. Functional identification of the actual and potential stem cell compartments in mouse spermatogenesis. *Dev Cell*. 2007; 12:195–206. [PubMed: 17276338]
- Peter IS, Faure E, Davidson EH. Predictive computation of genomic logic processing functions in embryonic development. *Proc Natl Acad Sci*. 2010; 109(41):16434–42.
- Qin J, Calabrese P, Tiemann-Boege I, Shinde DN, Yoon S-R, Gelfand D, Bauer K, Arnheim N. The molecular anatomy of spontaneous germline mutations in human testes. *PLoS Biol*. 2007; 5:e224. [PubMed: 17760502]
- Rulands S, Simons BD. Tracing cellular dynamics in tissue development, maintenance and disease. *Curr Opin Cell Biol*. 2016; 43:38–45. [PubMed: 27474807]
- Salzmann V, Inaba M, Cheng J, Yamashita YM. Lineage tracing quantification reveals symmetric stem cell division in *Drosophila* male germline stem cells. *Cell Mol Bioeng*. 2013; 6:441–448. [PubMed: 24465278]
- Schaub MA, Henzinger TA, Fisher J. Qualitative networks: a symbolic approach to analyze biological signaling networks. *BMC Syst Biol*. 2007; 1:1–21. [PubMed: 17408505]
- Schedl, T. *Germ Cell Development in C. elegans*. Springer Science & Business Media; 2013.
- Setty Y, Dalfó D, Korta DZ, Hubbard EJA, Kugler H. A model of stem cell population dynamics: in silico analysis and in vivo validation. *Development*. 2012; 139:47–56. [PubMed: 22147952]
- Sheng XR, Matunis E. Live imaging of the *Drosophila* spermatogonial stem cell niche reveals novel mechanisms regulating germline stem cell output. *Development*. 2011; 138:3367–3376. [PubMed: 21752931]
- Simons BD, Clevers H. Strategies for homeostatic stem cell self-renewal in adult tissues. *Cell*. 2011; 145:851–862. [PubMed: 21663791]
- Tomlin CJ, Axelrod JD. Biology by numbers: mathematical modelling in developmental biology. *Nat Rev Genet*. 2007; 8:331–340. [PubMed: 17440530]
- Tyson JJ, Novak B. Regulation of the eukaryotic cell cycle: molecular antagonism, hysteresis, and irreversible transitions. *J Theor Biol*. 2001; 210:249–63. [PubMed: 11371178]
- Yoon S-R, Choi S-K, Eboreime J, Gelb BD, Calabrese P, Arnheim N. Age-dependent germline mosaicism of the most common Noonan syndrome mutation shows the signature of germline selection. *Am J Hum Genet*. 2013; 92:917–926. [PubMed: 23726368]
- Yoshida S. From cyst to tubule: innovations in vertebrate spermatogenesis. *WIREs Dev Biol*. 2016; 5:119–131.

Figure 1. The “virtuous cycle” between modeling and experiment

An experiment suggests a new biological hypothesis, which can then be expressed as a model informed by the available data. The model helps to interpret the data, and generates predictions for experimental testing. A correct prediction increases confidence in the model and the hypothesis it represents, while an incorrect prediction demands that the model be refined.

Figure 2. Comparing different mechanisms of spermatogenic stem cell maintenance

A: Mammalian spermatogenesis. Stem cells in the basal layer of the seminiferous tubule divide, and their daughters gradually move toward the lumen as they develop into sperm. **B:** Clonal labeling results for mouse seminiferous tubule (reproduced, with permission, from Klein et al., 2010). Left, stained clones after 3 months; clone length was measured along the tubule. Right, the average clone length grows proportional to the square root of time. **C:** $P_n(t)$ denotes the probability that a clone hosts n stem cells at time t . This equation describes how $P_n(t)$ changes over time, assuming that labeled and unlabeled cells have equal proliferative capacity, and that symmetric stem cell division occurs in response to stem cell

loss (see text for details). **D**: The solution to the equation in 'C', given in terms of modified Bessel functions of the first kind, $I_n(x)$. The accompanying graph illustrates the behaviour of $I_n(x)$ for several values of n . **E**: Calculating the probability of a surviving clone hosting n stem cells at time t . **F**: Calculating $\langle n \rangle$, the average number of stem cells in a clone. At late times, $\langle n \rangle$ grows proportional to the square root of time.

Figure 3. Modeling the spread of a mutation among human spermatogenic stem cells

An illustration of the Qin et al. (2007) model concept. The testis is represented by a three-dimensional (3D) lattice containing dividing spermatagonial cells. If a mutation arises during the growth phase of the simulation, it spreads rapidly since the stem cell pool is still expanding by symmetric division (scenario 1). During the adult phase, a mutation cannot spread if stem cells divide strictly asymmetrically (scenario 2), whereas mutant clones can grow over time if symmetric divisions are permitted (scenario 3).

Figure 4. Determining the frequency of symmetric divisions in *Drosophila* spermatogenesis

A: *Drosophila* spermatogenesis. The hub, a collection of somatic cells, serves as a stem cell niche for both germ-line stem cells and somatic cyst stem cells. Germ-line stem cells divide, producing gonialblasts that are then surrounded by cyst cells. Mature sperm are generated from gonialblasts after four transit-amplifying divisions and spermiogenesis. **B:** An illustration of the Salzmann et al. (2013) model of germ-line stem cell self-renewal. The cartoon shows three ways in which a single GFP-labeled cell could be produced, starting from either a labeled singlet, doublet, or triplet at the previous cell cycle. The equation describes how the number of singlets will change over time, taking into account all three

possible scenarios. The full Salzmänn et al. model also includes equations for the number of labeled doublets and triplets, as well as a parameter, f_{ss} , representing symmetric stem-stem division frequency.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 5. Using modeling to address experimentally intractable questions

A: The *C. elegans* hermaphrodite germ line. A single adult gonad arm is shown (only the germ cells and distal tip cell are depicted; other somatic gonadal cells are not represented). The gonad arm contains a dividing germ cell population at its distal end, maintained by signals from the distal tip cell. As germ cells move down the gonad, they enter meiosis (mitotic and meiotic regions are not to scale). In late larvae, meiotic germ cells develop into sperm, while in adults they differentiate into oocytes; many oocyte-fated germ cells undergo apoptosis. **B-D:** A mathematical model of *C. elegans* germ cell movement (Mattingly et al., 2015). Equation 'B' expresses the assumption that at position x along the gonad, any

imbalance between cell production, S , and apoptosis, R , must cause a change in the germ cell flux, J . Equation 'C' then shows how cell velocity can be solved, in terms of s (the cell birth rate), N (the total number of distal cells), R , and $f(x)$ (the probability distribution of x). Finally, 'D' presents a formula for $t(x)$: the time taken for a cell to travel from the distal tip to position x . Here, r is the apoptosis rate and R^* is a rescaling of R . This equation can be used to estimate the time elapsed since a germ cell left the distal tip, using only static image data. **E:** A model of the *C. elegans* germ line (reproduced, with permission, from Chiang et al., 2015). The gonad tube is "unwrapped" and modeled as a rectangular lattice, in which each square can contain at most one germ cell (circles). Dividing cells fill the tube in a distal to proximal direction, pushing their neighbors aside as necessary. On entering the meiotic zone, cells queue up and eventually undergo apoptosis or become gametes.

Figure 6. Whole *C. elegans* germ line modeling, I

A: A lattice-based model in which each germ cell occupies a certain number of squares according to its differentiation status (modified and reproduced, with permission, from Setty et al., 2012). Germ cells regularly move to adjacent, unoccupied squares. Yellow, proliferative; orange, meiotic prophase; green, oocyte; blue, sperm; and grey, embryo. A labeled photomicrograph of a *C. elegans* gonad arm is shown alongside for comparison. **B:** Part of the statechart governing germ cell behavior in the model shown in 'A' (modified and reproduced, with permission, from Setty et al., 2012). Included regions controlling the cell cycle, differentiation, GLP-1 receptor state, and movement direction. **C:** A three-dimensional mechanical model of the germ line in which spherical cells repel each other when they overlap (reproduced, with permission, from Atwell et al., 2015). The gonad boundary grows over time, capturing larval development, while cell behavior is governed by an internal statechart. Yellow, mitotic (plus meiotic S); green, meiotic; pale blue, sperm-fated; pink, oocyte-fated; dark blue, sperm; and cyan, distal tip cell nucleus.

Figure 7. Whole *C. elegans* germ line modeling, II

A: A molecular dynamics model of the germ line (reproduced, with permission, from Beyer et al., 2012a). Germ cells are modeled as circles that exert force on each other when they overlap. Cell behavior is governed by an internal program that responds to the signaling environment associated with a cell's current location. **B:** A three-dimensional model of the adult germ line based on molecular dynamics (repulsion forces plus Brownian motion). (reproduced, with permission, from Hall et al., 2015). Pictured is a model snapshot, with cells colored as follows: red, proliferative; orange, early meiosis; green, diplotene; and blue, diakinesis. Below the snapshot is a diagram indicating which signals are associated with different gonad regions. **C:** A qualitative network governing germ cell behavior in the model

shown in 'B'. This network specifies how the activity of intracellular molecules changes in response to environmental signals.