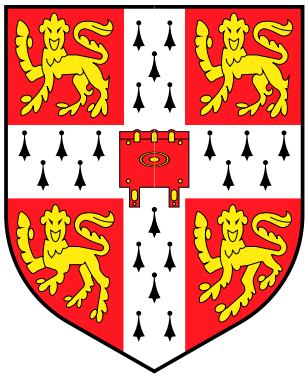
The laws of cell fate in epidermal maintenance

Allon M. Klein
Clare College, Cambridge



Dissertation submitted
for the degree of Doctor of Philosophy

February 2008

Abstract

A central goal of biology is to identify rules that govern cell fate in multicellular tissue: rules that determine whether a cell will divide or differentiate into a specialised cell type. One may look for a signature of such rules in the statistics of large populations of cells. Here we consider the case of progenitor cells in *adult mammalian epidermis*, where we infer a stochastic mechanism of cell fate during steady-state tissue turnover.

This work overturns the widely accepted paradigm of epidermal maintenance, which states that epidermis is supported by two proliferating cell populations — long lived stem cells that can renew themselves indefinitely, and their short-lived progeny that undergo only a limited number of rounds of division. We show that observations of (i) scaling of size distributions of cell families (clones) in transgenic mice, (ii) spatial correlations between proliferating cells, are consistent with a model of steady-state epidermal maintenance involving a *single* population of progenitor cells that undergo a critical birth-death process.

The resulting model establishes a quantitative platform for studying the effects of drugs and mutations on the epidermis, without engaging with the complex biochemical circuitry regulating cell fate. To exemplify the utility of the model, we use it to infer the effect of a non-disruptive drug (all-trans retinoic acid) on the rates of cell division and differentiation in transgenic mice, and we then use the revised rates to predict clone size distributions following an extended drug treatment.

Finally, drawing upon observations of spontaneous large-scale patterning of human epidermal cells, we infer a mechanism by which human epidermal stem cells spontaneously organise into quiescent clusters during homeostasis, whilst becoming active in response to spatial disruption of the tissue. This mechanism offers a unified model of epidermal maintenance in mice and men, consisting of a committed progenitor cell population whose stochastic behaviour enables stem cells to remain quiescent unless needed for tissue repair.

Preface

This dissertation describes work done between October 2005 and April 2008 at the Cavendish Laboratory, Cambridge, under the supervision of Prof. Benjamin D. Simons.

The introductory chapter 1 contains reviews of previous work on which this thesis is based. Chapters 2*, 3† and 5‡ contain material that has been published in the literature. The experiments described in chapter 2 and in chapters 4-7 were carried out by Elisabeth Clayton, David P. Doupé and Phil H. Jones at the Hutchinson laboratory in Cambridge[1].

* E. Clayton *et al.*, Nature **445**, 185–9 (2007).
† A. M. Klein *et al.*, Phys. Rev. E **76**, 021910 (2007).
‡ A. M. Klein *et al.*, Phys. Rev. E (accepted).

This dissertation, except where specifically indicated in the text, is the result of my own work and includes nothing which is the outcome of work done in collaboration. No parts of this thesis have been submitted for any other qualification.

Statement of length

This dissertation does not exceed 60,000 words, including tables, footnotes, references and appendices.

Allon M. Klein

[1] *MRC Cancer Cell Unit, Hutchison-MRC Research Centre, Cambridge CB2 2XZ, UK*

Acknowledgements

I would like to thank the wonderful people behind this work for their excitement and patience, and not least for their efforts to communicate with each other across the curious biological-physical cultural divide: special thanks go to David Doupé and Phil Jones for two years of "making Ben and Allon happy[2]"; for their clear exposition of cell biology; and for producing such exciting experimental results. Thanks also to Elizabeth Clayton for generating the outstanding data that started the entire collaboration on cell fate. Most of all I would like to thank my supervisor Prof. Ben Simons for his close mentoring and direction.

I am grateful to Prof. Sam Edwards for his insightful comments about the significance of scaling of distribution functions.

Finally, I would like to thank my wife Roohia for encouraging me to return to physics; for supporting me during these two years; and for caring about every obstacle and success that has followed. This thesis is dedicated to her.

[2]Phil Jones, *verbatim*.

Contents

Chapter 1

Introduction

1.1 Motivation: Cell tissue as a collective phenomenon

A central goal of biological research is to understand the laws governing cell fate in living tissue, laws that predict when a cell will divide or differentiate into a specialised cell type. Beyond their ability to explain how normal tissues develop and are maintained, such laws have important implications for explaining the mechanisms of aging, wound healing and cancer.

Central to the question of cell fate is the study of adult (or somatic) *stem cells*. These are cells that are specialised to different tissue types, and are considered to be responsible for generating new cells upon demand: either in response to wounding, or to replace cells lost during routine tissue turnover [1, 2]. Adult stem cells are believed to be long-lived, capable of sustaining their own population throughout the lifetime of an organism [1]. Due to this property of self-renewal, adult stem cells are endowed with unique control over their own fate: they may either divide to regenerate their own population, or they may undergo a process of differentiation, by which they lose their stem-like properties and become committed to an irreversible genetic program that endows them with new functional properties within the tissue.

While the possible channels of cell fate may in principle be defined in terms of the choice between self-renewal and differentiation, the process followed by individual cells in determining their final fate remains poorly understood. At a fundamental level, such "cellular choices" are the end result of complex molecular regulation. To understand the laws of cell fate, the biological community has traditionally worked to identify the central constituents of molecular regulation, i.e. specific genes, proteins and metabolites that control cell fate [3]. This form of analysis has been described as a "reductionist approach", as it supposes that, despite the potential complexity of the biochemical circuits, only a small number of molecular constituents determine cell fate [3].

Yet, although a wide range of biochemical pathways regulating adult stem cell behaviour have now been identified [4, 5, 6], one is challenged to find convincing concepts and methods to comprehend how these constituents interact to give rise to the cell fate choices of stem cells, and their progeny, at the cellular or tissue level [3]. As a result, the classification of cell fate in one system may fail to

explain cell behaviour in another, such as when comparing cell behaviour in healthy tissue to that seen during wound repair, cancer, or incubation in culture. More significantly, it is far from clear that one may consistently identify a set of individual genes that dominate the resulting function of cells in multicellular tissue. Rather, one may expect the observed properties of the cells to result from a complex interplay between different gene products.

Systemic and phenomenological approaches to modelling cell fate: Therefore, rather than following a reductionist viewpoint, the pluralism of causes and effects in biological systems must either be explicitly addressed, or carefully overcome. One is led to consider two alternative methods:

- One may attempt to observe, through quantitative measures, the multiple intra-cellular components simultaneously, viz. gene expression levels, chemical concentrations, etc.. Then, by rigorous data integration with mathematical models, one may develop a system-wide perspective (so-called *systems biology*) on component interactions, revealing the dynamical properties of the cell and tissue systems [7]. From here, large-scale characteristics such as a particular functional state of the cells and tissue, or robustness [8], can be quantitatively understood and rationally manipulated.
- Or, one may attempt to directly observe quantitative measures that characterise the *entire* cell tissue population, such as details of the cell lineage or the large-scale organisation and patterning of cells in the tissue. By developing mathematical models that capture the observed cell population behaviour, one may *infer* the relevant underlying dynamics of the cells. In contrast to both the reductionist approach and canonical systems biology, such an approach does not require prior knowledge of the relevant intra-cellular components, and it can assist in their identification.

Whereas the former (systems biology) approach has gained much popularity in recent years [9], it has been largely successful in cases where the relevant system components can been identified with comparative ease, such as in prokaryotes, simple single-celled eukaryotes, or other well-studied systems [10]. For multi-cellular tissue, on the other hand, the laws of cell fate may be significantly more complex. Yet, despite its apparent complexity, cell fate in these systems is heavily constrained by the very existence of a stable tissue capable of robust physiological function. Such large-scale constraints lead naturally to the latter (phenomenological) approach, the focus of this thesis.

1.2 A prototype tissue: Mammalian epidermis

The outer-most layer of mammalian skin, known as the epidermis, is a well-studied tissue that undergoes constant renewal throughout adult life, and is therefore a good prototype for understanding cell fate in regenerating adult tissues.

The skin and its appendages ensure a number of critical functions necessary for animal survival. Skin protects animals from water loss, temperature change, radiation, trauma, and infections, and it allows animals to perceive their environment through tactile sense. Overall, approximately 20 different cell types reside within the skin and contribute to its various functions [12] (see Fig. 1.1a)

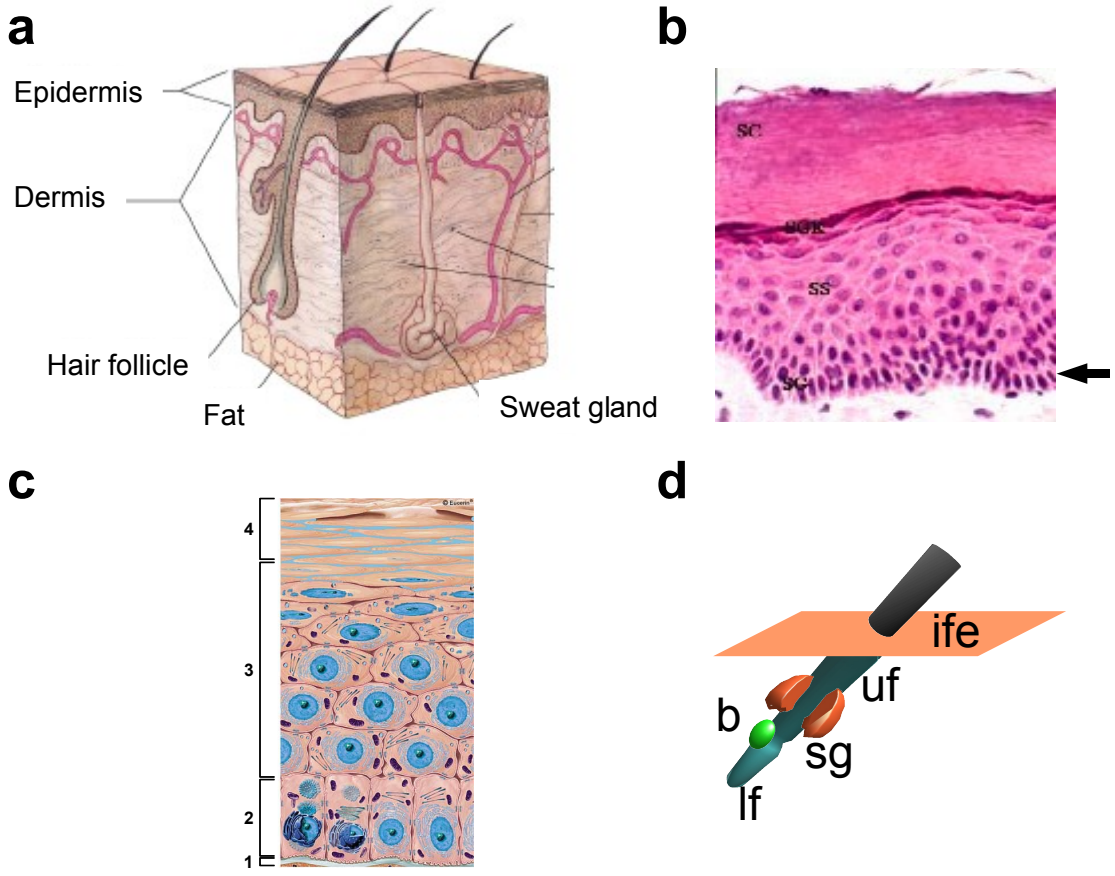


FIG. 1.1: Architecture of mammalian skin and epidermis. **a**, Mammalian skin consists of the sub-cutaneous layer (fat), the dermis and the epidermis that is interspersed with hair follicles. The dermis is a complex environment containing blood vessels, nerves, sweat glands and other cell types, whereas the interfollicular epidermis (IFE) is a simple tissue composed perdominantly of keratinocyte cells. **b, c**, Micrograph and schematic cross-sections of IFE. In the schematic, the labels correspond to (1) the basal membrane separating the epidermis from the dermis; (2) the basal layer, in which cell division is localised; (3) the supra-basal layers in which cells undergo differentiation, growing in size, filling with keratin and eventually flattening out and losing their nucleus. The cell rise to finally join (4) the water-tight cornified layer, from which they are ultimately shed. In **b**, the basal layer is indicated by the arrow. **d**, Schematic stucture of the hair follicle. Hair follicles contain stem cells located in the bulge (b, green), with the potential to generate lower hair follicle (lf), sebaceous gland (sg, orange), and upper follicle (uf) during normal maintenance, and to regenerate the entire epidermis upon wounding. These stem cells play no role in routine maintainance of the interfollicular epidermis (IFE) [11].

The epidermis, in this context, forms a water-tight layer consisting of layers of specialised cells known as keratinocytes, which are organised into hair follicles interspersed with *interfollicular* epidermis (IFE) [13] (see Fig. 1.1b-d). The IFE, in particular, consists of only *one* cell type (the keratinocyte).

In the following, we shall introduce key features of the epidermis, reserving additional details of epidermal morphology to later chapters. This section draws heavily upon reviews found at Refs. [14, 12, 13], as well as the following papers written by (or with) AMK [15, 16, 17, 18], and references therein.

Structure of the interfollicular epidermis

In the IFE, proliferating keratinocyte cells are confined to the basal epidermal layer (see Fig. 1.1c). The basal layer produces, secretes, and assembles an extracellular matrix, which constitutes much of the underlying basement membrane that separates the epidermis from the underlying dermis. As they differentiate into specialised skin cells, the basal cells withdraw from the cycle of cell proliferation and then leave the basal layer, migrating towards the epidermal surface. In the early stages of the differentiation program, the cells remain transcriptionally active. However, as they move up through the cell layers, the cells lose their nucleus and culminate in the production of dead flattened cells that form a water-tight cornified layer (Fig. 1.1c). In human epidermis, the cornified layer consists of 25-30 layers of such cells. Ultimately, these cells are shed from the skin surface. In total, estimates of the keratinocyte turnover time (the average time between cell division and shedding) are 39-45 days in humans [19], corresponding to approximately 1,500 cells turned over per day per mm^2 [20].

Progenitor cells of the IFE

To maintain the integrity of the IFE, new cells must be generated to replace those lost through shedding. To characterise this process, the progenitor cells found in the basal layer have been studied extensively. Early studies using a radio-active DNA marker that is diluted during cell division have revealed that a sub-population of progenitor cells retain their label for many weeks, whereas the bulk of cells lose their label more rapidly by dividing, on average, once every 7-8 days [12]. The so-called "label-retaining" cells have been interpreted as stem cells that divide infrequently, while the latter cells are seen to be a second, short-lived, progenitor cell population[1].

Further evidence for two progenitor cell compartments is seen in *in vitro* experiments, where cells isolated from the skin are plated in culture and allowed to form growing colonies[2] [24]. When plated at low cell density, it was found the cultured human keratinocytes can form (a) large macroscopic cell colonies (known as *holoclones*) that continue to grow upon further re-plating, (b) microscopic colonies (known as *paraclones*) in which all cells eventually undergo terminal differentiation, or (c) macroscopic but small, irregularly shaped colonies (known as *meroclones*) that, upon further re-plating, eventually give rise to paraclones, see Fig. 1.2. These experiments demonstrate a

[1] However, we will challenge this interpretation in chapter 2.

[2] The technology for such experiments was developed in the mid-1970s, when the culture conditions allowing the growth of human IFE *in vitro* were discovered [21]. As well as its application to the study of progenitor cell characteristics, this seminal discovery allows sheets of skin to be generated for use in treatment of severely burned patients [22, 23]. In the past 25 years, this technology has saved many lives.

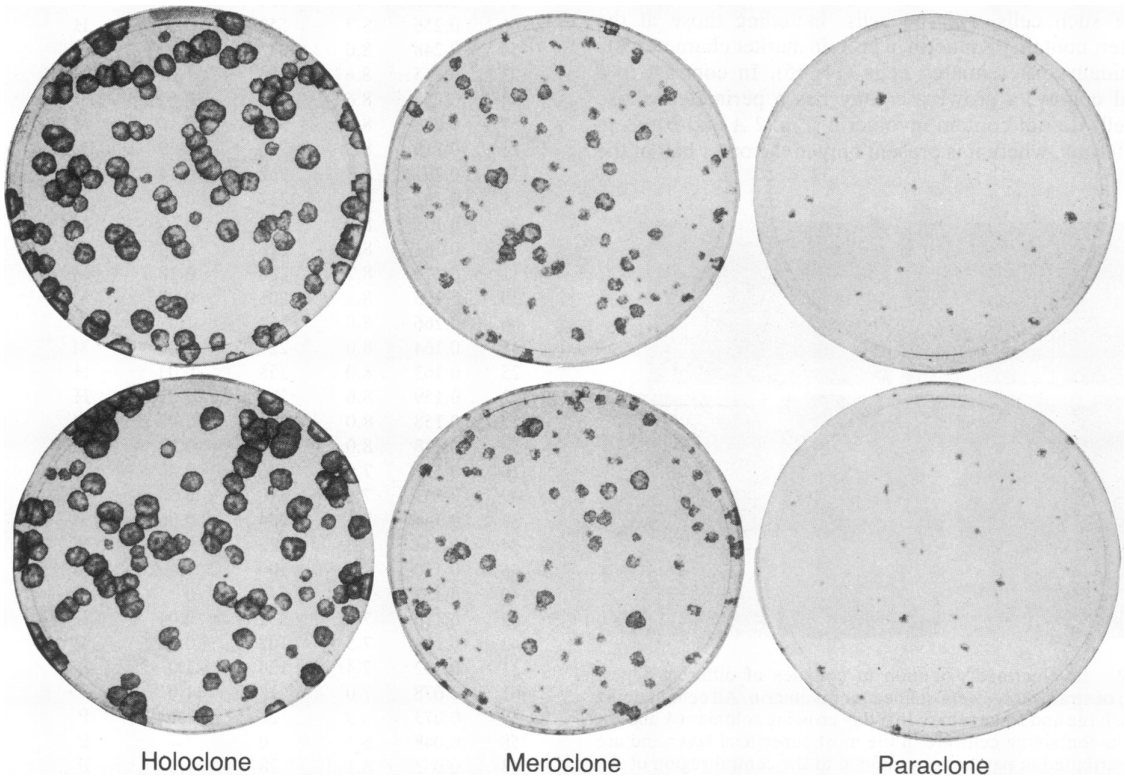


FIG. 1.2: Evidence for *in vitro* proliferative heterogeneity of human keratinocytes, after Ref. [24]. When the progeny (or clone) of a single cultured keratinocyte are replated after a 7 day incubation period, "holoclones" give rise to large and highly-proliferative colonies; "meroclones" give rise to large wrinkled colonies as well as a variable fraction of microscopic colonies in which all cells have undergone terminally-differentiation; and "paraclones" generate only microscopic colonies.

heterogeneity in the proliferation potential of epidermal progenitor cells: the highly-proliferative clonogenic cells that give rise to holoclones are considered to be stem cells, whereas paraclones must result from progenitor cells that are incapable of self-renewal through super-linear growth. Meroclones are thought to result from the proliferation of poorly-defined 'partially-differentiated' stem cells[3].

The stem/TA cell model of epidermal homeostasis

As a result of these experimental investigations, it has been thought for many years that interfollicular epidermis is maintained by two distinct progenitor cell populations in the basal layer. These comprise long-lived stem cells with the capacity to self-renew, and their progeny, known as transit-amplifying cells (TA), which go on to differentiate and exit the basal layer after several rounds of cell division [25], see Fig. 1.3. Stem cells are also found in the hair follicles, but whilst they have the potential to generate epidermis in circumstances such as wounding, they do not appear to contribute to maintaining normal epidermis [11, 26].

[3]However, it remains a challenge (which is addressed in this thesis) to identify the mechanism that gives rise to wrinkled morphology of meroclone colonies.

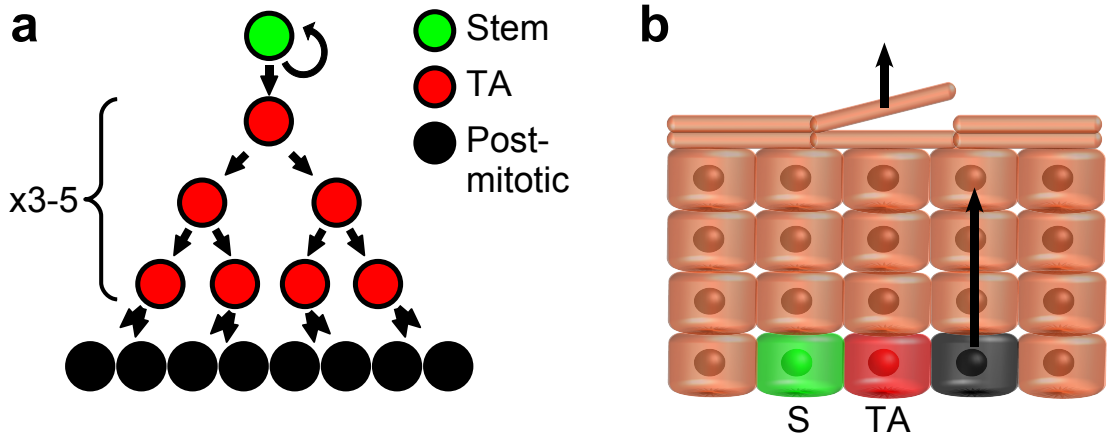


FIG. 1.3: The stem/TA hypothesis of tissue maintenance. **a** Self-renewing stem cells give rise to a population of transient-amplifying (TA) cells that undergo 3–5 rounds of cell division before differentiating into post-mitotic cells. **b** Progenitor cells are confined to the basal layer, whilst post-mitotic cells migrate through the supra-basal layers until they are shed from the skin surface.

Histological analysis has shown that mouse epidermis is organized in stacks of cells with a hexagonal surface area lying on a bed of ten basal cells [27, 28, 25], see Fig. 1.4. According to the "stem/TA" cell model of interfollicular homeostasis, this structure was hypothesized to indicate organisation of the tissue into regularly sized "epidermal proliferative units" or EPUs, in which a central stem cell supports a surrounding, clonal, population of transit amplifying cells, which in turn generate a column of overlying differentiated cells [27, 28] (Fig. 1.4c).

However, although several experimental approaches have been used to attempt to demonstrate the existence of EPUs, conclusive evidence for their existence is lacking. The EPU model predicts that slowly-cycling stem cells should be found in a patterned array in the IFE; cell labelling studies have failed to demonstrate such a pattern [30]. In chimaeric mice the EPU model predicts that the boundaries of mosaicism in the IFE should run along the boundaries of EPUs; instead boundaries were found to be highly irregular [31]. Genetic labelling studies using viral infection or mutation to activate expression of a reporter gene in epidermal cells have demonstrated the existence of long-lived, cohesive clusters of labelled cells in the epidermis, but these clusters do not conform to the predicted size distribution of the EPU [30, 32, 33, 25, 34, 35].

Thus, until recently the means by which homeostasis of IFE was achieved has been unclear. In the following chapters, drawing upon a range of experimental results, we shall challenge the EPU concept, and with it, the stem/TA model. We shall show that stem cells appear to remain largely quiescent during normal epidermal homeostasis, and that rapidly proliferating cells are capable of independently maintaining the IFE. These rapidly proliferating cells are not short-lived TA cells, but are instead regulated to behave stochastically, so that they are capable of a significant number of cell divisions before differentiating.

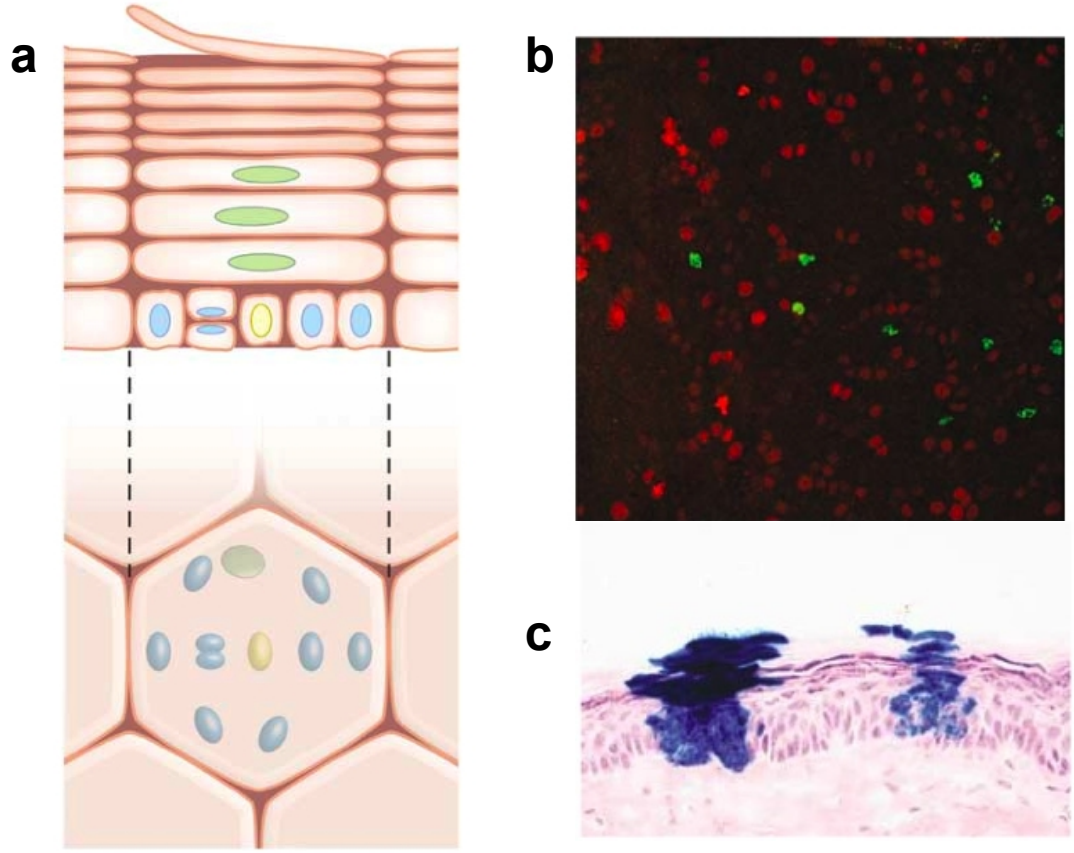


FIG. 1.4: Early evidence for the stem/TA hypothesis and epidermal proliferating units (EPUs), after Ref. [29]. **a** Schematic showing the observed structure of mouse ear epidermis [28], where the cornified-layer consists of hexagonally shaped cells that are stacked above basal layer cells including one centrally-located cell (yellow). It was first suggested that the central basal layer cell (yellow) is a stem-cell, which is surrounded by transit-amplifying progeny. However, it was later found that the central cell plays no role in epidermal maintenance. **b** Wholemount IFE stained for the cell proliferation marker Ki67, and for a label-retaining cell (LRC) marker BrdU (green). LRCs are found by introducing a marker for DNA synthesis (BrdU), which is then diluted through repeated cell divisions. Cells that retain the marker over extended periods of time are considered to be quiescent. The existence of LRCs is seen as evidence for a population of slowly cycling stem cells. **c** Genetic labelling of proliferating basal-layer cells allows the lineage of proliferating cells (clones) to be visualised. The cross-section shows column-like progeny, which were initially interpreted as evidence for EPUs. However, when imaged in 3D, the variation between the sizes of neighbouring clones calls into question the validity of the EPU model (as will be shown in chapter 2).

1.3 Theoretical models of epidermal cell fate — a brief survey

The importance of cell fate in biology and medicine has led to the development of several quantitative models of epidermal maintenance. However, as such models have largely been challenged to generate predictive laws, or to engage with new experimental data, this survey will be kept brief. Models of epidermal cell fate have attempted to engage with different aspects of the experimental system, including the basal layer *topology* and cell *age distributions*. Such models have either considered the *coarse-grained dynamics* of the cell populations, or individual *cell-based* dynamics, as described below.

An elegant example of quantitative cell fate modelling is presented in the topological approach of Dubertret and Rivier [36], who considered the basal layer as a surface of convex polygonal cells, known as a *Voronoi tesselation* (or a 'foam'), consisting of a single type of cell that is capable of both division and differentiation through upward migration[4]. The proposed model successfully reproduces the experimental *topology* of the basal layer, i.e. the distribution in the number of sides per cell, by relating the probabilities of cell division and upward migration to the number of nearest neighbours. Yet, despite its success (and elegance) in describing the steady-state topology, it is a challenge to apply the model to explain cell fate outside of the specific topological context.

Taking a different approach, Savill [39] considered the *age distribution* of cells in the basal layer. To develop a theoretical model of cell age distributions, he assumed that epidermal maintenance adheres strictly to the stem/TA cell model, which predicts that cell age should vary as a result of the differences between the stem, TA and post-mitotic cell populations, as well as due to the stochastic variability within each population. However, this study acknowledges a lack of adequate experimental data for comparison with the theory.

Other *mean-field theories* of cell fate have drawn upon variants of the *Fisher equation* in order to describe cell turnover during wound repair and tissue maintenance [40, 41]. Here, the dynamics of the cell population are described by a reaction-diffusion equation for the coarse-grained cell density, with the effects of cell division and differentiation accounted for by a logistic source term. Although the Fisher equation introduces an important regulatory principle, namely that of density-regulation of the cell division rates, current studies do not attempt to engage with the question of individual cell fate.

Beyond the mean-field, a range of individual cell-based (ICB) models have been developed to model epidermal maintenance [42, 43]. These are models that simulate the evolution of a cell population by defining characteristics and dynamical properties for each individual cell, and then simulating the dynamics of a large number cells. In the simplest of such models, the spatial distribution of cells is either ignored or modelled using a simple lattice, however several elaborate schemes exist to allow the volume and surface properties of cells to be accounted for *in silico* [44, 45]. Despite the superficial appeal inherent in the rich simulations that result from ICB models, many existing studies introduce a large number of (uncontrolled) parameters to support their simulation results[5]. As a result, such models are often challenged to develop predictive insights[6].

[4] Similar topological models have been developed in other tissue types [37, 38].

[5] Typical parameters may include cell adhesion and migration rates, local nutrient supplies, compressibility, and cell division and differentiation rates.

[6] By contrast, even very complex models may often be understood in terms of minimal key ingredients that

1.4 Outline of research

This thesis is organised around a series of publications by the author, as follows: In chapters 2 and 3, following closely to Refs. [15, 16], we describe the experimental system used for genetically labelling and tracking families of cells (clones) in mouse tail-skin, and we show how the experimental clone fate data can be used to both rule out the stem/TA cell model of epidermal maintenance, and to infer a *new* paradigm for epidermal maintenance.

In chapter 4, we then use the new model to characterise the effect of a topical drug treatment on the epidermis, allowing us to predict experimental clone size distributions in the drug-treated system.

In chapter 5, following closely upon Ref. [17], we extend the model of cell division in mouse tail-skin to account for the spatial distribution of cells within the basal layer. By drawing upon observations of clone cohesiveness to characterise the cell mobility, we predict that progenitor cells should *cluster*, in good quantitative agreement with experiment.

In chapters 6-7, we then address the question of epidermal maintenance in *humans*, drawing upon observations of steady-state stem cell patterning. We infer a patterning mechanism that is capable of accounting for a wide range of empirical observations reported in the literature, and provides a unified model of cell fate in both mice and men.

We conclude with a discussion of open questions and potential future work in chapter 8.

dominate their dynamics, such as may be understood by turning to the of field universal non-equilibrium processes and interface growth models, see e.g. [46, 47, 45].

Chapter 2

Lineage tracing experiments: challenging the stem/TA cell model

Chapter overview

The long-standing model of adult epidermal homeostasis supposes that skin tissue is maintained by two discrete populations of progenitor cells: self-renewing stem cells; and their progeny, known as transit amplifying (TA) cells, which differentiate after several rounds of cell division [48, 49, 50]. This chapter describes inducible genetic labelling experiments that may be used to test the stem/TA hypothesis. In these experiments, the fate of a representative sample of progenitor cells in mouse tail epidermis is tracked at single cell resolution *in vivo* at time intervals up to one year.

The primary aim of this chapter is to familiarise the reader with the lineage tracing experiments that will be analysed in the following four chapters of the thesis. We will demonstrate here that the resulting clone size distributions are *inconsistent* with the classical stem/TA hypothesis. The more significant implications of these results for understanding epidermal cell fate will be discussed in the following chapter (3).

This chapter draws significantly upon the work carried out by E. Clayton[1], D.P. Doupé[1] and P.H. Jones[1], also reported in Ref. [15]. Experimental work was performed by EC, DPD and PHJ, project planning by PHJ and D.J. Winton[2]. AMK and B.D. Simons conducted the data analysis. In particular, the experimental sections 2.1 and 2.2 are largely the authorship of PHJ. We have not included the full experimental methods nor the supplementary experimental results in this thesis. A reader interested in these experimental aspects is kindly referred to Ref. [15].

[1] *MRC Cancer Cell Unit, Hutchison-MRC Research Centre, Cambridge CB2 2XZ, UK*

[2] *Cancer Research UK Department of Oncology, University of Cambridge, Cambridge Institute for Medical Research, Cambridge CB2 2XY, UK*

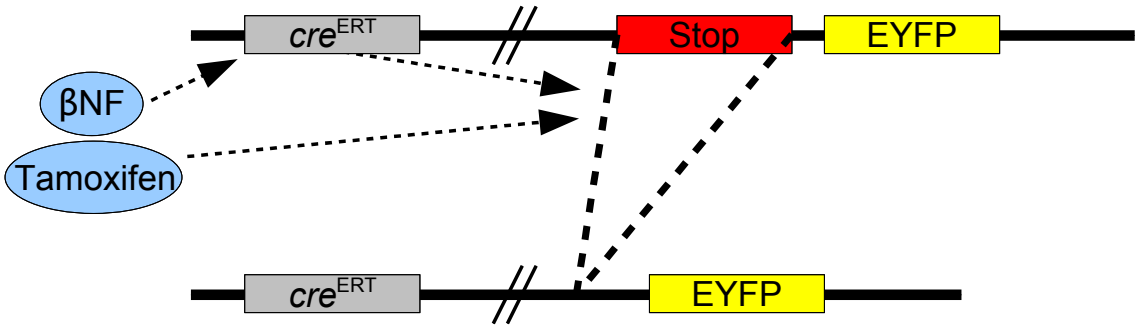


FIG. 2.1: The genetically modified mice (strain Ah*cre*$^{\mathrm{ERT}}$ R26$^{\mathrm{EYFP/wt}}$) carry two genetic modifications (top): firstly they carry the gene for an enhanced yellow flourescent protein (EYFP), whose expression was blocked by a 'stop' cassette (red). Therefore, the cells do not initially express EYFP. Secondly, the mice carry the gene for the DNA-modifying enzyme *cre*$^{\mathrm{ERT}}$ recombinase, which is targeted at the stop cassette. The recombinase is expressed following treatment with βnapthoflavone (βNF) (a drug that induces the *cre* promoter). In the presence of a second drug, Tamoxifen, *cre*$^{\mathrm{ERT}}$ mediates excision of the stop cassette (dashed lines) resulting in EYFP expression in the recombinant cell and its progeny (bottom). By administering a sufficiently low dosage of the activating drugs (βNF and Tamoxifen), the frequency of genetic labelling was kept very low (1 in 600 cells labelled). See Ref. [15] for further technical details.

2.1 The experimental system: genetically labelled clones

As described in the thesis introduction, the mammalian interfollicular epidermis (IFE) is maintained by proliferating epidermal progenitor cells (EPCs) that are found in the basal cell layer. Label retaining studies show that IFE contains slowly cycling basal cells which, following the stem/TA cell hypothesis, have been interpreted as representing stem cells that support clonal units of TA and differentiated cells [27, 28]. However, these studies are unable to reveal the dynamics of EPC behaviour during epidermal homeostasis. Previous genetic labelling studies to track the fate of proliferating cells have either required epidermal injury or have yielded too few labelled clones to permit quantitative analysis [32, 33, 34, 35].

To track the fate of EPCs in normal epidermis we have used inducible genetic marking to label a sample of cells and their progeny in adult mice. As shown schematically in Fig. 2.1, animals were genetically modified to carry the enhanced yellow flourescent protein gene (EYFP), whose expression is blocked by a removable genetic 'stop cassette'. EYFP is then expressed in response to a treatment with βNF and tamoxifen at 6-9 weeks of age, which leads to the removal of the stop-cassette and activation of the EYFP gene. Significantly, the fraction of initially labelled cells may be kept low by adjusting the treatment dosage (see caption of Fig. 2.1). Cohorts of mice were culled for analysis at intervals following a single injection of the inducing drugs. Cells expressing EYFP and their labelled progeny were detected by confocal microscopy of wholemount epidermis [30]. At two days post induction, only single labelled cells were seen, at a frequency of 1 in 600 cells in the basal layer, indicating that the clusters of cells encountered at later time points are clones, each derived from a single progenitor cell (Fig. 2.2). Analysis of subsequent cohorts of mice demonstrated clones that remained cohesive and expanded progressively in size (Fig. 2.2, [15]). Scoring clones

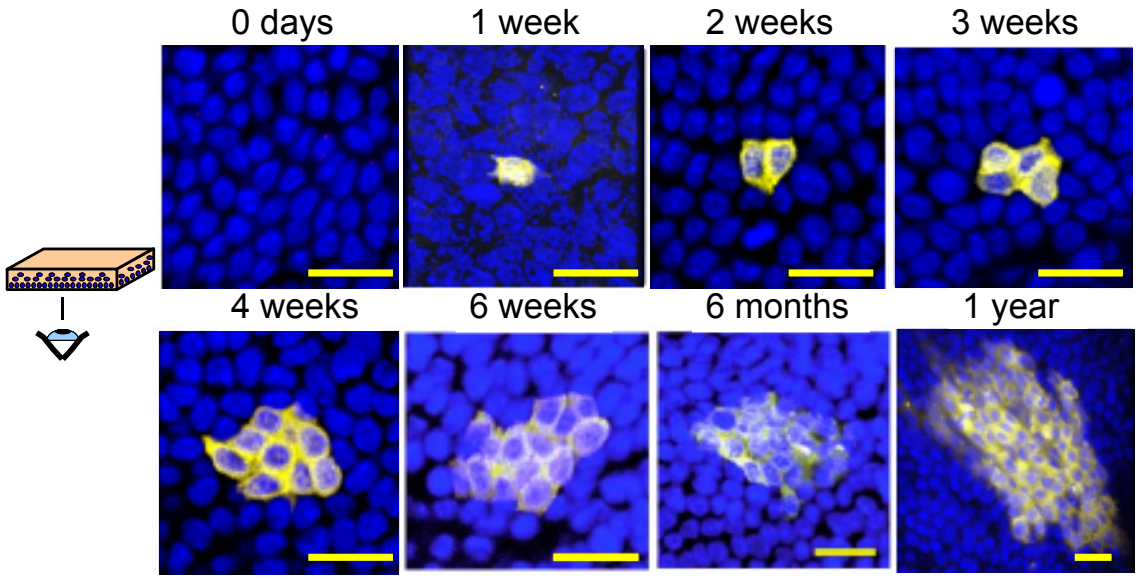


FIG. 2.2: *In vivo* clonal labelling of epidermal progenitor cells. Projected Z-stack confocal images of IFE wholemounts from genetically modified mice (strain Ah*cre*$^{\mathrm{ERT}}$ R26$^{\mathrm{EYFP/wt}}$) viewed from the basal surface at the times shown following induction. Yellow, EYFP; blue, DAPI nuclear stain. Scale bar, 20 μm.

that contained one or more basal cells, the observed clone size distribution (ie. the total number of nucleated cells per clone) up to six weeks post labelling, and the basal layer clone size distribution up to one year (see methods in Ref. [15]) are shown in Fig. 2.3.

The density of labelled clones containing at least one basal layer cell in tail epidermis rose from 2 days to a peak at two weeks after induction, as EYFP levels accumulated to detectable levels in all labelled cells. Clone numbers then fell to $7\% \pm 2\%$ (mean $\pm$ SD) of the peak value by three months, and $3\% \pm 2\%$ at one year; similar results were seen in back skin [51]. This decline was accompanied by the appearance of multi-cellular clones containing only suprabasal cells, consistent with clonal loss through differentiation [15]. Analysis of the spatial distribution of IFE clones indicates that labelled clones are not replaced by unlabelled clones migrating from hair follicles [15]. Moreover, none of the labelled clones can derive from bulge stem cells as this region is not labelled [15].

Before attempting to interpret the clone fate data, it is necessary to assess the extent to which they are influenced by tissue growth or apoptosis. Firstly, the rate of increase in epidermal surface area due to growth was low (estimated at less than 3.5%/month over the time course of the experiment) whilst apoptosis was undetectable in basal layer cells (see Ref. [15], supplementary results). Furthermore, the number of basal layer cells/unit area and the proportion of cycling cells (as assessed both by Ki67 and cdc6 immunostaining) showed no significant difference between two week and one year samples. Both techniques of assessing the proportion of cycling cells gave similar results, as did flow cytometry [52, 53]: 22%±3% (mean ± SD) for Ki67; 24%±4% for cdc6; and 22±1% for flow cytometry (see Ref. [15], supplementary results). Finally, there was no significant difference between the proportion of cycling cells in the labelled and unlabelled cell populations, either at 5

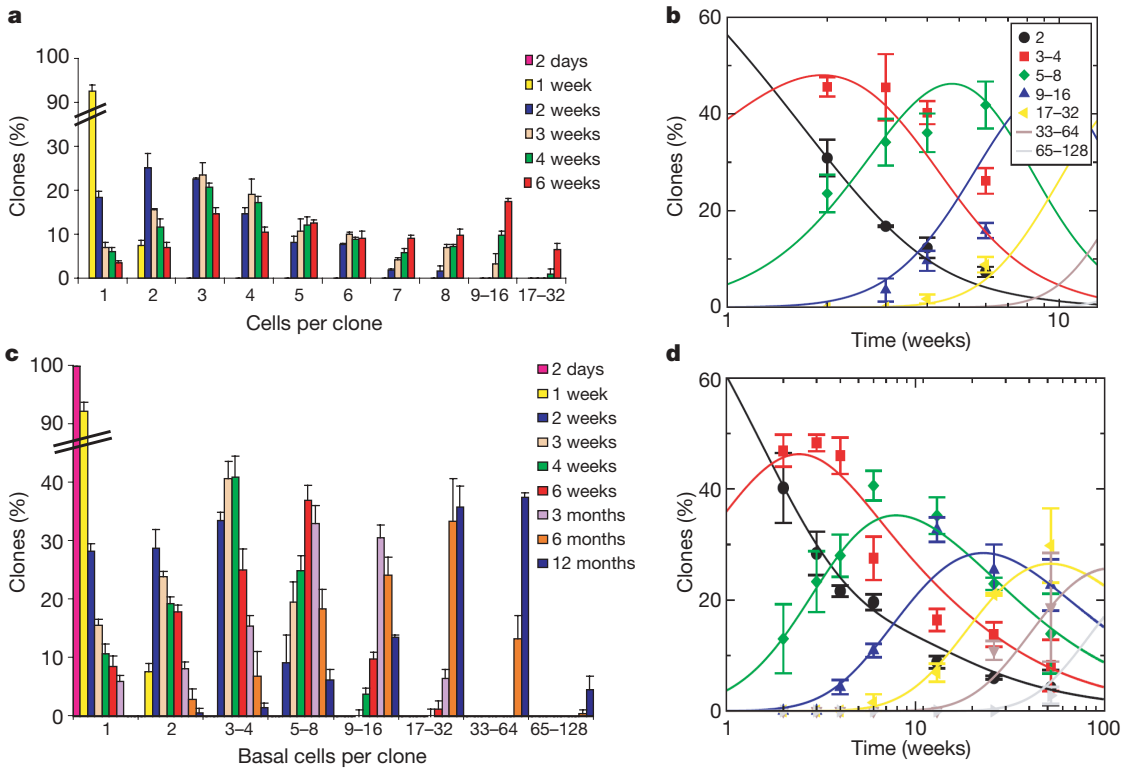


FIG. 2.3: Clone fate data. **a**, Distribution of clone size (total cells /clone) as a function of cell number, as measured at 2 days, 1, 2, 3, 4 and 6 weeks post induction. Error bars indicate standard error of the mean (SEM). **b** and **d**, Distribution of (**b**) clone size (total cells per clone) and (**d**) basal cells per clone as a function of time for different values of cell number (error bars indicate SEM). Here we have aggregated clone sizes in ranges increasing in size in powers of two (see legend within figure). To eliminate possible ambiguities due to labelling efficiency, single cell clones are eliminated from the distribution thereby removing the population of post-mitotic cells labelled at induction. We focus on timepoints of 2 weeks or more post induction when EYFP levels have stabilised. Continuous curves show the behaviour of the one progenitor cell model proposed in the following chapter (3). **c**, Distribution of basal cells per clone as a function of basal cell number, as measured at 2 days, 1, 2, 3, 4, 6 weeks, 3, 6 and 12 months post induction (error bars indicate SEM).

days or one year post induction (see Ref. [15], supplementary results). We therefore conclude that basal layer cells labelled at induction are typical of the entire basal cell population, and that the observed clonal evolution is representative of the adult system in epidermal homeostasis.

2.2 Failure of the stem/TA hypothesis

According to the stem/TA cell hypothesis, TA cells undergo a limited number of cell divisions followed by differentiation [25]. To test this prediction we examined clones at 3 weeks, over 90% of which are lost by 12 weeks post induction. Significantly, clones comprising three or more cells contained both basal and suprabasal cells, indicative of asynchronous terminal differentia-

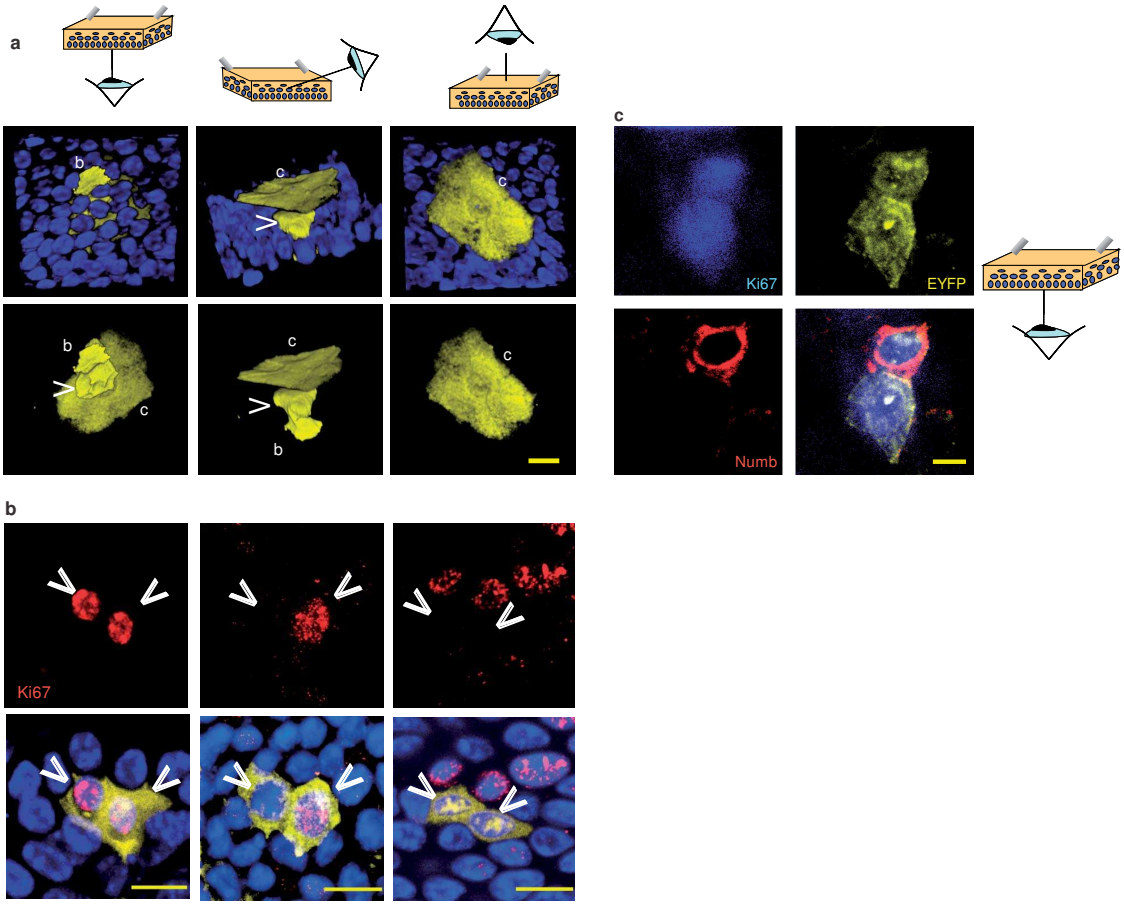


FIG. 2.4: Asymmetric cell fate in epidermal progenitors. **a**, Visualization of a three-cell clone exhibiting asynchronous terminal differentiation. Projected Z-stack images show one basal cell (labelled b), and two suprabasal cells: a cornified layer cell (labelled c), and a second suprabasal cell indicated by the arrowhead. Cartoon shows the angle of view. Upper panels show the clone marker EYFP, yellow, and the nuclear marker DAPI, blue; lower panels are corresponding images with only EYFP shown. Scale bar, 20 μm. **b**, Visualization of two-cell clones (both cells basal, 3 weeks post-recombination), showing the different proliferative fates of the daughter cells of a single division, providing evidence for symmetric and asymmetric cell fate. Clones are viewed from the basal epidermal surface, stained for the proliferation marker Ki67 (red), DAPI (blue), and EYFP (yellow); arrowheads indicate position of EYFP-labelled cells. Three types of clone are shown, with two, one and zero Ki67 positive cells. Scale bar, 10 μm. **c**, Two-cell clone (both cells basal, 3 weeks post-recombination, viewed from the basal epidermal surface,) stained for the proliferation marker Ki67 (blue), numb (red) and EYFP (yellow), showing asymmetric distribution of numb, providing evidence for asymmetric cell fate resulting from a planar division. Scale bar, 5 μm.

tion (Fig. 2.4a). Furthermore, the immunostaining of clones consisting of two basal cells reveals that a single cell division may generate either one cycling and one non-cycling daughter, or two cycling daughters, or two non- cycling daughters (Fig. 2.4b). This raises the question of whether there is asymmetric cell division within the basal plane as described in the Drosophila peripheral nervous system and Zebra Fish retinal precursors [54, 55]. Three-dimensional imaging of whole-mount epidermis revealed that only 3% of mitotic spindles lie perpendicular to the basal layer indicating that, in contrast to embryonic epidermis, the vast majority of EPC divisions generate two basal layer cells (see Ref. [15] supplementary results), [56, 57]. The observation of asymmetric partitioning of numb protein, which marks asymmetric division in neural and myogenic precursors, in clones containing two basal cells suggests that planar-orientated asymmetric division also occurs in the epidermis (Fig. 2.4c) [58, 59]. EPC behaviour thus differs substantially from that observed in committed precursors in other systems [60, 61].

We next considered the behaviour of the long lived clones that persist for over 3 months. Within the stem/TA cell hypothesis, the epidermis is organised into epidermal proliferative units (EPUs) comprising about 10 basal cells supported by a single self- renewing stem cell [28]. If individual stem cells retain their self renewal capacity, the stem/TA cell model predicts that the basal layer clone size distribution must become time-independent and characteristic of a single EPU (see section 2.A and Ref. [25]). Such behaviour is in stark contrast to the progressive increase in average clone size observed in the epidermis (Fig. 2.3).

Faced with this apparent contradiction, one could attempt to revise the stem cell/TA cell model staying within the general paradigm. This might include introducing the capacity for stem cell aging and/or migration [35]. Alternatively, one may try to exploit the range of experimental data to seek evidence for a new paradigm for epidermal homeostasis. Intriguingly, such evidence is found in scaling properties of the observed clone size distribution. In the following chapter, we will argue that the clone fate data are compatible with a model in which interfollicular epidermis is maintained by only one compartment of proliferating cells. Whether this model should be considered as an extreme variant of the stem/TA cell hypothesis or a new concept is arguably a matter of semantics, a point we will return to later.

Chapter appendix

2.A Classical EPU model and the predictions of the stem/TA cell hypothesis

In the 'classical EPU model', long-lived self-renewing stem-cells support a transit-amplifying cell compartment, which in turn divide into post-mitotic cells. Within this framework, we will show that:

- in steady-state, the clone size distribution associated with an EPU is strongly peaked around some average value with fluctuations that are typically small as compared to the mean size;
- the size distribution of a representative sample of labelled basal cells (such as addressed in the experiment) must eventually converge onto the same *stationary* distribution as that of the EPU.

According to the classical EPU model, stem cells can divide indefinitely, and are the source of all interfollicular keratinocytes in normal epidermis. Stem cells (labelled as S-type) divide into a daughter stem cell and a differentiated transit-amplifying (TA) cell which, in turn, is thought to undergo several rounds of division before exiting the cell cycle. Finally, post-mitotic cells (B) are free to exit the basal layer. Focussing on the basal layer cell distribution, the model may be summarised by the rate equations,

$$\begin{array}{lcl} \mathrm{S} & \xrightarrow{\lambda_\mathrm{S}} & \mathrm{S} + \mathrm{TA}_1 \\ \mathrm{TA}_n & \xrightarrow{\lambda_\mathrm{A}} & \mathrm{TA}_{n+1} + \mathrm{TA}_{n+1} \\ \mathrm{TA}_N & \xrightarrow{\lambda_\mathrm{A}} & \mathrm{B} + \mathrm{B} \\ \mathrm{B} & \xrightarrow{\Gamma} & \oslash, \end{array} \tag{2.1}$$

where TA_n indexes the n-th generation of the TA cell population, and $N \sim 3-5$ [25] represents the number of rounds of division of the TA cell population. Here λ_S and λ_A (assumed, for simplicity, to be independent of generation) represent the division rates of stem and TA cells respectively, and Γ denotes the post-mitotic cell transfer rate from the basal to the suprabasal layer. Although one can conceive of possible variations, while one assumes that an EPU is maintained by a single long-lived stem cell, the model encapsulated by (2.1) exemplifies the general properties of any model based on the stem cell/TA cell hypothesis.

To analyse the model, let us define $P_{n_\mathrm{S},n_{\mathrm{A}_1},\cdots,n_{\mathrm{A}_N},n_\mathrm{B}}$ as the probability for finding a given clone with n_S stem cells, $(n_{\mathrm{A}_1} + \cdots + n_{\mathrm{A}_N})$ TA cells, and n_B post-mitotic cells at some time t after induction. The time-evolution for this probability distribution is specified by the Master equation

$$\begin{aligned} \frac{dP}{dt} &= n_\mathrm{S}\lambda_S \left[P_{n_{\mathrm{A}_1}-1} - P\right] + \lambda_\mathrm{A} \sum_{k=1}^{N-1} \left[(n_{\mathrm{A}_k}+1)P_{n_{\mathrm{A}_k}+1,n_{\mathrm{A}_{(k+1)}}-1} - n_{\mathrm{A}_k}P\right] \\ &\quad + \lambda_\mathrm{A}\left[(n_{\mathrm{A}_N}+1)P_{n_{\mathrm{A}_N}+1,n_\mathrm{B}-1} - n_{\mathrm{A}_N}P\right] + \Gamma\left[(n_\mathrm{B}+1)P_{n_\mathrm{B}+1} - n_\mathrm{B}P\right], \end{aligned} \tag{2.2}$$

where, for clarity, we have suppressed indicies of $P_{\{n\}}(t)$ which remain unrevised in the respective terms.

The time evolution of $P_{\{n\}}$ depends upon the initial conditions: With $P_{\{n\}}(0) = \delta_{n_S,1}(\prod_{k=1}^{N}\delta_{n_{A_k},0})\delta_{n_B,0}$, i.e. with one initially labelled stem cell, the solution to Eq. (2.2) describes the time-evolution of a persisting clone destined to form an EPU. At long times, the corresponding probability distribution becomes stationary. Recasting Eq. (2.2) in terms of the moment-generating function (see, e.g., Ref. [62] for details), one finds a stationary distribution

$$P_{\{n\}} = \mathcal{P}(n_{\mathrm{A}_1}|\lambda_{\mathrm{S}}/\lambda_{\mathrm{A}})\left[\prod_{k=2}^{N}\mathcal{Q}(n_{\mathrm{A}_k}|2^{k-1}\lambda_{\mathrm{S}}/\lambda_{\mathrm{A}})\right]\mathcal{Q}(n_{\mathrm{B}}|2^{N}\lambda_{\mathrm{S}}/\Gamma), \tag{2.3}$$

where $\mathcal{P}(n|X)$ denotes the Poisson distribution with mean X, and $\mathcal{Q}(n|X)$ is a Poisson-like distribution,

$$\mathcal{Q}(n|X) = \frac{e^{-X}}{n!}\sum_{k=\mathrm{ceil}(n/2)}^{\infty}\frac{(2k)!X^k}{4^k k!(2k-n)!},$$

characterised by its mean X, and variance $3X/2$. From this result one may conclude that, while each EPU contains one and only one stem cell, the mean and variance of the relative populations of TA and post-mitotic cells in an EPU is given by

$$\begin{aligned}\langle n_A\rangle &= (2^N-1)\frac{\lambda_S}{\lambda_A}, \qquad \sigma_A^2 = \frac{2^N-\frac{4}{3}}{2^N-1}\frac{3}{2}\langle n_A\rangle \\ \langle n_B\rangle &= 2^N\frac{\lambda_S}{\Gamma}, \qquad \sigma_B^2 = \frac{3}{2}\langle n_B\rangle,\end{aligned} \tag{2.4}$$

with the mean and variance in the total size of each EPU equal to $\langle n\rangle = \langle n_S\rangle + \langle n_A\rangle + \langle n_B\rangle$ and $\sigma^2 = \sigma_S^2 + \sigma_A^2 + \sigma_B^2$. Each stem cell supports a population of TA and post-mitotic cells with a variance that is approximately proportional to the mean. Since the average total size of an EPU is equal to the inverse of the stem-cell density $\rho_S = 1/\langle n\rangle$, *the variance in size between EPUs is independent of the details of the TA cell compartment*. For example, taking a stem cell density of $\rho_S = 0.02$ (i.e. 2%), we find the average total number of cells per clone in the basal layer to be $\langle n\rangle = 1/\rho_S = 50$ with variance $\sigma \simeq 9$, so that all EPUs have 50 ± 9 cells in the steady-state, irrespective of the density, ρ_{A}, and number of generations, N, of TA cells.

On this platform, let us now consider the transient behaviour of a labelled clone population if we assume an equal labelling efficiency of cells in the basal layer. Referring to Eq. (2.1), one would expect all clones arising from initially labelled TA and B cells to inevitably detach from the basal layer. At the same time, a stem cell labelled at induction will give rise to a persisting clone which converges onto a fully-developed EPU. We conclude that, following a transient period, the size distribution of a randomly-labelled population will *converge onto the steady-state EPU size distribution* (2.3). This convergence is illustrated in figs. 2.5a and b where the clone size distributions predicted for the non-equilibrium process (2.1) are plotted against time for two values of the stem cell density. Notably, the steady-state behaviour differs markedly from that observed in experiment where the size of surviving clones continues to grow, and the distribution $P_{\{n\}}(t)$ does not settle, at least within the observed time period.

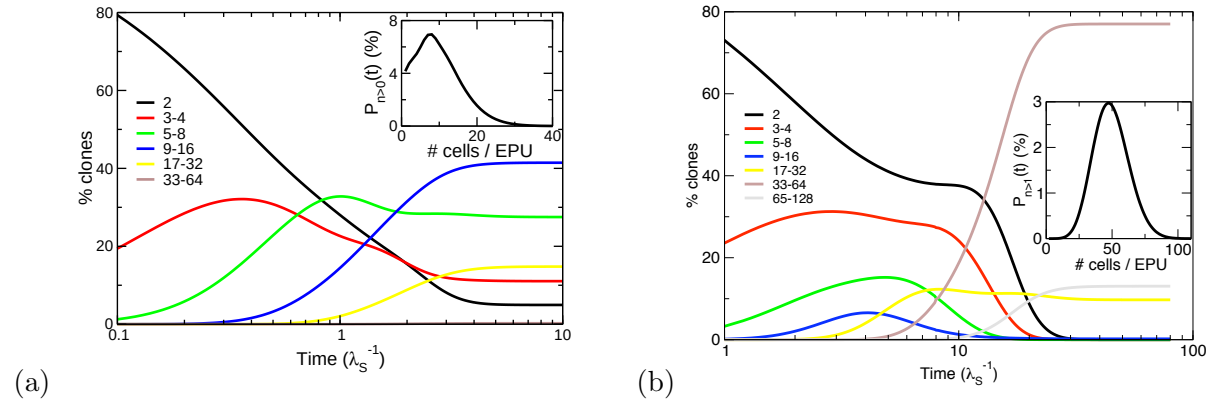


FIG. 2.5: Predicted clone size distributions for stem/TA cell model. The probability distribution $P_{n>1}^{\text{pers.}}(t)$ for finding a labelled clone with n basal cells is plotted against time (measured in units of the average stem-cell cycle time $1/\lambda_S$) for (a) $\rho_s = 0.1$ (i.e. 10%), and (b) $\rho_s = 0.02$ (i.e. 2%). For consistency with the main text, we have eliminated single cell clones from the statistical ensemble. The distribution is obtained for $N = 3$ generations of TA cells by integrating Eq. (2.2) starting from an initial condition corresponding to each clone consisting of a single, randomly-labelled cell. To perform the integration, we have used the respective values of stem cell number density, ρ_s, while keeping the overall density of proliferating cells at the experimentally measured value of $\rho_s + \rho_A = 0.22$ (i.e. 22%). The mean-field relations in Eq. (2.4) constrain the rates λ_A and Γ, leaving only one free parameter, the stem cell cycling rate, λ_S. For $\rho_s = 0.1$, the TA cell division rate is some 7 times faster than the stem cell division rate while for $\rho_s = 0.02$, the cycling rates are are comparable (note that larger N will lead to $\lambda_S \ll \lambda_A$). As expected, in the long-time limit, the probability distribution becomes time-independent. *Inset*: The full probability distribution, $P_{n>1}^{\text{pers.}}(t)$, is plotted as a function of n after a time (a) $t = 10/\lambda_S$ and (b) $t = 50/\lambda_S$ when both distributions have converged onto their steady-state.

Chapter 3

Mechanism of cell division in murine tail skin

Chapter overview

By studying the *scaling behaviour* of the clone fate data presented in chapter 2, we infer here a new model of homeostasis involving only *one* type of progenitor cell. These cells are found to undergo both symmetric and asymmetric division at rates that ensure epidermal homeostasis. We show that the rates of cell division and differentiation may be evaluated by considering the long-time and short-time clone fate data, and that the data is consistent with cells dividing independently rather than synchronously. These results raise questions about the potential role of stem cells in tissue maintenance *in vivo*.

Motivated by these findings, we then consider a mechanism for cancer onset based closely on the model for normal adult skin. By analysing the expected changes to clonal fate in cancer emerging from a simple two-stage mutation, we propose that clonal fate data may provide a novel method for studying the earliest stages of the disease.

The contents of this chapter draw significantly upon a recent publication [16] by AMK *et al.*.

3.1 Introduction

As described in chapter 2, by genetically labelling cells at a defined time in transgenic mice, we have obtained access to the kinetics of labelled cells *in vivo*, with single-cell resolution. From a theoretical perspective, the analysis of clonal fate data presents a challenging "inverse problem" in population dynamics: While it is straightforward to predict the time-evolution of a population distribution according to a set of growth rules, the analysis of the inverse problem is more challenging, open to ambiguity and potential misinterpretation.

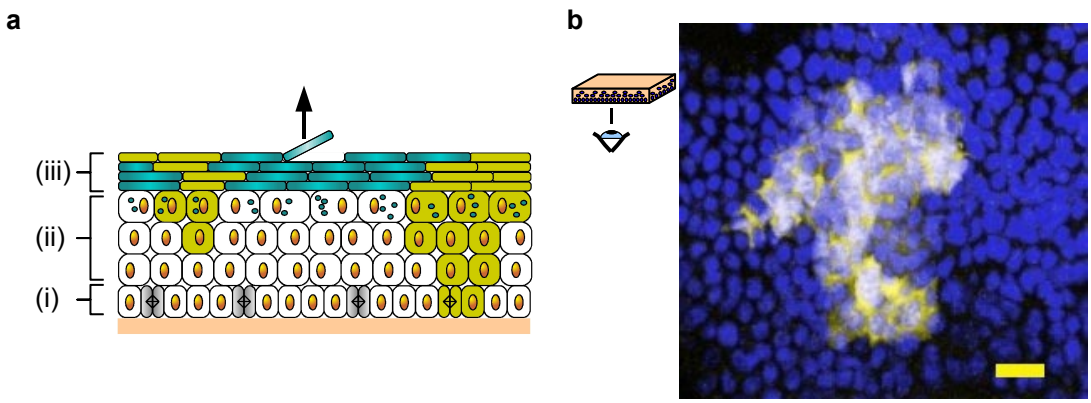


FIG. 3.1: (a) Schematic cross-section of murine interfollicular epidermis (IFE) showing the architecture of typical labelled clones. The shaded regions (yellow) indicate two distinct clones, the progeny of single basal layer cells labelled at induction. While the clone on the right retains at least one labelled cell in the basal layer, the clone on the left hand side has detached from the basal layer indicating that all of the cells have stopped proliferating. The former are designated as "persisting clones" and contribute to the clone size distributions, while the latter, being difficult to resolve reliably, are excluded from experimental consideration. See caption of Fig. 1.1 for a description of layers (i)–(iii). (b) Typical example of a clone acquired at a late time point, viewed from the basal layer surface. Cell nuclei are labelled blue; the hereditary clone marker (EYFP) appears yellow. Scale bar: 20μm

In the previous chapter, we showed that the range of clone fate data resulting from genetic labelling experiments undermines the basis of the stem/TA cell hypothesis of murine epidermal homeostasis. Expanding upon this preliminary theoretical finding, the aim of this chapter is to infer the true mechanism of epidermal homeostasis, and to describe the potential of the system as a method to explore early signatures of carcinogenic mutations.

3.1.1 Background: Experimental Methodology

To organise our discussion and to define the elements of the analysis, we begin with a concise review of the experiment described in the preceding chapter and Ref. [15]. To generate data on the fate of individual labelled cells and their progeny, hereafter referred to as clonal fate data, inducible genetic marking was used to label a sample of cells and their progeny in the epidermis of transgenic mice. The enhanced Yellow Fluorescent Protein (EYFP) label was then detected by confocal microscopy, which enables 3D imaging of entire sheets of epidermis. Low-frequency labelling of approximately 1 in 600 basal-layer epidermal cells at a defined time was achieved by using two drugs to mediate a genetic event which resulted in expression of the EYFP gene in a cohort of mice. This low efficiency labelling ensures that clones are unlikely to merge (see discussion in section 3.2.1). By analysing samples of mice at different time points it was possible to analyse the fate of labelled clones at single cell resolution *in vivo* for times up to one year post-labelling in the epidermis (see, for example, Fig. 3.1(b)) [30, 15].

With the gradual accummulation of EYFP levels, the early time data (less than two weeks) reveals a small increase in the number of labelled clones containing one or two cells. At longer times, clones

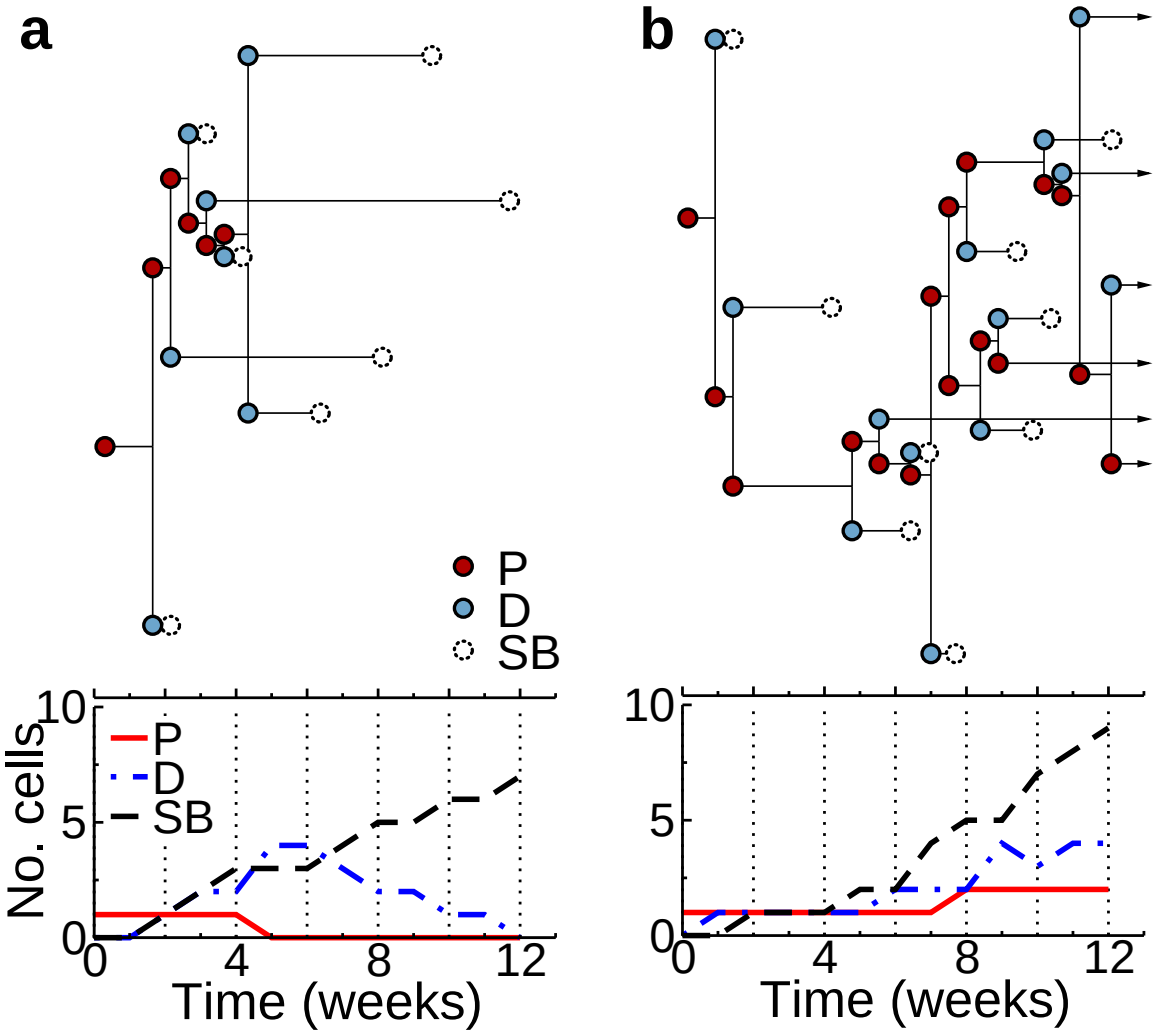


FIG. 3.2: (Color online) *Top:* Theoretical lineage for the first 12 weeks post-labelling of **a**, a detached clone in which all cells have undergone a transition to terminal differentiation by week 12, and **b**, a persisting clone in which some of the cells maintain a proliferative capacity, according to model (3.2). Circles indicate progenitor cells (P), differentiated cells (D), and suprabasal cells (SB). Note that, because the birth-death process (3.2) is Markovian, the lifetime of cells is drawn from a Poisson distribution with no strict minimum or maximum lifetime. The statistics of such lineage trees do not change significantly when we account for a latency period between divisions that is much shorter than the mean cell lifetime (see discussion in section 3.2.3). *Bottom:* The total number of proliferating, differentiated and supra-basal cells for the two clones as a function of time.

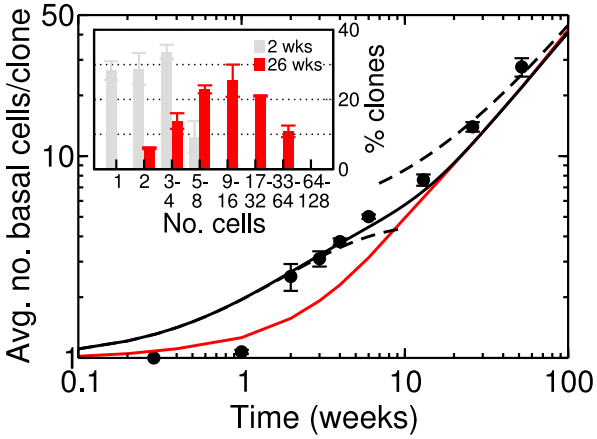


FIG. 3.3: (Color online) Mean number of basal layer cells in persisting clones. The experimental data (circles) show an inexorable increase in the size of persisting clones over the entire time course of the experiment. The behaviour at short times (from $2-6$ weeks) and at long times (beyond 13 weeks) follows the two simple analytical approximations described in the main text (lower and upper dashed curves). For times earlier than two weeks (referring to section 3.2.4), clones remain approximately one cell in size. The experimental data are consistent with the behaviour predicted by process (3.2) (black line) when it is assumed that only A-type cells are labelled at induction. In contrast, assuming that A and B type cells label in proportion to their steady-state population leads to an underestimate of average clone size between two and six weeks (lower curve, red online), as does the assumption that type B cells label with better efficiency (not shown). *Inset:* The underlying distribution of basal cells per clone at 2 weeks and 26 weeks post-labelling. The data is binned by cell count in increasing powers of 2.

increase in size while cells within clones begin to migrate through the suprabasal layers forming relatively cohesive irregular columns (see Fig. 3.1(a)).

The loss of nuclei in the cornified layer (fig. 3.1) makes determination of the number of cornified layer cells in larger clones by microscopy unreliable. Therefore, to identify a manageable population, attention was focused on the population of basal cells in "persisting clones", defined as those labelled clones which retain at least one basal layer cell, such as is exemplified in the theoretical lineage maps in Fig. 3.2. After two weeks, the density of persisting clones was seen to decrease monotonically indicating that the entire cell population within such clones had become differentiated and the clone detached from the basal layer (shown schematically in figs. 3.1(a) and 3.2(a)). However, as shown in Fig. 3.3, the average size of the persisting clone population showed a steady increase in size throughout the entire duration of the experiment.

In appendix 2.A we demonstrated that a non-stationary steady-state clone size distribution is incompatible with any model in which the IFE is supported by a population of long-lived stem cells. With the latter, one would expect the number density of persisting clones to reach a non-zero minimum (commensurate with the labelling frequency of stem cells) while the average clone size would asymptote to a constant value characteristic of a single epidermal proliferative unit. This inconsistency has led us to abandon, or at least substantially revise, the orthodox stem/TA cell hypothesis and look for a different paradigm for epidermal maintenance.

But, to what extent are the clone fate data amenable to further theoretical analysis? Indeed, the application of population dynamics to the problem of cell kinetics has a long history (see, e.g., Refs.[63, 64, 65, 66]) with studies of epidermal cell proliferation addressed in several papers [39, 67, 68, 20]. However, even in the adult system, where cell kinetics may be expected to conform to a "steady-state" behaviour, it is far from clear whether the cell dynamics can be modelled as a simple stochastic process. Regulation due to environmental conditions could lead to a highly nonlinear or even non-local dependence of cell division rates. Indeed, *a priori*, it is far from clear whether the cell kinetics can be considered as Markovian, i.e. that cell division is both random and independent of the past history of the cell. Therefore, instead of trying to formulate a complex theory of cell division, taking account of the potential underlying biochemical pathways and regulation networks [39], we will follow a different strategy looking for signatures of steady-state behaviour in the experimental data and evidence for a simple underlying mechanism for cell fate. Intriguingly, such evidence is to be found in the scaling properties of the clone size distribution [15].

3.1.2 Scaling

To identify scaling characteristics, it is necessary to focus on the basal layer clone size distribution, $P_n(t)$, which describes the probability that a labelled progenitor cell develops into a clone with a total of n basal layer cells at a time t after labelling. (Note that, in general, the total number of cells in the supra-basal layers of a clone may greatly exceed the number of basal layer cells.) With this definition, $P_0(t)$ describes the "extinction" probability of a clone, i.e. the probability that *all* of the cells within a labelled clone have migrated into the supra-basal layers. To make contact with the experimental data, it is necessary to eliminate from the statistical ensemble the extinct clone population (which are difficult to monitor experimentally) and single-cell clones (whose contribution to the total ensemble is compromised by the seemingly unknown relative labelling efficiency of proliferating and post-mitotic cells at induction), leading to a reduced distribution for "persisting" clones,

$$P^{\text{pers.}}_{n\geq 2}(t) \equiv \frac{P_n(t)}{1-P_0(t)-P_1(t)}\,.$$

Then, to consolidate the data and minimise fluctuations due to counting statistics, it is further convenient to *bin* the distribution in increasing powers of 2,

$$\mathcal{P}^{\text{pers.}}_k(t) = \sum_{n=2^{k-1}+1}^{2^k} P^{\text{pers.}}_{n\geq 2}(t)\,,$$

i.e. $\mathcal{P}^{\text{pers.}}_1(t)$ describes the probability of having two cells per clone, $\mathcal{P}^{\text{pers.}}_2(t)$ describes the probability of having 3-4 cells per clone, and so on. Referring to Fig. 3.4, one may see that, after an initial transient behaviour, the clone size distribution asymptotes in time to the simple scaling form,

$$\mathcal{P}^{\text{pers.}}_k(t) = f(2^k/t)\,. \tag{3.1}$$

This striking observation brings with it a number of important consequences: As well as reinforcing the inapplicability of the stem cell/TA cell hypothesis, such behaviour suggests that epidermal maintenance must conform to a simple model of cell division. The absence of further characteristic

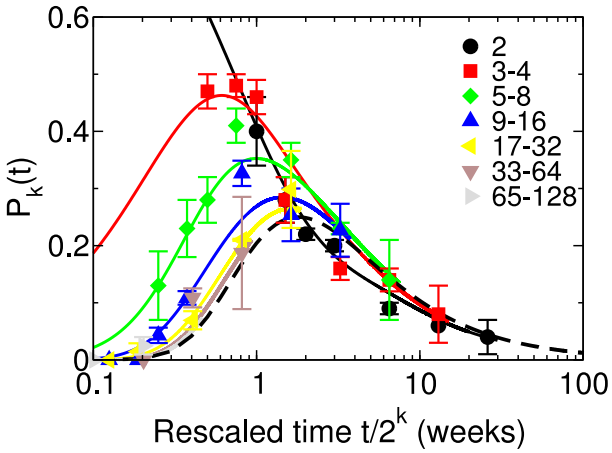


FIG. 3.4: (Color online) Time dependence of the grouped size distribution of *persisting* clones, $\mathcal{P}_k^{\text{pers.}}(t)$, plotted as a function of the rescaled time coordinate $t/2^k \mapsto t$. The data points show measurements (extracted from data such as shown in Fig. 3.3(inset), given fully in ref. [15]), while the solid curves show the probability distributions associated with the non-equilibrium process (3.2) for the basal-layer clone population as obtained by a numerical solution of the Master equation (3.3). (Error bars refer to standard error of the mean). At long times, the data converge onto a universal curve (dashed line), which one may identify with the form given in eq. 3.7. The rescaling compresses the time axis for larger clones, so that the large-clone distributions appear to converge much earlier onto the universal curve.

time-scales, beyond that of an overall proliferation rate, motivates the consideration of a simple kinetics in which *only one process dictates the long-time characteristics of clonal evolution.*

Moreover, from the scaling observation one can also deduce two additional constraints: Firstly, in the long-time limit, the average number of basal layer cells within a *persisting* clone *increases linearly with time*, viz.

$$\begin{aligned} P_{n\geq 2}^{\text{pers.}}(t) &\simeq \frac{d}{d(2^k)}\mathcal{P}_k^{\text{pers.}}(t) = \frac{1}{t}\, f'(2^k/t) \\ \langle n\rangle_{\text{pers.}} &\equiv \sum_{n\geq 2}^{\infty} n\, P_{n\geq 2}^{\text{pers.}}(t) \simeq \int_0^{\infty} dn \frac{n}{t}\, f'(n/t) \propto t\,. \end{aligned}$$

Secondly, if we assume that labelled progenitor cells are representative of *all* progenitor cells in the epidermis, and that the population of clones with only one basal layer cell is not "extensive" (i.e. $\lim_{t\to\infty} P_1(t) = 0$), this means that, in the long-time limit, the clone persistence probability must scale as $1 - P_0(t) \propto 1/t$ such that

$$\langle n\rangle = \sum_n n\, P_n(t) \stackrel{!}{=} \rho\,,$$

where the constant, ρ, is given by the fraction of proliferating cells in the basal layer. Without this condition, one is lead to conclude that the labelled population of basal layer cells either grows or diminishes, a behaviour incompatible with the (observed) steady-state character of the adult system.

Although the manifestation of scaling behaviour in the clone size distributions gives some confidence that the mechanism of cell fate in IFE conforms to a simple non-equilibrium process, it is nevertheless possible to conceive of complicated, multi-component, models which could asymptote to the same long-time evolution. To further constrain the possible theories, it is helpful to draw on additional experimental observations [15]: Firstly, immunostaining of clones with a total of two cells (using the proliferation marker Ki67 and, separately, the replication licensing factor cdc6) reveals that a single cell division may generate either one proliferating and one non-proliferating daughter through asymmetric division, or two proliferating daughters, or two non-proliferating daughters (cf. [56, 57, 58]). Secondly, three-dimensional imaging of the epidermis reveals that only 3% of mitotic spindles lie perpendicular to the basal layer indicating that divisions may be considered to be confined to the basal layer, confirming the results of earlier work that indicates a dividing basal cell generates two basal layer cells [57].

This completes our preliminary discussion of the experimental background and phenomenology. In summary, the clone fate data reveal a behaviour wholely incompatible with any model based on the concept of long-lived self-renewing stem cells. The observation of long-time scaling behaviour motivates the consideration of a simple model based on a stochastic non-equilibrium process and is indicative of the labelled cells being both a representative (i.e. self-sustaining) population and in steady-state. In the following, we will develop a theory of epidermal maintenance which encompasses all of these observations.

3.2 Theory of epidermal maintenance

3.2.1 Model

Taken together, the range of clonal fate data and the observation of symmetric and asymmetric division are consistent with a remarkably simple model of epidermal homeostasis involving only one proliferating cell compartment and engaging just three adjustable parameters: the overall cell division rate, λ; the proportion of cell divisions that are symmetric, r; and the rate of transfer, Γ, of non-proliferating cells from the basal to the supra-basal layers. To maintain the total proliferating cell population, a constraint imposed by the steady-state assumption, we have used the fact that the division rates associated with the two channels of symmetric cell division must be equal. Denoting the proliferating cells as type A, differentiated basal layer cells as type B, and supra-basal layer cells as type C, the model describes the non-equilibrium process,

$$\begin{aligned} \mathrm{A} &\xrightarrow{\lambda} \begin{cases} \mathrm{A}+\mathrm{A} & \text{Prob. } r \\ \mathrm{A}+\mathrm{B} & \text{Prob. } 1-2r \\ \mathrm{B}+\mathrm{B} & \text{Prob. } r \end{cases} \\ \mathrm{B} &\xrightarrow{\Gamma} \mathrm{C}\,. \end{aligned} \tag{3.2}$$

Finally, the experimental observation that the total basal layer cell density remains approximately constant over the time course of the experiment leads to the additional constraint that

$$\Gamma = \frac{\rho}{1-\rho}\lambda\,,$$

reducing the number of adjustable parameters to just two.

By ignoring processes involving the shedding of cells from the surface of the epidermis, the applicability of the model to the consideration of the *total* clone size distribution is limited to appropriately short time scales (up to six weeks post-labelling). However, if we focus only on the clone size distribution associated with those cells which occupy the basal layer, the model can be applied up to arbitrary times. In this case, the transfer process must be replaced by one in which $\mathrm{B} \xrightarrow{\Gamma} \emptyset$. In either case, if we treat all instances of cell division and cell transfer as independent stochastic events, a point that we shall revisit later, then the time evolution associated with the process (3.2) can be cast in the form of a Master equation. Defining $P_{n_\mathrm{A},n_\mathrm{B}}(t)$ as the probability of finding n_A type A cells and n_B type B cells in a given clone after some time t, the probability distribution evolves according to the Master equation:

$$\begin{aligned}\partial_t P_{n_\mathrm{A},n_\mathrm{B}} &= r\lambda\left[(n_\mathrm{A}-1)P_{n_\mathrm{A}-1,n_\mathrm{B}} - n_\mathrm{A}P_{n_\mathrm{A},n_\mathrm{B}}\right] \\ &+r\lambda[(n_\mathrm{A}+1)P_{n_\mathrm{A}+1,n_\mathrm{B}-2} - n_\mathrm{A}P_{n_\mathrm{A},n_\mathrm{B}}] \\ &+(1-2r)\lambda[n_\mathrm{A}P_{n_\mathrm{A},n_\mathrm{B}-1} - n_\mathrm{A}P_{n_\mathrm{A},n_\mathrm{B}}] \\ &+\Gamma[(n_\mathrm{B}+1)P_{n_\mathrm{A},n_\mathrm{B}+1} - n_\mathrm{B}P_{n_\mathrm{A},n_\mathrm{B}}]\,. \end{aligned} \tag{3.3}$$

If we suppose that the basal layer cells label in proportion to their population, the latter must be solved subject to the boundary condition $P_{n_\mathrm{A},n_\mathrm{B}}(0) = \rho\delta_{n_\mathrm{A},1}\delta_{n_\mathrm{B},0} + (1-\rho)\delta_{n_\mathrm{A},0}\delta_{n_\mathrm{B},1}$. Later, in section 3.2.4, we will argue that the clone size distribution is compatible with a labelling efficiency which favours A over B type cells. Either way, by excluding single cell clones from the distribution, this source of ambiguity may be safely eliminated. Although the Master equation (and its total cell number generalisation) is not amenable to exact analytic solution, its properties can be inferred from the consideration of the A cell population alone for which an explicit solution may be derived.

When considered alone, A type cells conform to a simple set of rate laws,

$$\mathrm{A} \xrightarrow{2r\lambda} \begin{cases} \mathrm{A}+\mathrm{A} & \text{Prob. } 1/2\,, \\ \emptyset & \text{Prob. } 1/2\,, \end{cases} \tag{3.4}$$

an example of a Galton-Watson process, long known to statisticians (see, e.g., Ref. [62]). In this case, the probability distribution, which is related to that of the two-component model through the relation $p_{n_\mathrm{A}}(t) = \sum_{n_\mathrm{B}=0}^{\infty} P_{n_\mathrm{A},n_\mathrm{B}}(t)$, can be solved analytically. (Here, we have used a lower case p to discriminate the probability distribution from its two-component counterpart.) For an initial distribution $p_{n_\mathrm{A}}(0) = \delta_{n_\mathrm{A},1}$ it may be shown that [62],

$$p_{n_\mathrm{A}}(t) = \left(1+\frac{1}{r\lambda t}\right)^{-(n_\mathrm{A}+1)} \times \begin{cases} 1 & n_\mathrm{A}=0\,, \\ \frac{1}{(r\lambda t)^2} & n_\mathrm{A}>0\,. \end{cases} \tag{3.5}$$

From this system and its associated dynamics, one can draw several key implications:

Epidermis is maintained through an ever-decreasing clonal population:

Starting with a single labelled cell, the Galton-Watson process predicts that the persistance probability of the resulting clone (i.e., in this case, the probability that the clone retains at least one proliferating cell), is given by

$$p_{n_\mathrm{A}>0} \equiv 1 - p_0(t) = \frac{1}{1+r\lambda t}\,,$$

i.e. as with the experiment, the persistance probability of a clone decays monotonically, asymptoting to the form $1 - p_0(t) \propto 1/t$ at time scales $t \gg 1/r\lambda$, the time scale for symmetric division. Applied to the experimental system, this suggests that labelled clones continue to detach from the basal layer indefinitely. At the same time, defining

$$p^{\text{pers.}}_{n_\text{A}>0}(t) = \frac{p_{n_\text{A}}(t)}{1 - p_0(t)}\,,$$

as the size distribution of *persisting* clones, the mean number of basal layer cells in a persisting clone grows steadily as

$$\langle n_\text{A} \rangle_{\text{pers.}} \equiv \sum_{n=1}^{\infty} n_\text{A}\, p^{\text{pers.}}_{n_\text{A}>0}(t) = 1 + r\lambda t\,,$$

such that the overall cell population remains constant, viz. $\langle n_\text{A} \rangle \equiv \sum_{n=0}^{\infty} n_\text{A}\, p_{n_\text{A}}(t) = 1$, i.e. the continual extinction of clones is compensated by the steady growth of persisting clones such that the average number of proliferating cells remains constant: given enough time, all cells would derive from the same common ancestor, the hallmark of the Galton-Watson process[1].

This linear increase in clone size may lead one to worry about neighbouring clones coalescing. Fortunately, the continual extinction of clones ensures that the fraction of clones conjoined with their neighbours remains small and of same order as the initial labelling density[2]. The fact that this fraction is constant is again indicative of the steady-state condition maintained throughout the experiment.

Larger clones begin to exhibit the stability of the macroscopic system:

If, at some instant, a clone is seen to have, say, N_A proliferating cells then, after a further time t, its size will fluctuate as

$$\frac{\langle (n_\text{A} - \langle n_\text{A} \rangle)^2 \rangle^{1/2}}{\langle n_\text{A} \rangle} = \sqrt{\frac{2r\lambda t}{N_A}}\,.$$

Thus clones (as defined by the A cell population) will maintain an approximately stable number of cells providing $t \ll N_A/r\lambda$. For larger clones this time may exceed the lifetime of the system. At the limit where macroscopic sections of the basal layer are considered, the statistical fluctuations are small. The increased stability of larger clones also explains the surprising prediction that, given enough time, all clones eventually become extinct (viz. $\lim_{t\to\infty} p_{n>0}(t) = 0$). Calculated explicity, the extinction probability for a clone of size $N_A \gg 1$ scales as $p_0(t) \approx e^{-N_A/r\lambda t}$ [62] approaching unity at long times. However, because this extinction probability is small when $t \ll N_A/r\lambda$, a large enough clone may easily persist beyond the lifetime of the system.

[1] Curiously, Galton and Watson first dealt with this question amidst the concern that aristocratic surnames were becoming "extinct" in Victorian Britain [69]!

[2] The fraction of clones in contact with their neighbours is estimated by assuming that all clones are randomly and independently distributed, and by noting the empirical observation that clones are oblique in shape, so that the distribution in clone area has the same form as $P^{\text{pers.}}_{n>0}(t)$.

The properties of the proliferating cell population dictates the behaviour of the entire clone size distribution:

At asymptotically long times, one may show[3] that the full probability distribution for finding $n = n_{\rm A} + n_{\rm B}$ cells within a persisting clone scales in proportion to $p^{\rm pers.}_{n_{\rm A}}(t)$, viz.

$$\lim_{t\gg 1/r\lambda} P^{\rm pers.}_{n>0}(t) = \frac{\rho}{r\lambda t}\exp\left[-\frac{\rho n}{r\lambda t}\right] , \tag{3.6}$$

and so

$$\lim_{t\gg 1/r\lambda} \mathcal{P}^{\rm pers.}_{k}(t) \simeq \exp\left[-2^k\frac{\rho}{2r\lambda t}\right] - \exp\left[-2^k\frac{\rho}{r\lambda t}\right] , \tag{3.7}$$

i.e. the probability distribution acquires the scaling form found empirically. Referring to Eq. (3.1), we can therefore deduce the form of the scaling function,

$$f(x) = \exp[-\rho x/2r\lambda] - \exp[-\rho x/r\lambda] . \tag{3.8}$$

As a result, at long times, the average basal layer population of persisting clones becomes proportional to the average number of proliferating cells per clone, $\langle n\rangle_{\rm pers.} = (1 + r\lambda t)/\rho$, a behaviour consistent with that seen in experiment (see Fig. 3.3).

The creation and transfer of differentiated cells dictates the short-time behaviour of the clone size distribution:

In fitting the model to the data (see below), we will find that the rates λ and Γ at which differentiated cells are created and then transferred into the super-basal region are significantly larger than the rate of symmetric division $r\lambda$, which dictates the long-time behaviour of the clone size distribution. In this case, at early times ($t \lesssim 1/\Gamma$), the clone size distributions are dominated by the differentiation and transfer rates, which remain prominent until the population of labelled differentiated cells associated with each proliferating cell reaches its steady-state value of $(1-\rho)/\rho$. One may therefore infer that, at short times, the mean number of basal layer cells in clones arising from proliferating cells is given by

$$\lim_{t\ll 1/\Gamma}\langle n\rangle_{\rm pers.} = 1/\rho - (1/\rho - 1)e^{-\Gamma t} ,$$

and that the early-time clone size distribution is Poisson-distributed, viz.

$$\lim_{t\ll 1/\Gamma} P^{\rm pers.}_{n\geq 2}(t) = \frac{(\langle n\rangle_{\rm pers.} - 1)^{n-1}}{\left(e^{\langle n\rangle_{\rm pers.}-1} - 1\right)(n-1)!} . \tag{3.9}$$

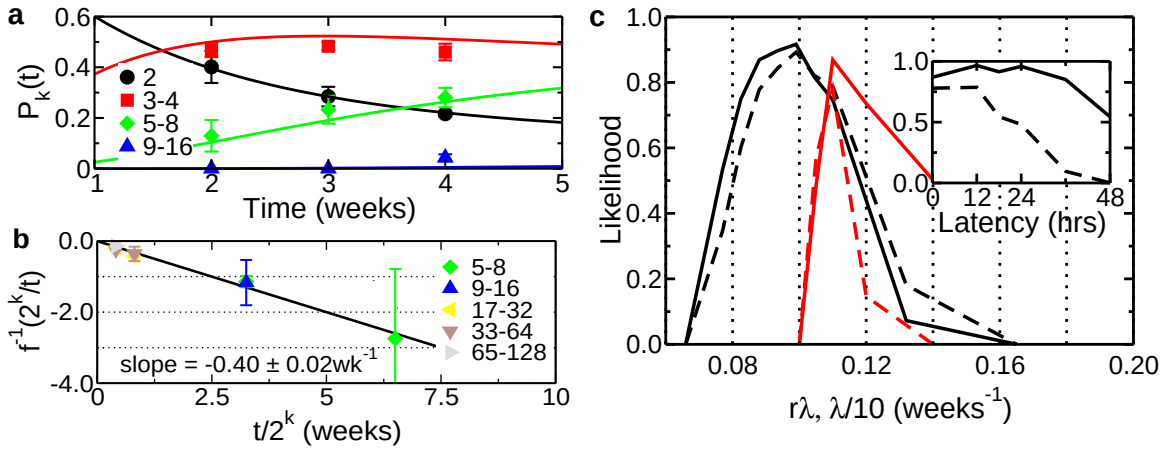


FIG. 3.5: (Color online) (a) Fit of Eq. (3.9) to the short-time clone size distributions. At early times (a), the data is optimally fitted by Eq. (3.9) using the value $\lambda = 1.1$/week and the empirical value $\rho = 0.22$ (solid lines show fit). To ensure integrity of the analysis, data for times earlier than week 2 have been excluded (see section 3.2.4). (b) Linearisation of the long-time asymptotic data using the "inverse" scaling function $f^{-1}(2^k/t)$ for $t \geq 13$ weeks and $k \geq 3$ (see main text). (c) Likelihood of the overall division rate λ (red) and the symmetric division rate $r\lambda$ (black), as assessed from a χ^2 test of the numerical solution to Eq. (3.3) [70]. A fit to the basal-layer clone size distribution alone (dashed) is less discriminatory than a simultaneous fit to both the basal-layer and total clone size distributions (solid curves). The likelihood of $r\lambda$ is shown for the optimal value of λ, and vice-versa. *Inset:* Referring to section 3.2.3, the likelihood is plotted against the duration of a latency period ($\tau_{\text{min.}}$) immediately following cell division, and assuming that division events are otherwise independent (see main text).

3.2.2 Fit to the data

With these insights it is now possible to attempt a fit of the model to the data. Referring to Fig. 3.5, one may infer the rate of cell division λ from the short-time data, and the symmetric division rate $r\lambda$ from the long-time scaling data. In particular, taking the fraction of proliferating cells in the basal layer to be $\rho = 0.22$, a figure obtained experimentally by immunostaining using Ki67 [15], a fit of Eq. (3.9) to the short-time data (fig. 3.5(a)) is consistent with a transfer rate of $\Gamma = 0.31$/week which, in turn, implies a rate of cell division of $\lambda = 1.1$/week. Furthermore, by plotting the long-time, large-k, size-distributions in terms of the "inverse" to the scaling function,

$$\begin{aligned} f^{-1}(2^k/t) &\equiv \left(2\ln\left[(1-(1-f(2^k/t))^{1/2})/2\right]\right)^{-1} \\ &= \left(2\ln\left[(1-(1-\mathcal{P}_k^{\text{pers.}}(t))^{1/2})/2\right]\right)^{-1}, \end{aligned}$$

the data converge onto a linear plot (Fig. 3.5(b)). The resulting slope takes the value $-r\lambda/\rho$, from which we may infer the symmetric division rate $r\lambda = 0.09 \pm 0.01$/week, and $r = 0.08 \pm 0.01$.

[3] To obtain Eq. (3.6), we treat n_{A}, n_{B} as continuous variables in Eq. (3.3) (a good approximation at large values of n). Then, making the ansatz that the B-cell population remains slave to the A-cell population viz. $n_{\text{A}} = \rho n$, $n_{\text{B}} = (1-\rho)n$, the Master equation simplifies to the approximate form

$$\partial_t P = \frac{r\lambda}{\rho}(n\partial_n^2 P + 2\partial_n P),$$

which is solved by $P_n(t) = (\rho/r\lambda t)^2 \exp(-\rho n/r\lambda t)$, leading to Eq. 3.6.

These figures compare well with an optimal fit of the *entire* basal layer clone size distribution (Fig. 3.4), obtained by numerically integrating the Master equation (3.3). The fitting procedure is shown in Fig. 3.5(c) (solid curves), where the likelihood of the model is evaluated for a range of values of λ and $r\lambda$, as assessed from a χ^2 test of the model solution [70]. One may see that the likelihood is maximised with an overall division rate of $\lambda = 1.1$/week and a symmetric division rate in the range $r\lambda = 0.1 \pm 0.01$/week, thus confirming the validity of the asymptotic fits. Moreover, the corresponding fit of both the basal layer distribution and the *total* clone size distribution, including both basal and supra-basal cells, is equally favourable (Fig. 3.5(c), dashed). Thus, in the following sections we shall use the asymptotically fitted value of $r = 0.08$, however any choice of the parameter in the range $r = 0.08 - 0.10$ gives similar results.

Although the comparison of the experimental data with the model leaves little doubt in its validity, it is important to question how discerning is the fit. By itself, the observed increase in the size of persisting clones is sufficient to rule out any model based on long-lived self-renewing stem cells, the basis of the orthodox EPU model. However, could one construct a more complicated model, which would still yield a similar fit? Certainly, providing the long-time evolution is controlled by a single rate-determining process, the incorporation of further short-lived proliferating cell compartments (viz. transit-amplifying cells) would not affect the observed long-time scaling behaviour. However, it seems unlikely that such generalisations would provide an equally good fit to the short-time data.

More importantly, it is crucial to emphasize that the current experimental arrangement would be insensitive to the presence of a quiescent, long-lived stem cell population. Yet, such a population could play a crucial role in *non*-steady state dynamics such as that associated with wound healing or development. We are therefore led to conclude that the range of clone fate data for normal adult IFE are consistent with a simple non-equilibrium process involving just a single progenitor cell compartment.

3.2.3 Stochastic behaviour of cell division

At this stage, it is useful to reflect upon the sensitivity of the model to the stochasticity assumption applied to the process of cell division. Clearly, the scaling behaviour (Eq. 3.6) depends critically on the statistical independence of successive cell divisions; each cell division results in symmetric/asymmetric cell fate with relative probabilities as detailed in (3.4). But, to what extent would the findings above be compromised if the cell *cycle-time*, i.e. the time between consecutive cell divisions, were not determined by an independent stochastic process? This question may have important ramifications, because the assumption of independent cell division, used in formulating the Master equation (3.3), introduces a manifestly unphysical behaviour by allowing cells to have arbitrarily short cycle times. Moreover, although a wide distribution of cell cycle-times has been observed for human keratinocytes *in vitro* [71], it is possible that that keratinocytes *in vivo* may divide in *synchrony*, giving a cell cycle-time distribution narrowly centered about the mean $(1/\lambda)$. In the following, we shall address both of these points: Firstly, we shall show that, up to some potential latency period (the time delay before a newly-divided cell is able to divide again), consecutive cell divisions occur independently as an asychronous, Poisson process. Secondly, while the data is insufficient to detect a latency period of 12 hours or less between consecutive cell divisions, the data does discriminates against a period lasting longer than 24 hours.

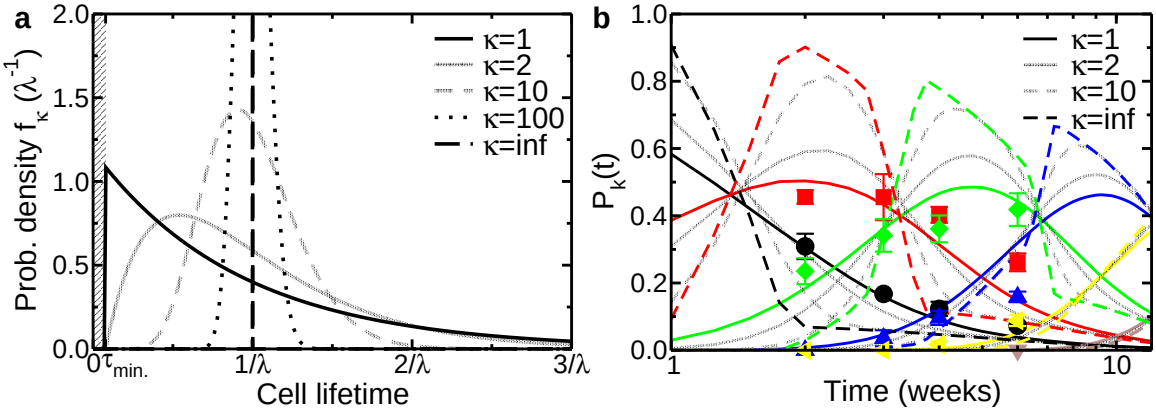


FIG. 3.6: (Color online). (a) Examples of progenitor cell cycle-time distributions with the same average cycle time $1/\lambda$, ($\lambda = 1.1$/week), and with a latency period of $\tau_{\text{min.}} = 12$ hours introduced between consecutive cell divisions (hashed region). The case $\kappa = 1$ corresponds to a model of independent cell division such as assumed in section 3.2.1, but now accounting for an initial latency period. The case $\kappa \to \infty$ (black dashed) corresponds to all cells having an exact cell cycle-time of $1/\lambda$. Note that small values of κ allow for both very short and very long cycle times. (b) Using Monte-Carlo simulations of process (3.2), the clone size distributions predicted by each of the different cycle-time distributions in (a) are compared with the empirical data. Data points show the size distribution of persisting clones including supra-basal layer cells over the first 6 weeks post-labelling (extracted from data given fully in Ref. [15]; for legend see Fig. 3.4), and the theoretical curves correspond to the same legend as in (a). All of the models give an optimal fit with the same value of $\lambda = 1.1$/week, $r = 0.08$.

To investigate the degree to which the model is sensitive to the particular cell cycle-time distribution, let us revisit the original model of independent cell division with several variations: Firstly, we introduce a latency period of τ_{min} immediately following cell division, in which daughter cells cannot divide. This biologically-motivated constraint renders a more complicated yet more realistic model of cell division than the idealised system studied in the previous section. Motivated by observations of the minimal cycle-time of (human) keratinocytes [71], where a latency period of $\tau_{\text{min}} \simeq 10$ hours was observed *in vitro*, we shall here consider the a range latency periods of up to 48 hours. Secondly, we compare the empirical clone size distributions with a model where all progenitor cells have a cycle-time of exactly $1/\lambda$, i.e. where cells within each clone divide in perfect synchrony. Finally, we shall investigate a range of intermediate models with different distributions of progenitor cell cycle-time (see Fig. 3.6(a)).

Technically, the resulting clone size distributions may be evaluated through Monte Carlo simulations of the non-equilibrium process (3.2) with the cycle-time τ of each proliferating cell selected at random from a Gamma distribution of the form

$$f_\kappa(\tau) = \begin{cases} 0 & \tau < \tau_{\text{min.}} \\ \frac{\kappa^\kappa (\tau - \tau_{\text{min.}})^{\kappa-1}}{\bar{\tau}^\kappa \Gamma(\kappa)} e^{-\frac{\kappa(\tau-\tau_{\text{min.}})}{\bar{\tau}}} & \tau \geq \tau_{\text{min.}} \end{cases},$$

where $\bar{\tau} = 1/\lambda - \tau_{\text{min.}}$ is the average time to division following the initial latency period $\tau_{\text{min.}}$, and κ is the "shape parameter" of the Gamma distribution. In particular, the choice of shape parameter $\kappa = 1$ corresponds to the exponential distribution which characterises the independent cell cycle-time distribution, whereas $\kappa \to \infty$ describes the case in which all A-cells have an exact cycle-time of $1/\lambda$ (see Fig. 3.6(a)). Then, to reflect the assumption that initially-labelled, spatially separated,

progenitor cells have uncorrelated cell cycles, the time to the initial division event post-labelling is adjusted by a random time $\tau \in [0, 1/\lambda]$. Finally, for an unbiased comparison of the models, we optimise the value of λ for each model separately against the empirical data, whilst keeping $r\lambda = \text{const.}$ to ensure an optimal fit of the long-time data, as discussed below.

The resulting clone size distributions are shown in Fig. 3.6(b), where the case of independent division following a 12-hour latency ($\kappa = 1$) and the exact cycle-time case ($\kappa \to \infty$) are compared to the empirical *total* clone size distribution, which includes both basal and supra-basal (type C) cells, over the first 6 weeks post-labelling. Two intermediate cases are also shown for comparison ($\kappa = 2, 10$). Focusing first on the results for the case $\kappa = 1$, which bears closest resemblance to the Markovian model analysed using the Master Equation (3.3), one may see by inspection that the quality of the fit to the data remains good even when the effects of a latency period between cell divisions is taken into account. More rigorously, a likelihood analysis reveals that the two cases are statistically indistinguishable (see Fig. 3.5(c), inset), which indicates that the duration of a latency period of $\tau_{\text{min.}} \lesssim 12$ hours is beyond the current empirical resolution. However, referring to Fig. 3.5c (inset), a similar analysis of longer latency periods reveals that for periods of $\tau_{\text{min.}} \gtrsim 24$ hours, the fit to the data is significantly poorer.

Turning next to the predicted basal-layer clone size distributions at late times ($t \gtrsim \rho/r\lambda$) (not shown), one may see that all of the proposed distributions asymptotically converge: Starting with exactly one cell, then the moment-generating function $G(q, s) = \sum_{n=0}^{\infty} p_n(s) q^n$ associated with the A cell population distribution $p_n(s)$ after s cell cycles satisfies the recursion relation [72]:

$$G(q, s+1) - G(q, s) = r \left(G(q, s) - 1\right)^2 ,$$

which asymptotes to the continuous master equation $\lim_{s \gg 1} \partial_s G(q, s) = r \left(G(q, s) - 1\right)^2$, with the relative magnitude of the leading-order correction dropping off as $1/s$. But with $s = \lambda t$, this equation is simply the master equation for the moment-generating function associated with the original model, Eq. 3.5, and so the two models converge. One may therefore conclude that, beyond the first several weeks of the experiment ($t \gg 1/\lambda$), the fit to the data is sensitive only to the average cycle time of progenitor cells. With this in mind, we note that for the case of perfectly synchronous cell division, an optimal (albeit poor) numerical fit was obtained when $\lambda = 1.2/\text{week}$, a figure that compares well with the fit for the independent case. It appears therefore that *the predicted average cell division rate (λ) is insensitive to the shape of the cell cycle distribution.*

Finally, let us turn to the early time behaviour ($t \sim 1/\lambda$), where the predicted distributions are distinct. Referring to Fig. 3.6(b), one may see, at 2-4 weeks post-labelling, that relatively large clones (5−8 cells) appear earlier than expected by a model assuming synchronous division, and that, compared with the same model, a sizeable proportion of small clones (e.g., 2 cells) lingers on for far longer than expected. The same behaviour is observed for the basal layer clone size distribution (not shown). One may therefore infer that cell division conforms to a model of *independent* rather than *synchronous* division, allowing for some progenitor cells to divide unusually early, and for others to remain quiescent for an unusually long period of time.

In summary, we have established that, following division, progenitor cells do not divide for a period that is likely to last up to 12 hours, and not more than 24 hours. After this latency period, the data is consistent with cells switching to a mode of independent, asynchronous, cell division. These results shed light on why the simple model of independent cell division presented in section 3.2.1 succeeds in producing such a remarkable fit to the data.

3.2.4 Labelling efficiency and EYFP accumulation in basal cells

Although the integrity of the fit of the model to the data provides some confidence in its applicability to the experimental system, its viability as a model of epidermal homeostasis rests on the labelled clone population being representative of all cells in the IFE. Already, we have seen that the model, and by inference, the labelled clone population, has the capacity to self-renew. However, the slow accumulation of EYFP after induction, together with the question of the relative labelling efficiency of the two basal layer cell types, leaves open the question of the very short-time behaviour. Accepting the validity of the model, we are now in a position to address this regime.

In doing so, it is particularly useful to refer to the time evolution of clone size as measured by the average number of basal cells in a persisting clone. As expected from the scaling analysis discussed in section 3.1.2, a comparison of the experimental data with that predicted by the proposed cell kinetic model shows a good agreement at long times (Fig. 3.3). However, comparison of the data at intermediate time-scales provides significant new insight. In particular, if we assume *equal* labelling efficiency of progenitor and differentiated cells, i.e. that both cell types label in proportion to their steady-state population (shown as the lower (red) curve in the Fig. 3.3), then there is a substantial departure of the predicted curve from the experimental data for times of between two and six weeks. Intriguingly, if we assume that differentiated cells simply don't label, then the agreement of the data with theory is excellent from two weeks on! We are therefore lead to conclude that, at least from two weeks, all labelled clones derive from progenitor cells labelled at induction.

With this in mind, we may now turn to the average clone size as inferred from the data at two days and one week. Here one finds that the model appears to substantially over-estimate the clone size. Indeed, Fig. 3.3 suggests that the average clone size is pinned near unity until beyond the first week post-labelling, i.e. the relative population of single-cell clones is significantly *larger* than expected at one week, yet falls dramatically to the theoretical value at two weeks. Referring again to the slow accumulation of EYFP, can one explain the over-representation of single-cell clones at one week post-labelling? At one week, two-cell clones are observed soon after cell division, and thus express lower concentrations of EYFP compared to single-cell clones. As a result they may be under-represented. At later times, all labelled clones become visible as EYFP concentration grows, explaining the coincidence of experiment and theory at two weeks. It follows, of course, that the size distributions at later time points would be unaffected by slow EYFP accumulation. However, a full explanation of this effect warrants further experimental investigation, and is beyond the scope of this paper.

3.3 Manifestation of mutations in clonal distributions

Having elucidated the mechanism of normal skin maintenance, it is interesting to address its potential as a predictive tool in clonal analysis. Conceptually, the action of mutations, drug treatments or other environmental changes to the tissue can effect the non-equilibrium dynamics in a variety of ways: Firstly, a revision of cell division rates or "branching ratios" (i.e. symmetric vs. asymmetric) of *all* cells may drive the system towards either a new non-equilibrium steady-state or towards a non-steady state evolution resulting in atrofication or unconstrained growth of the tissue. (The

development of closed non steady-state behaviour in the form of limit cycles seems infeasible in the context of cellular structures.) Secondly, the stochastic revision of cell division rates or branching ratios of *individual* cells may lead to cancerous growth or extinction of a sub-population of clones. The former may be referred to as a "global perturbation" of the cell division process while the second can be referred to as "local". In both cases, one may expect clonal analysis to provide a precise diagnostic tool in accessing cell kinetics. To target our discussion to the current experimental system, in the following we will focus on the action of a local perturbation in the form of a carcinogenic mutation, reserving discussion of a global perturbation, and its ramifications for the study of drug treatment, to a separate publication.

Let us then consider the action of a local perturbation involving the activation of a cancer gene in a small number of epidermal cells, which leads to the eventual formation of tumours. In the experimental system, one can envisage the treatment coinciding with label induction, for example by simultaneously activating the EYFP and the cancer gene. In this case, clonal fate data should simply reflect a modified model of cell proliferation leading to the eventual failure of the steady-state model of tissue maintenance.

3.3.1 A simple model of carcinogenesis

To quantify the process of cancer onset, we start by establishing the simplest possible changes to process (3.2) which may be associated with tumour growth. Cancer is widely held to be a disease caused by genetic instability that is thought to arise when a progenitor cell undergoes a series of mutations [73, 74, 75]. As a result, cells within the mutant clone prefer to proliferate, on average, over processes leading to terminal differentiation or death. In this investigation we shall consider a "simple" cancer resulting from *two* rate-limiting mutations: Referring to our proposed labelling experiment, the controlled induction of a cancer-causing mutation during label induction defines the first mutation; a second, rate-limiting step then occurs with the stochastic occurrence of a second cancer causing mutation. Examples of the first type of mutation may be genes that affect the ability of a cell to respond to genetic changes of the cell, e.g. *p53*, whilst the second mutation may be of a gene that affects clone fate such as the *Ras* oncogene [74]. We may therefore distinguish between "stage one" mutated cells, which maintain the steady-state, and "stage two" cells, which have the capacity for tumour formation.

The resulting process of cell proliferation is set by three parameters: The overall rate of mutation ν from a stage one A cell into a cancerous stage two cell; the division rate μ of the stage two cells; and the degree of imbalance Δ between their stochastic rate of proliferation and differentiation. In summary, focusing on the proliferating cell compartment only, and denoting the stage two mutated cells as type A*, then the revised cell proliferation model includes the additional non-equilibrium processes

$$\begin{array}{lcl} A & \xrightarrow{\nu} & A^* \\ A^* & \xrightarrow{\mu} & \begin{cases} A^* + A^* & \text{prob. } (1+\Delta)/2 \\ \emptyset & \text{prob. } (1-\Delta)/2\,. \end{cases} \end{array} \tag{3.10}$$

The rate ν may be interpreted as the mean rate with which a stage-one cell acquires an additional mutation necessary to activate a second oncogene. The mutated cells then give rise, on average, to an exponentially growing cell lineage with growth rate $\Delta\mu$.

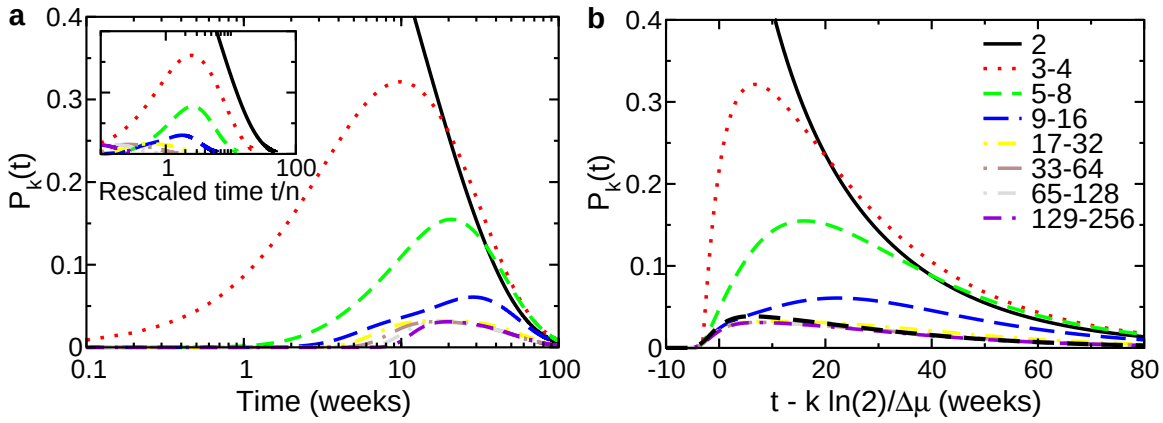


FIG. 3.7: (Color online) (a) The total number of basal layer cells per labelled clone during the onset of cancer according to process (3.10). The figure was plotted by numerically integrating Eq. (3.11) using the empirical value $r\lambda = 0.088$/week found for normal skin, together with hypothetical values of the cancer growth parameters $\nu = 0.1\,r\lambda$, $\mu = 10\,r\lambda$, and $\Delta = 0.5$. To compare with normal skin, the predicted clone size distributions are replotted against the rescaled time coordinate $t/2^k \mapsto t$ in (a) inset. In contrast with Fig. 3.4, here the curves fail to converge. In (b), the same curves are shown converge onto the universal form given in Eq. (3.11) (dashed) when they are plotted against a new rescaled time $t \mapsto t'_k = t + k \ln 2/\Delta\mu$. Note that the large-clone distributions converge rapidly, whereas the distributions for smaller clones are affected by the non-negligible contribution of non-cancerous (A) cells to the small-clone size distribution.

This nonequilibrium process was originally addressed by Kendall, who predicted the distribution in the number of tumours detected at time t after mutation [76]. His focus on tumour statistics may reflect the experimental limitations in clonal analysis at the time: Until recently it was not possible to reliably detect clones at all, let alone to count the number of cells per clone. Experimentally, however, the clone size distributions are a more efficient measure of cell kinetics than the tumour number distributions, because they result in a far richer data set, and are accessible within weeks rather than months. We shall therefore extend Kendall's approach to predict the clone size distributions at times far earlier than tumour appearance.

3.3.2 Clonal behaviour during early-stage cancer

To familiarise ourselves with the modified model, consider the evolution of the average clone size with time. Focusing on the proliferating cell compartment with n type A cells and n^* type A* cells in a clone, the relevant mean-field equations are

$$\partial_t \langle n \rangle = -\nu \langle n \rangle \,,$$
$$\partial_t \langle n^* \rangle = \nu \langle n \rangle + \Delta\mu \langle n^* \rangle \,,$$

which give the expected shift from linear growth of clones in normal skin to that of exponential growth, $\langle n+n^* \rangle = (\nu e^{\Delta\mu t} + \Delta\mu e^{-\nu t})/(\nu+\Delta\mu)$. More interestingly, referring to the Master equation below, one may show that the variance in clone size also changes qualitatively: Whereas for normal skin the RMS variance in clone size grows as $t^{1/2}$, here the variance in the long-time limit is *finite*,

$$\lim_{t\to\infty} \frac{\langle (n^* - \langle n^* \rangle)^2 \rangle^{1/2}}{\langle n^* \rangle} = \sqrt{1 + \Delta^{-1}} \,.$$

That is, the relative broadening of the clone size distribution observed in normal skin is halted by the introduction of an exponentially growing cell population.

These observations may already provide a crude method for identifying carcinogenesis through clonal analysis. To do better, it becomes necessary to solve for the full size distribution by extending the Master equation (3.3) to include process (3.10). If we neglect the fate of differentiated cells, then the Master equation now describes the evolution of the probability $P_{n,n^*}(t)$ for finding n type A cells and n^* type A* cells in a clone,

$$\begin{aligned}\partial_t P_{n,n^*} = r\lambda\left[(n-1)P_{n-1,n^*} - nP_{n,n^*}\right] + r\lambda\left[(n+1)P_{n+1,n^*} - nP_{n,n^*}\right] + \nu\left[(n+1)P_{n+1,n^*-1} - nP_{n,n^*}\right] \\ + \frac{1+\Delta}{2}\mu\left[(n^*-1)P_{n,n^*-1} - n^*P_{n,n^*}\right] + \frac{1-\Delta}{2}\mu\left[(n^*+1)P_{n,n^*+1} - n^*P_{n,n^*}\right] ,\end{aligned}$$

subject to the experimental boundary condition $P_{n,n^*}(0) = \delta_{1,0}$ corresponding to exactly one "stage one" cell per clone at $t = 0$. As for the case of normal skin, we shall later be interested in the distribution of persistent clones, defined as,

$$\mathcal{P}_{2^k}^{(\text{canc.})}(t) = \sum_{N=2^{k-1}+1}^{2^k}\sum_{n=0}^{N}\frac{P_{n,N-n}(t)}{1-P_{0,0}(t)-P_{1,0}(t)-P_{0,1}(t)} .$$

While it is not possible to solve Eq. (3.11) analytically, progress may be made when we allow for the widely-accepted view that tumours are *monoclonal*, that is they arise from a single "stage two" mutated cell [74]. This assumption conveniently limits us to the parameter space $\nu \ll \Delta\mu$, for which an approximate long-time solution for the full clone size distribution may be found.

Referring to the appendix for details, we find that the binned clone size distribution takes the long-time asymptotic scaling form,

$$\mathcal{P}_k^{(\text{canc.})}(t) \simeq \mathcal{N}\left[\mathcal{I}_{\beta,a}\left(\frac{1}{2\phi_k(t)}\right) - \mathcal{I}_{\beta,a}\left(\frac{1}{\phi_k(t)}\right)\right] , \tag{3.11}$$

where $\phi_k(t) = (1+\Delta^{-1})e^{\Delta\mu t}/2^k$

$$\mathcal{I}_{\beta,a}(x) = \int_1^\infty d\zeta\,\frac{\zeta^{-1-\beta}e^{-x\zeta}}{(1+a\zeta^{-\beta})^2} ,$$

$\mathcal{N} = \frac{4r\lambda\chi^2}{\Delta\mu(\chi+\nu/2r\lambda)}$, $\chi^2 = \left(\frac{\nu}{2r\lambda}\right)^2 + 2\frac{\Delta\nu}{(1+\Delta)r\lambda}$, $\beta = 2\chi r\lambda/\Delta\mu$, and $a = \frac{2r\lambda\chi-\nu}{2r\lambda\chi+\nu}$. Despite its apparent complexity, this distribution is characterised by a simple scaling behaviour: Referring to Fig. 3.7(a), the predicted clone size distributions are plotted using the scaling appropriate to the normal (unperturbed) system (cf. Fig. 3.4). In this case, it is apparent that the scaling $t \mapsto t/2^k$ fails. By contrast, from the expression for $\phi_k(t)$, it is clear that the size distributions should scale according to the time translation, $t \mapsto t'_k = t + k\ln 2/\Delta\mu$ as confirmed by the results shown in Fig. 3.7(b). Further consideration of the size distribution exposes several additional features, which may provide further access to the new model parameters:

(i) The long-time distribution decays with a rate $\beta\Delta\mu$: Expanding $\mathcal{I}_{\beta,a}(x)$ for small x gives us the asymptotic form of the universal decay curve. For $\beta < 1$, consistent with the monoclonicity requirement $\Delta\mu \gg \nu$, we find

$$\lim_{t\gg\Delta\mu}\mathcal{P}_k^{(\text{canc.})}(t) = \mathcal{N}\,\Gamma(-\beta)(2^{-\beta}-1)\phi_k(t)^{-\beta}, \tag{3.12}$$

where $\Gamma(x)$ denotes the Gamma function. This expression allows us to estimate β from the rescaled clone size distributions, providing access to the cell division and mutation parameters of the observed cells.

(ii) The probability of tumour formation is finite: This is a well-known feature of the simple non-critical birth-death process (3.10) [62]. Referring to the appendix, we find that the probability p_T for any given clone to survive and form a tumour is finite,

$$p_T = 1 + \frac{\nu}{2r\lambda} - \sqrt{\left(\frac{\nu}{2r\lambda}\right)^2 + \frac{2\Delta\nu}{r\lambda(1+\Delta)}}\,.$$

As a result, the onset of cancer will halt the steady decrease in the density of labelled clones that is a hallmark of the unperturbed system.

These properties, and especially the change in scaling behaviour, allow the onset of early-stage cancer to be identified from observations of clones less than one hundred cells in size. This may provide a dramatic improvement both in speed and accuracy over current experimental models, which rely on much later observations of tumours (or hyperplasias) in order to deduce the cell kinetics at early-stages.

3.4 Chapter conclusions: A new paradigm for epidermal maintenance

To summarize, we have shown that the range of clone fate data obtained from measurements of murine tail epidermis are consistent with a remarkably simple stochastic model of cell division and differentiation involving just one proliferating cell compartment. These findings overturn a long-standing paradigm of epidermal fate which places emphasis on a stem cell supported epidermal proliferative unit. As well as providing significant new insight into the mechanism of epidermal homeostasis, these results suggest the utility of inducible genetic labelleling as a means to resolve the mechanism of cell fate in other tissue types, and as a means to explore quantitatively the effects of drug treatment and mutation.

To conclude, we note that the analysis above has focused on the dynamics of the clonal population without regard to the spatial characteristics. Indeed, we have implicitly assumed that any model capable of describing the cell size distributions will also succeed in maintaining the near-uniform areal cell density observed in the basal layer. However, it is known that, when augmented by spatial diffusion, a simple Galton-Watson birth-death process leads to "cluster" formation in the two-dimensional system whereupon local cell densities diverge logarithmically [77, 78]. Significantly, these divergences can not be regulated through a density-dependent mobility. In chapter 5 we shall examine how the Galton-Watson process emerges from a two-dimensional reaction-diffusion type process.

From a practical perspective, there is also the significant question of how the cell kinetic model might be generalised to describe other forms of epidermis. In particular, although it is not feasible

to repeat these experiments *in vivo* in humans, in chapters 6-7 we shall examine new evidence for cell fate regulation in human epidermis, a system of obvious medical significance.

Lastly, our analysis of the cancer system referred to the relatively simple case of a two-stage mutation. It is, of course, well-known that tumour formation is usually the result of multiple mutations. Understanding whether clonal fate data can be used to probe the kinetics of *multi-stage* mutation remains an interesting future challenge.

Chapter appendix

3.A Clone size distributions in the two-stage cancer model

To derive the clone size distribution given in Eq. (3.11), we start by quoting the known result for the probability distribution $\Pi_{n^*}(t;\tau)$ of finding n^* stage-two A* cells at time t starting from a single A* cell at time τ [62],

$$\Pi_{n^*}(t;\tau) = \begin{cases} \frac{(1-\Delta)\left(1-e^{-\Delta\mu(t-\tau)}\right)}{(1+\Delta)-(1-\Delta)e^{-\Delta\mu(t-\tau)}} & \text{for } n^*=0 \\ \left(\frac{2\Delta}{(1+\Delta)\left(1-e^{-\Delta\mu(t-\tau)}\right)}\right)^2 e^{-\Delta\mu(t-\tau)}\left(1-\frac{2\Delta e^{-\Delta\mu(t-\tau)}}{(1+\Delta)-(1-\Delta)e^{-\Delta\mu(t-\tau)}}\right)^{n^*-1} & \text{for } n^*\geq 1 \end{cases} \tag{3.13}$$

When $\Delta\mu(t-\tau) \gg 1$, this distribution asymptotes to the form

$$\Pi_{n^*}(t;\tau) \simeq \begin{cases} \frac{(1-\Delta)}{(1+\Delta)} & \text{for } n^*=0 \\ \left(\frac{2\Delta}{(1+\Delta)}\right)^2 e^{-\Delta\mu(t-\tau)}\times & \\ \quad \exp\left(-n^*\frac{2\Delta}{(1+\Delta)}e^{-\Delta\mu(t-\tau)}\right) & \text{for } n^*\geq 1 \end{cases} \tag{3.14}$$

From the value of Π_0 we see that even when a cell has mutated, it is not guaranteed to result in a tumour: This will only occur with a probability of $f = 1-\Pi_0(t\to\infty) = (2\Delta)/(1+\Delta)$. The value of f plays an important role in determining the statistics of tumour formation, as will be seen below.

We now make two approximations: First, we take the long-time clone size distribution to be dominated by the statistics of A* cells. This is a safe assumption at times $t \gtrsim 1/\nu$ and $t \gg 1/\Delta\mu$, as may be seen by considering the behaviour of the mean-field equations in section 3.3.2. This approximation allows us to focus on the size distribution of A* cells only, $p_{n^*}(t)$, which is related to the full clone size distribution by the sum $p_{n^*}(t) = \sum_{n=0}^{\infty} P_{n,n^*}(t)$. Secondly, we assume that the entire population of type A* cells in each clone arises from the first mutated cell that gives rise to a stable, exponentially growing lineage of cells. This corresponds to the condition $\nu \ll \Delta\mu$, as discussed in the main text.

With these two approximations, the probability of finding a labelled clone containing $n^* > 0$ mutated cells is given by the population distribution of the *first surviving cell lineage* of A* cells,

$$p_{n^*}(t) \simeq \mathcal{N}\int_0^t \Pi_{n^*}(t-\tau)\sum_{m=1}^{\infty}(1-f)^{m-1}r_m(\tau)d\tau\,, \tag{3.15}$$

where $\mathcal{N}$ is some normalisation constant, and we have introduced the probability $r_m(\tau)d\tau$ for the m-th lineage of mutated cells within a given clone to be created during the interval $\tau\to\tau+d\tau$ through the mutation process $A\to A^*$. The weight factor $(1-f)^{m-1}$ gives the probability that the first $m-1$ cell lineages of A* cells within a clone will become extinct — a situation necessary to make the m-th cell line relevant to the distribution according to the monoclonal approximation.

The rates $r_m(\tau)$ may be accessed by considering the probability $w_{n,m}(t)$ that a clone containing n type A cells at time t also contains m *independent lineages* of mutated A* cells, each arising from a separate mutation event. (Later we shall treat the evolution of each of these cell lines post-creation). To solve for $w_{n,m}(t)$ we must introduce its moment-generating function $G(q,Q^*;t) \equiv \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} w_{n,m}(t) q^n (Q^*)^m$, which evolves (from Eq. 3.11) according to the dynamical equation,

$$\dot{G} = \left[r\lambda(q-1)^2 + \nu(Q^*-q)\right]\partial_q G\,. \tag{3.16}$$

Solving this equation subject to the initial condition of one "stage-one" (A) cell per clone, we find the solution

$$G(q,Q^*;t) = \xi_{Q^*} - \frac{2\xi_{Q^*}}{1+\frac{\xi_{Q^*}+(q-1-\frac{\nu}{2r\lambda})}{\xi_{Q^*}-(q-1-\frac{\nu}{2r\lambda})}e^{-2\xi_{Q^*}r\lambda t}} + 1 + \frac{\nu}{2r\lambda}\,, \tag{3.17}$$

with $\xi_{Q^*} \equiv \sqrt{(\frac{\nu}{2r\lambda})^2 + \frac{\nu}{r\lambda}(1-Q^*)}$. Eq. (3.17) describes the evolution of a single A cell as it proliferates and eventually gives rise to a set of internal lines of mutated cells.

Before we proceed to find $p_{n^*}(t)$, note that setting $q=1, Q^*=1-f$ in Eq. (3.17) gives us the result (quoted in the main text) for the asymptotic fraction p_T of clones in which all mutated cell lines become extinct,

$$p_T = 1 + \frac{\nu}{2r\lambda} - \sqrt{\left(\frac{\nu}{2r\lambda}\right)^2 + \frac{\nu f}{r\lambda}}\,.$$

On the other hand, setting $q=1$ only in Eq. (3.17) gives the moment-generating function for (yet another) distribution $W_m(t) = \sum_{n=0}^{\infty} w_{n,m}(t)$ of a clone containing m independent lines of A* cells irrespective of the number of normal cells in the clone. Finally, noting that $\dot{W}_m(t) = r_m - r_{m+1}$ then gives:

$$r_m(t) = -\sum_{n=0}^{m-1} \dot{W}_m(t)\,, \tag{3.18}$$

which we may substitute into Eq. (3.15) to find (for $n^* > 0$)

$$p_{n^*}(t) \simeq -\mathcal{N}\int_0^t d\tau \Pi_{n^*}(t-\tau)\dot{G}(1,1-f;\tau)\,. \tag{3.19}$$

From this expression, simplified by the large-n^* approximation ($\sum_{n^*} \simeq \int dn$), we obtain the final form of the binned size distribution given in Eq. (3.11).

Chapter 4

Effects of drug treatment on the mechanism of epidermal homeostasis

Chapter overview

Understanding the complex phenotypes produced by drugs *in vivo* requires insights into how drugs alter the behaviour of proliferating and differentiating cells that function to maintain normal tissue homeostasis. In chapter 3 we established a method to access the laws of *in vivo* cell kinetics through the analysis of the fate of a representative sample of clones of genetically labelled cells and their progeny. Here, we explore theoretically and test experimentally the predictive power of clonal analysis in evaluating the effects of topical drug action on the epidermis. Using *all-trans retinoic acid* as a test drug, we show that the results of an extended drug course may be successfully predicted by the single compartment model. As well as providing validation of the model of normal epidermal homeostasis, this chapter establishes a platform for characterising the effects of drugs on the epidermis.

This chapter demonstrates a very close overlap between experiment and the theory of epidermal maintenance, and it has been written to allow publication at some future date. While the theoretical analysis is completely in place, the experiment reported here requires a repeat of one of the control experiments (as described in the body of the text). Experimental work was performed by E. Clayton[1], D.P. Doupé[1] and P.H. Jones[1], while AMK conducted the theoretical work and is the author of the text.

4.1 Introduction

A major challenge in studying the effects of drugs on tissues is the ability to dissect the effects of the drug on cell behaviour from complex tissue phenotypes that may follow drug treatment.

[1] *MRC Cancer Cell Unit, Hutchison-MRC Research Centre, Cambridge CB2 2XZ, UK*

Normal adult tissues are maintained in a homeostatic state by proliferating cells which replace cells lost from the tissue. Drugs may act to alter the rates of cell proliferation, differentiation or cell death, but it is impossible to measure these changes and their effect on cell behaviour with precision using conventional histology. A powerful technique in solving this problem is clonal analysis, the genetic labelling of a representative sample of cells within the tissue to enable their fate and that of their progeny to be tracked. This approach gives information on the proliferation, migration, differentiation and cell death (apoptosis) of the labelled cells. In chapters 2 and 3, this approach was used to demonstrate that cell proliferation and differentiation in adult murine epidermis conform to a remarkably simple model. Motivated by the success of this method in resolving the rules of normal epidermal maintenance, the aim of the present chapter is to explore theoretically the action of a drug on the epidermis through signatures in the resulting clone size distributions. To test the ideas developed in this work, we will also present the results of a pilot investigation exploring the influence of all trans retinoic acid (ATRA) on mouse tail epidermis.

The chapter is organised as follows: Using the results obtained in chapter 3 for normal (wildtype) epidermis as a platform, in section 4.2 we will investigate the action of drug treatment on epidermal maintenance focusing on how revisions in the clone fate data can be used to infer changes in the respective cell division rates and "branching ratios". To illustrate the integrity of the model, its utility as a diagnostic tool, and its predictive power, in section 4.3 we will describe the results of a pilot study of ATRA treatment on mouse tail epidermis. Finally, in section 4.4, we will close with a discussion of the methodology and future challenges.

4.2 Effects of drug treatment on the epidermis

We turn now to consider how clonal analysis may be used to infer the effect of drug treatment or other environmental changes on the rates and channels of cell proliferation and differentiation in the epidermis. As described in section 3.3 for the case of mutations, a drug treatment can affect tissue in a variety of ways:

- Firstly, a revision of cell division rates of *all* cells may drive the system towards either a new non-equilibrium steady-state behaviour or towards a non-steady state evolution resulting in atrofication or unconstrained growth of tissue[2].
- Secondly, the random revision of cell division rates or "branching ratios" (symmetric vs. asymmetric) of *individual* cells may lead to cancerous growth or extinction of a sub-population of clones.

The former may be referred to as a "global perturbation" of the cell division process while the second can be referred to as "local". In both cases, one may expect clonal analyses to provide a precise diagnostic tool in accessing cell kinetics. To target our discussion to the current experimental system, in the following we will focus on the action of a global perturbation. The effects of a local

[2]The development of closed non steady-state behavior in the form of limit cycles seems infeasible in the context of cellular structures.

perturbation, and its ramifications for the study of cancer development, have already been briefly discussed in the previous chapter (section 3.3).

Let us then consider the action of a global perturbation resulting from a drug treatment of the epidermis and leading to new steady-state behavior. Referring to the birth-death process (3.2), it is clear that, in the absence of new channels of cell division, the revised steady state behaviour may be characterised in full by the change in the overall cell division rate, λ, the branching ratio, r, and the fraction of A cells in the basal layer, ρ. However, if we restrict attention to the long-time scaling regime, referring to Eq. (3.6), one may see that, when coupled to direct observations of ρ, the change in the rate of *symmetric* division, $r\lambda$, is sufficient to fully specify the time-evolution of the resulting clone size distribution. A further numerical analysis of the transient short-time behaviour of the system may then be used to determine the independent changes to r and λ.

In the experimental system, one can envisage treatment either preceding or following label induction, each leading to different transient behaviours of observed clonal fate. If treatment is initiated early enough, one may expect the tissue to establish a revised steady-state by the time of induction. In this case, clonal fate data must simply reflect a modified model of cell proliferation. If, however, drug treatment is applied some time T after label induction, then the observed clone size distributions will involve a continuous, if seemingly abrupt (see below), transition between the wildtype and drug-treated systems. When taken together, such experiments can reveal the effect of drug treatments on cell kinetics as well as providing a mutual verification of any proposed proliferation model for the treated system.

Staying within the framework of the critical birth-death process (3.2), of the three adjustable parameters, λ, r and Γ, only two can be considered independent, subject to observations of the relative population of A and B cells. To address the action of a non-disruptive drug treatment on the properties of the clone size distribution, it is convenient to refer below to changes in the symmetric cell division rate,

$$r\lambda \to (r\lambda)' = \alpha r\lambda \,,$$

and the relative population of A and B cells,

$$\rho \to \rho' = \beta\rho \,.$$

For the latter, although ρ is itself slave to the relative change in the rates of cell division and transfer, viz.

$$\frac{\Gamma'}{\Gamma} = \frac{\lambda'}{\lambda}\frac{\beta(1-\rho)}{1-\beta\rho} \,,$$

by addressing experiments within the long-time limit we may focus solely on the timescale $r\lambda/\rho$ (Eq. 3.6), so that the underlying dependence of ρ on the division and transfer rates becomes unimportant. Let us therefore emphasize this point as a motivation for focusing on the behaviour at late times ($t > 1/r\lambda$): here the relative change β, which is readily accessible by immunostaining of the basal layer, reflects the effective impact of the change in the rates of cell division and transfer, which are themselves difficult to access.

To begin let us consider the simpler case in which the relative A and B cell populations are left unchanged by drug treatment (i.e. $\beta = \rho'/\rho = 1$, and $\Gamma'/\Gamma = \lambda'/\lambda$). Focussing on the A-cell

population and referring to the two experimental arrangements above, when drug treatment is initiated *prior* to label induction it is easy to see that the modified clonal distribution is

$$p_n^{(\mathrm{DRUG})}(t) = p_n^{(\mathrm{WT})}(\alpha t)\,,$$

where $p_n^{(\mathrm{WT})}(t)$ is given by Eq. (3.5). By contrast, when labelling is induced at some time T before drug treatment, the size distribution will switch continuously such that

$$p_n^{(\mathrm{DRUG})}(t \geq T) = p_n^{(\mathrm{WT})}(T + \alpha(t-T)).$$

With $\rho' = \rho$, in the long-time limit, one may then recover the full basal layer clone size distribution through the relation (3.6). As a result, one may deduce that the average size of persisting clones changes from its wildtype dependence, $\langle n^{\mathrm{WT}}(t)\rangle = (1 + r\lambda t)/\rho$ (shown in fig. 3.3, inset), to

$$\langle n^{\mathrm{DRUG}}(t > T)\rangle = \langle n^{\mathrm{WT}}(T)\rangle + \alpha r\lambda(t-T)/\rho\,,$$

after drug treatment, i.e. for $\alpha > 1$, there is an enhanced long-term growth in the size of persisting clones.

Let us now consider the more general situation in which we allow the relative proportion of A and B cells to change, i.e. $\beta = \rho'/\rho \neq 1$. Here, such a revision demands some qualification of the underlying model. Referring to Eq. (3.2), one may note that the birth and death of A cells takes place at equal rates, implying that the overall population of labelled A cells must remain unchanged during drug treatment. Keeping in mind that the observed total basal layer cell density remains approximately constant throughout the time course of the experiment, a change in the relative weight of A and B cell populations must therefore invoke processes which lie beyond the model: In particular, if $\beta > 1$, there must exist a transient, *non-steady-state*, period of adaptation when new progenitor cells are recruited into the IFE to replace those differentiated cells shed in the adjustment. Conversely, if $\beta < 1$, there must be a period when progenitor cells are expelled from the IFE.

Whether this non-steady-state behaviour directly influences the labelled clone population or, instead, rebalances only the unlabelled clone population, is unclear (see below) and may be drug-dependent. In general, one may conceive of at least three mechanisms for achieving a shift from the wildtype to the drug-treated steady-state, involving

(a) a (transient) change in the balance of symmetric division rates (in process (3.2)),

(b) the recruitment of type A cells from other parts of the tissue (such as the hair follicle), or

(c) the activation of an otherwise quiescent population of progenitor cells (such as stem cells) to restore the appropriate A cell population.

Fortunately, whatever the transient mechanism for adjusting ρ, at sufficiently long times after initiating drug treatment, one may show that, subject to an overall rescaling of the time coordinate, the form of the asymptotic steady-state distribution (3.6) is invariant leading to the more general relation,

$$P_n^{(\mathrm{DRUG})}(t \geq T) = P_n^{(\mathrm{WT})}(T/\beta + \tau + \frac{\alpha}{\beta}(t-T))\,, \tag{4.1}$$

if drug treatment is initiated at a time T after label induction. That is, a transient period of accelerated growth/attrophication in the A-cell compartment displaces the time coordinate of the probability distributions by some value τ, which is set by the mechanism for changing the A cell population, e.g. $\tau = 0$ for case (b) above, and $\tau \simeq T(1 - \beta^{-1})$ for cases (a) and (c). On the other hand, if drug treatment is initiated sufficiently early prior to label induction to allow a complete transition into the new steady-state, then the distribution maintains the wild-type form, adjusted to reflect the modified steady-state viz.

$$P_n^{(\mathrm{DRUG})}(t) = P_n^{(\mathrm{WT})}(\frac{\alpha}{\beta}t)\,. \tag{4.2}$$

Note that Eq. 4.1, which predicts an instantaneous change in the clone size distribution upon drug application at time T, is rigorous in the limit that the cell populations adjust instantaneously to their new non-equilibrium values. However in the real case, where there is no abrupt change, Eq. 4.1 is only exact for times t beyond the (brief) transient period of population rebalancing.

To illustrate this behaviour, let us consider the outcome when drug treatment is initiated at a sufficiently long time post-label induction so that, upon treatment initiation, the wildtype system may be considered to have reached the scaling limit (3.6), i.e. we are interested in times in excess of the symmetric cell division time, $t > 1/r\lambda$. The action of drug treatment on the basal layer clone size distribution is shown in Fig. 4.1 for a typical parameter set (see caption) alongside that of the corresponding distribution of the unperturbed system. Here we have supposed that A cells are recruited from elsewhere, and do not contribute to the labelled clone distribution — option (b) above. Upon drug application, modelled here by an abrupt change of the cell division and transfer rates, the size distributions show a continuous, but rapid, adjustment before the system enters a smooth long-time evolution. With $\tau = 0$, the period of rapid change can be ascribed to the "rebalancing" of the labelled B cell population, giving a prediction for the abrupt drop in average clone size,

$$\langle n^{\mathrm{DRUG}}(t > T)\rangle \simeq \frac{1}{\beta}\langle n^{\mathrm{WT}}(T)\rangle + \frac{\alpha}{\beta}\frac{r\lambda}{\rho}(t - T)\,, \tag{4.3}$$

which, from the properties of Eq. 4.1, is rigorously true when the transient period is negligibly short.

Referring then to the numerical results in fig. 4.1(inset), one may indeed confirm that, following drug treatment at $T = 13$ weeks, the average basal cell population for persisting clones grows linearly with time, and in keeping with Eq. 4.3, the growth rate increases by a factor of $\alpha/\beta = 20/3$ from its pre-treatment value (of $r\lambda/\rho = 0.48$/week, see figure caption). Also, one may see from the figure that the transient period associated with rebalancing the relative A and B cell populations lasts less than 0.1 weeks, consistent with the abrupt activation of the drug assumed for the calculation, and with the excess labelled B cells transferring from the basal layer with the modified transfer rate of $\Gamma' = 29.16$/week required to maintain the steady-state in the drug-treated system.

To characterise the evolution of the full clone size distributions in fig. 4.1, one may identify from Eq. 4.1 that, following initiation of the drug treatment, the solution conforms to the universal scaling behaviour of the wildtype system, but with the time coordinate *translated* as well as rescaled due to the drug treatment, such that $[t - T(1 - \alpha^{-1})]/n \mapsto t$. This scaling behaviour is seen in fig. 4.2, where, by replotting the data from fig. 4.1 for times $t > T$, one may compare the wild-type and drug-treated systems. In particular, the two systems map onto *distinct manifestations* of the

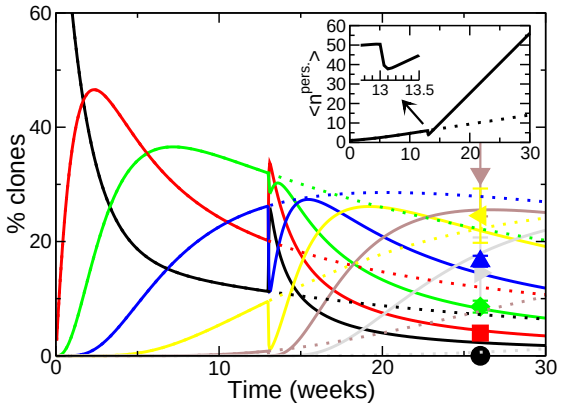


FIG. 4.1: Probability distribution showing the total number of basal layer cells per labelled clone, with drug treatment starting from $T = 13$ weeks post-labelling. The solid curves show the grouped time-dependent probability distribution of *persisting* clones containing two or more cells (see caption and legend of fig. 3.4). The effect of drug treatment was modelled by increasing the overall cell division rate from its initial value of $\lambda = 1.1$/week by a factor of 20 from time T onwards, and by increasing the transfer rate Γ to effect a change in the relative population of A cells from an initial value of 19% to a final value of 57% post-drug treatment (giving $\Gamma = 0.258$/week before drug treatment, and $\Gamma' = 29.16$/week after). The probability of symmetric cell division $2r = 0.16$ was held constant throughout. The dotted curves show the evolution of the same clone size distributions in the absence of drug treatment, and, for comparison, the data points show the results of the preliminary investigation into the effects of ATRA treatment. Error bars show the standard error of the mean. To obtain the figure we numerically integrated the full master equation (Eq. 3.3) starting with exactly one cell per clone at $t = 0$. The initiation of drug treatment was then modelled by an abrupt change to the rates of cell division and transfer at $T = 13$ weeks.
Inset: The average basal-layer population of persistent clones is shown for the same numerical experiment.

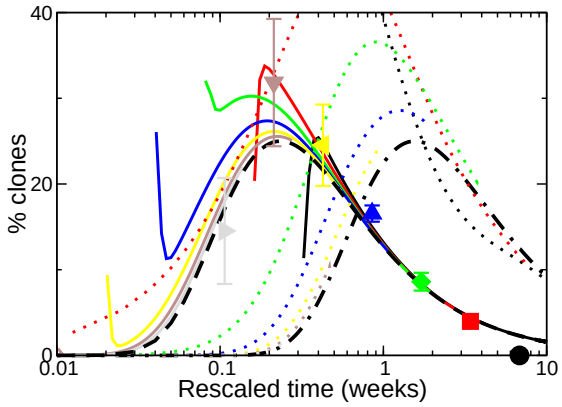


FIG. 4.2: The clone size distribution from fig. 4.1 plotted using the rescaled time coordinate $[t - T(1 - \alpha^{-1})]/n \mapsto t$ for the drug-treated system (solid curves plotted for $t \geq T$), and using $t/n \mapsto t$ for the wild-type system (dotted curves). See caption of fig. 3.4 for further discussion of rescaling. For the drug-treated case, the system rapidly converges onto the long-time limit, giving an excellent fit to the universal curve (dashed), compared to the slow convergence to the universal curve for the wild-type case (dashed-dot). The data points, showing the preliminary results for ATRA treatment (see caption of fig. 4.1), are in good agreement with the universal scaling curve.

universal curve (Eq. 3.6), with a peak in the universal curve located at $t'_{\rm WT} = \rho/(2r\lambda \ln 2)$ for the wild-type case, and at $t'_{\rm DRUG} = (\beta/\alpha)t'_{\rm WT}$ for the drug-treated case. Thus, one may again quantify the effect of the drug from the location of the two maxima, with $t'_{\rm WT}/t'_{\rm DRUG} = 1.6{\rm weeks}/0.24{\rm weeks}$, which gives precisely the expected value of $t'_{\rm WT}/t'_{\rm DRUG} = \alpha/\beta = 20/3$.

4.3 Experiment

To illustrate the application of these ideas, we have undertaken a pilot study of the clone size distribution in mouse tail after treatment by all trans retinoic acid (ATRA). From earlier studies, it is known that ATRA treatment alters substantially the proliferative capacity of progenitor cells, resulting in epidermal thickening, particularly of the suprabasal layers [79]. Yet, observations of mitoses within the affected tissue indicate no qualitative change in the division processes as compared to those observed in the wildtype system. Clonal analysis gives unique insight into the effects of the ATRA treatment, by relating the changes in clone size distributions to quantitative changes in the cell division and differentiation rates. Following the same experimental procedure for tracking clone fate data as that used in normal tail-skin [15], two protocols were used: In the first, mice were treated daily for two weeks with either 0.5ml of 1mM ATRA (Sigma) in ethanol and propanediol 70:30 or vehicle alone applied topically to the tail skin, beginning immediately after clone label induction. In the second, mice were induced, and, after a three month interval, were treated similarly for five days each week with either 0.33mM ATRA or vehicle alone for a further three months. For the former case of 2 weeks' treatment, the control resulted in smaller clones compared to the wildtype, possibly due to transient effects associated with application of the

ethanol and propanediol vehicle. In the latter case, the clone size distribution for the control was consistent with the same division and transfer rates evaluated for the wild-type system in Ref. [15], adjusted for the slight change observed in the concentration of proliferating cells, as assessed by Ki67 immunostaining ($19 \pm 2\%$ for the control compared to 22% in Ref. [15]. This change, which may reflect some small influence of the ATRA vehicle, may also result from statistical variation between different wild-type cohorts of mice. On the other hand, clone size was seen to increase substantially in ATRA treated animals, in keeping with an observed increase in the concentration of proliferating cells (ca. $57 \pm 3.7\%$ in the ATRA treated system as compared to 19% in the control). The corresponding clone size distributions of the ATRA treated system are shown in Figs. 4.3a and 4.3b.

Before attempting to analyse the clone size data quantitatively, we must first address the three-fold increase in the relative population of basal layer progenitor cells on drug treatment, i.e. $\beta \equiv \rho'/\rho = 3$. Referring to the discussion in the previous section, we are faced with at least three possible scenarios (a-c) none of which are ruled out by the available data. However, as we will see below, since measures are restricted well into the long-time limit, the results of the pilot study are largely insensitive to the mechanism of progenitor cell recruitment. Therefore, taking each of the experiments in turn, let us consider first the problem of ATRA treatment at induction. If we assume that the transition to the drug treated steady-state system is immediate (see below) and that, after two weeks, the system has reached its asymptotic behaviour, we are lead to conclude from the fit of Eq. 4.2 to the clone size distributions (fig. 4.3a) that the symmetric cell division rate is changed by a factor of $\alpha \equiv (r\lambda)'/r\lambda \simeq 20$! With $r\lambda = 0.08 \times 1.1/\text{week}$ for the wildtype system, we would therefore expect the system to reach a steady-state behaviour after a time $t = 1/(r\lambda)' \simeq 4\,\text{days}$, a result consistent with our assumptions above. For the control, on the other hand, the system at two weeks is far from the long-time limit ($t < 1/r\lambda$), and the distribution is fitted by numerical integration of wild-type master equation (3.3).

Then, turning to the second of the two experiments in which ATRA treatment occurs 3 months after label induction, Fig. 4.3b shows that the same parameter choice ($\alpha = 20$ and $\beta = 3$) shows an excellent quantitative agreement between the predictions of theory and the experimental data at the six month time point, if we assume that progenitor cells are recruited from outside the labelled clone population (see the following section), viz. $\tau = 0$. Equally, the control is fitted well by the long-time asymptotic distribution for the wild-type system (3.6).

Finally, to consolidate the discussion from the previous section, the fit of the theory to the six month data for the ATRA-treated system is shown in the context of the full time-evolution of the clone size distributions in fig. 4.1, and again in the context of the long-time universal scaling in fig. 4.2.

4.4 Predictive clonal analysis: discussion and conclusions

To conclude, we have shown that clonal analysis provides a powerful tool to explore both qualitatively and quantitatively the action of drug treatment on the interfollicular epidermis. Referring to the new model of epidermal homeostasis introduced by Ref. [15], we have elucidated the clonal characteristics associated with the action of a "benign" drug treatment leading to a global revision

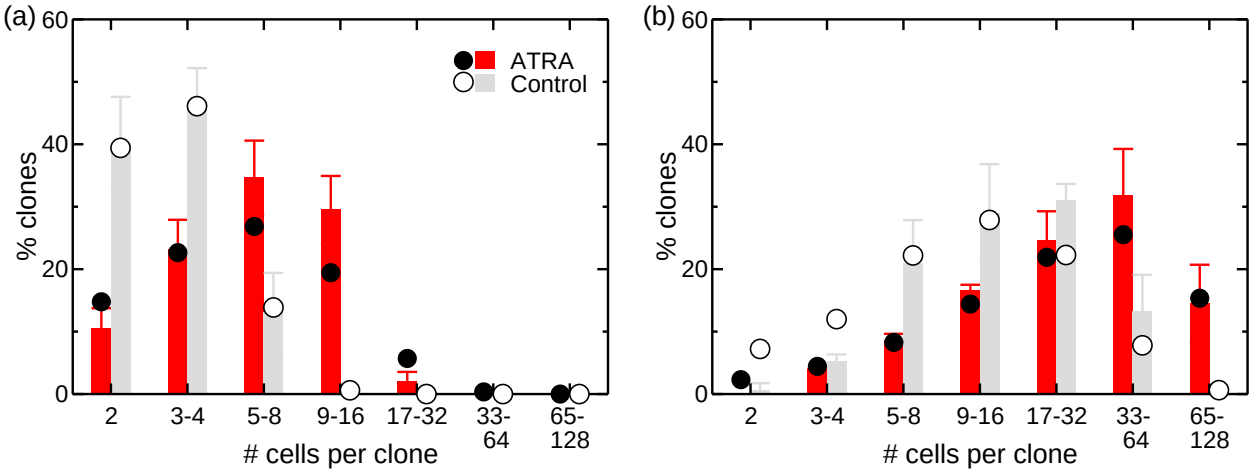


FIG. 4.3: Basal layer clone size distributions for wild-type IFE (grey) or ATRA-treated IFE (red), (a) at two weeks post-labelling, with treatment initiated at label induction, and (b) at six months post-labelling, with treatment initiated at three months post-labelling. Bars show the results of experiment (for details on method, see Ref. [15]), and circles show the theoretical predictions obtained from Eq. 4.2 in (a), and Eq. 4.1 in (b) for the ATRA-treated systems (black), assuming both a 20-fold increase in the rate of symmetric cell division when ATRA is applied ($\alpha = 20$), and that the observed 3-fold increase in the relative number of A cells is accounted for by an introduction of new unlabelled A cells (i.e. $\beta = 3$, $\tau = 0$). For the control, theoretical predictions (white) were obtained by direct numerical integration of Eq. 3.3.

of the cell division rates and branching ratios. These results have been used to assess quantitatively the action of ATRA treatment on cell division rates, revealing that ATRA accelerates the rate of symmetric cell division by a factor of $\alpha = 20$. The findings have been used to successfully predict the long-time clonal evolution in the drug treated system.

Despite the success of the single progenitor cell compartment model in resolving clone fate data of the normal and drug-treated system, the observed changes in the volume fraction of progenitor cells can not be accounted for entirely within the framework of the non-equilibrium process (3.2). Unfortunately, so fast is the acceleration of the symmetric cell division rate due to the ATRA treatment, the origin of progenitor cell recruitment at a transient stage can not be resolved with the existing dataset. (In this context, one may note that a favourable fit of the model to the experimental data can be obtained for a range of growth scenarios, viz. $0 < \tau \simeq T(1-\beta)^{-1}$ addressed in section 4.2.) Moreover, given the period required to fully affect a drug treatment, it is difficult to conceive of a reliable method to probe quantitatively the transient behaviour experimentally.

Finally, turning again to the long-term effects of ATRA treatment, there is an open question regarding the origin of the observed change in the symmetric division rate $r\lambda$. It is unkown whether the change can be completely accounted for by an increase in the overall rate of cell division λ, or whether there is also a change to the symmetric branching ratio r. Although one could in principle determine these effects separately by considering the very early-time clonal evolution following ATRA treatment, in practice the acceleration in symmetric cell division makes access to this time period difficult. Thus, assessment of these quantities in future studies may require independent methods.

Chapter 5

Spatial correlations in murine epidermal maintenance

Chapter overview

In chapters 2 and 3 we showed that in mouse tail-skin, where proliferating cells are confined to a two-dimensional layer, cells proliferate and differentiate according to a simple stochastic model of cell division involving just one type of proliferating cell that may divide both symmetrically and asymmetrically. Curiously, these simple rules provide excellent predictions of the cell population dynamics without having to address their spatial distribution. Yet, if the spatial behaviour of cells is addressed by allowing cells to diffuse at random, one deduces that density fluctuations destroy tissue confluence, implying some hidden degree of spatial regulation in the physical system.

To infer the mechanism of spatial regulation, we consider here a two-dimensional model of cell fate that preserves the form of the empirical clone size distributions. By identifying the resulting behaviour with a three-species variation of a non-equilibrium lattice model known as the *Voter model*, we predict that proliferating cells in the basal layer should *cluster*. Analysis of empirical correlations of cells stained for proliferation activity confirms that the expected clustering behaviour is indeed seen in nature. As well as explaining how cells maintain a uniform two-dimensional density, these findings present an interesting experimental example of voter-model statistics in biology.

The contents of this chapter draw significantly upon Ref. [17], written primarily by AMK, with experimental images provided by David P. Doupé.

5.1 Introduction

To gain insight into the processes of cancer onset, aging and wound healing, biologists have long recognised that the spatial organisation of cells in tissue provides indirect access to the underlying

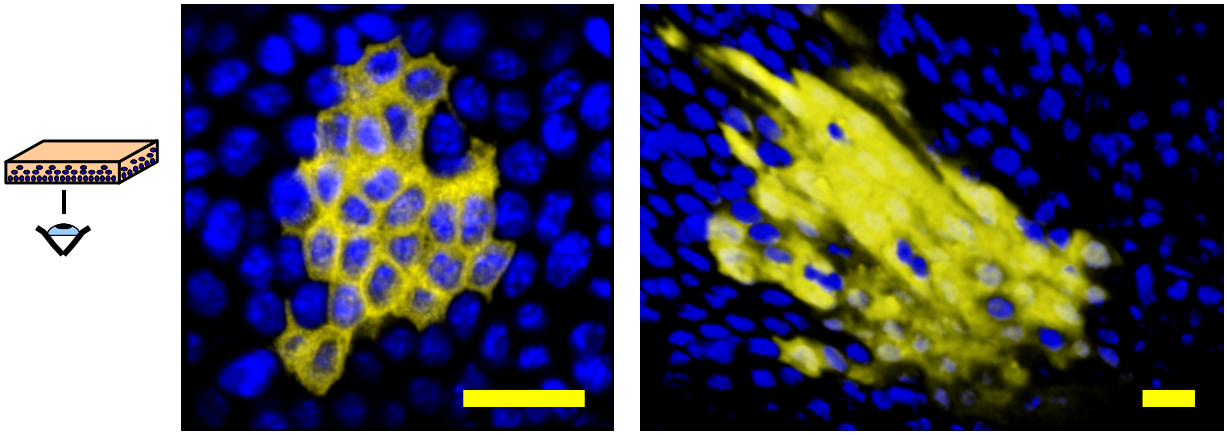


FIG. 5.1: Cross-sections of two typical clones acquired at a late time point, showing clone cohesiveness. The clones are viewed from the basal layer surface as indicated. Cell nuclei are labelled blue; the hereditary clone marker (EYFP) appears yellow. Scale bar: 20μm. See also Figs. 2.2 and 3.1 for further examples.

cell behaviour. Some tissues, such as the auditory hair cells of the inner ear are arranged into repeating units containing groups of specialised cells essential for the function of the tissue [80]. In contrast, in other tissues cells do not organize into coherent structures that reflect their cooperative function. Indeed, the arrangement of some cell types appears random [30]. Inferring the rules of cell behaviour in these apparently unstructured tissues appears challenging. One may ask, therefore, how cell behaviour in such tissues is regulated in the absence of well defined spatial roles.

In this context, let us turn again to the case of mouse tail-skin epidermis, for which the laws governing cell behaviour have now been resolved (chapters 2 and 3). As summarised by the stochastic birth-death process (3.2), the model of epidermal maintenance describes a single population of progenitor cells capable of both symmetric and asymmetric division, which give rise to a population of non-proliferating cells that transfer from the basal layer to the suprabasal layers.

Referring to the question of the spatial organisation of cells in tissue, it is an interesting fact that process (3.2) is capable of fitting the wide range of clone fate data within a "zero-dimensional" framework, i.e. without having to address the spatial orientation of cells within the basal layer. Yet, the observed uniformity of cell density implies a degree of regulation beyond that which can be addressed in the zero dimensional framework. In particular, when augmented by spatial diffusion, the proposed model leads to "cluster" formation in the two-dimensional system whereupon local cell densities are predicted to diverge logarithmically [77, 78]. In biological terms, such behaviour would correspond to a severe disruption of the epidermis, with much of the tissue dying away, leaving only a few isolated and very thick clusters of epidermis. Significantly, the observation that labelled families of cells remain largely *cohesive* (see for example Fig. 5.1), reveals that cell mobility must be small, so that such divergences would be significant within a mammalian lifetime. These divergences can not be regulated through a local density-dependent mobility.

Thus, the success of the zero-dimensional fit, despite the predicted divergence in two dimensions, leaves us with an interesting challenge that is the focus of the present chapter: Can we uncover, from the spatial distribution of basal layer cells, the mechanism by which cells regulate a uniform

cell density without compromising the integrity of the zero-dimensional fit, as embodied in process (3.2)?

In order to identify the underlying rules of cell division and differentiation, we shall draw upon the results of two types of experiment. First, we shall revisit the clone fate data used in the original zero-dimensional analysis in chapters 2 and 3, in order to examine the previously-discarded spatial distribution of labelled basal layer cells for signatures of underlying regulation. Second, we shall consider the statistics of the entire population of basal layer cells. In particular, by immunostaining basal layer cells for markers of cell proliferation, it is possible to analyse the spatial distribution of all progenitor cells.

Thus, the aim of this chapter is to elucidate how the experimental observations constrain any proposed theory of spatial behaviour in the basal layer. In summary, we shall show that the dynamics predicted by process (3.2) are indeed consistent with the constraint on uniform cell density, provided that cell division occurs only upon the migration of a nearby type B (i.e. differentiated) cell into the basal layer. Moreover, we confirm that the clone fate data is consistent with a restricted degree of cell mobility, whereby cell motion is *not* diffusive and random. Instead, differentiated cells only migrate laterally as a response to fluctuations in the local density. Finally, to test the validity of the proposed spatial model, we use it to predict that, while maintaining a uniform *total* areal cell density, the population of progenitor cells should *cluster* over time. By considering the radial correlation function for the spatial distribution of progenitor cells, we find that this prediction is in good qualitative agreement with experiment. Quantitatively, the comparison reveals that the experimental degree of progenitor cell clustering is slightly higher than that expected for the parameter value of $r = 0.08$ determined previously through clonal analysis. Although several technical difficulties may challenge the reliability of these quantitative results, we speculate that such excess clustering may be a signature of spatial regulation of cell fate during asymmetric division. These results also shed light on previous observations of clustering of cells undergoing mitosis in the epidermis [81, 82]. Such observations have been interpreted in the biological community to be a signature of regulation that leads to coordinated cell division. By contrast, this work shows that the tendency of proliferating cells (and therefore mitoses) to cluster is in fact consistent with cells dividing independently and stochastically — indeed it is the hallmark of the proposed spatial process.

This chapter is organised as follows. In section 5.2.1 we develop a phenomenological model of cell behaviour that incorporates the experimental constraint on uniform cell density. We identify the proposed model as a variation upon the so-called *monomer-monomer* model of surface catalysis. We then analyse this simpler model in section 5.2.2, including an exact solution for the two-point density correlations in the closely-related monomer-monomer model. In section 5.3 we test the uniform-density model against a range of experimental data, first through a qualitative comparison of the model with the empirical clone shape data (section 5.3.1), and then by analysing the spatial correlations observed for proliferating cells (section 5.3.2). We conclude with a discussion of cell clustering in section 5.4.

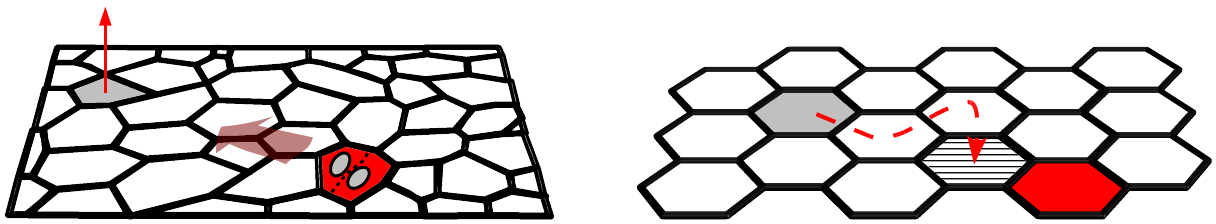


FIG. 5.2: Schematic motivating the proposed lattice model defined in Eq. (5.1). *Left:* A cartoon of cells within the basal layer, showing the exit of a cell (light grey) through upward migration into the suprabasal layers, concurrent with the division of a nearby progenitor cell (red online). To ensure continuity of the basal layer, it is postulated that cells rearrange to maintain a uniform cell density (wide arrow). *Right:* In a simple lattice model that captures the essence of the steady-state dynamics, the exit of a cell from the basal layer gives rise to a vacant lattice site (light grey), which rapidly diffuses by exchanging position with adjacent cells (white). Upon coming into contact with a proliferating cell, the latter may divide and replace the vacancy with a daughter cell (striped).

5.2 A spatial model of cell kinetics

To understand how the experimental observations constrain any proposed theory of spatial behaviour in the basal layer, we shall first address the constraint imposed by the observed uniform cell density. To this end, we postulate that proliferating cells only divide upon the migration of a nearby differentiated cell into the supra-basal layers (see Fig. 5.2). This requirement is purely phenomenological, as it ensures a uniform density without specifying the mechanism by which it is implemented. Indeed, a range of regulatory pathways can be seen to give rise to the same phenomenology, for example by coupling cell division processes to the local stress [83, 38] or using short-ranged morphogen gradients as well as feedback by cell-cell communication. Once this initial constraint is accounted for in this section, we shall draw upon the further observations of clone and basal layer morphology in order to identify additional rules governing cell behaviour.

In general, it is a challenge to couple the division and migration of basal layer cells while still allowing for some degree of cell compressibility, whereby a dividing cell may compensate for the exit of a non-adjacent cell through lateral motion (see Fig. 5.2, top). Several approaches have been used in the past to overcome this problem. In the context of tissue development, one may treat the cell tissue as an elastic medium, whereby the local cell density is coupled to the cell division process through the stress in the surrounding tissue [83]. Such an approach is capable of accounting for a range of realistic properties of two-dimensional cell tissue growth, such as the distribution of cell sizes, as well as cell compressibility. Yet, for the simple problem of steady-state tissue maintenance, involving no net growth, it is unlikely that the complexity of the elastic tissue model is required to explain much of the experimental data. A second approach treats the two-dimensional basal layer as a 'foam' of cells (a Voronoi tesselation), for which the steady-state condition may be used to relate between the local cell topology (or more precisely, the number of nearest-neighbours in the basal layer) and the likelihood of division or migration. Applied to the problem of epidermal maintenance, this approach successfully predicts the steady-state topology of epidermal basal-layer cells [36], and it identifies that cells with a larger number of neighbours are more likely undergo division, whereas cells with a small number of neighbours are more likely to migrate in to the

supra-basal layers. However, it is a challenge to extend this model to allow for two distinct cell populations that are exclusively committed to either division or migration, such as described by process (3.2). Yet a third approach draws upon simulations in which cells are modeled as quasi-spherical particles that deform during cell division [43, 44]. However, as this approach draws upon a wide range of (uncontrolled) parameters to describe the cell-cell interactions, it is more complex than required for this case.

Therefore, in the following we shall suffice with a simpler description of the basal layer, by drawing upon the non-equilibrium lattice models discussed in the recent literature. In particular, we shall model the basal layer as a lattice in which each site is occupied by one cell. Cell compressibility is then modeled by a population of lattice vacancies, which are created upon cell migration ($B \to \oslash$) and then diffuse rapidly as compared to the cell kinetic rates (λ, Γ), before annihilating upon the division of an adjacent cell (see Fig. 5.2). Since this rule-based model is not primarily based on a direct physical representation of individual cells, it may overlook certain physical effects. For example, the migration of type B cells out of the basal layer may be facilitated by mechanical forces exerted by neighbouring cells — a situation that is hard to represent with a cellular automaton. However, recalling that the stochastic rules embodied by process (5.1) have been experimentally verified, it is reasonable to start by considering a similar stochastic process in two dimensions. Later, we shall further justify the use of the lattice model by showing that the lattice geometry does not effect the qualitative behaviour of the system (see section 5.2.2(e)).

Although a significant advantage of the lattice description of the basal layer is the ease by which it may be simulated, a range of analytic results are also made accessible by showing that in certain limits, the model reduces to a simple two-component model that belongs to the generalised *voter model* universality class [84, 85, 86]. The voter model universality class describes lattice processes that undergo phase separation in two dimensions in the absence of surface tension. In the context of basal layer kinetics, this "phase separation" corresponds to the clustering of proliferating cells. Curiously, in the special case $r = 1/4$, the basal layer lattice model reduces to the reaction-limited *monomer-monomer* model of surface catalytic reactions, in which the classical voter model dynamics are augmented by infinite-temperature Kawasaki exchange dynamics [87]. As well as drawing upon a range of existing results afforded by these models, we shall derive here an exact solution for the two-point spatial correlation function of the monomer-monomer model. These results reveal the continuous transition between voter-like behaviour and diffusive behaviour as the relative rates of symmetric vs. asymmetric cell division (or reactant deposition in surface catalysis) are adjusted.

5.2.1 The lattice model

As mentioned above, we are interested in constructing a spatial model that recovers the behaviour of process (3.2), and maintains a uniform cell density. To account for the steric repulsion of basal layer cells, we will characterise the basal layer as a two-dimensional lattice, where each site is host to one of the two cell types, or it remains vacant. When the vacancy fraction is very low, then such a lattice description presents a reasonable approximation of the observed near-uniform arrangement of basal layer cells during normal adult skin maintenance. Then, to regulate the cell density, progenitor cells (A) are allowed to divide only when neighbouring a site vacancy, while the migration of post-mitotic (B) cells from the basal layer leads to the creation of vacancies which are

free to diffuse in the basal layer through the displacement of neighbouring cells (see Fig. 5.2). In summary, denoting a site vacancy with the symbol $\oslash$, the lattice model may be written in terms of the non-equilibrium process:

$$\begin{aligned} &\mathrm{A}\ \oslash \quad \xrightarrow{\lambda'} \quad \begin{cases} \mathrm{A\ A} & \text{Prob. } r \\ \mathrm{A\ B} & \text{Prob. } \frac{1}{2} - r \\ \mathrm{B\ A} & \text{Prob. } \frac{1}{2} - r \\ \mathrm{B\ B} & \text{Prob. } r \end{cases} \\ &B \quad \xrightarrow{\Gamma} \quad \oslash \\ &\mathrm{X}\ \oslash \quad \xrightarrow{\sigma} \quad \oslash\ \mathrm{X}\,, \end{aligned} \tag{5.1}$$

where the site hopping rate σ reflects the capacity of vacancies to diffuse within the basal layer, and X is used to denote either a type A or a type B cell. To gain some initial insight into the dynamics, and to identify constraints on the parameter space, one may consider the steady-state mean-field cell densities associated with this process. It is straightforward to show that the mean-field equation for the vacancy fraction $n_\oslash$ is given by

$$\partial_t n_\oslash = \sigma S \nabla^2 n_\oslash + \Gamma(1 - \rho - n_\oslash) - \lambda' \rho n_\oslash\,, \tag{5.2}$$

where ρ is the (constant) A-cell fraction, and S is the area per lattice site, corresponding to the average areal cross-section of a basal layer cell. In the following we shall work in units of the average cell cross-section ($S = 1$). From here one may see that the near-uniform cell density in the basal layer ($n_\oslash \ll 1$ and uniformly distributed) constrains us to the region of parameter space $\lambda' \gg \Gamma$, such that any vacancy created through the migration of a type B cell out of the basal layer is rapidly removed upon coming into contact with a proliferating cell. In this limit, the numerical value of the parameter λ' becomes irrelevant, as may be seen by calculating the effective local division rate $\lambda \equiv \lambda' n_\oslash$, which takes the value $\lambda = \Gamma(1-\rho)/\rho$, independent of λ'. This relationship between the rate of cell division and migration is identical to that obtained in process (3.2) [16].

Although the uniform vacancy density is a stable fixed point of the mean-field dynamics (Eq. 5.2), one may worry whether fluctuations about the mean-field solution are capable of compromising the uniform density of the basal layer cell lattice in actual manifestations of process (5.1). To eliminate this concern, one requires rapid dissipation of density fluctuations independently of the cell kinetics, from which we infer a that the lattice vacancy population must diffuse rapidly compared to the time-scale of cell division and upward migration, viz. $\sigma \gg \Gamma$, as stated earlier. Biologically, this condition corresponds to the assumption that cells are largely incompressible, so that local density fluctuations lead to the rearrangement of cells on a time-scale that is significantly faster than that of cell division (cf. Ref. [83]). In section 5.3.1 we shall show that the empirical clone fate data further constrain the parameter space to the region $\lambda' \gg \sigma$, whereby the exit of a type B cell is compensated for by a nearby cell division.

With these definitions, in the parameter space $\lambda' \gg \sigma \gg \Gamma$ one may see that the spatial model introduces *no new relevant parameters* compared to process (3.2). That is, the model behaviour depends only on the (known) zero-dimensional parameters (r, λ, Γ), with the contribution of the new parameters (λ', σ) entering through the dimensionless ratios Γ/λ', Γ/σ and σ/λ', all of which may be made arbitrarily small. But, as a precondition on its validity, does the model also reproduce the observed 'zero-dimensional' basal layer clone size distributions? This is by no means obvious, given the critical (and therefore delicate) nature of process (3.2) [16]. For example, a high density of progenitor cells in the lattice model may lead to vacancy depletion and jamming, an effect that

has no analog in the zero-dimensional system. Therefore, in section 5.3 we shall show, using Monte-Carlo simulations, that the proposed lattice model indeed succeeds in reproducing the empirical clonal statistics from Ref. [15]. With this basic confidence in the validity of the model, we now proceed to analyse its behaviour in more detail.

In general, the cell kinetics in process (5.1) describe a hard-core non-equilibrium system involving three cell species. Recent progress in the field of non-equilibrium statistical mechanics has resulted in several possible formalisms with which to study such systems [47, 88, 89]. However, for the case at hand, these approaches are unnecessarily complex. Rather than analysing the current microscopic model, it is convenient to recast the cell kinetics into a simpler form that describes the same phenomenology. In particular, for the parameter space of interest ($\lambda' \gg \sigma \gg \Gamma$), it is sufficient to consider a lattice fully occupied by A and B cells, without addressing the population of lattice vacancies. To see this, one may see from Eq. (5.2) that the vacancy dynamics occur on a fast time scale compared to that of cell migration ($1/\Gamma$), and division ($1/\lambda$). Therefore, referring back to the lattice process (5.1), we may (heuristically) eliminate the vacancy population by replacing the cell division process (A $\oslash \rightarrow$ X Y) with a direct cell-cell 'reaction' process (A B $\rightarrow$ X Y), and by replacing the A-cell division rate λ' with the effective rate $\lambda(\mathbf{x}) = \lambda' n_{\oslash}(\mathbf{x}) = \Gamma[1 - \rho(\mathbf{x})]/\rho(\mathbf{x})$. Here, the local A-cell fraction $\rho(\mathbf{x})$ refers to the A-cell number density coarse-grained over the nearby lattice neighbourhood. For example, denoting the number of type A cells on a lattice site as $n_{\mathbf{x}}$, we have $\rho(\mathbf{x}) = \sum_{\mathbf{x}'} n_{\mathbf{x}'} w(|\mathbf{x} - \mathbf{x}'|)$ where $w(x)$ is some suitably chosen normalised envelope function. Within this framework, one may then replace process (5.1) with the more simplified form

$$\text{A B} \quad \xrightarrow{\lambda(\mathbf{x})} \quad \begin{cases} \text{A A} & \text{Prob. } r \\ \text{A B} & \text{Prob. } \frac{1}{2} - r \\ \text{B A} & \text{Prob. } \frac{1}{2} - r \\ \text{B B} & \text{Prob. } r \end{cases} . \tag{5.3}$$

The degree to which this heuristic simplification is indeed justified will be discussed at the end of section 5.2.2, together with a quantitative comparison of the behaviour of the exact and simplified models. It is already clear that process (5.3) cannot describe the explicit upward migration of post-mitotic cells from the basal layer. However, in the physically relevant limit $\lambda' \gg \sigma$, eliminating the vacancy population has no qualitative effect on the statistics of the progenitor cell compartment, and therefore processes (5.1) and (5.3) are expected to result in the same basal layer phenomenology.

Interestingly, process (5.3) is closely related to the model of monomer-monomer surface catalysis [90, 91]. In particular, when the coarse-grained distribution of type A cells is effectively uniform, such that $\lambda(\mathbf{x}) \simeq$ const., then one may identify the symmetric branches of process (5.3) with the classical zero-temperature *voter model*, while the asymmetric division channel A B $\rightarrow$ B A, describes the Kawaski dynamics of an infinite-temperature Ising spin model. Later it will become clear that for the empirical value of $r = 0.08$, the significant contribution of the latter will justify the approximation of near-constant λ.

This analogy provides access to several known results. In Ref. [87], Krapivsky showed that starting from random initial conditions, the classical voter model (i.e. with $r = 1/2$) will lead to a lattice of N sites becoming completely saturated with either type A or type B cells after a time $\lambda t \sim N \ln N$. Moreover, Frachebourg and Krapivski gave an exact solution for the two-point spatial correlations in this case [90], from which they inferred that, in the time leading up to saturation ($\lambda t \ll N \ln N$), the different cell types separate into domains of ever-increasing size, with a typical lengthscale

growing as $L \sim \ln \lambda t$, and with the density of interfaces $c_{\rm AB}$ between type A and type B cells dropping as $c_{\rm AB} \sim 1/\ln \lambda t$. As the system approaches saturation ($t \sim N \ln N$), one of the cell types comes to dominate. The classical voter model is an example of domain growth in the absence of surface tension [91]. Therefore, the boundaries between domains rich in A and B cells are completely unstable, leading to strikingly different and irregular domain morphologies, compared to the smooth phase-separated shapes resulting from surface-tension mediated domain growth.

Qualitatively, the results found for the classical voter model ($r = 1/2$) allow us to make several interesting predictions relating to the spatial distribution of A and B cells. In particular, some degree of clustering of proliferating cells is to be expected in adult mice, resulting from the growth of domains rich in progenitor cells. Moreover, the ongoing growth of the domain size L suggests that larger clusters are expected in old vs. young epidermis. Yet, to make full contact between process (5.1) and the empirical data, it becomes necessary to calculate the model properties whilst allowing for the relatively low value of $r = 0.08$ found in the experimental system. Therefore, in the following, we shall extend the analysis of Frachebourg and Krapivsky to obtain results valid for arbitrary r. Indeed, with $r = 1/4$, the following analysis results in the exact solution to the reaction-limited monomer-monomer surface catalysis model.

5.2.2 Exact solution for two-point correlations

In the following, we will follow the same approach as taken in Refs. [87, 90] for the $r = 1/2$ case, but we generalise to allow for arbitrary r and different lattice geometries. For completeness, we include here aspects of the solution that were also described in some detail in Ref. [87], such as the Master equation and the dynamical equations required to define the problem. We start by identifying type A cells with state 1 and type B cells with state 0, so that a lattice with site index $\mathbf{i}$ may be described in terms of the Ising variables $\Phi = \{n_{\mathbf{i}}\}$, $n_{\mathbf{i}} \in \{0, 1\}$. Referring to Ref. [87], the Master equation for the probability distribution $P(\Phi, t)$ for the system to occupy state Φ at time t is given by,

$$\begin{aligned}\frac{d}{dt}P(\Phi,t) &= \frac{\lambda}{2}\sum_{\mathbf{i},\,\mathbf{e}}\Bigg\{ r\left[U_{\mathbf{i}}^{(\mathbf{e})}(\hat{F}_{\mathbf{i}}\Phi)P(\hat{F}_{\mathbf{i}}\Phi,t) + U_{\mathbf{i}}^{(\mathbf{e})}(\hat{F}_{\mathbf{i}+\mathbf{e}}\Phi)P(\hat{F}_{\mathbf{i}+\mathbf{e}}\Phi,t)\right] \\ &\quad +(\frac{1}{2}-r)U_{\mathbf{i}}^{(\mathbf{e})}(\hat{F}_{\mathbf{i}}\hat{F}_{\mathbf{i}+\mathbf{e}}\Phi)P(\hat{F}_{\mathbf{i}}\hat{F}_{\mathbf{i}+\mathbf{e}}\Phi,t) - (\frac{1}{2}+r)U_{\mathbf{i}}^{(\mathbf{e})}(\Phi)P(\Phi,t)\Bigg\}. \qquad (5.4)\end{aligned}$$

Here, $\{\mathbf{e}\}$ represent the nearest-neighbour lattice vectors ($|\mathbf{e}| = 1$), and $U_{\mathbf{i}}^{(\mathbf{e})}(\Phi) \in \{0, 1\}$ indicates whether the cells at sites $\mathbf{i}$ and $\mathbf{i} + \mathbf{e}$ are a 'reactive' pair, viz.

$$U_{\mathbf{i}}^{(\mathbf{e})}(\Phi) = n_{\mathbf{i}} + n_{\mathbf{i}+\mathbf{e}} - 2n_{\mathbf{i}}n_{\mathbf{i}+\mathbf{e}}\,.$$

The "spin-flip" operator is defined by $\hat{F}_{\mathbf{i}}\Phi = \{n_{\mathbf{j}} \text{ for all } \mathbf{j} \neq \mathbf{i};\ 1 - n_{\mathbf{i}}\}$, so that $\hat{F}_{\mathbf{i}}\Phi$ and $\hat{F}_{\mathbf{i}+\mathbf{e}}\Phi$ correspond to the symmetric division channels and $\hat{F}_{\mathbf{i}}\hat{F}_{\mathbf{i}+\mathbf{e}}\Phi$ corresponds to an asymmetric division in which the location of the type A and B cells is reversed (viz. A B $\rightarrow$ B A).

From here, recalling that $\langle n_{\mathbf{i}}n_{\mathbf{j}}\rangle = \sum_{\Phi} n_{\mathbf{i}}n_{\mathbf{j}}P(\Phi, t)$, it is simple to show that the two-site correlation function evolves according to the discretized diffusion equation for non-neighbouring sites,

$$\frac{d}{dt}\langle n_{\mathbf{i}}n_{\mathbf{j}}\rangle = \frac{\lambda}{2}(\Delta_{\mathbf{i}} + \Delta_{\mathbf{j}})\langle n_{\mathbf{i}}n_{\mathbf{j}}\rangle, \qquad (5.5)$$

where $\Delta_{\mathbf{i}}$ is the discrete Laplacian operator, defined by $\Delta_{\mathbf{i}} n_{\mathbf{i}} = \sum_{\mathbf{e}} (n_{\mathbf{i}+\mathbf{e}} - n_{\mathbf{i}})$. However, the diffusion equation is modified for nearest-neighbour correlations, giving

$$\begin{aligned} \frac{d}{dt}\langle n_{\mathbf{i}} n_{\mathbf{i}+\mathbf{e}}\rangle &= \frac{\lambda}{2}(\Delta_{\mathbf{i}} + \Delta_{\mathbf{i}+\mathbf{e}})\langle n_{\mathbf{i}} n_{\mathbf{i}+\mathbf{e}}\rangle \\ &+ (\frac{1}{2} - r)\lambda[2\langle n_{\mathbf{i}} n_{\mathbf{i}+\mathbf{e}}\rangle - \langle n_{\mathbf{i}+\mathbf{e}}\rangle \\ &- \langle n_{\mathbf{i}}\rangle], \end{aligned} \tag{5.6}$$

and the on-site moment is trivially $\langle n_{\mathbf{i}} n_{\mathbf{i}}\rangle = \langle n_{\mathbf{i}}\rangle$.

Making the simplifying assumption that the initial distribution $P(\Phi, 0)$ is translationally invariant, then the fraction of type A cells is given by $\langle n_{\mathbf{i}}\rangle = \rho = \text{const.}$, and $\langle n_{\mathbf{i}} n_{\mathbf{j}}\rangle$ depends on $(\mathbf{i} - \mathbf{j})$ at all times. Therefore, introducing the correlation function $C_{\mathbf{i}} = \langle n_{\mathbf{j}} n_{\mathbf{j}+\mathbf{i}}\rangle$, we can rewrite Eqs. (5.5) and (5.6) as follows:

$$\frac{d}{dt} C_{\mathbf{i}} = \lambda \Delta_{\mathbf{i}} C_{\mathbf{i}} - \sum_{\mathbf{e}} \delta_{\mathbf{i},\mathbf{e}} (1 - 2r)\lambda(\rho - C_{\mathbf{e}}) \,, \tag{5.7}$$

for $|\mathbf{i}| \geq 1$, subject to the constraint $C_{\mathbf{0}} = \rho = \text{const}$. That is, the correlation function evolves according to a discrete diffusion equation with sink terms at the nearest-neighbour sites and with a fixed boundary condition at the origin. The linear nature of the problem allows one to seek a solution in terms of the relevant Green's function, e.g. $\hat{\Delta}_{\mathbf{i}}^{-1} \equiv G_{(i_x,i_y)}(t) = e^{-4\lambda t} I_{i_x}(2\lambda t) I_{i_y}(2\lambda t)$ for a square lattice $\mathbf{i} = (i_x, i_y)$. For uncorrelated initial conditions, viz. $C_{\mathbf{i}}(t = 0) = \rho^2 + \delta_{\mathbf{i},\mathbf{0}}(\rho - \rho^2)$, one may write down a general solution in the form

$$\begin{aligned} C_{\mathbf{i}}(t) &= \rho^2 + (\rho - \rho^2) G_{\mathbf{i}}(t) \\ &+ \sum_{\mathbf{j}} \int_0^t d\tau J_{\mathbf{j}}(\tau) G_{\mathbf{i}-\mathbf{j}}(t - \tau) \,, \end{aligned} \tag{5.8}$$

where $J_{\mathbf{i}}(t)$ is the source distribution required to both maintain $C_{\mathbf{0}} = \text{const}$, and also to incorporate the sink terms from Eq. (5.7). Thus, $J_{\mathbf{i}}(t) = 0$ for $|\mathbf{i}| > 1$, $J_{\mathbf{e}} = -(1 - 2r)\lambda(\rho - C_{\mathbf{e}})$, and $J_{\mathbf{0}} = z\lambda(\rho - C_{\mathbf{e}})$, where z is the number of nearest neighbours, and the initial conditions imply that $C_{\mathbf{e}}$ is the same for all nearest neighbours. For now we will explicity consider the square lattice ($z = 4$), however the effect of changing lattice geometry will be discussed near the end of this section. Making use of Eq. (5.8), one may write down a set of self-consistent equations for the source terms, viz.

$$\begin{aligned} J_{\mathbf{0}}(t) &= 4\lambda \Big\{ \rho - \rho^2 - \int_0^t [J_{\mathbf{0}}(t - \tau) G_{(0,1)}(\tau) \\ &+ J_{\mathbf{e}}(t - \tau) \left(G_{(0,0)}(\tau) + G_{(0,2)}(\tau) + 2G_{(1,1)}(\tau) \right)] \Big\} \end{aligned} \tag{5.9}$$

$$J_{\mathbf{e}}(t) = -\frac{1 - 2r}{4} J_{\mathbf{0}}(t) \,, \tag{5.10}$$

which, upon taking the Laplace transform $j(p) = \mathcal{L}[J(t)]$, $g(p) = \mathcal{L}[G(t)]$, gives the expression for the source term,

$$j_{\mathbf{0}}(p) = \frac{4\lambda(\rho - \rho^2)}{p(1 + 4\lambda g_{(0,1)} - (1 - 2r)\lambda h)}, \tag{5.11}$$

where we have defined $h(p) = g_{(0,0)}(p) + g_{(0,2)}(p) + 2g_{(1,1)}(p)$ for the square lattice.

Note that in the classical voter model ($r = 1/2$), the sink terms vanish ($j_{\mathbf{e}} = 0$), and a single source is located at the origin. The calculation may then proceed exactly as described in Ref. [90]. On the other hand, for the infinite-temperature Kawasaki dynamics ($r = 0$), the source and sink terms create no net correlation ($j_{\mathbf{e}} = -j_{\mathbf{0}}/4$), and they serve only to maintain the stationary state $C_{\mathbf{0}} = \rho$, $C_{\mathbf{i}\neq\mathbf{0}} = \rho^2$.

The discussion so far has been exact. We now turn to the long-time asymptotic solution of the correlation function, for which it is sufficient to consider the behaviour of $j_{\mathbf{0}}, j_{\mathbf{e}}$ at small p ($p \ll \lambda$). In the following we shall make use of the following expansions:

$$\lim_{p/\lambda\to 0} g_{(0,0)}(p) = \frac{1}{4\pi\lambda}\ln(32\lambda/p) + \mathcal{O}[p\ln(p)]\,,$$

and (for the same limit $p/\lambda \to 0$)

$$\begin{aligned} g_{(0,0)}(p) - g_{(0,1)}(p) &= \frac{1}{4\lambda} + \mathcal{O}[p\ln(p)] \\ g_{(0,0)}(p) - g_{(1,1)}(p) &= \frac{1}{\pi\lambda} + \mathcal{O}[p\ln(p)] \\ g_{(0,0)}(p) - g_{(0,2)}(p) &= \frac{1-2/\pi}{\lambda} + \mathcal{O}[p\ln(p)]\,. \end{aligned}$$

With these expansions, the source terms take the long-time asymptotic values

$$\lim_{p/\lambda\to 0} \quad j_{\mathbf{0}}(p) = \quad \frac{4\pi\lambda(\rho-\rho^2)}{p\,[\pi(1-2r) + 2r\ln(32\lambda/p)]}\,. \tag{5.12}$$

The corresponding long-time behaviour is therefore

$$\lim_{t\gg 1/\lambda} J_{\mathbf{0}}(t) = \frac{4\pi\lambda(\rho-\rho^2)}{\pi\,(1-2r) + 2r\ln(32\lambda t)}\,, \tag{5.13}$$

giving the expected $1/\ln t$ decay. Now the effect of Kawasaki dynamics becomes clear: one may see that reducing r from its maximal value of $1/2$ has the opposing effect of weakening the magnitude of the net source ($J_{\mathbf{0}} - 4J_{\mathbf{e}} \propto 2r$) while extending the time over which the source decays, $t_r \sim (t_{r=1/2})^{1/2r}$. In the trivial limit $r = 0$, the system becomes stationary as expected, whereas when $r = 1/2$ one retrieves the same expression as in Ref. [90].

Taken together, Eqs. (5.8), (5.10) and (5.13) give the solution for the long-time asymptotic behaviour of the two-point correlation function. With a mind towards experiment, as well as to compare with the known results for $r = 1/2$, we now summarise the features of this solution:

(a) The density of interfaces c_{AB} between A and B cells drops asymptotically as $c_{\mathrm{AB}} \sim 1/2r\ln t$. The density of interfaces between adjacent A and B cells, $c_{\mathrm{AB}}(t) \equiv 2(\rho - C_{\mathbf{e}}(t))$, is an order parameter used to describe the transition from an uncorrelated initial state with $c_{\mathrm{AB}} = 2(\rho-\rho^2)$, to the jammed absorbing state $c_{\mathrm{AB}}(t\to\infty) = 0$. For $r = 1/2$, it was previously shown that $c_{\mathrm{AB}} \sim 1/\ln t$.

From its definition, we identify the order parameter to be proportional to the source term, viz. $c_{\mathrm{AB}}(t) = J_{\mathbf{0}}/2\lambda$, so that the long-time asymptotic behaviour is found from Eq. (5.13). This expression reveals the continuous transition between Kawasaki dynamics ($r = 0$, $c_{\mathrm{AB}} = \text{const}$) and voter

dynamics ($r = 1/2$, $c_{\rm AB} \sim 1/\ln t$). In particular, as $t \to \infty$, the absorbing-state phase transition occurs at $r = 0$.

(b) The spatial correlation function $C_{\mathbf{i}}(t)$ decays as $a(t) - b(t)\ln|\mathbf{i}|$ at short distances, and as a Gaussian at long distances.
Away from the origin, where a continuum description suffices, then the correlation function depends only on the distance $x = |\mathbf{i}|$, and all sources appear to be located at the origin, viz. $J_{\mathbf{i}}^{(\text{eff.})} = \delta_{\mathbf{i},\mathbf{0}}(J_{\mathbf{0}}+4J_{\mathbf{e}})$. Replacing the discrete problem with a continuous one simplifies the Green's function, with $\lim_{x\gg 1} G_{\mathbf{i}}(t) \to G(x,t) = e^{-x^2/4\lambda t}/(4\pi\lambda t)$ for a square lattice. As a consequence, the long-time asymptotic correlation function may be approximated as $\lim_{t\gg 1/\lambda, x\gg 1} C_{\mathbf{i}}(t) = C(x,t)$, with

$$C(x,t) = \rho^2 - \frac{2r(\rho-\rho^2)}{\pi\,(1-2r) + 2r\ln(32\lambda t)}\,\mathrm{Ei}\left(-\frac{x^2}{4\lambda t}\right), \tag{5.14}$$

where $\mathrm{Ei}(x) = -\int_{-x}^{\infty} dt\, e^{-t}/t$ is the exponential integral. From the small-argument expansion of this integral, we find

$$\lim_{t\gg 1/\lambda,\ x^2\ll\lambda t} \frac{C(x,t)-\rho^2}{\rho-\rho^2} = a(t) - b(t)\ln x\,, \tag{5.15}$$

with

$$a(t) = 1 - \frac{1-2r(\pi-\gamma_e-\ln 8)}{2r\ln(32\lambda t) + \pi\,(1-2r)}\,,$$

$b(t) = 4r/[2r\ln(32\lambda t) + \pi\,(1-2r)]$, and γ_e is the Euler constant.

At long times, $C(1,t) \geq C_{\mathbf{e}}(t)$, so one may infer that the correlation function is concave near the origin. Away from the origin, it is useful to define a correlation length ξ to characterise the short-range correlations, viz.

$$\xi^{-1} \equiv (\rho-\rho^2)\left.\frac{\partial C}{\partial x}\right|_{x=1} = b\,, \tag{5.16}$$

corresponding to the typical size of A-cell rich domains growing as $\xi \sim \ln t$. One may see that when $r > 0$, a variation in r merely adjusts the correlation length by a constant, in contrast to its effect on the order parameter $c_{\rm AB}$. For $r = 0$, the lattice configuration is random and $C(x,t) = \rho^2$, corresponding to an "infinite" correlation length.

(c) The time to saturation T of a lattice of N sites is approximately $T \sim N(\ln N + 1/(2r))/\lambda$.
As mentioned earlier, the classical voter model predicts that any finite-sized system inevitably approaches an absorbing state, in which the lattice is either completely saturated by type A cells or else they have become extinct. Fortunately, the time-scale T in which the absorbing state is reached is $T \sim (N\ln N)/\lambda$ for $r = 1/2$, which, for a biological system with $1/\lambda \sim 1$ week and $N \gg 1$, far exceeds the lifetime of a mammalian organism.

For arbitrary r, the time-scale T may be estimated by repeating the calculation in Ref. [87]: The saturation condition is $\sum_{\mathbf{i}} C_{\mathbf{i}}(T) = N\rho$. Replacing the summation by integration, $\sum_{\mathbf{i}} C_{\mathbf{i}} \to \int_0^{\infty} xC(x,t)dx$, we arrive at the asymptotic relation

$$\lim_{\lambda t\gg 1} \sum_{\mathbf{i}} \frac{C_{\mathbf{i}}-\rho^2}{\rho-\rho^2} \simeq \frac{4r\lambda T}{\pi(1-2r)+2r\ln(32\lambda T)}\,,$$

and the result for T follows. Interestingly, this result predicts that in even a reasonably large system and finite r, when $N \gg e^{\pi/2r}$, then the time to saturation is insensitive to the value of r. Indeed, for such systems the cross-over to non-voter-like behaviour only occurs at very small values of $r \sim 1/\ln N$.

(d) The product of the correlation length and interface density gives a time-invariant characteristic of voter-like coarsening.
It has been previously noted that in voter-like coarsening, the characteristic length-scale of domains is inversely proportional to the interface density c_{AB} [85]. An important implication of this observation is that at asymptotically long times, one may identify a time-invariant characteristic of the correlations, which we define as $\Omega \equiv (\rho - \rho^2)/(c_{\mathrm{AB}}\xi) = 2r/\pi$.

To characterise the Ω-constant, one may identify its definition as the ratio between the domain size ξ and domain circumference, as calculated from the number of A-B interfaces associated with each domain $c_{\mathrm{AB}}\xi^2$. Thus, the domain perimeter has a trivial fractal dimension of one, with Ω indicating the perimeter roughness, or curvature. Values of $\Omega \simeq 1$ may be associated with cohesive and smooth domains (i.e. with a large area-to-interface ratio), whereas systems with $\Omega \to 0$ have highly fragmented, or rough, domains. Not surprisingly, smaller values of r lead to rougher domains as a result of the Kawasaki dynamics. Yet, one may see that even at the maximum value of $r = 1/2$, the voter model predicts rough domains ($\Omega < 1/2$) — an observation readily seen in simulations, see e.g. Ref. [91].

With an eye to the empirical analysis in the next section, let us note that the Ω-constant allows us to characterise *static* observations of type A cell correlations, and thus provides a valuable test of whether a given data set is consistent with the long-time behaviour of process (5.3).

(e) The lattice geometry influences only the timescale of clustering, with higher coordination number corresponding to a faster timescale.
For lattice geometries with different coordination number $z \neq 4$ (and one site per unit cell), the calculation proceeds as for a square lattice, but now three modifications must be made: First, as mentioned above, the central source term is $J_{\mathbf{0}} = z\lambda(\rho - C_{\mathbf{e}})$, and there are now z sink terms $J_{\mathbf{e}}$ at nearest neighbouring sites. Second, the appropriate Green's function now has the (non-separable) form

$$G_{\mathbf{i}}(t) = \frac{1}{4\pi^2}\int d^2\mathbf{q}\exp\left[i\mathbf{q}\cdot\mathbf{i} - \lambda t\left(z - \sum_{\mathbf{e}} e^{i\mathbf{q}\cdot\mathbf{e}}\right)\right],$$

from which we obtain the general form of the small-p (long-time) expansion of the Laplace transform $\lim_{p/\lambda\to 0} g_{\mathbf{0}}(p) \sim \ln(\lambda/p) + d_{\mathbf{0}}/2\lambda$ and $g_{\mathbf{i}}(p) = g_{\mathbf{0}}(p) - d_{\mathbf{i}}/2\lambda$, where $d_{\mathbf{i}}$ are numerical constants, e.g. $d_{\mathbf{e}} = 2/z$. Third, the sum over sink terms in Eq. (5.9) is revised to reflect the lattice geometry. With these three modifications, one finds that the order parameter c_{AB} takes the general form $c_{\mathrm{AB}} = 2\pi(\rho - \rho^2)/[\alpha(1-2r) + 2r\ln(\beta\lambda t)]$, where α and β are geometry-dependent constants. One is led to conclude that the lattice geometry only serves to rescale the time variable by some constant, $t \to \beta e^{\alpha(1-2r)}t$. It is interesting to contrast this with the more significant effect of modifying the branching ratio r, which instead rescales $\ln t$.

With regards to the experimental system, the apparent insensitivity of the results to details of the

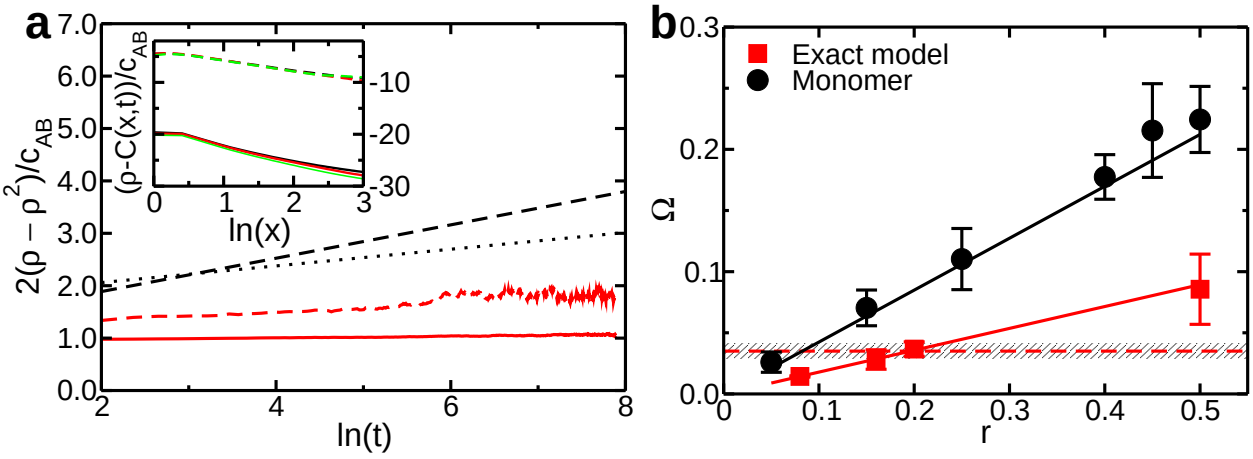


FIG. 5.3: Comparison of the exact basal layer lattice model (5.1) (red curves) with the simplified model (5.3) (black). (a) The order parameter $1/c_{\mathrm{AB}}$ is plotted against $\ln t$. The curves correspond to the simplified model with $r = 1/2$ (dashed black) and $r = 1/4$ (dotted), and to the exact model with $r = 1/2$ (dashed red) and $r = 0.08$ (solid curve). For the simplified model, the analytical expression for c_{AB} was used. For the exact model, the results were obtained from numerical simulation on hexagonal lattices of size $N = 1024 \times 1024$, with $\rho = 0.22$ and setting $\lambda = 1$. *Inset*: The correlation function plotted against logarithm of the distance $\ln x$, evaluated from numerical simulation and Eq. (5.17). The time-invariant ratio $(\rho - C(x,t))/c_{\mathrm{AB}}$ is plotted at $\lambda t = 50, 100, 1000$ for the classical voter model (dashed), and at $\lambda t = 350, 3500, 7000$ for the exact model with $r = 1/2$ (solid curves). Consistent with the expected behaviour (see Eq. (5.14)), the curves overlap and are convex near $x = 1$, becoming linear at $x > 1$. (b) The long-time asymptotic roughness constant Ω plotted against r for both models. Data points correspond to the values calculated using Eq. 5.17 from numerical simulations such as shown in Fig. 5.4, using the algorithm described in section 5.3.1. Error bars result from fluctuations due to finite-size effects, as evaluated by considering the random variation in Ω over the time of the simulation. The black curve gives the theoretical value of Ω for the simplified (monomer) model on a hexagonal lattice, while the red curve gives the best-fit to the numerical results for the exact model.

lattice geometry reinforces the validity of a lattice-based description of the basal layer. Namely, while it is clear that the basal-layer is not a periodic lattice of uniform cell size and coordination number, it may nonetheless be modeled as such.

Finally, because the average coordination number of $z = 6$ is expected for the biological system, we have calculated the order parameter exactly for a hexagonal lattice, giving $\alpha^{(\mathrm{hex})} \simeq 0.98\pi$, and $\beta^{(\mathrm{hex})} = 51$.

This completes our theoretical discussion of process (5.3), and of the monomer-monomer surface catalysis model. It now remains to be seen whether the exact model of cell division (5.1) is indeed described by the properties of the approximate process (5.3) (with uniform λ). Let us recall that the latter is a reasonable approximation assuming that the density of type A cells is approximately uniform. Thus, although the approximation must fail on the time-scale $\lambda T \sim N \ln N$ associated with the jamming transition, the logarithmically slow growth of correlations suggests that at shorter times ($t \ll T$) the degree of progenitor cell clustering should be sufficiently low as to make the analysis self-consistent.

To test whether the model indeed satisfies the expected behaviour, in Fig. 5.3 we compare the spatial correlation of type A cells, as given by the properties derived above, with the results of process (5.1) as obtained by numerical simulation using a Gillespie-like algorithm (described below in section 5.3.1). An example of the cellular automata simulations used for the comparison is shown in Fig. 5.4, where the distribution of type A cells (black) is shown on a hexagonal lattice. Qualitatively, one may see from this figure that process (5.1) does indeed give rise to some clustering. By extracting the correlation function from such figures, we obtained the quantitative comparison shown in Fig. 5.3. In Fig. 5.3a, we compare the evolution of the inverse order parameter $1/c_{\mathrm{AB}}$ against $\ln t$ for the two models. One may see from the linear behaviour of the exact model (red online), that the order parameter indeed shows the expected $1/\ln t$ decay. Equally, we may confirm that the radial correlation function $C(x,t)$ has the functional form $a - b\ln x$, with parameters $(1-a)$, b both proportiaonal to c_{AB}, as demonstrated by plotting $(C(x,t)-\rho)/c_{\mathrm{AB}}$ in Fig. 5.3a (inset). Finally, we may confirm that both models have the same long-time r-dependence by plotting Ω against r in Fig. 5.3b. Here, one may see that indeed Ω depends linearly on r.

Yet, although Fig. 5.3 reveals that processes (5.1) and (5.3) result in the same functional dependence of the correlation function on r and t, it is striking that the exact model (5.1) results in a slower growth in correlations compared to the simplified model, characterised both by a slower decay of c_{AB} seen in Fig. 5.3a, and by rougher domains (or lower Ω) seen in Fig. 5.3b. How might one explain this change? Referring to the discussion in section 5.2.1, let us recall that the simplified model is connected to the exact model by relating the local division rate to the mean-field vacancy density ($\lambda(\mathbf{x}) = \lambda' n_{\oslash}(\mathbf{x})$), with the two models becoming equivalent when $n_{\oslash}(\mathbf{x}) = \text{const}$ for an arbitrarily small degree of coarse-graining. This condition is satisfied in the limit $\sigma \gg \lambda'$ with the process $\mathrm{A} \oslash \to \oslash \mathrm{A}$ removed. However, for the physically meaningful case $\sigma \ll \lambda'$ one can no longer treat the vacancy density as uniform, leading to quantitative (but not qualitative) differences in the behaviour of the two models. As shown schematically in Fig. 5.5, the vacancy population has the effect of accelerating cell division on the edge of smooth progenitor cell clusters, as a result of the higher concentration of nearby post-mitotic cells migrating into the super-basal layers. Conversely, cell division on rough cluster edges is slowed down. In total, rough A-B interfaces remain stable over a longer period of time, leading to the observed differences between the two models.

5.3 Empirical analysis of cell distributions in the basal layer

We are now in a position to turn to the experimental analysis of the IFE. In order to identify the underlying rules of cell division and differentiation, one can envision two types of experiment:

- First, one may track the fate of *individual cells* and their progeny, and then look for a cell kinetic description compatible with their observed behaviour. Such an approach was used with considerable success in the definition of the zero-dimensional process (3.2) through clonal analysis, and we shall extend it to the analysis of the spatial process in section 5.3.1.
- Second, in section 5.3.2 we shall consider the statistics of the *entire population* of basal layer cells. In particular, by immunostaining basal layer cells for markers of cell proliferation, it is possible to analyse the spatial distribution of all progenitor cells within the layer. Then, referring to the theoretical discussion in section 5.2, one may look for a signature of the

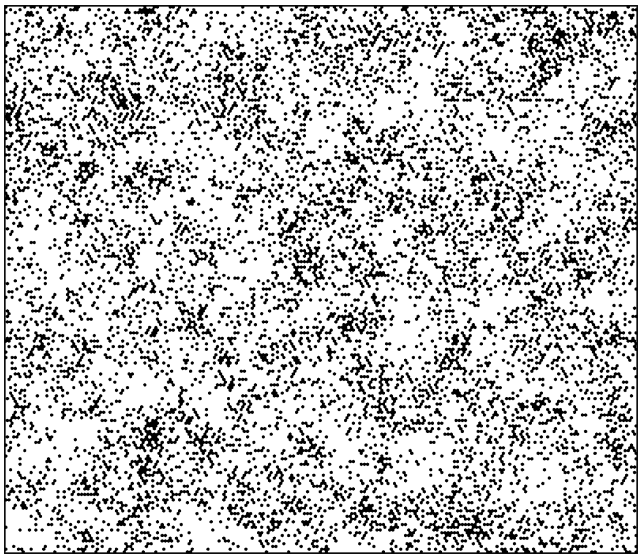


FIG. 5.4: Cellular automaton simulation of process (5.1), showing the distribution of progenitor cells (black hexagons) on a lattice of $N = 200 \times 200$, and using the experimental branching ratio $r = 0.08$ and progenitor cell fraction $\rho = 0.22$. The frame shown corresponds to an evolution time of $t = 30/\lambda$, where $t = 0$ corresponds to random initial conditions. White areas are fully occupied by post-mitotic cells.

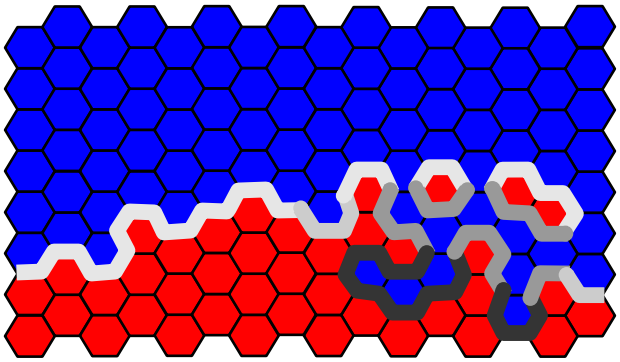


FIG. 5.5: Schematic demonstrating the variation in the effective division rate in process (5.1), accounting for the quantitative variation in behaviour between processes (5.1) and (5.3). The reactive interface between progenitor cells (light grey, red online) and post-mitotic cells (dark grey, blue online) is indicated by the thick grey line, with a faster division rate shown in light grey and slower division rate shown in dark grey. On the left, a smooth interface results in faster cell division, as few progenitors on the boundary must compensate for the migration of many post-mitotic cells in the bulk. On the right, the "rough" interface results in a variation between fast and slow division rates, due to restricted access of vacancies into the rough interfacial regions. Referring to the discussion at the end of section 5.2.2, this leads to increased stability of rough compared to smooth interfaces.

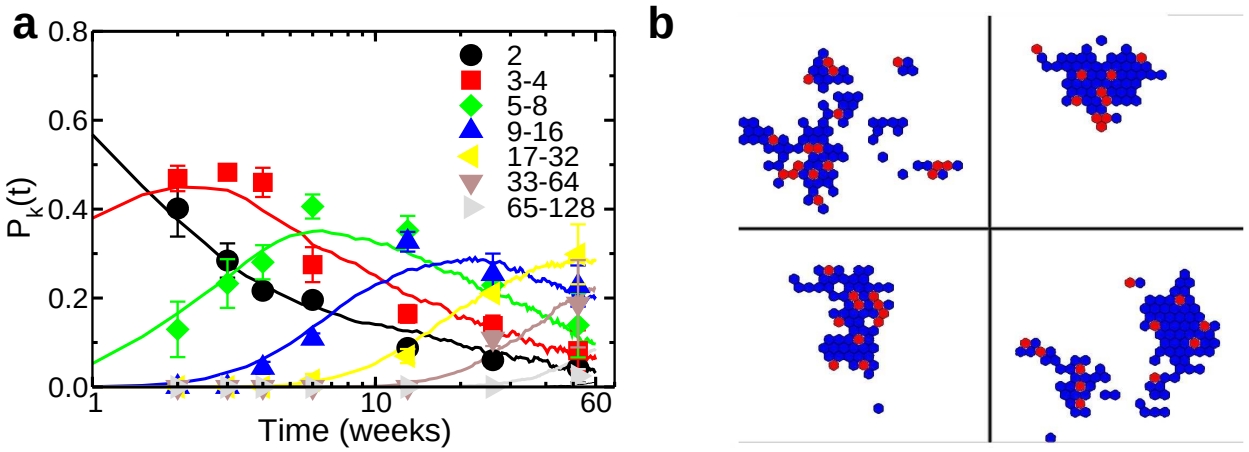


FIG. 5.6: (a) Comparison of the empirical clone size distribution (data points) to predictions of process (5.1) (solid curves), as obtained from Monte-Carlo simulations of 10^4 labelled clones. The distributions are plotted in terms of the probability $\mathcal{P}_k(t)$ for a clone to have between $2^{k-1}+1$ and 2^k basal layer cells at time t after labelling, normalised to include only clones with 2 or more cells in the basal layer ($k \geq 1$). The empirical data is reproduced from Ref. [15]. (b) Examples of the basal layer structure of several large late-stage clones evolving according to process (5.1), starting from a uniform random distribution of unlabelled cells, with one single type-A cell labelled at $t = 0$. Light and dark grey hexagons (red and blue online) indicate sites occupied by type A and B cells, respectively. White areas are populated by unlabelled cells. Each frame corresponds to the progeny of one initially labelled cell.

underlying cell kinetics in the spatial correlation of proliferating cells.

Note that the two types of experiment give access to independent aspects of cell behaviour: The former probes the temporal evolution of cell lineages, whereas the latter reveals the static basal layer morphology. As such, the experiments provide a significant degree of mutual verification of any proposed theory of cell behaviour. Yet, even the best of such experiments leave room for some ambiguity: For example, it is far from clear what importance should be assigned to the embryonic development of the IFE in predetermining the spatial distribution of cells. Moreover, one may in principle conceive of regulatory pathways that leave no signature on the spatial distribution of cells, or which may not be distinguished from an independent stochastic process. Thus, in the following we will look for the simplest possible model of cell behaviour that succeeds in capturing the known biological constraints.

5.3.1 Clonal analysis

To begin our analysis of the empirical data, we start by considering the fate of individual labelled cells and their progeny (clonal fate data), using the data reported in chapter 2: As discussed in section 5.2, we must first establish, for any proposed model of spatial behaviour, that the zero-dimensional clone size distributions are faithfully reproduced, together with the long-time scaling form

$$\lim_{r\lambda t \gg 1} P_{n>0}(t) = \frac{\rho}{r\lambda t} \exp\left(\frac{\rho\, n}{r\lambda t}\right)$$

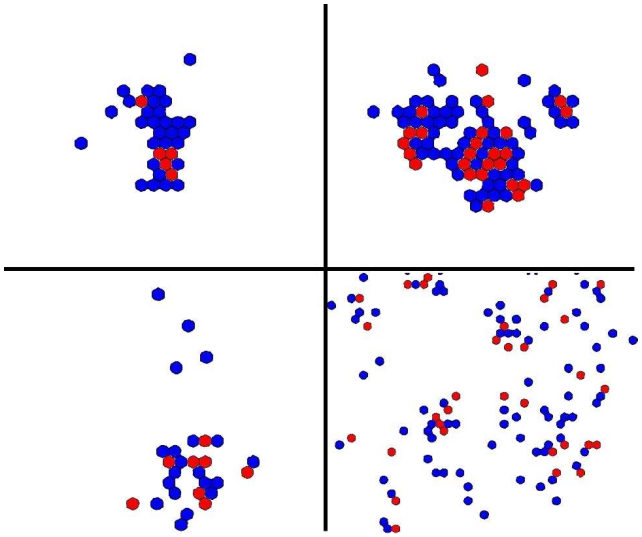


FIG. 5.7: The loss of cohesiveness with increasing the relative hopping rate σ/λ' is shown through examples of late-stage clone simulations. Light and dark grey hexagons (red and blue online) indicate sites occupied by type A and B cells, respectively, from the same clone. White areas are populated by unlabelled cells. The simulations used the same parameter set as described above for the clones in Fig. 5.6, but using the parameter values: $\sigma = 20$/week, $\sigma = 200$/week with $\lambda' = 10,000$/week for the top-left and top-right clones respectively; $\sigma = 200$/week, $\sigma = 2000$/week with $\lambda' = 200$/week for the bottom-left and bottom-right clones. For the latter, which is clearly unphysical, the scale has been reduced twofold to demonstrate the wide dispersion of labelled cells. The examples demonstrate that clones are cohesive when $\sigma/\lambda' \ll 1$, but dispersive otherwise.

identified in chapter 3.

To this end, we conducted multiple simulations of process (5.1) as an asynchronous cellular automata evolving on a hexagonal lattice of $N = 60 \times 60$ sites over a period corresponding to $T = 60$ weeks in the experimental system (recall that $\lambda = 1.1$/week [15]). At $t = 0$, the lattice was fully occupied by randomly-placed type A and type B cells. A single, randomly chosen, type A-cell was assigned a hereditary 'label' at the start of each simulation. In effect, each such simulation mimics the evolution of one labelled clone, so that repeated simulations may be used to sample the full clonal statistics, viz. Monte-Carlo sampling. In particular, by tracking the number of labelled cells as a function of time over 10^4 such clone simulations, we could compare the clonal statistics predicted by process (5.1) with those expected from the zero-dimensional process (3.2). In keeping with the empirical fit from the previous section, we used a B-cell migration rate of $\Gamma = 0.31$/week, and an initial A-cell fraction of $\rho = 0.22$. The 'fast' rates of cell division and hole diffusion were set to be $\lambda' = 10^4$/week and $\sigma = 200$/week, although the precise values are unimportant provided that $\lambda' \gg \sigma \gg \Gamma$, as discussed in section 5.2. The results are plotted in Fig. 5.6(a), where we compare the clone size distribution $P_{n>0}(t)$ to the empirical data. One may see that the fit is remarkably good, and, within the current empirical resolution is indistinguishable from the predictions of zero-dimensional model (cf. Ref. [16]). In summary, the fit provides a first validation of process (5.1) as a viable model of spatial behaviour in the basal layer.

We may now revisit the clone fate data with a view to study spatial structure. Unfortunately, although the clone size data was stored for all clones, images showing their spatial structure were

retained only for a small number of clone samples (see, for example, Fig. 5.1). Therefore we are not in a position to conduct a comprehensive *quantitative* analysis of clone shape evolution. Nevertheless, a striking qualitative feature of clones is that they remain largely cohesive, as shown in the example in Fig. 5.1. (Indeed, without this property the very enterprise of clonal analysis would have proved difficult). Therefore, to challenge the validity of the proposed lattice model of cell division, we test, using additional Monte-Carlo simulations, the ability of the model to produce cohesive clones over the one-year time period of the experiment. Representative results are shown in Fig. 5.6 for large clones at 60 weeks post-labelling. When $\lambda' \gg \sigma \gg \Gamma$, clones were seen to remain largely cohesive throughout the period of the simulation. We may therefore conclude that, at least with respect to the existing clonal fate data, process (5.1) presents a reasonable phenomenological description of the spatial behaviour of basal layer cells. In particular, the cohesive nature of clones may be completely explained in terms of an independent stochastic process, with no evidence for further forms of regulation.

Yet, to what extent is the observation of cohesiveness sensitive to modifications of the basal layer lattice model? To test the degree to which cohesiveness constrains the model, we allowed for a degree of cell mobility within the basal layer, by introducing the additional exchange process $AB \to BA$ with rate $\gamma \sim \Gamma$. The exchange process is a reasonable candidate for cell behaviour, motivated by the observation that keratinocytes in culture are highly motile, which leads one to postulate whether cells *in vivo* are capable of independent lateral migration in the basal layer. We found that for any non-small value of γ, (i.e. $\gamma \gtrsim \Gamma$), the progeny of labelled clones rapidly dispersed, as shown in Fig. 5.7. We are led to conclude that *epidermal cells in vivo move only in response to a local density gradient* resulting from cell division and migration. Moreover, a second investigation in which the hard-core mobility σ was made larger than λ' (and keeping $\gamma = 0$) again led to a loss of clone cohesiveness. We are therefore led to conclude that, within the framework of the non-interacting lattice model, *progenitor cells undergo division rather than lateral migration as a response to a local drop in density.*

In conclusion, the observation of clone cohesiveness imposes a severe constraint on cell behaviour, which allows us to rule out several variations of the basic lattice model. Yet, at least qualitatively, there is no clear evidence either *for* or *against* additional regulation of cell division in the clone shape data. So to what extent is the lattice model truly capable of shedding new light on the mechanism of cell fate regulation? To do better, we must look for quantitative data for comparison. Fortunately, such data can be found in the readily-available spatial distribution of progenitor cells.

5.3.2 Correlation analysis

As mentioned earlier, an intriguing feature of the non-interacting lattice model is the prediction of clustering of proliferating cells. How does this prediction compare with experiment? Before we turn to consider new results for mouse tail-skin, it is interesting to first consider results presented in past work. In a detailed study of proliferating cells in Hamster cheek epidermis by Gibbs and Casarett [81], a subset of progenitor cells were labelled using a radioactive marker for DNA synthesis (S-phase). By measuring the number of unlabeled cells separating consecutive labeled ones in one-dimensional basal layer cross-sections, it was possible to access the full radial distribution for the separation between adjacent cells undergoing S-phase. Remarkably, while the large-interval distribution decayed exponentially as expected for an uncorrelated random distribution, the proba-

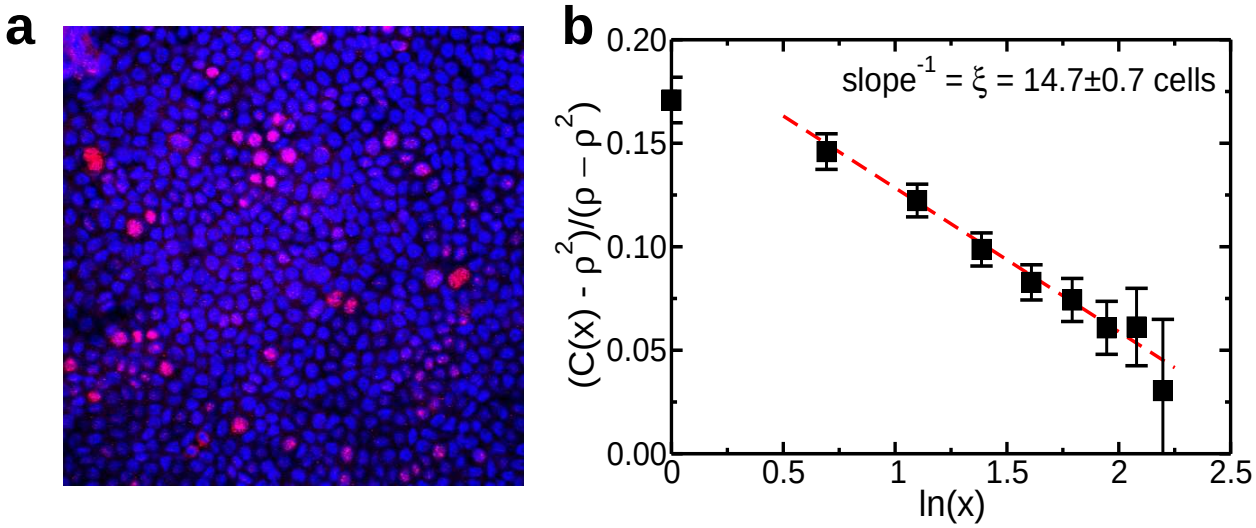


FIG. 5.8: (a) Confocal micrograph of wholemounted mouse tail skin IFE, showing the two-dimensional basal layer immunostained for the nuclear marker DAPI (dark grey, or blue online) and the proliferation marker Ki67 (light grey, or red online). The area surrounding the stained nuclei is occupied by unstained cell cytoplasm (black). (b) The empirical radial correlation function $C(x)$, as defined in Eq. (5.18). Data points show results obtained by analysis of the Ki67-stained epidermal wholemounts exemplified in (a), taken from a mouse aged 8 weeks; the dashed line shows the fit to the analytical form of the correlation function predicted by process (5.1), $C(x,t) = a(t) - \xi^{-1}(t) \ln x$ (see Eqs. 5.14, 5.16), where the single time point t is fixed by the age of the mouse and by initial conditions (see discussion in section 5.4). From the fitted slope one may extract the correlation length $\xi(t)$ defined in Eq. (5.16), giving $\xi = 14.7$ cell diameters.

bility of finding a nearby cell in S-phase was significantly higher at short distances, indicating that proliferating cells were clustered. A second study addressing the distribution of S-phase cells in mouse esophagus also revealed identical qualitative results [82].

Not surprisingly, in the absence of the intuition afforded by the voter model, such clustering has been interpreted in the biological community as evidence of an underlying regulatory process, which leads to the synchronous division of nearby cells [82]. On the other hand, analysis of clone fate data (chapter 3) has revealed that cell division within clones occurs independently rather than synchronously. It is therefore satisfying to note that the tendency of proliferating cells to cluster *is* in fact consistent with independent division — indeed it is the hallmark of voter-model dynamics.

To extract the spatial correlation between progenitor cells in mouse tail skin, we analysed confocal micrographs of basal layer cross-sections of IFE that were stained for the proliferation marker Ki67, such as shown in Fig. 5.8(a). Cells bright in Ki67 are designated as type A (progenitor) cells, whereas Ki67-dull cells were designated as type B (i.e. differentiated) cells, with no capacity to divide. Using image analysis software (ImageJ), the coordinates of each Ki67-bright cell were extracted. This data allows a full statistical analysis of the spatial distribution of progenitor cells. In particular, we shall focus on the radial correlation function

$$C(x) = \left\langle \frac{1}{A} \int_0^{2\pi} \frac{d\theta}{2\pi} \int_A d\mathbf{x}'\, n(\mathbf{x}')\, n(\mathbf{x}' + \mathbf{x}(x,\theta)) \right\rangle \tag{5.17}$$

where A denotes the area of each sample, $n(\mathbf{x})$ denotes the areal density of proliferating cells at position $\mathbf{x}$, and the brackets $\langle\cdot\rangle$ indicate averaging over all basal layer samples.

Using Eq. (5.17), the aim of the correlation analysis is to assess whether basal layer progenitor cells in adult mice do indeed cluster, and if so, to assess whether the experimental correlation function is consistent with the predictions made in section 5.2.2. In principle, one may look to the data for signatures of the expected spatial dependence $C(x,t) \sim a(t) - b(t)\ln x$ (for a given value of t), as well as for evidence of increased clustering with time, viz. $c_{\mathrm{AB}}(t) \sim 1/(2r\ln t)$. For the latter, however, the temporal analysis is difficult to implement due to the sensitivity required to resolve the $1/\ln t$ decrease in c_{AB} at long times. In particular, the predicted increase in clustering over a biologically-relevant period of $\lambda t \sim 10^1 - 10^2$ cell cycles corresponds to a decrease in c_{AB} of $< 10\%$, whereas variations in the efficiency of Ki67 labelling between different mice introduce systematic errors of the same order. Therefore, in the following we shall restrict ourselves to the quantitative analysis of tissue samples taken from a single adult mouse (aged 8 weeks).

To incorporate the empirical coordinates of proliferating cells into Eq. (5.17), we replace the product $n(\mathbf{x}')n(\mathbf{x}'+\mathbf{x})$ with the sum $\sum_{i,j} \delta(\mathbf{x}' - \mathbf{R}_i) f(\mathbf{x}' + \mathbf{x} - \mathbf{R}_j)$, where $\delta(\mathbf{x})$ is the Dirac delta function, $\{\mathbf{R}_i\}$ is the set of all progenitor cell coordinates, and $f(\mathbf{X})$ is a two-dimensional Gaussian envelope with width w. With this substitution, the integrals in Eq. (5.17) can be solved exactly. Then, explicitly accounting for the averaging procedure over sample of variable size, we obtain the expression

$$\begin{aligned} C(x) \;=\; & \left\langle \frac{\rho}{2\pi w^2 N_{\mathrm{A}}(x)} \sum_{i=1}^{N_{\mathrm{A}}} \sum_{j\neq i}^{N_{\mathrm{A}}(x)} \exp\left(-\frac{x^2 + R_{ij}^2}{2w^2}\right) \right. \\ & \left. \times I_0\left(-\frac{xR_{ij}}{w^2}\right) \right\rangle . \end{aligned} \tag{5.18}$$

Here, the sum i, j over all progenitor cell coordinates arises from the empirical expression for $n(\mathbf{x}')n(\mathbf{x}'+\mathbf{x})$ given above, and we have defined $R_{ij} \equiv |\mathbf{R}_i - \mathbf{R}_j|$. The prefactor, exponential factor and modified Bessel function (I_0) result from solving the integrals in Eq. (5.17). The total number of progenitor cells found in each sample is N_A, and $N_\mathrm{A}(x)$ is the number of progenitor cells at a distance x or more from the sample edges. As defined, the correlation function avoids errors resulting from edge effects by averaging each sample over the $N_\mathrm{A}(x)$ progenitor cells that are unaffected by the finite sample size. To correctly average over different samples, as indicated by $\langle\cdot\rangle$, the sample results are weighted by $N_A(x)$, e.g. $C^{(1+2)}(x) = [N_A^{(1)}(x)C^{(1)}(x) + N_A^{(2)}(x)C^{(2)}(x)]/[N_A^{(1)}(x) + N_A^{(2)}(x)]$.

Making use of Eq. (5.18), the experimental correlation between progenitor cells was averaged over 10 samples of approximately 30×30 cells each, see Fig. 5.8(b). Remarkably, the data indeed shows a significant degree of progenitor cell clustering, in good agreement with the linear decay in $\ln x$ expected from Eq. (5.14). In real terms, one may infer from the value of the nearest-neighbour correlation $C(1)$ that each progenitor cell is in contact, on average, with approximately *two* adjacent progenitor cells, compared to 1.3 progenitor cells expected for an uncorrelated random distribution (at $\rho = 0.22$).

For a more careful test of the theory, one may extract from Fig. 5.8(b) the order parameter $c_\mathrm{AB} = 0.29 \pm 0.02$, and the correlation length of $\xi = 14.7 \pm 0.7$ cells, from which we find the empirical roughness constant, $\Omega = 0.04 \pm 0.01$. Repeating the analysis with samples taken from different mice results in the same value of c_AB, but with values of Ω in the range $\Omega = 0.02 - 0.04$. Referring to Fig. 5.3(b), where the empirical value of Ω is compared to the model predictions (dashed), we see that the model indeed recovers the correct order of magnitude of the roughness constant. Qualitatively, then, it appears that the model is consistent with the observed clustering. Let us emphasise that this fit requires *no additional parameters*, and is purely a result of mapping the cell kinetics onto a lattice.

5.4 Spatial regulation: discussion and conclusions

To summarise, we have demonstrated that the zero-dimensional model of cell division is consistent with the maintenance of the basal layer at uniform density, and we have shown that the size distribution and qualitative shapes of labelled clones are consistent with a simple stochastic model of cell division and differentiation on a two-dimensional lattice. These results explain why the "zero-dimensional" model of chapters 2-4 gives such an excellent fit to the clone fate data despite taking no account of additional regulatory pathways. Significantly, despite the many forms of cell fate regulation known to exist in development and adult tissue, the *only* extra-cellular regulation required to understand the existing observations of clone fate in normal IFE is *steric*, viz. the coupling of cell division to the local cell density.

Beyond the success of the model in explaining the observed clone fate data, we have also identified that the degree of progenitor cell clustering, as measured by Ki67 staining, is in good qualitative agreement with the predictions of the spatial process. In particular, there are two features that allow us to characterise the spatial process. First, the empirical roughness constant $\Omega = 0.04 \pm 0.01$ has the expected order of magnitude predicted by the model (Fig. 5.3b), and second, the correlation function is in excellent agreement with the expected decay form $a(t) - b(t)\ln x$ (at fixed t), as seen

in Fig. 5.8.

Taken together, these results have two important implications for future investigations of epidermal cell fate regulation: First, by demonstrating that the zero-dimensional process (3.2) is indeed capable of maintaining a uniform total basal layer cell density, the spatial process consolidates the proposed stochastic model as a robust platform for investigating biochemical constituents in future work. For example, by over/under-expressing specific genes and then studying the resulting change in the empirical parameters $(r,\ \lambda,\ \rho)$, one may attempt to identify the role of each constituent in regulating cell behaviour. Second, the new model introduces a set of spatial measures (such as the roughness constant Ω) that yield further information in such investigations of the biochemistry, beyond that which may be obtained through empirical evaluation of the zero-dimensional parameters alone.

Beyond the qualitative features of the dynamics, one may ask whether there are any implications to the quantitative features of the correlation analysis. Unfortunately, as mentioned earlier, variations in the efficiency of Ki67 labelling prevent us from generating the comprehensive statistics necessary to accurately quantify the universal values of the system parameters. Nevertheless, taken at face value, it appears that the empirical value of Ω obtained from the sample analysed in section 5.3.2 is only consistent with process (5.1) when one imposes a branching ratio $r = 0.19 \pm 0.04$, which differs significantly from the value of $r = 0.08$ established through clonal analysis (see fig. 5.3b). This value is further consistent with the observed AB interface concentration of $c_{\mathrm{AB}} = 0.29$ and the correlation length of $\xi = 14.7$ cells, which correspond to a long evolution time of $\lambda t \approx 10^4 - 10^6$ for $r = 0.08$, but a realistic biological time-scale of $\lambda t \approx 10^1 - 10^2$ for $r = 0.19$.

Although it is possible that the discrepancy in the inferred value of r results from errors in progenitor cell classification as discussed in section 5.3.2, the difference is large enough to call into question the reliability of the correlation analysis. Beyond the issue of progenitor cell labelling efficiency, there is also the more basic question of whether Ki67 is at all effective as a marker of progenitor cells. In particular, it is widely accepted that Ki67 is a marker of cell growth as well as proliferation, which may imply that differentiated cells remain Ki67-bright for some time after division, or that progenitor cells may fail to continuously express Ki67 [92, 93]. However, assuming that Ki67 is indeed a faulty marker, it becomes difficult to explain why the analysis nevertheless results in an excellent fit to the '$a(t) - b(t) \ln x$' decay form of correlations (at fixed t).

We may also challenge our assumption of using "random initial conditions" to model steady-state maintenance. If initial conditions are at all relevant, then the observed clustering may be a signature of development and growth processes that are no longer active in the adult system. To test the importance of initial conditions, we have compared the order parameter c_{AB} between adult mice aged 8 and 60 weeks (corresponding to yound and old adult mice). If clustering is inherent to initial conditions, then one would expect c_{AB} to grow over time as the memory of initial conditions erodes. On the other hand, as mentioned in section 5.3.2, steady-state maintenance results in $< 10\%$ change in the order parameter over this time period, which would be undetectable given the systematic errors in Ki67 labelling. Indeed, we find no change from $c_{\mathrm{AB}} = 0.29 \pm 0.02$ at 60 weeks, so there is no indication that the observed clustering is a signature of tissue development.

Finally, it is interesting to speculate whether additional forms of regulation may be capable of reconciling the higher value of $r = 0.19$ found for the spatial process, with the lower value of

$r = 0.08$ found from the zero-dimensional analysis. We propose a simple revision of the model that is capable of reconciling the spatial and zero-dimensional data: Referring to process (5.1), if the two channels of asymmetric division in process (5.1) occur with different probabilities (viz. $P_{\mathrm{A}\oslash\to\mathrm{AB}} \neq P_{\mathrm{A}\oslash\to\mathrm{BA}} \neq 1/2 - r$, and $P_{\mathrm{A}\oslash\to\mathrm{AB}} + P_{\mathrm{A}\oslash\to\mathrm{BA}} = 1 - 2r$), then the effective value of the branching ratio r is effectively renormalised in the spatial process, while leaving it unchanged in zero-dimensions. To reconcile the two empirical values of r, it is simple to show that one requires $P_{\mathrm{A}\oslash\to\mathrm{AB}} \simeq 7P_{\mathrm{A}\oslash\to\mathrm{BA}} \simeq (1 - 2r)/8$. Thus, one might speculate that the increase in clustering is associated with a spatial asymmetry in daughter cell fate during asymmetric division. It would be an interesting challenge to devise an experiment with which to test this asymmetry.

Chapter 6

Patterning as a signature of human epidermal stem cell regulation

Chapter overview

While the preceding chapters have focused on the question of epidermal maintenance in mouse tail-skin, the discovery of a new paradigm of tissue maintenance compels us to consider whether similar rules may govern the behaviour of human keratinocytes. Yet, despite the similarities between the epidermis of mice and men, the two systems also differ sufficiently to demand a fresh look for evidence of a simple process governing cell fate in human epidermis.

In the basal layer of human epidermis, it has been shown that quiescent stem cells spontaneously organise into clusters separated by proliferating and differentiating keratinocytes. In this chapter (and the next), we will show that a wide range of experimental phenomenology, including such stem cell patterning, can be explained by attributing stem cells with just two non-trivial characteristics, which are the end result of intricate molecular regulation. The organisational principles that result from these characteristics can be captured by a coarse-grained or hydrodynamic theory of cell fate.

This chapter is organised as follows: In section 6.1 we survey the existing experimental phenomenology in human epidermis. With this background, in section 6.2 we propose a new model of cell fate, consisting of two populations of progenitor cells that undergo a stochastic birth-death process. Although this model provides only a caricature of the underlying complex molecular circuitry involved in the regulation of division, differentiation, and cell migration, it is sufficiently rich to fully characterise the large-scale dynamics. To relate the proposed model to the patterning phenomenon, in section 6.3 we develop a hydrodynamic theory for the coarse-grained cell densities in the basal layer, with the properties of the system equations, including predictions for the steady-state pattern morphology, discussed in sections 6.4 and 6.5.

The early sections of this chapter (6.1-6.3) have been written to form the basis of a paper aimed

at the biological community. In particular, sections 6.1 and 6.2 were co-written with PHJ[1]. The experimental results presented in section 6.2 were generated by DPD[2] and PHJ.

6.1 Introduction: Human IFE

Historically, human interfollicular epidermis has been thought to be maintained by slow-cycling self-renewing stem cells, which generate a population of transit amplifying (TA) cells that differentiate after a limited number of cell divisions (see chapter 1 and Refs. [27, 28, 25]). However, the cumulative work of chapters 2–5, describing the analysis of clonal fate data in the tail skin of adult mice, has significantly challenged this classical stem/TA cell model. As discussed in chapter 2, although this work does not preclude the existence of a traditional stem cell population within IFE it implies that, if present, such a population remains quiescent in homeostatic tissue.

Experimental phenomenology in human IFE
Although the general architecture of human IFE parallels that of mouse, there is strong evidence for proliferative heterogeneity within the basal layer cell population. As described in chapter 1, a pioneering study has demonstrated that, when single cell-derived colonies of cultured human keratinocytes are subcloned, three types of colony develop [24]; large circular colonies with a very high proliferative potential (holoclones), small irregularly shaped colonies with very limited proliferative potential (paraclones); and colonies with intermediate properties (meroclones). Subsequent studies showed that cultured keratinocytes could be fractionated on the basis of their expression of the β1 integrin family of extracellular matrix receptors [94]. Cells expressing high levels of β1 integrin formed large actively growing colonies, whilst those expressing lower levels formed small abortive colonies in which all cells underwent terminal differentiation. β1 integrin-bright cells can be isolated directly from human epidermis as they adhere rapidly to extracellular matrix proteins [95]. These cells are able to found large actively growing colonies in vitro and regenerate human epidermis when grafted onto immunocompromised mice. By contrast slowly adhering cells, low in β1 integrin expression, form small colonies that undergo terminal differentiation and are unable to regenerate epidermis in xenografts [95].

Subsequent studies of integrin expression in human epidermis [95] revealed that the basal layer was organised into irregular clusters of keratinocytes expressing high levels of β1 integrin, interspersed with regions of lower integrin expression (Fig. 6.2). Moreover, cells expressing high levels of the notch ligand Delta and the cell surface proteins MCSP and LRIG1 are also clustered, and co-localise with β1 integrin-bright cells [96, 97, 98, 99]. In contrast, the desmosomal protein, desmoglein3, has a reciprocal distribution localising in regions of low integrin expression [100]. Consistent with clustering, Delta and MCSP have both been shown to promote the cohesiveness of cultured human keratinocytes [96, 98].

A striking feature of the clusters is that the great majority of the constituent cells appear quiescent, as evidenced by immunostaining for the proliferation markers Ki67 and BrdU (Fig. 6.2), [101, 98]. Both the proliferating cell population, and post-mitotic basal layer cells (identified by expression

[1]Phil H Jones
[2]David P. Doupé

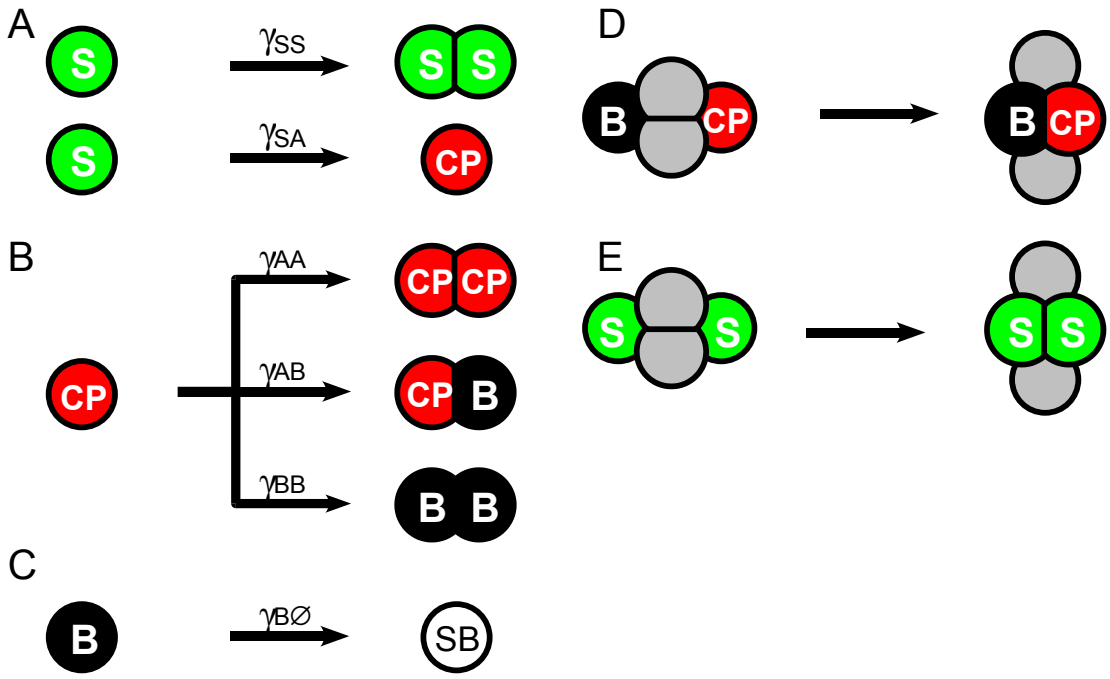


FIG. 6.1: The Stem/CP cell hypothesis. **a**, Stem cell division and differentiation: stem cells (S) give rise to two daughter stem cells, or they may differentiate into progenitor cells, committed to terminal differentiation (CP). **b**, CP cell division and differentiation: CP cells give rise to either two daughter progenitor cells, or one progenitor cell and one post-mitotic (B) cell, or two post-mitotic cells. With the symmetric division rates equal ($\gamma_{AA} = \gamma_{BB}$), as observed in mouse tail skin, the CP cell compartment is self-sustaining allowing a stem cell population to remain quiescent during homeostasis. An imbalance in these rates ($\gamma_{AA} < \gamma_{BB}$) activates the stem cell compartment and limits the lifetime of CP cell-derived clones. **c**, Stratification: Post-mitotic cells migrate out of the basal layer. The parameters γ_{XY} denote the bare average rates of cell division, differentiation, and stratification. **d,e**, Lateral displacement and stem cell adhesion: Local rearrangement of cells may result from several distinct mechanisms of lateral mobility (see main text and appendix 6.B), leading to an effective diffusion (grey cells indicate unspecified neighbouring cells). Stem cells are constrained in their motion by their tendency to aggregate (**e**).

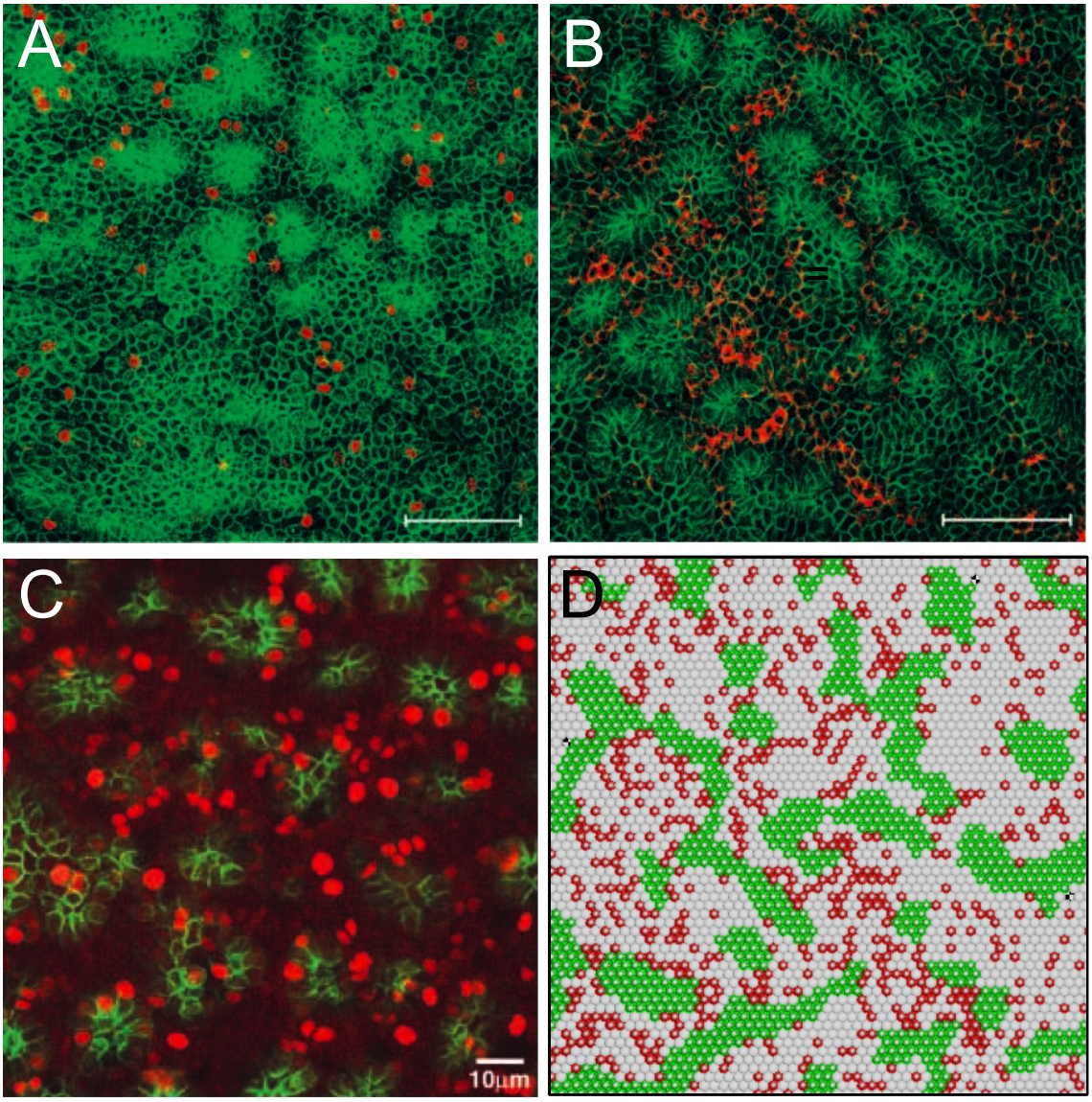


FIG. 6.2: In vivo patterning and quiescence of stem cells. **a** Confocal wholemount of human breast epidermis stained for the stem cell marker $\beta 1$ integrin (green) and BrdU (red). **b** Similar wholemounts of human foreskin epidermis stained for the stem cell marker $\beta 1$ integrin (green) and Keratin 10 (red), a protein that is expressed in cells undergoing terminal differentiation. Both panels (**a, b**) are reproduced from Ref. [101]. The wholemounts reveal that stem cells lie in cohesive quiescent clusters surrounded by cycling and differentiating cells. Scale bars = 100μm. **c** Further evidence for patterning is seen in human breast epidermis wholemounts stained for the stem cell marker MCSP (green) and the Ki67 proliferation marker (red), reproduced from Ref. [98]. **d**, Typical realisation of the stem/CP cell model (Fig. 6.1) obtained through a cellular automata simulation (see section 7.1). The simulation shows that the simple rules in Fig. 6.1 give rise to irregular yet distinct clusters of stem cells (green) within a sea of CP (red) and post-mitotic cells (white).

of the terminal differentiation marker Keratin 10), appear to lie between the clusters [101]. Taken together, these results suggest that, in human IFE, stem cells aggregate into cohesive clusters of near-quiescent cells, interspersed with cycling and differentiating cells with a much lower proliferative potential [101, 95].

Superficially, the evidence for stem cell clustering within a reciprocal pattern of proliferating and differentiating cells in human IFE is in marked contrast to murine epidermis where proliferating and differentiating cells appear to be scattered randomly throughout the IFE [30]. Whether the apparent differences between mouse and human IFE are capable of revealing new aspects of stem cell behaviour depends, in part, on whether stem cells are responsible for clustering, or whether the pattern results from some external process that is not influenced by stem cell behaviour. It is therefore significant that, when placed in culture, keratinocytes spontaneously reconstitute the in vivo pattern even when seeded at clonal density[3] [95].

Historically, the experimental phenomenology of human epidermis has been interpreted within the framework of the classical stem/TA hypothesis [95]. Cells residing between stem cell clusters have been thought to represent the short-lived TA cell population that is continuously replenished by the stem cell compartment. However, this interpretation has been challenged recently by a powerful experiment that provides access to clone fate data in xenografted human IFE. In this experiment, human skin was grafted onto nude mice and then transduced with lentiviral reporter vectors. When the epidermis was examined six months later, the persisting cell clones were found to have a wide range of size and shape, and appeared to originate not only from basal cells within $\beta 1$ integrin-bright clusters, but also from cells between clusters [102]. While one assumes that long-lived clones can only arise from the stem cell population, a tenet of the classical stem/TA cell hypothesis, these findings appear to conflict with the evidence for stem cell clustering. However, if ordinary progenitor cells have the potential to undergo a significant number of cell divisions prior to differentiation, as seen in murine IFE [15], long-lived clones could arise from both stem cell clusters and the intervening cells.

6.2 The Stem/CP cell model

Although the classical stem/TA cell model has been used widely as a platform to interpret experiment, it does not attempt to engage with the apparent spatial organisation (clustering) and regulation (quiescence) of cells in human tissue. Yet, the regeneration of stem cell patterning and quiescence in culture, which occurs without external signals from other cell types (such as dermal fibroblasts), is suggestive of a simple organisational principle involving the cooperative behaviour of the keratinocyte population. As well as their defining property of long-term self-renewal, we will show that the wide range of experimental phenomenology can be explained by attributing stem cells with just two further characteristics, which are the end result of intricate molecular regulation. If we suppose that symmetric stem cell division is inhibited by the accumulation of neighbouring cells and, as observed, stem cells adhere more strongly to each other than to other cell types, one can:

[3] A sufficiently low seeding density such that the resulting macroscopic colonies may be identified as clones deriving from single cells.

- Elucidate the origin of spatial stem cell patterning and quiescence in homeostatic tissue;
- Explain the facility of stem cells to reconstitute their niche and regenerate patterning and quiescence in culture;
- Reveal how the maintenance of human and murine epidermal homeostasis can be embraced within a single framework.

In the following, we will show that the organisational principles that underpin epidermal maintenance can be captured by a coarse-grained or hydrodynamic theory of cell fate. To identify the central ingredients of the theory, we will orient our discussion around a particular stochastic model, a 'caricature' of the underlying complex molecular circuitry.

Motivated by the experimental phenomenology outlined above, and the known properties of murine IFE, we will suppose that the basal layer of human IFE is characterised by a stem and CP cell population. To characterise the properties of these populations, we can draw upon the observed behaviour of keratinocytes in culture:

Stem cell compartment: As a largely quiescent cell population in homeostatic tissue, the division and differentiation properties of the epidermal stem cell compartment are difficult to discern and open to debate. However, an objective definition of a stem cell must include the capacity for self-renewal, and the potential to generate differentiated progeny [49]. These characteristics are embodied by stem cells that either undergo symmetric division into two daughter stem cells, or differentiate to form a progenitor cell (CP) committed to terminal differentiation (Fig. 6.1a). Although one could conceive of other channels of stem cell fate, such as asymmetric division, we will find that such characteristics do not effect the large-scale organisation of the tissue, the focus of the present investigation (see appendix 6.D).

In addition to their division and differentiation potential, we must also address the observed tendency of stem cells to aggregate into clusters. As mentioned above, stem cells are more adherent to underlying extracellular matrix proteins than other basal cells as they express high levels of functional $\beta 1$ integrins, restricting their mobility [101, 95, 94]. Stem cells also adhere more tightly to each other than to other basal cells by virtue of expressing factors that promote cohesiveness, such as the Notch receptor Delta and the cell surface proteoglycan MCSP [96, 98, 103]. In the following, we will suppose that the adhesiveness of stem cells to the underlying basement membrane and to their neighbours constrains the motion of stem cells and promotes clustering (Fig. 6.1e).

Committed progenitor cell compartment: In addition to the stem cell population, we will suppose that human IFE contains a compartment of CP cells whose behaviour resembles that found in murine IFE in that they undergo division that is stochastic, independent of past history, and with cell fate (symmetric versus asymmetric) chosen at random from a fixed distribution (Fig. 6.1b) [104]. Several lines of evidence support this assertion: Firstly, immunostaining two cell clones for the proliferation marker Ki67 in vitro reveals that, as in the murine system, progenitor cell division may lead to either symmetric or asymmetric cell fate (Fig. 6.3). Secondly, earlier studies have shown that the cell cycle time in vitro is broadly distributed, consistent with the stochastic

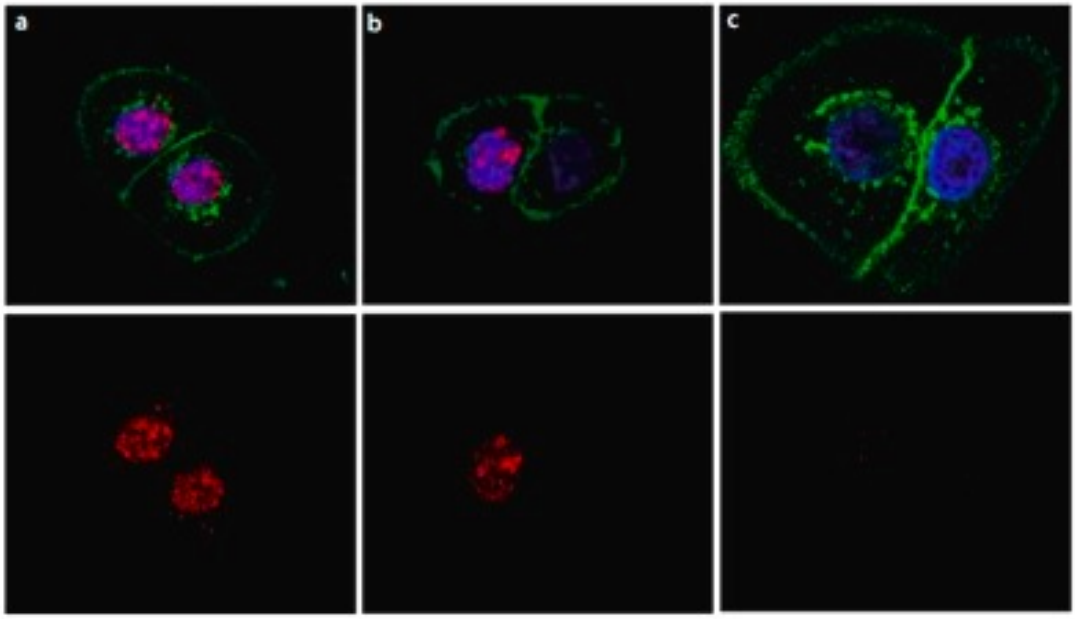


FIG. 6.3: Two-cell clones of human keratinocytes in vitro, stained for the proliferation marker Ki67 (red), nuclear stain DAPI (blue), and membrane stain WGA (wheat germ agglutinin, green), show three different proliferative outcomes of division. From the distribution of Ki67, one may see that division gives rise to two proliferating cells (**a**), one proliferating and one non-proliferating cell (**b**), or two non-proliferating cells (**c**). Images courtesy of David P. Doupé and Phil H. Jones.

and independent behaviour of CP cells [71, 16]. Thirdly, although not conclusive, the observation of long-lived clones residing outside stem cell clusters in human epidermal xenografts [102] is typical of cells capable of an unlimited number of rounds of division before terminal differentiation [15, 16]. Fourthly, in the same xenograft experiments, the observation of a wide range of clone sizes six months after labelling, which seems incompatible with the "clonal unit" predicted by the classical stem/TA cell model, is consistent with a CP cell population [104, 105]. Based on extensive analysis of mouse IFE, we will assume that, in adult epidermis, the vast majority of basal cell divisions are in-plane generating two basal cells [15, 106, 57]. The existence of mitoses perpendicular to the basement membrane would not influence the structure of the hydrodynamics.

Regulation of cell division: Homeostasis of tissue places severe constraints on cell behaviour. Within the basal layer, the cell density remains approximately uniform indicating that proliferation is tightly regulated during normal tissue turnover: On average, for each cell division, one cell must leave the basal layer through upward migration. One approach to analysing this form of regulation is to relate the average cell cycle time to the local cell density, with proliferation being halted when the density becomes too high. Indeed, this form of regulation can be motivated by a range of biochemical regulatory mechanisms, such as cell-cell trans-membrane signalling, gradients of short-range diffusible signalling factors, or mechanical stress-based control [83].

This completes our definition of the stochastic stem/CP cell model as applied to human IFE. In seeking to address the origin of stem cell patterning and quiescence, previous works have introduced alternative stochastic models of cell fate that place emphasis on adhesion and regulation [42]. Here we have focussed on the simplest model that stays within the general paradigm laid down by the organisational principles outlined above, and is consistent with the observed properties of the experimental system.

Although one can conceive of numerous variants of the model that stay within the same general paradigm, we have focussed on the simplest model that is consistent with the observed properties of the experimental system, and which characterises the hydrodynamics in full.

6.2.1 The stem/CP cell principle of epidermal maintenance

To understand the organisational principle of stem cell regulation, let us show how the stem/CP cell model can explain patterning of tissue. To this end, let us first consider a limit of the model in which stem cell division and differentiation are fully suppressed. In this case, the symmetric division rates associated with the CP cell population must be balanced, as observed in the murine system, so that the migration of post-mitotic cells from the basal layer is wholly compensated by the production of cells through CP cell division. If present, the adhesive properties of (quiescent) stem cells would lead to their gradual aggregation into dense stem cell-rich clusters through the process of spinodal decomposition [107]. Without stem cell division or differentiation, this process would continue leading to the formation of ever-larger clusters.

By contrast, if symmetric CP cell division is biased towards terminal differentiation, stem cells must differentiate to maintain the CP cell population. This has the effect of arresting the growth of stem cell clusters: While small clusters form through the combined effects of stem cell adhesion and proliferation, when clusters become too large, stem cell differentiation leads to their fragmentation. Together, the cell division and differentiation processes impose a characteristic cluster size set by the balance between the depletion of stem cells within a cluster through differentiation, and their self-renewal through symmetric division. As a result of the mutual adhesion of stem cells, newly created CP cells are expelled from clusters. Combined with the effect of upward migration of post-mitotic cells, we will show that this exclusion leads to an effective repulsion between neighbouring clusters leading to large-scale pattern formation (section 6.4).

In summary, the rules of cell division and differentiation provide a mechanism for stem cells to aggregate into stable cohesive clusters that is robust to changes in the rates of division, differentiation and migration. This architecture provides a means to regulate stem cell division by allowing the majority of stem cells (contained within clusters) to remain quiescent. In the normal (steady-state) system, epidermal homeostasis is maintained predominantly through the turnover of CP cells with a small contribution from the slow turnover of stem cells on the boundaries of stem cell-rich clusters and their differentiation. Disruption of tissue serves to activate the stem cell population until the spatial pattern is restored.

However, although these qualitative properties of the tissue (the spatial organisation of cells and their activity) can be inferred from simple arguments, the power of the stem/CP model lies in its potential to provide new predictive insights into cell behaviour: How sensitive is the large-scale

behaviour of the stem and CP cell populations (such as the stem cell cluster size) to changes in stem cell behaviour? What predictions can be made about the function of stem and CP cells when tissue is damaged, or driven far from steady-state (as in culture)? To address issues such as these, one must develop a more formal theory.

6.3 Stem/CP cell hydrodynamic theory

Although the stochastic stem/CP cell model (Fig. 6.1) provides insight into the organisational principle of epidermal homeostasis, its potential range of validity is limited. Further channels of division or differentiation, or different regulatory processes may alter the detailed characteristics of the tissue. To overcome the potentially complex and stochastic behaviour of individual cells, we may develop a predictive theory of cell fate by addressing the coarse-grained cell dynamics, i.e. the dynamics of local basal layer cell densities, $c(r,t)$, averaged over several cell diameters (and measured in units of the cross-sectional area of a typical basal cell). To discriminate between different cell types, the density may be subdivided into the sum, $c(r,t) = c_S(r,t)+c_A(r,t)+c_B(r,t)$, of local stem, (S), committed progenitor (A), and post-mitotic (B) cell densities. Changes in the local cell densities can then be expressed by the continuity equations for each of the three cell types, $\mathrm{X} = \mathrm{S}, \mathrm{A}$ or B,

$$\partial_t c_{\mathrm{X}} = R_{\mathrm{X}} - \nabla \cdot \mathbf{J}_{\mathrm{X}} \tag{6.1}$$

Processes that change the overall cell population give rise to rates, R_{X}, which incorporate the average rates of cell division, differentiation, and upward migration. Changes to the local density may also result from the lateral motion of cells within the basal layer. The resulting redistribution of cell density is associated with a flux $\mathbf{J}_{\mathrm{X}}$. In steady-state, the local cell densities become stationary, $\partial_t c_{\mathrm{X}} = 0$, implying that the local division and differentiation rates are exactly balanced by the flux (see Fig. 6.4). To relate the contributions of the rate and flux terms in (6.1) to the behaviour of stem and CP cells in human epidermis, it is once again helpful to begin by considering a case that is similar but well understood, that of murine IFE that has been discussed extensively in the preceding chapters.

6.3.1 Hydrodynamics of the murine CP cell model

If present, stem cells in mouse tail IFE do not appear to contribute to tissue homeostasis. Instead, CP cells maintain tissue through ongoing division, leading to symmetric or asymmetric cell fate (Fig. 6.1b). Let us therefore begin by considering a tissue where stem cells are altogether absent, $c_{\mathrm{S}}(r,t) = 0$. While CP cell division serves to increase the overall cell density, $c(r,t)$, the migration of post-mitotic cells from the basal layer leads to its reduction (Fig. 6.1c). Together, referring to Fig. 6.1, the processes of progenitor cell division and post-mitotic cell migration translate, respectively, to the rates $R_{\mathrm{A}} = -\Delta c_{\mathrm{A}}$ and $R_{\mathrm{B}} = \Gamma c_{\mathrm{A}} - \gamma_{\mathrm{B}\oslash} c_{\mathrm{B}}$. Here $\Delta = \gamma_{\mathrm{BB}} - \gamma_{\mathrm{AA}}$ represents the effective differentiation rate of CP cells into post-mitotic cells, $\Gamma = 2\gamma_{\mathrm{BB}} + \gamma_{\mathrm{AB}}$ denotes the net rate at which CP cells generate post-mitotic cells, and $\gamma_{\mathrm{B}\oslash}$ is the rate of migration of post-mitotic cells from the basal layer. Where stem cells make no contribution to homeostasis, maintenance of the CP cell population is achieved by balancing the symmetric division rates, $\Delta = 0$.

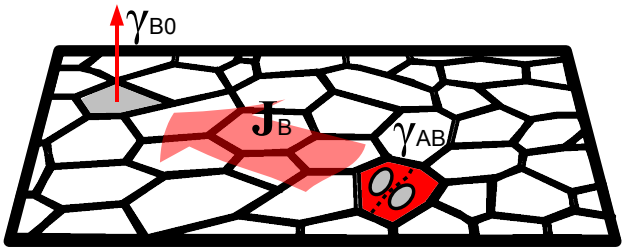


FIG. 6.4: Stratification, proliferation, and lateral displacement in homeostatic epidermis. Schematic of the basal epidermal layer illustrating the steady-state dynamics of Eq. (6.1). To maintain uniform coarse-grained cell density, the upward migration of a post-mitotic cell must be compensated by a cell division in a neighbouring part of the tissue combined with a flow, or flux, of cells. In the schematic, an asymmetric cell division (red) leads to a net flux of post-mitotic cells (wide arrow, $\mathbf{J}_{\mathrm{B}}$) towards the vacancy left by migration (grey). In steady-state, the local division and differentiation rates are exactly balanced by their flux, ($R_{\mathrm{B}} = \nabla \mathbf{J}_{\mathrm{B}}$). Referring to Fig. 6.1, the average rates contributing to the local kinetic term $R_{\mathrm{B}}(r,t)$ are denoted by γ_{AB} and γ_{B0} for the processes of CP cell division and post-mitotic cell migration, respectively.

The observed uniformity of the basal layer cell density requires regulation of the net cell division rate, Γ. As mentioned above, one approach to incorporating this constraint is to ensure that cells stop dividing when the local cell density becomes too high. Such regulation can be enforced by defining the local division rate, $\Gamma(r,t) = [1 - c(r,t)]\Gamma^*$, where Γ^* denotes the bare rate. As the total density approaches unity, the division rate shuts down. Although one can not exclude further nonlinear dependencies of the division rate on cell density, $c(r,t)$, their presence would not affect the qualitative behaviour. Moreover, this form of regulation is consistent with recent studies of murine tail IFE, where the observation that (genetically-labelled) clones remain highly cohesive rules out long-distance lateral cell motion that would be required to maintain density in an unregulated environment [17].

Several mechanisms are responsible for slow lateral migration of cell density (see appendix 6.B): firstly, basal layer cells must continuously undergo local rearrangement to ensure that upward migration of post-mitotic cells is compensated by growth and division of progenitor cells (compare this with the rearrangement resulting from hole motion in chapter chapter 5); secondly, variations in daughter cell fate following division leads to a local random rearrangement of cell density (cf. Eq. (5.1)); thirdly, small-scale variations in tissue density give rise to an elastic response inducing motion of nearby cells (again, note the correspondence with rapid hole diffusion in chapter 5). Taken together, at leading order, such processes contribute to a slow diffusive dynamics. To incorporate the both fast (elastic) and slow (rearrangement) diffusion processes, it is useful to define a free energy $F_0[\{c_{\mathrm{X}}\}]$,

$$F_0[\{c_{\mathrm{X}}\}] = \int d\mathbf{r} \left\{ \sum_{\mathrm{X}} c_{\mathrm{X}} \ln c_{\mathrm{X}} + \chi(1-c)\ln(1-c) \right\}, \tag{6.2}$$

from which the flux is obtained as,

$$\mathbf{J}_{\mathrm{X}} = -\sum_{\mathrm{Y}} M_{\mathrm{XY}} \nabla \frac{\delta F_0}{\delta c_{\mathrm{Y}}}.$$

(see appendix 6.B), with the Onsager mobility $M_{\mathrm{XY}} = \sigma c_{\mathrm{X}}(\delta_{\mathrm{XY}} - c_{\mathrm{Y}})$. The parameter χ gives the relative contribution of the fast elastic response compared to the slower diffusion processes (with

$\chi \gg 1$). Here the sum over Y=S, A, and B ensures that the different cell types move in contrary directions so that the local cell density remains bound ($c(r,t) < 1$). For simplicity, we will suppose that the mobilities of the constituent CP and post-mitotic cell densities are equal with a value σ. Generalisations to include variations between different cell types, which require terms higher order in density, do not change the qualitative behaviour (see appendix 6.B).

With all elements of the continuity equation (6.1) defined, we are now in a position to explore the large-scale properties of the murine epidermal system and compare them with the known properties of the CP cell model and the experimental system:

- *Murine CP cells follow a critical birth-death process:* As mentioned above, maintenance of the cell population demands that the symmetric CP cell division rates are equal, $\Delta = 0$. An imbalance in these rates would lead rapidly to a disruption of the tissue, with the basal layer cell population becoming extinct ($\Delta > 0$), or saturated by an ever-increasing population of progenitor cells ($\Delta < 0$).
- *Murine IFE admits a stable spatially-uniform mean-field solution:* The diffusion of the CP and post-mitotic cell populations results in a uniform steady-state distribution of the average basal layer cell densities c_{A} and $c_{\mathrm{B}} \approx (1 - c_{\mathrm{A}})$ allowing the cell division and migration rates to be related as $\Gamma \approx (1 - cA)\gamma_{\mathrm{B}\oslash}/c_{\mathrm{A}}$. Even without stem cells, the murine system is capable of maintaining a stable and seemingly uniform solution for the average cell densities.
- *Critical fluctuations are significant in murine epidermal maintenance:* The exponential dependence of the dynamics on the imbalance in division rates, Δ, translates to an extreme sensitivity on local cell density. While the average density of CP and post-mitotic cells indeed remains uniform, their actual values exhibit fluctuations that cannot be captured by the long-range hydrodynamics implied by Eq. (6.1). Such critical fluctuations are a generic feature of theories that exhibit the type of balance found here, and they have been studied extensively in the literature [47]. Indeed, it is such critical fluctuations that are responsible for the hallmark scaling behaviour observed in the clone size distributions of murine IFE (chapters 2 and 3), as well as a weak clustering of the CP cell population (chapter 5).

In summary, when stem cells play no role in normal IFE maintenance, the large-scale dynamics result in a stable, spatially uniform distribution of the average CP and post-mitotic cell densities.

6.3.2 Human epidermis and the stem/CP cell model

Using the murine system as a platform, we are now in a position to address cell dynamics in human epidermis, and the role of the stem cell compartment. Here, the process of slow stem cell differentiation demands the development of a small imbalance in symmetric CP cell division rates (i.e. $\Delta > 0$) to maintain a constant population of CP cells. In this case, referring to Fig. 6.1, stem cell division and differentiation lead to a generalisation of the rate terms with,

$$\begin{aligned} R_{\mathrm{S}} &= (\gamma_{\mathrm{SS}} - \gamma_{\mathrm{SA}})c_{\mathrm{S}}, \\ R_{\mathrm{A}} &= \gamma_{\mathrm{SA}}c_{\mathrm{S}} - \Delta c_{\mathrm{A}}, \\ R_{\mathrm{B}} &= \Gamma c_{\mathrm{A}} - \gamma_{\mathrm{B}\oslash}c_{\mathrm{B}}. \end{aligned}$$

As with the murine system, we will further suppose that the cell division rates are regulated by the local cell density such that $\gamma_{\rm SS}(r,t) = [1 - c(r,t)]\gamma^*_{\rm SS}$ and $\Delta(r,t) = [1 - c(r,t)]\Delta^*$. In the absence of stem cell adhesion, it is then straightforward to show that the continuity equations exhibit a stable, spatially uniform, steady-state solution for the stem, CP and post-mitotic cell densities (this solution will be given below in section 6.4). Moreover, in this case, as a result of the imbalance in the symmetric CP division rates, the stability of the steady-state solution is not compromised by the type of critical fluctuations that influence the long-time behaviour of murine IFE.

To fully specify the dynamics of human IFE, we must finally incorporate the effects of stem cell adhesion (Fig. 6.1e). To do so, it is useful to define a revised free energy $F_0[\{c_{\rm X}\}] \mapsto F[\{c_{\rm X}\}]$, from which the flux is obtained as above, $\mathbf{J}_{\rm X} = -\sum_{\rm Y} M_{\rm XY}\nabla\frac{\delta F}{\delta c_{\rm Y}}$. Both the diffusive cell motion and the effect of stem cell adhesion can be captured by a free energy density of Cahn-Hilliard type,

$$F[\{c_{\rm X}\}] = F_0[\{c_{\rm X}\}] - \frac{J}{2}\int d\mathbf{r}\Big\{c_{\rm S}^2 - \alpha(\nabla c_{\rm S})^2\Big\}, \tag{6.3}$$

Here the diffusive dynamics are augmented by terms that capture the effect of stem cell adhesion [107]. The dimensionless parameter, J, characterises the strength of adhesion (generally $J \gg 1$), and the constant α is a measure of the surface tension of stem cells, i.e. a measure of the smoothness of the boundaries of stem cell clusters. Once again, for simplicity, we have assumed that the different cell populations are characterised by the same mobility, σ.

This completes the definition of the hydrodynamic theory of stem/CP cell dynamics in epidermal maintenance. With all of elements of the continuity equation (6.1) now in place, we are now in a position to explore the large-scale properties of the human epidermal system.

6.4 Properties of the hydrodynamic model in human IFE

To understand how the processes encoded in Eq. (6.1) relate to the principles of epidermal maintenance and pattern formation described in section 6.2.1, one may proceed by isolating separate components of the dynamics. Firstly, noting that the stem cell population evolves on a slow time scale relative to CP and post-mitotic cells ($\gamma_{\rm SA} \ll \Gamma$), then the survival of the cell population demands that the symmetric division rates associated with the CP cell population are nearly equal, viz. $\Delta/\Gamma \ll 1$. Therefore, on short time scales the cell population conforms to the stem/CP model of murine IFE maintenance outlined above, as described in earlier chapters. Secondly, as we will soon show, the adhesion properties of the stem cell compartment lead to a gradual segregation of cells through the development of dense, stem cell-rich, clusters of ever-increasing size through the process of spinodal decomposition [107]. In the absence of stem cell differentiation, this segregation would, in principle, proceed unchecked until phase separation between a stem cell-rich domain and the remaining cells was complete. Crucially, the inclusion of cell differentiation processes ($\gamma_{\rm SA} \neq 0$) has the effect of arresting cluster growth and leads to pattern formation according to a Turing-like mechanism that is reminiscent of patterning in chemically reactive mixtures [108]. For a more quantitative description of this process, we proceed to analyse Eq. (6.1) in more detail.

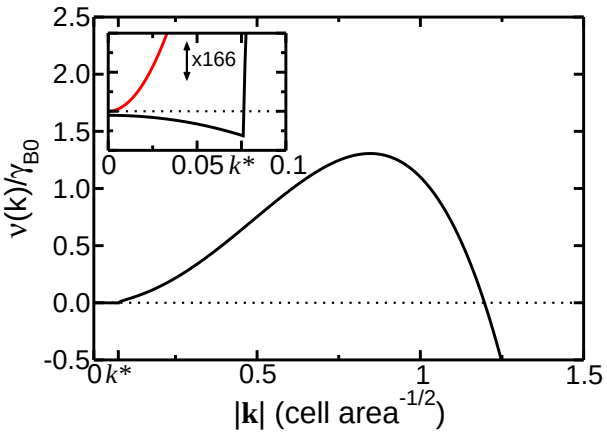


FIG. 6.5: Dispersion relation for the response of the system to fluctuations about the fixed point, showing stability at small k and spinodal-like instability at larger values of k. *Inset:* Focus on the long-wavelength region of the dispersion relation, with vertical axis magnified $\times 166$. The system shows a weak stable response at long wavelengths, in contrast to the spinodal-like case when reactions are turned off (red curve). The wavelength k^* indicates the cross-over to spinodal-like behaviour. Parameters used for plotting: $\Gamma = 200,\ \Delta = 2,\ \gamma_{\mathrm{SA}} = 0.02,\ \gamma_{\mathrm{SS}} = 4,\ \gamma_{\mathrm{B}\varnothing} = 1,\ J = 12,\ \alpha = 0.33$.

6.4.1 Stability and pattern formation

To explore the stability properties of the kinetic theory embodied by Eq. (6.1), one must first identify the fixed points of the dynamics. As well as the trivial fixed points of the empty system, $c = 0$, and the "jammed" system, $c = c_{\mathrm{A}} = 1$[4], Eq. (6.1) admits the uniform steady-state solution $R_{\mathrm{X}} = 0$ with $\bar{c}_{\mathrm{A}} = \frac{\gamma_{\mathrm{SS}}}{\Delta}\bar{c}_{\mathrm{S}}$, $\bar{c}_{\mathrm{B}} = \frac{\Gamma\gamma_{\mathrm{SA}}}{\Delta\gamma_{\mathrm{B}\varnothing}}\bar{c}_{\mathrm{S}}$, and $\bar{c} = \bar{c}_{\mathrm{S}} + \bar{c}_{\mathrm{A}} + \bar{c}_{\mathrm{B}} = 1 - \frac{\gamma_{\mathrm{SA}}}{\gamma_{\mathrm{SS}}}$.

Linearising the mean-field equations in fluctuations $\delta c_{\mathrm{X}} \propto e^{\nu(\mathbf{k})t + i\mathbf{k}\cdot\mathbf{r}}$, one obtains an eigenvalue equation for the growth rate of modes $\mathbf{k}$. Although the formal analytical expressions for the eigenvalues, $\nu(\mathbf{k})$, are unwieldy, the stability properties can be easily inferred for the physical system. In particular, for $\mathbf{k}$ large, the stability behaviour is dictated by the adhesion properties of the stem cell compartment alone. As a result, one finds that $\nu(\mathbf{k}) \simeq \sigma\mathbf{k}^2[J\bar{c}_{\mathrm{S}}(1-\bar{c}_{\mathrm{S}})(1-\alpha\mathbf{k}^2)\mathbf{k}^2 - 1]$ with the most unstable mode set by

$$k_{\max} \simeq \frac{J\bar{c}_{\mathrm{S}}(1-\bar{c}_{\mathrm{S}}) - 1}{2\alpha J\bar{c}_{\mathrm{S}}(1-\bar{c}_{\mathrm{S}})}\,.$$

However, for small $\mathbf{k}$, the stability properties are instead dictated by division/differentiation where the fixed point is stable (see Fig. 6.5). Such a cross-over between reaction-induced stability and adhesion-induced instability appears to be typical for systems involving both reactions and spinodal decomposition [108, 109]. Referring to the schematic in Fig. 6.6, one may qualitatively interpret

[4]One may worry whether the existence of a jammed fixed point compromises the validity of the model, however in practice it is irrelevant: Although the "jammed" fixed point is clearly unphysical, it is unstable with respect to density fluctuations (barring those in n_{S} only). It is therefore irrelevant when focussing on density configurations far from the jammed fixed point.

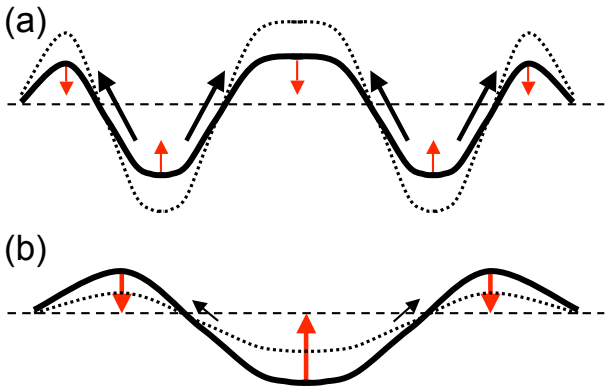


FIG. 6.6: Schematic showing of (in)stability of the homogeneous stem-cell density with respect to fluctuations at **a**, short and **b**, long wavelengths. Adhesive stem cells flow from low-density to high-density regions (black arrows), while cell division and differentiation act to restore the density to its homogeneous value. Starting with an initial fluctuation (solid curve), the system response is shown by the dotted curve. At short wavelengths the flux is large, leading to phase separation, whereas at long wavelengths the system remains stable due to a small flux.

the behaviour as follows: The adhesiveness of stem cells always favours instability at short wavelengths, which leads to spinodal decomposition. However the division/differentiation of cells has the effect of replenishing (diminishing) the local cell populations back towards their equilibrium values. As a result, at large wavelengths, the flux resulting from stem cell adhesion is slower than the division/differentiation activity of the cells, so that patterns do not form. The same processes are responsible for arresting the coarsening as the instability matures.

Next we turn to consider the nature of the adhesion-mediated transition from the homogeneous to the patterned state. In the absence of stem cell division and differentiation, the clustering of adhesive stem cells may be mapped directly onto the Ising model (Kawasaki dynamics). This mapping reveals a first-order phase transition as the adhesion strength is gradually increased[5], with a single stable homogeneous solution when $J < J^{(0)}_{\text{min.}}$, ($J^{(0)}_{\text{min.}} \geq 4$)[6], which becomes unstable when $J > J^{(0)}_c = 1/[\bar{c}_{\text{S}}(1-\bar{c}_{\text{S}})]$. As seen in Fig. 6.7, the inclusion of reactions does not alter the first-order nature of the transition, but merely revises the conditions for phase coexistence: Firstly, by stabilising the long-wavelength solution ($\text{Re}[\nu(k \to 0)] < 0$), the reactions raise the value of the critical adhesiveness required to destabilise the homogeneous solution, viz. $J_c > J^{(0)}_c$, and they consequently shift the critical instability to a finite wavelength $k_{\text{max}} > 0$[7]. Secondly, the fact that the average stem cell density is no longer conserved demands consideration of fluctuations in which the average stem cell density (in some area A), $\langle c_{\text{S}} \rangle = \frac{1}{A}\int_A c_{\text{s}} d^2r$, is modified from the homogeneous fixed-point density $\bar{c}_{\text{S}}$. When the average stem cell density is revised towards the maximally unstable value of $\langle c_{\text{S}} \rangle = 1/2$, then the bare critical adhesion strength is reduced towards its minimum value of $J = 4$, allowing for phase separation to proceed even when the adhesion strength drops below its homogeneous-state critical value. Conversely, depending on the details

[5]In the special case $\bar{c}_{\text{S}} = 1/2$, the transition is second-order by nature.

[6]In the coexistence region $J^{(0)}_{\text{min.}} < J < J_c$, stable stem-cell rich clusters first nucleate with density c^*_{S}, where c^*_{S} and $J^{(0)}_{\text{min.}}$ are given by the simultaneous solution of the equations $F[\bar{c}_{\text{S}}] = F[c^*_{\text{S}}]$, $\frac{\delta F}{\delta \bar{c}_{\text{S}}} = \frac{\delta F}{\delta c^*_{\text{S}}}$.

[7]The numerical values of J_c, k_{max} can be easily obtained by solving the conditions for criticality $\nu = 0$, $\frac{d\nu}{dk} = 0$.

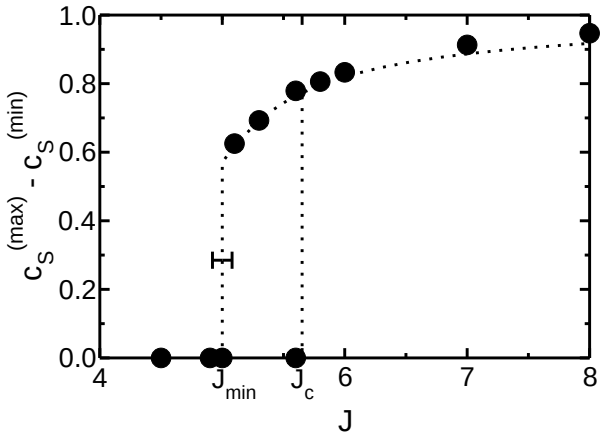


FIG. 6.7: A first-order transition occurs from a homogeneous to an ordered state as the stem cell adhesion strength J is varied. The plot shows the difference in stem-cell concentration between stem cell-rich and depleted domains $m = c_{\rm S}^{\rm max} - c_{\rm S}^{\rm min}$ as a function of J. For $J > J_c$, the spatially-homogeneous state is unstable, so that the only solutions allowed contain periodic stem cell clusters ($m > 0$). For $J < J_{\rm min.}$, only the homogeneous fixed point solution is stable ($m = 0$). In the nucleation (or co-existence) interval $J_{\rm min} < J < J_c$, both the homogeneous and the patterned solutions are stable. Data points show results obtained by numerically integrating Eq. (6.1) as described in section 6.4.2. Dotted curves are included as a guide to the eye. The parameters used for the simulations are $\sigma = 1$, $\alpha = 0.33$, $\gamma_{\rm SA} = 0.01$, $\gamma_{\rm B\varnothing} = 0.5$, $\Gamma = 200$, $\gamma_{\rm SS} = \Delta = 2.0$, corresponding to the fixed point $2\bar{c}_{\rm S} = 2\bar{c}_{\rm A} = \bar{c}_{\rm B} = 0.4975$, and the critical point value $J_c = 5.652$. The the nucleation transition is estimated from the numerical solution to occur at $J_{\rm min} = 5.0 \pm 0.05$, and $c_{\rm S}^* \simeq 0.57$, which varies from the transition values $J_{\rm min}^{(0)} = 4.4$, $c_{\rm S}^* = 0.75$ evaluated for the reaction-free system for the same value of $\bar{c}_{\rm S}$.

of the cell kinetics, the average stem cell density may change away from $\langle c_{\rm S} \rangle = 1/2$, so that $J_{\rm min.}$ grows (see Fig. 6.7). Applied to the experimental system, the robustness of stem cell adhesion observed in experiment suggests that $J \gg 1$, so that the possibility of phase coexistence is largely academic.

To complete our discussion of the instability, it is insightful to compare the long-time behaviour of the system with that of a system undergoing spinodal decomposition in the absence of reactions. For the latter, the short-time instability gives rise to long-time coarsening of the phase-separated domains, in which the domain size scales with time. By contrast, consider the effect of reactions deep within a large phase-separated domain. In this phase-separated region, where the local curvature of the densities $c_{\rm X}$ is small, the cell populations evolve according to the behaviour of the homogeneous system for which the attractive fixed point is $(\bar{c}_{\rm S}, \bar{c}_{\rm A}, \bar{c}_{\rm B})$. It follows that the system is incapable of fully separating into large-scale, locally-uniform domains of stem cells and non-stem cells. If we rule out the existence of a limit cycle, then we are led to conclude that the only possible steady-state solution is one that accommodates some form of spatial modulation. This mechanism for pattern formation is illustrated in Fig. 6.8. Here, the evolution of the stem-cell and A-cell density is shown as a set of flow lines in the phase diagram. For clarity, the density of B-cells and lattice vacancies are not shown. Irrespective of the initial homogeneous density, one may see that the flow lines (dashed) always flow to the fixed point. However, we have seen that the adhesiveness of stem cells destabilises the homogeneous system, instead driving phase separation into the metastable region

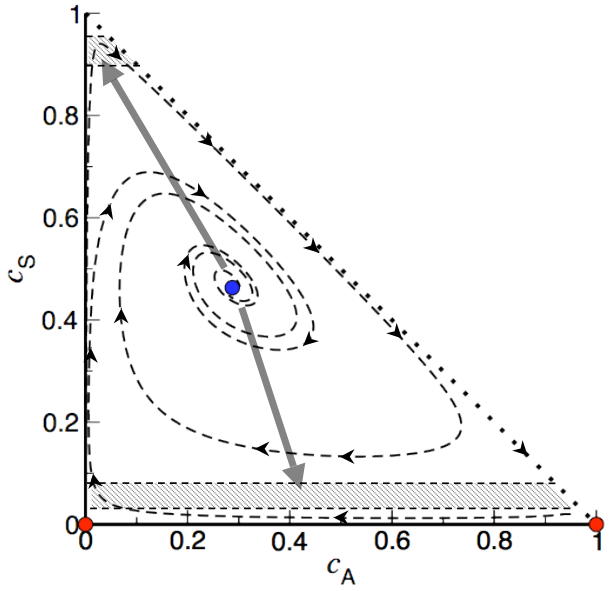


FIG. 6.8: The state diagram for the homogeneous-density solution to Eq. (6.1), projected onto the (c_S, c_A) plane. The system exhibits two unstable fixed points (red) and one stable fixed point (blue). Spatial fluctuations induce phase separation (grey arrows) into the metastable regions (hashed), which are themselves incapable of sustaining an extended homogeneous domain. One is led to conclude that the system is characterised by a spatially-modulated steady-state (see main text).

of the phase diagram (hashed regions), which flow, in return, back to the fixed point. The nature of this steady-state is the focus of sections 6.4.2–6.4.3.

6.4.2 Interlude: Numerical integration of a 4-th order reaction-diffusion equation

To verify the analysis of the mean-field coarse-grained model, we numerically integrated Eq. (6.1) in order to calculate the spatial and temporal evolution of the cell densities c_X. The method used for this numerical integration is described in this technical section, much of which may be skipped at first reading.

Equation (6.1) describes a set of three non-linear coupled 4-th order reaction-diffusion equations. To numerically solve these equations, we discretized the time and space variables in Eq. (6.1) onto a square lattice with space coordinates (i, j) and time coordinate t_n, viz. $c_X(\mathbf{r}, t) \mapsto c_X^{(i,j)}(t_n)$, and then used an explicit second-order Runge-Kutta method to evaluate the density fields at progressive time points [70].

According to this procedure, the discretized dynamical equations acquire the form[8]:

$$c_{\rm X}^{(i,j)}(t_{n+1}) = c_{\rm X}^{(i,j)}(t_n) - [\nabla \cdot \mathbf{J}_{\rm X}(t_n)]_{i,j}\delta t + R_{\rm X}[\{c_{\rm X}^{(i,j)}(t_n)\}]\delta t\,, \tag{6.4}$$

where δt was the variable time increment, and the local rate terms $R_{\rm X}$ are the same as defined above in section 6.3. Spatially periodic boundary conditions were used throughout. To explicitly ensure flux conservation, the flux divergence $\nabla \cdot \mathbf{J}_{\rm X}$ was discretized using a symmetric derivative expansion, $\partial_x c_{\rm X}^{(i,j)} \mapsto \frac{1}{2}(c_{\rm X}^{(i+1,j)} - c_{\rm X}^{(i-1,j)})$, which explicitly satisfies $\sum_{i,j}[\nabla \cdot \mathbf{J}_{\rm X}]_{i,j} = 0$. For example, for $\chi = 1$ and a lattice spacing w, we have

$$[\nabla \cdot \mathbf{J}_{\rm S}]_{i,j} \;=\; \frac{1}{2}\Big[\hat{\Delta}_{i,j}\left(M_{\rm SS}^{(i,j)}\mu_S^{(i,j)}\right) + M_{\rm SS}^{(i,j)}\hat{\Delta}_{i,j}\left(\mu_S^{(i,j)}\right) - \hat{\Delta}_{i,j}\left(M_{\rm SS}^{(i,j)}\right)\mu_S^{(i,j)}\Big] + \sigma\hat{\Delta}_{i,j}\left(c_S^{(i,j)}\right)\,,$$

and,

$$\begin{aligned}
\mu_{\rm S}^{(i,j)} &= -J\left[c_{\rm S}^{(i,j)} + \alpha\hat{\Delta}_{i,j}(c_{\rm S}^{(i,j)})\right]\,, \\
\hat{\Delta}_{i,j}(x) &= \frac{1}{w^2}\sum_{k=1}^{4}\left(x^{(i,j)+\mathbf{e}_k} - x^{(i,j)}\right)\,,
\end{aligned}$$

where $\mu_{\rm S}^{(i,j)}$ is the adhesive component of the chemical potential, $M_{\rm SS}^{(i,j)} = \sigma c_{\rm S}^{(i,j)}\left(1 - c_{\rm S}^{(i,j)}\right)$ is the local stem cell mobility, and $\hat{\Delta}_{i,j}(x)$ is the discrete Laplacian operator, with $\mathbf{e}_{\{k\}} = \{(\pm 1, 0),\ (0, \pm 1)\}$.

Although the explicit (forward) integration scheme shown in Eq. (6.4) suffers from instability when δt and the spatial grid spacing δx are large, it is nonetheless effective (but slow) for sufficiently small space and time intervals [70]. In particular, solutions are stable provided that $\delta x \ll W \sim \sqrt{\alpha}$ (the thickness of a domain wall) and $\delta t \ll \max_{i,j,{\rm X}}[||(\nabla \cdot \mathbf{J}_{\rm X})_{i,j} - R_{\rm X}||/\min[c_{\rm X}, (1 - c_{\rm X})]]^{-1}$. For the solutions presented in Figs. 6.9, 6.10 and 6.13 below, we used $\delta x = 0.3 - 0.4$, and we adapted δt between time steps to satisfy the stability condition. For added accuracy, we implemented Eq. (6.4) to second order in time using the Runge-Kutta method [70].

To observe the dynamics of instability and pattern formation described in the previous section, we applied the numerical integration scheme (6.4) starting with initial conditions $c_{\rm X}^{(i,j)}(t_0) = \bar{c}_{\rm X} + \eta_{\rm X}^{(i,j)}$, corresponding to the homogenous fixed point with added random noise $|\eta_{\rm X}^{(i,j)}| \leq 0.01\bar{c}_X$, $\sum_{\rm X}\eta_{\rm X}^{(i,j)} < 1 - \bar{c}$. An example of the results are shown in Fig. 6.9, where the stem cell density is plotted over a region corresponding to 60×60 cell diameters for a time series that converges upon the patterned steady-state. The figure shows spinodal-like instability at short-times, leading to formation of cohesive stem cell-rich domains. At progressive time points, domain growth is arrested, and the pattern forms as the domains repel each other and undergo ordering.

6.4.3 Properties of the steady-state basal layer

Although linear stability analysis (section 6.4.1) describes the onset of the transition from a uniform to a patterned state, it says little about the morphology of the steady-state. From the numerical

[8] This form is *first-order* in time; it is straightforward to obtain the associated second-order Runge-Kutta equation, see for example Ref. [70].

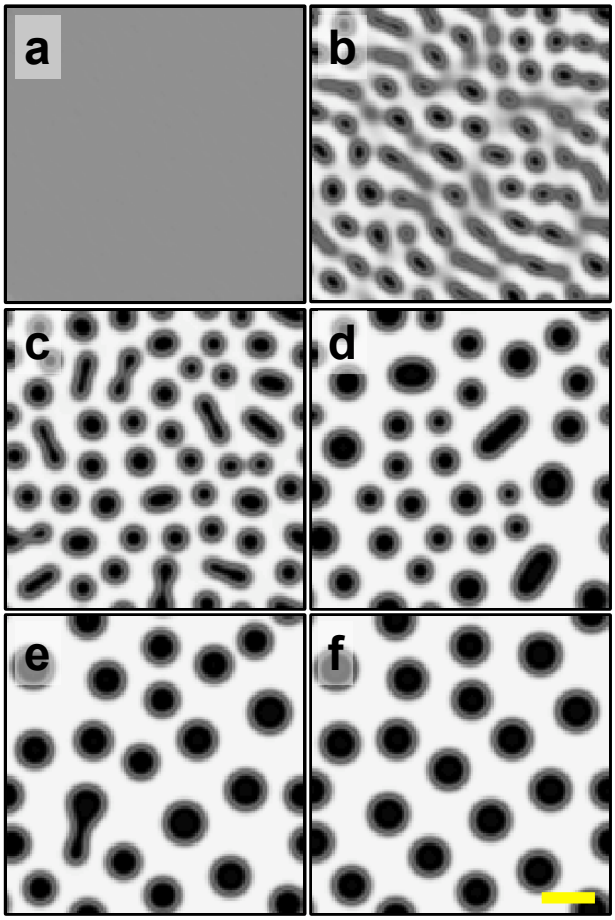


FIG. 6.9: Example of numerical solution to Eq. (6.1), showing the two-dimension stem cell density profile starting from the homogenous fixed point ($\bar{c}_{\mathrm{X}}$) with random fluctuations at $t = 0$, and converging onto the patterned steady-state solution. Frames **a**-**f** correspond to state of the basal layer at times $\gamma_{\mathrm{B}\varnothing}t = 0,\ 3,\ 8,\ 50,\ 350$ and $1,250$, respectively. Bright regions correspond to high stem cell density. The short-time behaviour **a**-**d** is reminiscent of spinodal decomposition of phase-separating mixtures, however the stem-cell population undergoes ordering at long times **e**-**f**. In **f**, the pattern has not yet fully ordered. See caption of Fig. 6.10 for system parameters. Scale bar = 10 cell diameters.

solution described above for the reaction-diffusion equations (6.1), shown in Fig. 6.10, one may see that the steady-state involves a near phase-separated periodic structure of hexagonal lattice symmetry and wavelength Λ. Within each lattice period, the system separates into a homogeneous circular stem cell-rich domain of area $a_{\rm S}$ and density $\bar{c}_{\rm S} \simeq 1$, and a stem cell-depleted region of relative area $a_{\rm AB} = \Lambda^2 - a_{\rm S}$ and density $\bar{c}_{\rm S} \simeq 0$, i.e., in the steady state, stem cell proliferation is limited to the boundary region around the stem cell-rich clusters (as seen directly in Fig. 6.10a).

To maintain each stem cell-rich domain at its steady-state size, the rate of stem cell differentiation within the domain is matched by stem cell creation at the domain wall. Since we are in the deep-quench limit $J \gg 1$, then the width, W, of the domain wall is controlled only by the adhesion properties of stem cells, with $W \simeq \sqrt{\alpha} \sim \mathcal{O}[1]$. As a result, one obtains a natural condition for the total density $\bar{c}$ inside the stem cell-*depleted* domain,

$$W\sqrt{a_{\rm S}}\Omega\,\gamma_{\rm SS}\,(1-\bar{c}) = a_{\rm S}\,\gamma_{\rm SA}\,, \tag{6.5}$$

where $\Omega = \frac{2\sqrt{\pi}}{W}\int_W dr\, c_{\rm S}(r)[1-c_{\rm S}(r)]$ is a dimensionless number of order unity, which is defined by integrating along a path $\int_W dr$ perpendicular to the domain wall (see Fig. 6.11). Similarly, integrating the system equations (6.1) over the near-uniform stem-cell depleted region up to the domain wall, one obtains the following relations for the remaining effective transition rates,

$$\begin{aligned} a_{\rm AB}\Delta\,\bar{c}_{\rm A}(1-\bar{c}) &\simeq a_{\rm S}\gamma_{\rm SA} \\ \Gamma\,\bar{c}_{\rm A}(1-\bar{c}) &\simeq \gamma_{\rm B\oslash}\bar{c}_{\rm B}\,, \end{aligned} \tag{6.6}$$

where $\bar{c}_{\rm A}$ and $\bar{c}_{\rm B}$ ($\bar{c} = \bar{c}_{\rm A} + \bar{c}_{\rm B}$) denote the (constant) densities of progenitor and post-mitotic cells inside the stem cell-depleted domain.

To estimate the size of the stem cell-rich domains, we note that the dynamics within the stem-cell rich regions are dominated by the processes of stem cell differentiation and diffusion. From dimensional analysis, we therefore expect the growth of the stem cell-rich domains to be arrested at a typical size of $a_{\rm S} \propto \sigma/\gamma_{\rm SA}$, which is the size when the type A cells created within the domain begin to destabilise the domain. The experimental observation that clusters have a diameter of 6-9 cells suggests that the timescale of stem cell differentiation $\gamma_{\rm SA}$ is small, viz. $\gamma_{\rm SA}/\sigma \ll 1$. Therefore, the steady-state solution may be associated with a constrained minimum of the free energy $F[\{c\}]$ (i.e. we may treat the system as if it were in equilibrium), from which we obtain the relationship (see section 6.C and Fig. 6.12),

$$a_{\rm S} \simeq \pi W\left(2\sqrt{\frac{\sigma}{\gamma_{\rm SA}}}e^{(1-J)/4} + \frac{W}{2}\right) + 4\pi\frac{\sigma}{\gamma_{\rm SA}}e^{(1-J)/2}\,. \tag{6.7}$$

Taken together Eqs. (6.5), (6.6) and (6.7) characterise key features of the steady-state morphology, giving access to the stem cell cluster size, the periodicity of the pattern $\Lambda^2 = a_{\rm S} + a_{\rm AB}$, as well as the fraction of progenitor cells and post-mitotic cells in the stem-cell depleted regions ($\bar{c}_{\rm A}$, $\bar{c}_{\rm B}$), as shown in Fig. 6.12. For example, a reduction in the ratio $\gamma_{\rm SA}\Gamma/\Delta\gamma_{\rm B\oslash} \simeq a_{\rm AB}/a_{\rm S} + \mathcal{O}[a_{\rm S}^{-1/2}]$ has the capacity to invert the steady-state pattern, leading to the formation of a lattice of stem cell-depleted domains within a sea of stem cells (Fig. 6.13).

Qualitatively, the mechanism of stem-cell patterning is reminiscent of pattern formation predicted for decaying particles in solution [110]. Once clusters are formed, they order into a periodic lattice due to a repulsive interaction that is mediated by the non-stem cell compartment, as shown

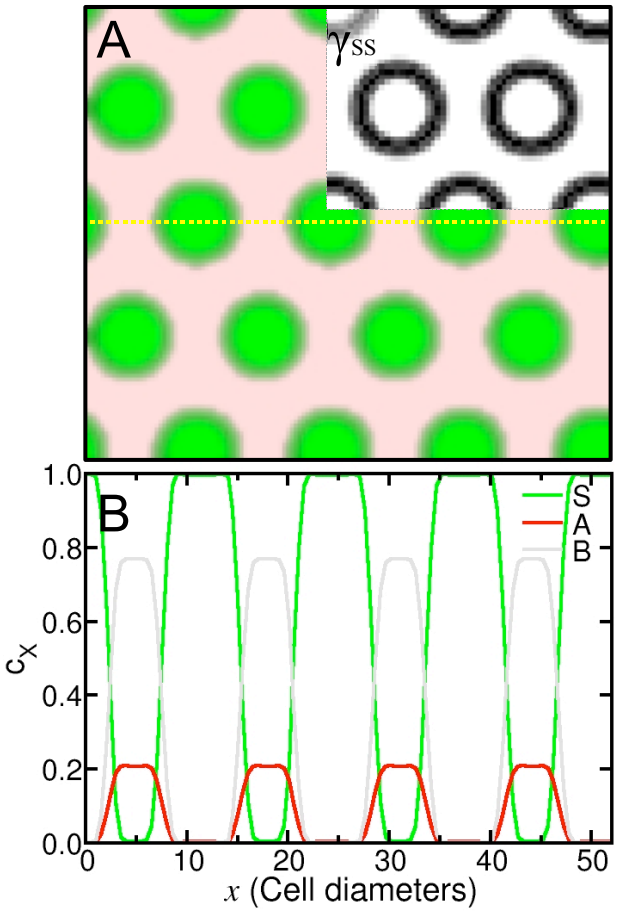


FIG. 6.10: **a** Solution for the stationary stem cell density as obtained from the numerical solution of the hydrodynamic equations (6.1). Stem cells aggregate into a periodic array of dense clusters (green), within a sea of committed progenitor and post-mitotic cells (pink). The proliferative activity of stem cells, as assessed by the local division rate, is shown inset (black) for the same stationary state revealing that stem cells within each cluster are quiescent. The slow creation of new stem cells through division compensates for the loss of stem cells through differentiation in the bulk of the cluster, and gives a natural mechanism for maintaining the cluster size. **b** The variation of local density of stem (S), committed progenitor (A) and post-mitotic (B) cells along the cross-section shown dashed. The results were evaluated by numerically integrating a discrete form of Eq. (6.1) until a steady-state was achieved (see section 6.4.2). To reduce the required computation time, the system was seeded with homogenous cell densities $c_X = \bar{c}_X$ and a hexagonal density modulation corresponding to the wavelength estimated from Eqs. (6.5), (6.6) and (6.7). The results appear to converge onto the same solution starting from random initial conditions (see Fig. 6.9). Parameters used for this calculation are $\sigma = 2,\ J = 12,\ \alpha = 0.33,\ \chi = 1,\ \Gamma = 200,\ \gamma_{SS}/\Gamma = 0.02,\ \Delta/\Gamma = 0.01,\ \gamma_{B\varnothing} = 1$ and $\gamma_{SA} = 0.01$.

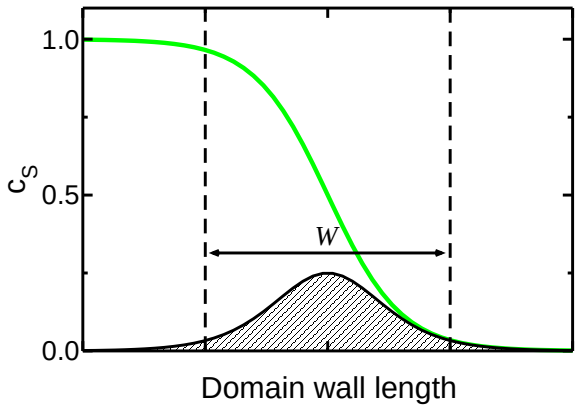


FIG. 6.11: One-dimensional cross-section perpendicular to the domain wall separating stem cell-rich and stem cell-depleted regions. Stem cell division is localised to the boundary region W. The green curve shows the stem cell density, and the effect of stem cell division enters the steady-state equations through the integral $\Omega \sim \mathcal{O}[1]$ (see main text), shown as the shaded area under the black curve.

schematically in Fig. 6.14. In effect, each stem cell-rich domain acts as a source of type A cells, which emerge into the stem cell-depleted region to maintain the total cell density $\bar{c}$. When two stem-cell domains become too close, then the density between the domains is higher than the surroundings, leading to a local decrease in the stem cell division rate $\gamma_{\rm SS} = \gamma^*_{\rm SS}(1-c)$. The resulting gradient in the stem cell division rates biases the creation of new stem cells away from the neighbouring cluster, leading the clusters to drift away from each-other. As a result, the period of time associated with long-range ordering of the mean-field pattern is significantly longer than the effective stem cell division rate.

To verify the steady-state analysis, we applied the numerical integration scheme described in section 6.4.2 with a range of different system parameters, starting from initial conditions that closely mimicked the patterned steady-state predicted from Eqs. (6.5), (6.6) and (6.7). The numerical solution was considered to reflect a steady-state of the system when the resulting pattern was seen to remain stable over an extensive period of time $t \sim 10/\gamma_{\rm SA}$. The parameter dependence of the stable pattern morphology was found to be in good agreement with analytical predictions, as shown by the data points in Fig. 6.12.

6.4.4 Parameter estimation in the stem/CP model

Turning to existing experimental data, such as embodied in the confocal micrographs in Fig. 6.2 and earlier studies, to what extent may we constrain the model parameters?

First, motivated by the *in vivo* murine system, we may estimate the CP cell cycle time to be $\Gamma \simeq 1/\text{week}$. Second, a rough estimate of the non-proliferating fraction of integrin-dull cells in the basal layer as approximately $50-80\%$ (see e.g. Fig. 6.2), allows us to estimate the upward migration

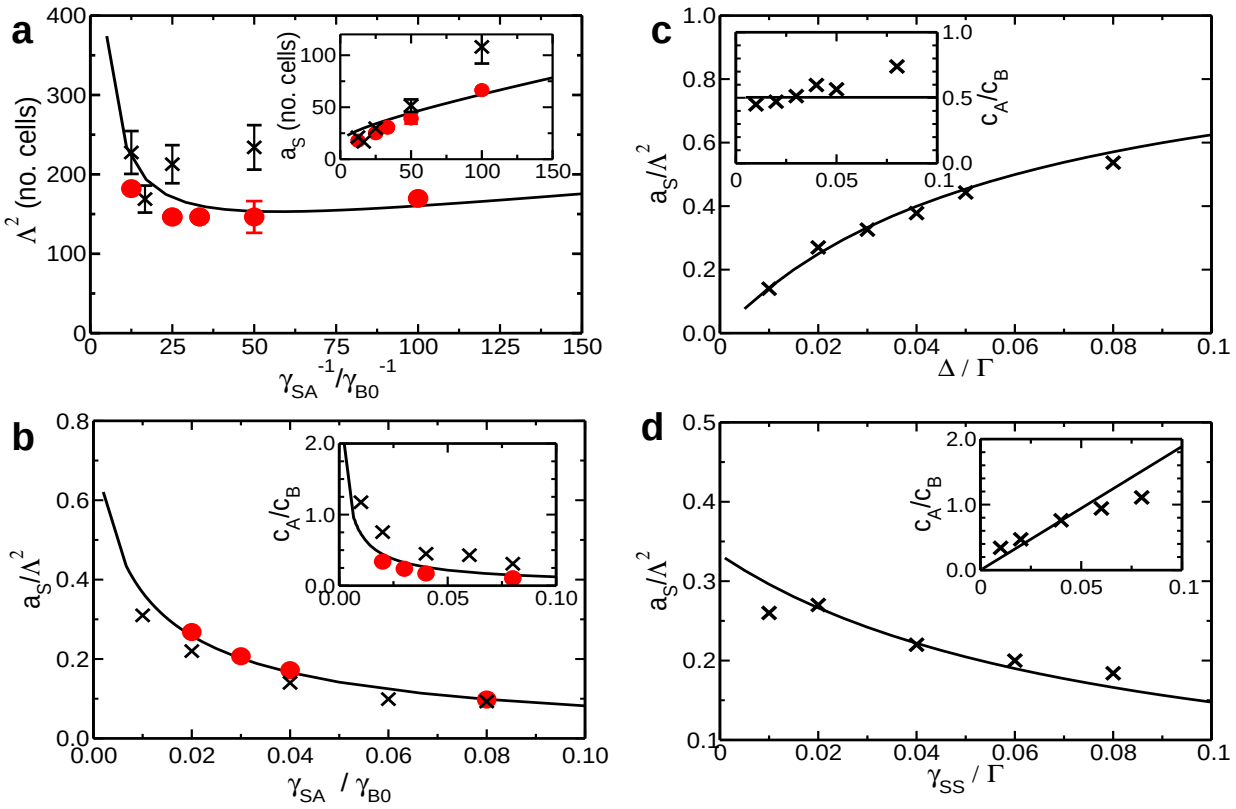


FIG. 6.12: Parameter dependence of the patterned steady-state morphology, demonstrated by plotting **a**, Λ^2, a_S, and **b-d**, a_S/Λ^2, $\bar{c}_A/\bar{c}_B$ against variations in the system parameters γ_{SA} (for **a, b**), Δ (**c**) and γ_{SS} (**d**). The solid curves correspond to the analytic approximations given by Eqs. (6.5), (6.6) and (6.7). The solid data points (red) were obtained from the numerical solution to the system equations (6.1), as described in the main text. Error bars indicate the tolerance of the system to variations in the pattern wavelength, as estimated by varying the initial conditions and then testing the stability of the patterned state. Referring to section 7.1, the crosses ($\times$) indicate the results obtained from cellular automata simulations of a specific manifestation the stem/CP model that is capable of accounting for the effects of fluctuations, as described in the following chapter (7). The plots correspond to the parameter sets described in the captions of Fig. 6.10(for the coarse-grained model) and Fig. 7.1 (for the cellular automata). To evaluate the analytical solution, the values of W and Ω were estimated from the numerical results to be $W = 2.5,\ \Omega = 1.5$.

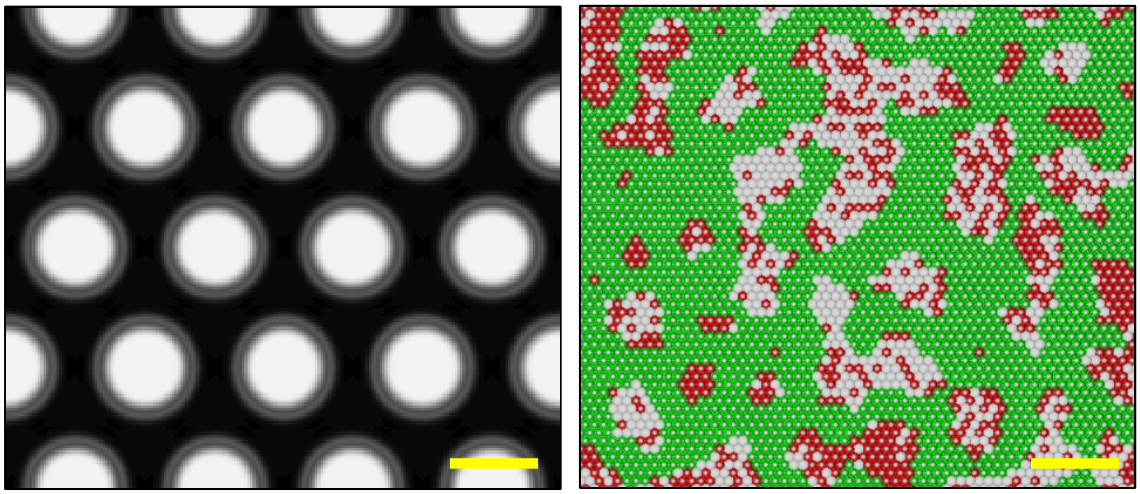


FIG. 6.13: *Left:* Example of steady-state pattern inversion, showing a pattern of stem cell-depleted clusters (bright) in a stem cell-rich background (dark). The figure shows the two-dimensional steady-state stem cell density profile, as obtained from numerical integration of Eq. (6.1) (see caption Fig. 6.10). The growth in stem cell fraction and the resulting pattern inversion occur due to a revision of the model parameters: the same parameters were used as in Fig. 6.10, but setting $\Delta = 4$ and $\gamma_{SA} = 0.01$. *Right:* The steady-state pattern associated with the same parameter set, obtained from cellular automata simulations that capture the effect of fluctuations about the mean-field. The figure shows stem, CP and post-mitotic cells in green, red and white respectively (see section 7.1). One may see that, as a result of their small size, the isolated domains of CP and post-mitotic cells are susceptible to fluctuations that result from jamming by CP cells. This fluctuation effect results in a deviation from the predicted mean-field solution, as may be seen by the increased stem cell fraction compared to the hydrodynamics.

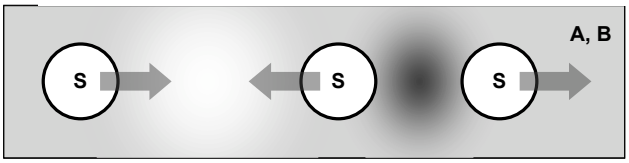


FIG. 6.14: Schematic of stem-cell domain repulsion. Stem cell domains (S) act as uniform sources of A cells, which in turn generate post-mitotic (B) cells to support a uniform cell density in the stem cell-depleted region (grey). When a the distance between stem-cell domains is higher or lower than the equilibrium separation, then the cell density between the domains becomes saturated (dark) or it drops (light), respectively. As a result, the effective density-dependent stem cell division rate $\gamma_{SA} = \gamma^*_{SA}(1 - c)$ at the domain wall becomes non-uniform, which leads to an effective motion of the stem cell clusters back to their equilibrium positions.

rate, $\gamma_{B\oslash} = (\bar{c}_A/\bar{c}_A)\Gamma \simeq 0.25 - 1/\text{week}$. These rates describe properties of the stem-cell depleted region of the epidermis, which dominate the short-time dynamics of epidermal maintenance.

Next, drawing upon previous assessments of the size of $\beta 1$ integrin-bright clusters as 6-9 cells diameter (Fig. 6.2(a,b) and Ref. [101]), one may estimate the ratio between stem cell division and differentiation rates in the steady state to be $\gamma_{SS}/\gamma_{SA} = (W\Omega)/\sqrt{a_S} \sim 5 - 10$. This assessment is in qualitative agreement with that obtained by examining stem cell clusters stained for the second stem cell marker MCSP (Fig. 6.2c and Ref. [98]).

Finally, drawing upon earlier estimates of the basal layer stem cell volume fraction $a_S/\Lambda^2 = 10 - 40\%$ [101], we may infer the ratio of stem and CP differentiation rates, viz. $\Delta/\gamma_{SA} \simeq (\Lambda^2/a_S - 1)/\left(1 + \frac{\bar{c}_A}{\bar{c}_B}\right) \simeq 1 - 10$, revealing that the rate of CP cell differentiation is comparable or one order of magnitude faster than that of stem cells. Although these estimates are crude, they constitute clear (if preliminary) predictions of the cell kinetics, and they may readily be tested and refined in future measurements of the basal layer morphology.

6.5 Chapter summary: Predictions and limitations of the hydrodynamic theory

In summary, we have developed a stem/CP cell model of epidermal maintenance that provides a unified framework in which one may address the maintenance of both murine and human epidermis. This model introduces two minimal ingredients for stem cell behaviour in the IFE, beyond their defining properties of self-renewal and differentiation: We have postulated that self-renewal through symmetric stem cell division is strongly inhibited by the local (coarse-grained) cell density; and that stem cells adhere to one another more strongly than to other cell types, consistent with observations of stem cell aggregation. With these additional ingredients, we have developed a hydrodynamic theory that reveals how stem cells spontaneously create a robust patterned morphology, independent of the detailed rates of cell division, differentiation and migration.

Although this coincidence of experiment and theory is encouraging, the strength of the theory hinges on its predictive power. In particular, the relations derived above between the pattern morphology and the rates of cell division and differentiation suggest a range of possible experiments in homeostatic tissue analogs such as organotypic cultures. One may characterise the role of individual biochemical constituents by studying their effect on the steady-state pattern in organotypic cultures or epidermal xenografts. For example, experiments leading to an increase in the rate of stem cell division relative to differentiation should result in an increase in the size of cohesive stem cell clusters (see Fig. 6.12). Similarly, an increase in the rate of CP cell differentiation, Δ, should result in a corresponding increase in the stem cell fraction.

Yet, despite the predictive power of the hydrodynamics, several open questions remain beyond the scope of the hydrodynamic theory:

- *How do fluctuations affect the dynamics in comparison to the coarse-grained description?* Referring to Fig. 6.2, it appears that the biological patterns do not exhibit long-range ordering

of the type seen for the mean-field hydrodynamic solution. Therefore, to what extent can we indeed apply the hydrodynamic theory to the biological system? To answer this question, in the following chapter we shall estimate the importance of fluctuations by considering the dynamics of a microscopic model of cell fate.

- *What are the implications of the stem/CP model for cell fate in tissue that is perturbed from the steady state?* In practice, the range of possible experimental systems for studying steady-state human epidermis is restricted. On the other hand, it is far easier to conduct experiments on non-steady-state systems such in culture. Therefore, it is important to understand how the current homeostatic theory may be applied in such situations. This will be the focus of section 7.2.

- *What are the implications of the stem/CP cell model for clonal analysis experiments?* Ultimately, perhaps the most direct and reliable method for exploring cell fate in homeostatic tissue is through lineage tracing experiments. Although it is obviously not possible to design controlled *in vivo* lineage tracing experiments in human epidermis, xenografts of human epidermis on mouse offer the potential for quantitative analysis [102]. We will briefly address the question of clonal analysis in section 7.3.

Engaging with these questions will be the focus of the following chapter.

Chapter appendix

6.A Large-scale equivalence of regulatory mechanisms

In this section we address the validity of the coarse-grained form of the cell division rates, which have the general form $\gamma_{\mathrm{XY}} = \gamma^*_{\mathrm{XY}}(1-c)$ during homeostasis. When coarse-grained over an area of several cells (smaller than the typical stem cell cluster size), the intra- and inter-cellular biochemical circuitry may lead, in principle, to regulation of cell division processes that depends on the local areal cell density, but also on intrinsic and extrinsic factors that are independent of the cell density. By definition, however, the latter factors cannot account for the near-uniform cell density in the basal layer, nor for stem cell quiescence within stem cell-rich domains. Therefore, to incorporate the constraint on uniform density and to explain the patterning phenomena, we introduce the coarse-grained division rates, $\gamma_{\mathrm{XY}} = \gamma_{\mathrm{XY}}(c, c_{\mathrm{S}}, c_{\mathrm{A}})$ that incorporate the effect of density-independent processes, while varying explicitly with the local coarse-grained cell densities[9].

It is important to emphasise that a large number of detailed mechanisms may result in density-dependence of the cell kinetics. For example, provided that cell division depends on cell growth, then regions where cell growth is constrained will contain a high density of small cells that are incapable of division. In a more direct example, trans-membrane signalling pathways may allow extrinsic control of cell fate, where the strength and effect of the trans-membrane signals are set by the number and type of neighbouring cells. Such effects have been observed in the context of Delta-Notch signalling [103, 97, 42]. Further examples of density regulation may be found in the stress-related mechanism proposed by Shraiman [83], and also in the topological mechanism proposed by Dubertret and Rivier [36].

Accepting that some form of density regulation exists, we must then address the specific form of density regulation that was introduced in the main text, viz. $\gamma_{\mathrm{XY}} = \gamma^*_{\mathrm{XY}} \times (1-c)$. When the epidermal tissue is in homeostasis, then the cell density is observed to be near-uniform, say $c \approx 1$ corresponding to one cell per unit area. In this case, despite the possibly complex and non-linear form of the density dependence $\gamma_{\mathrm{XY}}(c, c_{\mathrm{S}}, c_{\mathrm{A}})$, we are justified in considering a leading-order (linear) expansion of the density-dependence, $\gamma_{\mathrm{XY}} = \gamma^*_{\mathrm{XY}} \times (1-c) + \sum_{\mathrm{Z}} \gamma^{(\mathrm{Z})}_{\mathrm{XY}} \times (\bar{c}_{\mathrm{Z}} - c_{\mathrm{Z}})$. Applied to the particular case of stem cells, the observation of stem cell quiescence within clusters justifies the expansions $\gamma_{\mathrm{SS}} = \gamma^{(\mathrm{S})}(\bar{c}_{\mathrm{S}} - c_{\mathrm{S}}) + \gamma^*_{\mathrm{SS}}(1-c)$. However, noting that $\bar{c}_{\mathrm{S}} \simeq c = 1$ within the clusters, the effect of both terms would be to allow stem cell division only on the cluster boundaries, and so the full range of known homeostatic behaviour may be described by considering the second term alone. Similar arguments follow for the CP cell compartment; see [17] for a comprehensive discussion of the latter in the context of murine tail-skin maintenance.

To conclude, we present a hypothetical example that demonstrates a possible form of density regulation using a rapidly-diffusing trans-membrane signalling molecule. Consider a molecule that is generated by cells of type X and which remains confined to the basal layer before degrading into inactive constituents. Denoting the concentration of the molecule as x, one may conceive of cell

[9] In principle, one may incorporate non-local effects through the introduction of derivative terms, viz. $\gamma_{\mathrm{XY}} = \gamma_{\mathrm{XY}}(c, c_{\mathrm{S}}, c_{\mathrm{A}}, \nabla c, \nabla c_{\mathrm{S}}, \nabla c_{\mathrm{A}}, \nabla^2 c, \ldots)$.

division processes conditional on the signalling molecule such that $\gamma_{\rm XY} = \gamma_{\rm XY}(x)$. However, the local concentration $x(\mathbf{r},t)$ depends on the detailed position of type X cells within the basal layer at some time t, so that $\gamma_{\rm XY}$ may be considered to be a function of $c_{\rm X}$. To identify the explicit density-dependence of the cell division rate, it is necessary to relate the concentration profile $x(\mathbf{r},t)$ to the density profile $c_{\rm X}(\mathbf{r},t)$. For illustrative purposes, let us assume that a simple reaction-diffusion equation governs the dynamics of the signalling molecule, viz. $\partial x/\partial t = D\nabla^2 x + \alpha c_{\rm X} - \beta x$, where the molecular diffusion rate D greatly exceeds the cell mobility σ, and likewise the rate of creation α and degradation β of the signalling molecule greatly exceed the average rates of cell division and differentiation. In this case, $x(\mathbf{r},t)$ responds instantaneously to changes in the cell concentration profile, so that $\partial x/\partial t = 0$, and the spatial concentration profile of the molecule is $x(\mathbf{r},t) = (\alpha/2\pi D)\int d^2\mathbf{r}' c_{\rm X}(\mathbf{r}',t) K_0(|\mathbf{r}-\mathbf{r}'|/\xi)$, where $K_0(x)$ is the zeroth-order modified Bessel function of the second kind, and $\xi = \sqrt{D/\beta}$ is the RMS diffusion distance over the lifetime of the signalling molecule. Due to the inevitable degradation of the molecule, $K_0(x)$ decays exponentially[10] at distances much larger than ξ, so that the local concentration profile $x(\mathbf{r},t)$ is simply proportional to the density of type X cells, averaged over a radius ξ about $\mathbf{r}$, viz. $x(\mathbf{r},t) \propto \langle c_{\rm X}\rangle_\xi$, where the averaging procedure is $\langle f(\mathbf{r})\rangle_\xi \equiv \xi^{-2}\int d^2\mathbf{r}' f(\mathbf{r}) K_0(|\mathbf{r}-\mathbf{r}'|/\xi)$. However, provided that ξ is much shorter than the typical pattern size, then this average simply mimics the coarse-graining procedure, and $x(\mathbf{r},t) \propto c_{\rm x}(\mathbf{r},t)$. In summary, a rapidly-diffusing signalling molecule may directly relate the local division rate to the local cell density.

6.B Diffusive cell motion in basal layer IFE

There is a considerable literature discussing models of cell diffusion in a range of scenarios, including keratinocyte motion during wound healing processes [40, 111, 112, 113, 114], and the explicit effects of cell-cell adhesion [115]. Conceptually, lateral cell motion in the basal layer may result from (i) external forces exerted on a cell by the surrounding tissue, (ii) rearrangement during division and stratification of nearby cells, and (iii) autonomous cell motion. While the former two are likely to play a significant role in basal layer maintenance, there is no evidence (to our knowledge) that self-mobilisation of keratinocytes contributes significantly to cell motion in confluent tissue. Either way, provided that these mechanisms result in cell displacement that is uncorrelated (at long times) with the displacement of neighbouring cells, then their effect on the coarse-grained cell density is diffusive.

In this context, the aim of the current section is to consider the mechanisms leading to cell displacement (i, ii), in order to motivate the diffusive approximations for the coarse-grained cell densities $c_{\rm X}$. We shall highlight what we believe to be the dominant physics for each of these mechanisms, with a more systematic analysis lying beyond the scope of this work.

To begin, let us denote the local lateral displacement of a coarse-grained section of the basal layer at position $\mathbf{r}$ as $\mathbf{u}(\mathbf{r})$, so that the local velocity is $\mathbf{v}(\mathbf{r}) = \dot{\mathbf{u}}(\mathbf{r})$. We can then express the total cell flux, $\mathbf{J} = \sum_{\rm X}\mathbf{J}_{\rm X}$, in terms of the coarse-grained displacement velocity, viz. $\mathbf{J} = c\,\mathbf{v}$. To account for the effect of external forces exerted upon a cell by the surrounding tissue (i.e. mechanism (i)), it is then a matter of solving for the velocity $\mathbf{v}$ in response to pressure gradients within the tissue. As one may expect such pressure gradients to relax over a period of time much shorter than

[10] To be precise, $\lim_{x\to\infty} K_0(x) = e^{-x}/\sqrt{x}$.

the average cell cycle time, we may access the velocity profile by treating the basal layer tissue as a two-dimensional compressible viscous fluid, for which the dynamics are given by a Navier-Stokes equation. An approach similar to this one is detailed in Refs. [83, 38], in the context of (two-dimensional) wing disc development in Drosophila. However, to avoid entering a technical discussion at this stage, let us merely outline the expected form for $\mathbf{v}$ and $\mathbf{J}$ in confluent tissue: One expects the network of cell-cell interfaces and adhesion to the basement membrane to localise the velocity response and introduce significant damping, so that, approximately, the velocity will depend linearly on the pressure gradient, $\mathbf{v} \propto -\nabla p$, where p is the local two-dimensional pressure in the basal layer. To estimate the pressure, we assume that the volume of basal layer cells is largely uniform, so that variations in the pressure will result from compression that accompanies variations in the cell density c, viz. $p(c) \approx p_0 - p_1(1-c) + \cdots$. Significantly, the cell density remains close to unity at all times, so that only the leading order terms in $p(c)$ are relevant. From here we find $\mathbf{J} \approx \eta\, c\nabla(1-c) \simeq \eta\nabla(1-c)$, which is diffusive with a diffusion constant η. The separation of time-scales of cell division and relaxation implies $\eta \gg \Gamma$. Note that outside of confluence these approximations are no longer valid, and it is not clear whether cell motion results in a simple diffusion.

Next we turn to the effect of cell division and stratification on the local cell densities (ii). Here one must distinguish between two different mechanisms leading to a flux in the coarse-grained cell density: Firstly, referring to Fig. 6.4, the ongoing processes of cell division and upward migration (stratification) give rise to short-ranged pressure gradients that lead to the displacement and re-arrangement of nearby cells in accordance to the fluid dynamics outlined above. Without such motion, tissue confluence would be rapidly disrupted [17]. Secondly, a flux results from the the stochastic fate of daughter cells during CP cell division (Fig. 6.1), and from the random orientation of the mitotic spindle that sets the relative position of the two daughter cells within the basal layer. We shall label this a 'false' flux as it does not arise from the actual displacement of cells within the basal layer, but is instead a consequence of the (assumed) stochastic nature of cell division [17].

For the former mechanism, the combination of division and stratification of nearby cells implies that a given cell is subject to a large number of short-lived density gradients within its lifetime, each resulting in a small displacement (less than or equal to one cell diameter, see schematic in Fig. 6.4), and with a direction that is set by the relative positions of the nearby dividing and stratifying cells. This situation results in an effective Brownian motion of the cells, so that cell motion over sufficiently long periods of time resembles a diffusive random walk. Noting that the maximum displacement of each cell is one cell diameter per division/migration event, we are led to conclude that the magnitude of the effective diffusion constant that results (say β) is of same order as the cell division rate (in units of the average cell area), viz. $\beta \sim \Gamma$.

For the latter mechanism leading to a 'false' flux, the combined effect of symmetric CP cell division into either two CP cells or two post-mitotic cells results in an effective diffusion of the CP cell density, viz. $\partial c_{\mathrm{A}}/\partial t = 2\gamma_{\mathrm{AA}}\nabla^2 c_{\mathrm{A}}$, as may be seen by considering the basal layer as a lattice of non-overlapping cells, see Ref. [17]. Moreover, asymmetric CP cell division into one CP daughter cell and one post-mitotic daughter may result in further effective diffusion of the CP cell density as a result of randomly placing the daughter CP cell on one side of the mitotic spindle or the other [17]. Therefore, to the extent that the fate of the two daughter cells may be considered stochastic, the result of cell division is to randomly mix the cell populations even in the absence of further cell motion, resulting in diffusive dynamics. In the case $\gamma_{\mathrm{AA}} = \gamma_{\mathrm{BB}}$, the combined diffusion constant

can be shown to be Γ (in units of the average cell area) [17].

In summary, we have identified one mechanism for 'fast' diffusion of the overall cell density, which results from short-lived pressure gradients in the basal layer, and two mechanisms for 'slow' diffusion of cells relative to their neighbours as a result of the processes of cell division and stratification. To combine these processes within the framework of a single set of fluxes $\mathbf{J}_\mathrm{X}$ that is capable of accounting for the variation between the diffusion constants (or mobilities) of the different cell types, and to restore the effect of stem cell adhesion, let us define the coarse-grained effective free energy,

$$F = \int d^2\mathbf{r} \left\{ \sum_\mathrm{X} c_\mathrm{X} \ln c_\mathrm{X} + \chi(1-c)\ln(1-c) + \frac{J}{2}\left[c_\mathrm{S}^2 + \alpha\nabla^2 c_\mathrm{S}\right] \right\},$$

where $\chi = \eta/(\beta+\Gamma) \gg 1$ gives the relative magnitude of "fast" diffusion compared to the time-scale of the cell cycle. With this definition, the flux is $\mathbf{J}_\mathrm{X} = -\sum_\mathrm{Y} M_\mathrm{XY}\nabla(\delta F/\delta c_\mathrm{Y})$, where

$$M_\mathrm{XY} = m_\mathrm{X}(c_\mathrm{X})\left(\delta_\mathrm{XY} - \frac{m_\mathrm{Y}(c_\mathrm{Y})}{\sum_\mathrm{Z} m_\mathrm{Z}(c_\mathrm{Z}) + m^*(1-c)}\right)$$

is an Onsager-type mobility constructed from the individual cell mobilities, m_X, and from a mobility of the total density, m^* [116].

Although this construction is sufficiently general to allow the properties of each of the cell types to be addressed individually, the collective patterning phenomenon in which we are interested does not depend on detailed variations between the species. Therefore, for simplicity we shall consider a uniform cell mobility $m_\mathrm{X}(c_\mathrm{X}) = \sigma c_\mathrm{X}$, $m^*(1-c) = \sigma(1-c)$, where $\sigma \simeq \beta + \Gamma$ serves as the effective diffusion constant. With this approximation, M_XY reduces to the simple form of the Onsager mobility $\sigma c_\mathrm{X}(\delta_\mathrm{XY} - c_\mathrm{Y})$. Making one final simplification $\chi = 1$, we find $\mathbf{J}_\mathrm{X} = -\sigma\nabla c_\mathrm{X} +$ stem cell-adhesion terms, as presented in the main text.

6.C Stem cell cluster size

We calculate here the mean-field stem cell cluster size a_S for the steady-state patterned solution of Eq. (6.1). For strong stem cell adhesion and a slow differentiation rate ($\beta J \gg 1$, $\gamma^*_\mathrm{SA}/\sigma \ll 1$), the center of the stem cell-rich domain will be near saturation, $1 - c_\mathrm{S} \ll 1$. Expanding Eq. (6.1) to first order in $(1 - c_\mathrm{S})$ and imposing the steady-state condition $\partial_t c_\mathrm{X} = 0$, we obtain a simple Poisson's equation for the stem-cell density within the domain,

$$\nabla^2 c_\mathrm{S} = \frac{\gamma_\mathrm{SA}}{\sigma} c_\mathrm{S}\,, \tag{6.8}$$

subject to the symmetry condition $\nabla c_\mathrm{S} = 0$ at the center of the domain. One may see that stem cell motion is entirely diffusive far from the domain walls, and $c = c_\mathrm{S} + c_\mathrm{A} = 1$ in the stem cell-rich domain. The radial density profile within the stem cell-rich domain is therefore

$$c_\mathrm{S}(r) = c_\mathrm{S}(R)\frac{I_0\left(\sqrt{\gamma_\mathrm{SA}/\sigma}\, r\right)}{I_0\left(\sqrt{\gamma_\mathrm{SA}/\sigma}\, R\right)}, \tag{6.9}$$

where R is the domain radius (defined such that $a_{\rm S} = \pi R^2 + \frac{1}{2}\times$ domain wall area) and $I_0(x)$ is the zeroth-order modified Bessel function of the first kind. Consistent with the condition $\gamma_{\rm SA}/\sigma \ll 1$, the saturation condition $c_{\rm S}(0) \sim c_{\rm S}(R) \sim 1$ allows us to expand Eq. 6.9, viz. $c_{\rm S}(r) \simeq 1 + \frac{\gamma_{\rm SA}}{4\sigma}(r^2 - R^2)$.

To evaluate the domain radius R, it now remains to minimise the free energy $F[\{c_{\rm X}\}]$ associated with a fixed area of the basal layer. From the definition of the free energy given in the main text, and the observation that $a_{\rm AB}/a_{\rm S} = \Gamma^*\gamma^*_{\rm SA}/\Delta\gamma_{\rm B\oslash} + \mathcal{O}[1/R]$ is constant to leading order in $1/R$ (Eqs. 6.5 and 6.6), we have an average energy density

$$f \sim [F_{\rm S}(R)]/R^2 + f_{\rm AB}\,,$$

where

$$F_{\rm S} \simeq \int_{a_{\rm S}} d^2r \left[c_{\rm S}\ln c_{\rm S} + (1-c_{\rm S})\ln(1-c_{\rm S}) + \frac{J}{2}c_{\rm S}(1-c_{\rm S})\right]$$

is the free energy associated with each stem cell-rich domain (with an added symmetrising factor $-J\,c_{\rm S}/2$, equivalent to shifting the chemical potential by a constant factor $J/2$); $f_{\rm AB}$ is the free energy density in the stem cell-depleted region, which is independent of R. Using the steady-state form of $c_{\rm S}(r)$ as an ansatz, we find that the free energy is minimised when $R \simeq 2\sqrt{\sigma/\gamma_{\rm SA}}e^{(1-\beta J)/4}$. Finally, estimating the stem cell population in the domain wall area to be $2\pi(R + W/2)W \times \frac{1}{2}$, we obtain the result shown in Eq. (6.7).

6.D Channels of stem cell fate

As well as the channels of symmetric stem cell division into two daughter stem cells ($\rm S \to S + S$) and direct stem cell differentiation ($\rm S \to A$), one may conceive of additional channels of stem cell fate including asymmetric division into one stem cell and one CP cell ($\rm S \to S + A$), or symmetric division into two CP cells ($\rm S \to A + A$). We shall denote the coarse-grained average division rates for these channels as $\lambda_{\rm SA}$ and $\lambda_{\rm AA}$ respectively (not to be confused with the rates $\gamma_{\rm SA}$, $\gamma_{\rm AA}$, for direct stem cell differentiation and symmetric CP cell division, see Fig. 6.1).

With these rates, we may now compare the full coarse-grained cell kinetics with the simpler theory presented in the main text. Addressing the stem cell compartment first, the relevant kinetic term becomes $R_{\rm S} = (\gamma'_{\rm SS} - \gamma_{\rm SA})c_{\rm S}$, where the stem cell growth rate is now renormalised to account for the effect of symmetric stem cell differentiation, viz. $\gamma'_{\rm SS} = \gamma_{\rm SS} - \lambda_{\rm AA}$. That is, other than a numerical revision of the division rate, there is no effective change to the coarse-grained dynamics of the stem cell population. Turning next to the CP cell compartment, the new channels of stem cell fate result in a similar revision of the kinetic rates, viz. $R_{\rm A} = (\gamma_{\rm SA} + \lambda'_{\rm SA})c_{\rm S} - \Delta c_{\rm A}$, where the combined rate $\lambda'_{\rm SA} = 2\lambda_{\rm AA} + \lambda_{\rm SA}$ explicitly accounts for the density-dependent rate of stem cell differentiation through division at the cluster boundaries[11]. The kinetic term $R_{\rm B}$ is unchanged.

With these revised cell kinetics, and allowing for the same form of density regulation of the division rates, $\lambda_{\rm XY} = (1-c)\lambda^*_{\rm XY}$, it is then straightforward to re-derive the steady-state relations for the

[11] More generally, the rate $\gamma_{\rm SA}$ is set by processes that lead to direct stem cell differentiation independently of the local density, whereas the density-dependent rate $\lambda'_{SA} = (\lambda'_{\rm SA})^*(1-c)$ reflects the contribution of density-dependent processes. For stable pattern formation we require $\gamma_{\rm SA} > 0$, and $-\gamma_{\rm SA} < (\lambda'_{\rm SA})^*$.

basal layer morphology (Eqs. 6.5, 6.6 and 6.7). One may see that the condition on the average steady-state cluster size $a_{\rm S}$ is unchanged (Eq. 6.7), as it depends on no kinetic rates other than the differentiation rate within the quiescent stem cell clusters ($\gamma_{\rm SA}$). On the other hand, the steady-state condition ($\partial c_{\rm X}/\partial t = 0$) undergoes minor revisions, with Eq. (6.5) unchanged save for the relabelling $\gamma_{SS} \to \gamma'_{SS}$, and Eq. (6.6) gaining an additional term,

$$a_{\rm AB}\Delta\,\bar{c}_{\rm A} \simeq a_{\rm S}\gamma_{\rm SA} + W\sqrt{a_{\rm S}}\Omega\,\lambda'_{\rm SA}\,,$$

reflecting the creation of CP cells at the cluster boundaries through stem cell division (or other density-dependent pathways). However, drawing upon Eq. (6.5), the latter may be rewritten as

$$a_{\rm AB}\Delta\,\bar{c}_{\rm A} \simeq a_{\rm S}\gamma_{\rm SA}\left(1 + \frac{\lambda'_{\rm SA}}{\gamma'_{\rm SS}}\right),$$

which has the same functional form as the original steady-state condition (6.6), but now replacing $\gamma_{\rm SA}$ with $\gamma_{\rm SA}(1+\lambda'_{\rm SA}/\gamma'_{\rm SS})$. Finally, referring to the steady-state relations for the volume fractions $\phi_{\rm X}$ (${\rm X} = {\rm S}, {\rm A}, {\rm B}$) that are given in the main text, it is clear that these do not change barring a revision of the model parameters, viz. $\phi_{\rm A}/\phi_{\rm S} = \gamma_{\rm SA}(1+\lambda'_{SA}/\gamma'_{SS})/\Delta$, and $\sqrt{\phi_{\rm S}} \propto \gamma'_{\rm SS}/\gamma_{\rm SA}$.

Chapter 7

Beyond hydrodynamics: Implications of the stem/CP model

Chapter overview

In the preceding chapter we proposed a universal model of epidermal maintenance in mice and men, involving a cycling population of committed progenitor cells that maintain epidermis by undergoing a stochastic birth-death process, whilst a second population of stem cells remains largely quiescent until activated by spatial disruption of the steady-state tissue.

In this chapter, we look beyond the steady-state hydrodynamics of the proposed stem/CP model, in order to generae further predictions that may be compared with experiment. In particular, we engage with three questions raised in section 6.5: How does the discrete and stochastic nature of the cells affect the dynamics in comparison to the coarse-grained description? What are the implications of the stem/CP model for the fate of cell colonies incubated in culture? What are the implications of the stem/CP cell model for clonal analysis experiments in human epidermal xenografts? These questions are discussed, respectively, in sections 7.1, 7.2 and 7.3. Finally, in section 7.4 we discuss the implications of the stem/CP model for understanding adult tissue maintenance in epidermis and in other tissue types.

The experimental results presented in section 7.2 were generated by DPD[1] and PHJ.

7.1 Microscopic model and fluctuations

An important aspect of the hydrodynamic model developed in section 6.3 is the fact that it identifies the minimal ingredients required to explain the mechanism of large-scale patterning, making minimal assumptions regarding the large number of microscopic interactions that may give rise to

[1] David P. Doupé

the large-scale behaviour. Such a "phenomenological" approach makes contact between mean cell division and differentiation rates and the average features of the steady-state pattern. To establish the degree to which the steady-state characteristics of the basal layer morphology are reliable, one may in principle consider three factors that contribute to fluctuations: First, the discrete nature of the cells (and therefore of the cell density c) compels one to consider small-number (and extinction) statistics. Second, the effectively stochastic nature of the processes of cell division, differentiation and migration, as seen in experiment [16] result in a wide distribution of cell kinetic rates that directly contributes to stochastic fluctuations. Third, a wide range of microscopic processes may contribute to variations in behaviour at the scale of a single cell without affecting the mean-field hydrodynamics. Such effects may include slowly-changing intra-cellular conditions, local inter-cellular interactions, or external factors such as variations in nutrient supplies.

To develop a microscopic model of cell behaviour, one may draw upon a range of different approaches that have been developed to model cell-based processes. These include off-lattice models, in which cells are modeled as quasispherical particles that deform during cell division (e.g. [43]), and space-filling (Voronoi) tesselation models, in which cells are modeled as a 'foam' of convex polygons [117, 118, 119, 120, 121, 38]. These models treat cells as discrete entities capable of growth, division, differentiation, even memory, and that are influenced by the interaction between their internal and external environments.

However, in the absence of further constraints on cell behaviour at the single-cell level, one may conceive of a considerable number of valid microscopic models within the framework of the existing literature — models which are consistent with the large-scale hydrodynamics, but that differ dramatically at the microscopic level. Without further experimental investigation, any microscopic model will be at best over-specified, and at worst uncontrolled. Therefore, rather than proposing a specific microscopic mechanism for cell behaviour that would lie beyond the current experimental resolution, in the following we shall conduct a qualitative study of fluctuations by considering a simple and pedagogical lattice model of cell behaviour, which, while following closely the lattice model described in chapter 5, demonstrates the effects of the discrete and stochastic nature of biological cells on the hydrodynamics.

7.1.1 Lattice model

Referring to the organisational principles described in section 6.2 and to the lattice model introduced in chapter 5 to model murine epidermal maintenance, we shall incorporate the observed near-uniform cell density by modelling the basal layer as a two-dimensional (hexagonal) lattice where each site may host one of the three cell types or it may be vacant ($\oslash$) — note that, without vacancies, cell proliferation and migration becomes impossible. To further maintain a uniform cell density through density-dependent cell division rates, we will suppose that progenitor cell are only capable of division when neighbouring a vacancy. Thus, symmetric and asymmetric division of committed-progenitor compartment may be summarised by the processes

$$\begin{aligned} \mathrm{A} \oslash &\xrightarrow{\gamma_{\mathrm{AA}}} \mathrm{A}\ \mathrm{A} \\ \mathrm{A} \oslash &\xrightarrow{\gamma_{\mathrm{AB}}} \mathrm{A}\ \mathrm{B}\,, \\ \mathrm{A} \oslash &\xrightarrow{\gamma_{\mathrm{BB}}} \mathrm{B}\ \mathrm{B} \end{aligned} \tag{7.1}$$

while post-mitotic cells may migrate out of the basal layer,

$$\mathrm{B} \quad \xrightarrow{\gamma_{\mathrm{B}\oslash}} \quad \oslash\,. \tag{7.2}$$

Here γ_{XY} represents the respective reaction rates, related to the hydrodynamic parameters through the identity $\Gamma = z(\gamma_{\mathrm{AB}} + 2\gamma_{\mathrm{BB}})$, and $\Delta = z(\gamma_{\mathrm{BB}} - \gamma_{\mathrm{AA}})$, where $z = 6$ is the number of nearest neighbours for cells on the lattice.

To accommodate for the different processes of lateral cell mobility, we include the diffusion process,

$$\begin{aligned} \mathrm{A}\ \mathrm{B} \quad &\xrightarrow{\sigma'} \quad \mathrm{B}\ \mathrm{A} \\ \mathrm{X}\ \oslash \quad &\xrightarrow{\chi\sigma'} \quad \oslash\ \mathrm{X}\,. \end{aligned} \tag{7.3}$$

where $\mathrm{X} = \mathrm{A}, \mathrm{B}$, and the hopping rate $\sigma' \propto \sigma/(1 \text{ cell area})$ is related to the hydrodynamic mobility through a constant numerical factor of order unity (for example, $\sigma = \frac{3}{2}\sigma'$ for a hexagonal lattice[2]).

Without stem cells, when $\gamma_{\mathrm{AA}} = \gamma_{\mathrm{BB}}$, processes (7.1), (7.2) and (7.3) describe a critical birth-death process belonging to the Generalised Voter universality class, which recovers the observed population dynamics in murine IFE [17].

Turning to the stem cell compartment, and allowing for symmetric stem cell division or differentiation, the lattice model is related to the hydrodynamics through the processes

$$\begin{aligned} \mathrm{S}\ \oslash \quad &\xrightarrow{\gamma_{\mathrm{SS}}/z} \quad \mathrm{S}\ \mathrm{S} \\ \mathrm{S} \quad\ \ &\xrightarrow{\gamma_{\mathrm{SA}}} \quad \mathrm{A}\,. \end{aligned} \tag{7.4}$$

Finally, turning to the processes of stem cell mobility and clustering, there does not exist (to our knowledge) any microscopic model for which the Cahn-Hilliard dynamics of Eqs. (6.1), (6.3) gives the exact macroscopic description [122]. Nevertheless, the same qualitative behaviour (and much of the quantitative behaviour, see below) is captured by the constrained diffusion processes

$$\begin{aligned} \mathrm{S}_i\ \mathrm{X}_j \quad &\xrightarrow{w_{ij}} \quad \mathrm{X}_i\ \mathrm{S}_j \\ \mathrm{S}_i\ \oslash_j \quad &\xrightarrow{\chi w_{ij}} \quad \oslash_i\ \mathrm{S}_j \end{aligned} \tag{7.5}$$

with the exchange rate w_{ij} reflecting the change in surface energy from the initial state to the target. The exact form of w_{ij} is to some extent arbitrary, but it should satisfy the detailed balance condition: $w_{ij}e^{-\beta E_i} = w_{ji}e^{-\beta E_j}$, where $\beta E_i = \beta J \sum_{k\in \mathrm{N}(i)} n_{\mathrm{S},k}$ is the relative surface energy of the system when the stem cell is at its initial site i, βE_j is the energy evaluated when the stem cell has hopped to its neighbouring site, and w_{ji} is the hopping rate in the reverse direction. Here $\mathrm{N}(i)$ denotes the neighbours of lattice site i and $n_{\mathrm{S},k}$ is the stem cell occupancy at site k. A convenient choice of transition rates satisfying detailed balance is Glauber's hyperbolic tangent rule $w_{ij} = \frac{\sigma'}{2}\{1 - \tanh[\beta(E_j - E_i - J)]\}$. With respect to stem cell motion, this model is simply the dynamic Ising model with conserved particle number (the Kawasaki model) [123].

The lattice model encoded by processes (7.1)-(7.5) constitutes a non-equilibrium process for which the average site occupancies c_{X} for each of the respective cell types are closely related to the

[2]To see this, compare the long wavelength limit of the discrete and continuous diffusion equations.

hydrodynamics in Eq. (6.1). In particular, identifying the vacancy fraction as $c_\oslash = 1 - c$, it is clear that processes (7.1), (7.2) and (7.4) translate directly to the mean-field reaction terms R_X. Equally, near the critical point for stem-cell segregation (e.g. $\beta J \to 1/[\bar{c}_\mathrm{S}(1-\bar{c}_\mathrm{S})]$ for $\chi = 1$), the migration processes (7.3), (7.5) result in a mean-field dynamics that maps directly to the hydrodynamic flux $\mathbf{J}_\mathrm{X}$ [124, 125]. In the deep-quench limit ($\beta J \gg 1$), then the Kawaski dynamics of the stem cell compartment no longer maps rigourously onto Eq. (6.1), however there is no change to the qualitative behaviour of the lattice model, and the long-time coarsening dynamics of the stem cell clusters is quantitatively unaffected [124]. Therefore, by simulating the lattice model as an asynchronous cellular automaton, we are now in a position to assess the importance of fluctuations in determining the steady-state basal layer morphology.

To implement the basal layer lattice simulation, we use a Gillespie algorithm as follows. We define a hexagonal lattice of $N \times N$ sites, starting with each lattice site occupied at random by a cell of S, A or B at $t = 0$. At each time step, we evaluate the stochastic "reaction" rates $r_{i,j}$ associated with each lattice site i and with each process j given by processes (7.1)-(7.5), as determined from the occupancy of each site and of its nearest-neighbours. The next event to occur (viz. division, differentiation, etc.) is then chosen at random, with probability $p_{i,j} = r_{i,j}/\sum_{i,j} r_{i,j}$ for event j to occur at site i, and the time Δt leading up to the event is determined by drawing from an inverse exponential distribution ($\Delta t = \ln(\Theta)/\sum_{i,j} r_{i,j}$, where θ is a uniformly-distributed random variable $\Theta \in \{0, 1\}$). The lattice is then updated and the process repeated using the new lattice configuration.

Using this algorithm, the system was allowed to evolve over a time $t \gg \gamma_\mathrm{SA}$ until a steady-state was reached, as determined by requiring that the average fraction f_X of cells of type X $\in$ S, A and B become stationary[3]. By varying the choice of system parameters, it was possible to study the nature of the steady state associated with the microscopic model.

Results of the lattice simulation are demonstrated in figs. 7.1, 7.2 and 7.3 . In the figures, we see that S cells form obtuse clusters; in between the S-cell clusters there is a region of A and B cells. The A-cells appear to be localised more closely to the S-cell cluster boundaries, however many of them are also dispersed in the inter-cluster region.

Qualitatively, we may conclude that although the long-ranged hexagonal symmetry is lost in the lattice model, the patterning predicted by the hydrodynamics is nonetheless robust with respect to the fluctuations that arise from discrete cell statistics, as well as from the apparently stochastic nature of the events of cell division and differentiation. Strikingly, a comparison of Figs. 6.2 and 7.1 reveals that once fluctuations are accounted for by the simple lattice model, the resulting stem cell patterns indeed resemble the experiment results.

To confirm the validity of the analysis, we studied the variation in the pattern morphology with system parameters, and compared the results to the hydrodynamic theory in Fig. 6.12(b). In the figure, the cellular automata results for the steady-state fraction of stem, CP and post-mitotic cells (a_S/Λ^2, $\bar{c}_\mathrm{A}$ and $\bar{c}_\mathrm{B}$) are shown by the data points ($\times$). One may see that for the same parameter set (see caption of Fig. 7.1 for details), the lattice simulations appear to be in reasonable agreement

[3]More accurately for smaller system sizes, a second condition was added, with the system considered stationary when the total time of evolution was much longer than the coherence time for fluctuations in the the average fraction of cells of type S, A and B.

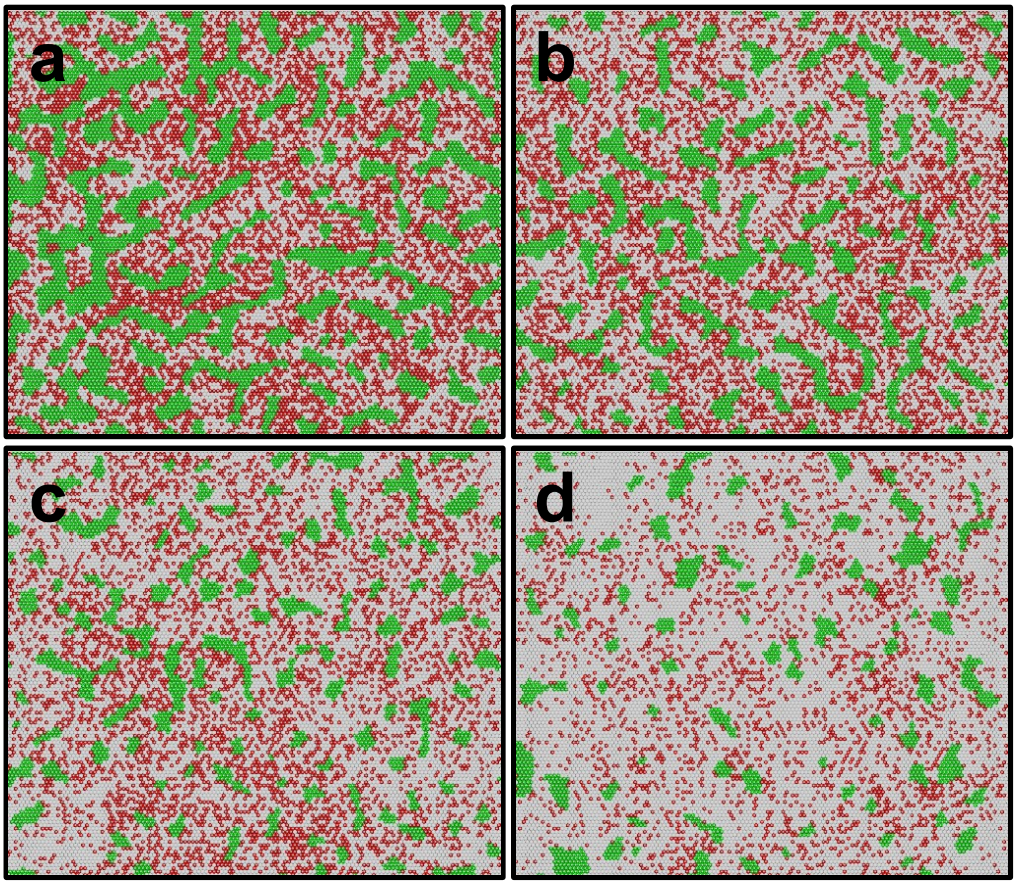


FIG. 7.1: Examples of steady-state basal layer morphology obtained from cellular automata simulations of processes (7.1)-(7.5) on a lattice of 128×128 cells, showing the effects of increasing the stem cell differentiation rate. Panels **a-d** correspond to the values $\gamma_{\rm SA}/\gamma_{\rm B\oslash} = 0.01$ (**a**), 0.02 (**b**), 0.04 (**c**) and 0.08 (**d**). Stem cells (green) form irregular domains within a background of committed progeniter (type A) cells (red) and post-mitotic (type B) cells (grey). The panels correspond to the data points ($\times$) shown in Fig. 6.12(a, b). Parameters used for the simulation are the same as in the caption of Fig. 6.10, except for the following changes that take advantage of the faster simulation time but do not affect the predicted basal layer morphology: The vacancy diffusion rate is restored to the physical limit $\chi = 100(\gg 1)$ and we set $\gamma_{\rm AA}/\Gamma = 0.1$. The 'bare' division rate Γ is increased to $\Gamma = 6000$, while the ratios of $\gamma_{\rm SS}/\Gamma$, Δ/Γ are unchanged. The latter changes have the effect of minimising the vacancy density $(1-c)$ without otherwise altering the dynamics.

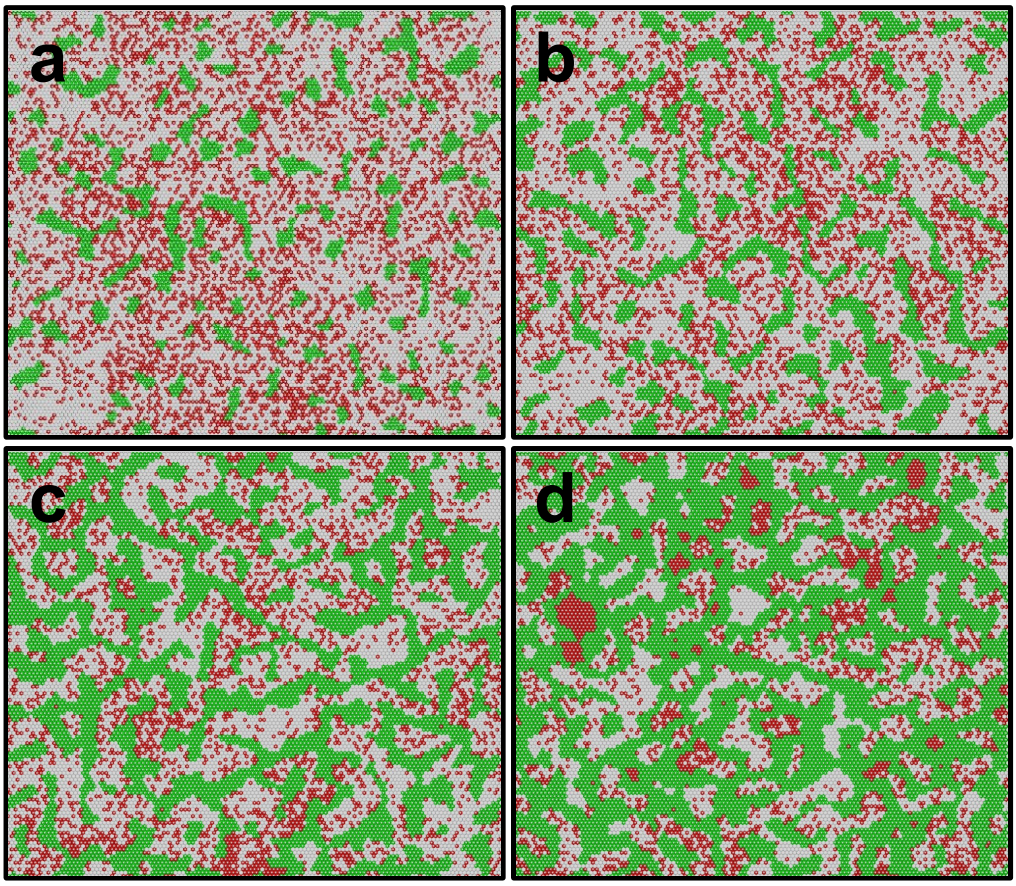


FIG. 7.2: Further examples of steady-state basal layer morphology obtained from cellular automata simulations, showing the effects of increasing the CP cell differentiation rate Δ/Γ, with $\Delta/\Gamma = 0.01$ (**a**), 0.02 (**b**), 0.04 (**c**) and 0.08 (**d**). The panels correspond to the data points ($\times$) shown in Fig. 6.12(c). The stem cell differentiation rate was held constant at $\gamma_{\mathrm{SA}}/\gamma_{\mathrm{B}\oslash} = 0.04$. See caption of Fig. 7.1 for legend and full parameter set.

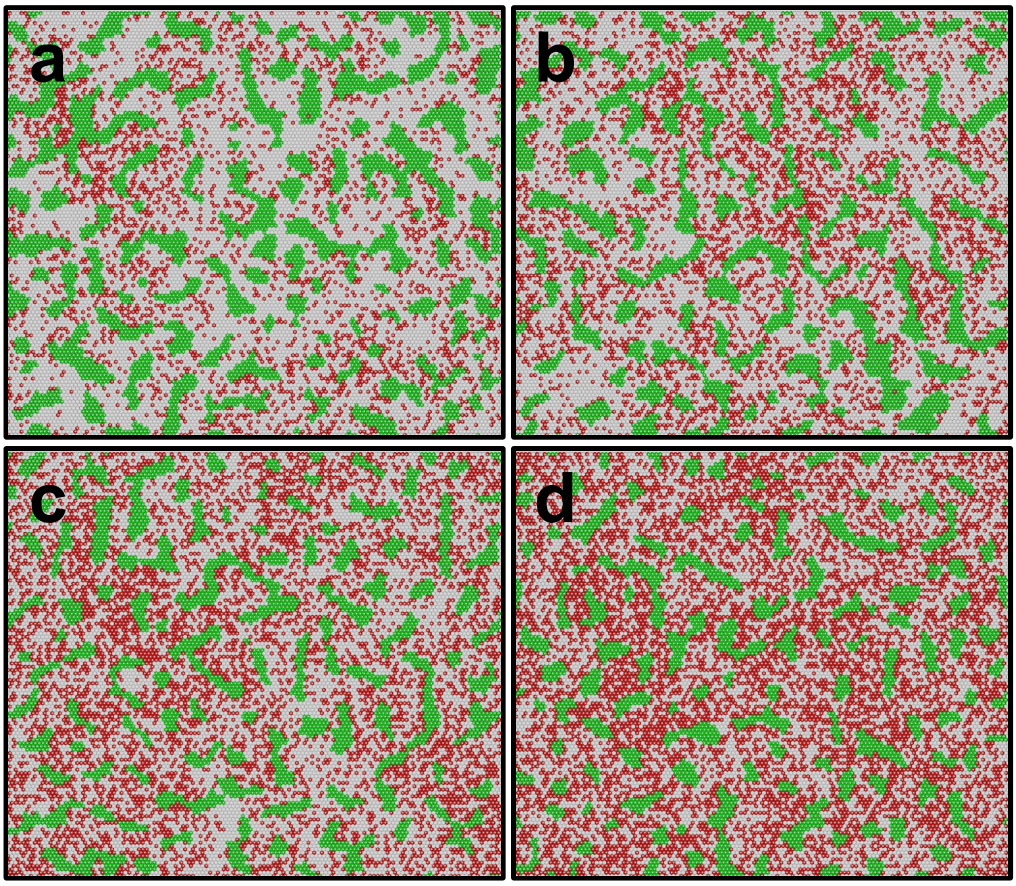


FIG. 7.3: Further examples of steady-state basal layer morphology obtained from cellular automata simulations, showing the effects of increasing the bare stem cell division rate γ_{SS}/Γ, with $\gamma_{SS}/\Gamma = 0.01$ (**a**), 0.02 (**b**), 0.04 (**c**) and 0.08 (**d**). The panels correspond to the data points ($\times$) shown in Fig. 6.12(d). The stem and CP cell differentiation rates were held constant at $\gamma_{SA}/\gamma_{B\varnothing} = 0.04$ and $\Delta/\Gamma = 0.02$ respectively. See caption of Fig. 7.1 for legend and full parameter set.

with the hydrodynamic theory. Significantly, the nature of the response to parametric variations is unchanged.

7.2 The stem/CP cell model in clonal cultures

Although the model, as it is currently embodied in Eq. (6.1), has been designed to describe cell division, differentiation and stratification during homeostasis, the same underlying principles can nonetheless be applied when the tissue is disturbed through injury, or artificially, as in culture. As mentioned earlier, a hallmark of the stem/CP cell model is its ability to restore patterning when driven from steady-state. While this prediction is difficult to study *in vivo*, experiments in culture offer the potential to study both the dynamics of pattern formation within a growing colony, and the developing morphology at a clone edge. This provides a strong motivation to turn to colony growth experiments.

Previous studies have revealed that, when cultured at clonal density, isolated human epidermal stem cells generate large growing colonies, which grow exponentially in area for an initial period lasting several days, before switching to quadratic growth as proliferation becomes localised to the outer-most cell layers of the clone [126]. In the subcloning experiments discussed in section 1.2, the appearance of secondary colonies sheds light on the proliferative potential of the cell that founded the primary clone, with circular holoclones having a higher proliferative potential than irregularly shaped meroclones [24]. How do these observations compare with that expected within stem/CP cell model?

Although one may conceive of both stem and CP cell proliferation on the perimeter, the latter are capable of only linear colony growth and are not expected to contribute to clonal expansion. Indeed, focusing on the rapidly-proliferating clone perimeter, we will show below that the clone edge is sensitive to the accumulation of CP and post-mitotic cells on the perimeter, both of which obstruct stem cells and lead to the *in vitro* growth of smaller clones with a wrinkled morphology. Applying the principles of the stem/CP model in this case, we find that such accumulation of CP and post-mitotic cells occurs when the stem cell capacity for self-renewal is reduced by even a modest fraction (of order 10-20%, see Fig. 7.5). It is striking that such behaviour sheds light on the long-standing puzzle surrounding meroclones and their wrinkled indicator colonies [24].

Yet, beyond a description of the edge morphology in holoclones and meroclones, it is not possible to describe the detailed proliferative activity on the clone perimeter without a further characterisation of the non-steady state behaviour, which is beyond the limits of the hydrodynamic theory. Instead, one may focus on the clone centre, where cell confluence restricts division. Here, provided that the rules embodied by Eq. (6.1) again become valid, one would expect the colony centre to contain clusters of quiescent stem cells within an active CP cell and post-mitotic population. Below (section 7.2.2) we report on preliminary experiments that explore the dynamics of pattern formation within the clone centre.

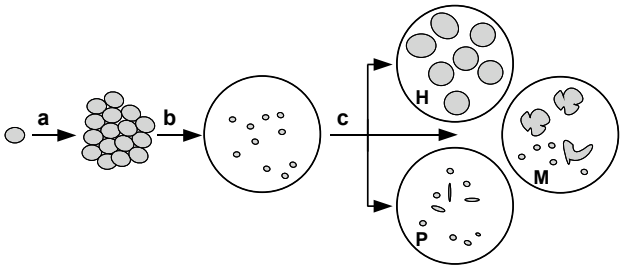


FIG. 7.4: Schematic of experiment used to define cell proliferative capacity (after Barrandon and Green [24]). Single keratinocytes are cultured at clonal density for seven days **a**, giving rise to an exponentially growing cell colony. The cells are then transferred to indicator dishes at clonal density **b**, where they are allowed to grow for a further twelve days **a**. The fraction of terminally-differentiated colonies ($f_{\rm TD}$) in the indicator dishes is used to characterise the proliferative capacity of the original cell. Clones with $f_{\rm TD} \leq 0.05$ are designated *holoclones*; these typically give rise to macroscopic and smooth indicator colonies (H). Clones with $f_{\rm TD} = 1$ are designated as *paraclones* (P). The range of clones with $0.05 < f_{\rm TD} < 1$ are designated as *meroclones*, and typically give rise macroscopic indicator colonies that are smaller and wrinkled in shape (M).

7.2.1 Edge morphology of clonal cultures

In subcloning experiments, isolated keratinocytes are cultured for a preliminary period (seven days) and the resulting colonies are counted and then re-plated into indicator dishes for a further twelve days. The experiments reveal a distribution in the size of the seven-day clones, which is reflected in both the fraction ($f_{\rm TD}$) of terminally-differentiated indicator colonies resulting from each clone, and in the edge morphology of the surviving indicator colonies. Let us now use the stem/CP cell model to relate between the proliferative potential of a clone and its observed edge morphology.

Following the classification scheme introduced in Ref. [24], the proliferative potential of the initially-isolated cells is defined in terms of the fraction of progeny ($1 - f_{\rm TD}$) that themselves give rise to macroscopic colonies in the indicator dishes. Holoclones score $f_{\rm TD} < 0.05$, i.e. over 95% of their seven-day progeny give rise to macroscopic clones. Meroclones score $0.05 < f_{\rm TD} < 1$, and paraclones $f_{\rm TD} = 1$. Upon re-plating one finds that holoclone progeny have similarly high proliferative potential, with secondary clones forming smooth and round cell colonies. By contrast, meroclones give rise to smaller (but still macroscopic) colonies, which are characteristically wrinkled in shape. To understand this variation in edge morphology, let us a consider a simple coarse-grained model of clone growth.

Coarse-grained model of clone growth

The observation that the clone radius grows linearly at long times, and that proliferating cells are localised at the cell perimeter [126], implies that either the clone centre becomes devoid of cycling stem cells at late times, or alternatively, that the clone centre converges onto the steady-state for normal tissue maintenance. Either way, to make contact with experimental results describing late-stage colony growth it is sufficient to consider the cell kinetics of the proliferating outer-most layers of the colony. When the different cell types are homogeneously mixed within these outer layers, then the average rate of clone growth is determined by the fraction $f_{\rm X}$ (X = S, A, B) of the

different cell types on the clone perimeter. Denoting the average number of proliferating layers on the clone perimeter as n_l, and the clone radius as R, the dynamical equations for clone growth are

$$\begin{aligned} \lim_{R \gg n_l} \frac{dR}{dt} &= \lambda n_l (f_{\rm S} + f_{\rm A}) \\ \frac{d}{dt}(f_{\rm X}\, R) &= \sum_{\rm Y} r_{\rm XY}\, f_{\rm Y} R - \Omega(t) f_{\rm X} R\,, \end{aligned} \tag{7.6}$$

where $\Omega(t)$ is the rate at which cells are transferred from the proliferating layers into the clone interior. The average kinetic rates are $r_{\rm SY} = \delta_{\rm YS}(\lambda - \gamma_{\rm SA})$, $r_{\rm AS} = \gamma_{\rm SA}$, $r_{\rm AA} = -\Delta$, $r_{\rm AB} = 0$, and $r_{\rm BY} = \delta_{\rm YA}(\lambda + \Delta)$. To avoid confusion with the steady-state system described in the main text, here we have introduced λ to be the average progenitor cell division rate on the clone edge (assumed to be identical for both stem and CP cells)[4]; $\gamma_{\rm SA}$ is again the rate of stem cell differentiation, and $1/\Delta \gg 1/\lambda$ is the average lifetime of a clone of committed progenitor cells *in vitro*. From the condition $\sum_{\rm X} f_{\rm X} = 1$ and Eq. (7.6), one may evaluate the transfer rate

$$\Omega(t) = \lambda(f_{\rm S} + f_{\rm CP})(1 - n_l/R).$$

Analysis of large clones stained for the proliferation marker Ki67[5] reveals that the number of proliferating layers is $n_l \simeq 5 - 10$.

Referring to Eq. (7.6), one may see that the condition for late-stage steady-state growth corresponds to the constraint $f_{\rm X}(t) \equiv \bar{f}_{\rm X} = {\rm const}$. Imposing this condition (with $\Delta \ll \lambda$), one finds the steady-state fractions $\bar{f}_{\rm S} = (1 - \gamma_{\rm SA}/\lambda)^2$, $\bar{f}_{\rm B} = \gamma_{\rm SA}/\lambda$, $\bar{f}_{\rm A} = 1 - \bar{f}_{\rm S} - \bar{f}_{\rm B}$, and the steady-state radial growth rate of the clone is

$$\dot{R} = n_l(\lambda - \gamma_{\rm SA}) = \lambda n_l \sqrt{\bar{f}_{\rm S}}.$$

These results merely reflect the obvious relation that exists between the growth rate and the fraction of stem cells on the clone perimeter. Comparing to the experimental results in Ref. [24], where the late-stage growth of circular colonies proceeds with $\dot{R} = 0.21{\rm mm/day} \simeq 8$ cell layers/day, we see that the the late stage growth is consistent with the values of $\lambda \simeq 1/{\rm day}$ and $n_l \simeq 10$ cell layers obtained in earlier studies [126].

Edge 'wrinkling' as a failure of the coarse-grained dynamics

Although the steady-state relationships give access to the cell division and differentiation rates in large clones, it is important to recall that they derive from a mean-field description of the clone perimeter, which is strictly valid when the stem, CP and post-mitotic cells are uniformly distributed within the n_l proliferating layers. Should the mean-field description fail through fluctuations of the stem cell fraction $f_{\rm S}$ on the perimeter, then one might expect the clone to become non-circular and irregular due to variations in the radial growth rate (c.f. Eq. 7.6).

To assess the importance of fluctuations, one may compare the average survival time of an individual stem cell within the proliferating layer $t_s = 1/(\gamma_{\rm SA} + \Omega)$, with the 'homogenisation' time t_h, corresponding to the time required for a stem cell to diffuse around the perimeter by the average distance separating stem cells. Migration is expected to be diffusive due to random lateral rearrangement upon division of adjacent cells. As we shall now show, the ratio $\chi \equiv t_h/t_s$ serves as a

[4] Direct measurement of clone growth rates in earlier studies reveals that the cell division rate is $\lambda \simeq 1$ day [126].

[5] David P. Doupé and Philip H. Jones, unpublished data.

good measure for the magnitude of fluctuations. When $\chi \ll 1$ fluctuations in $f_{\rm S}$ may be neglected as stem cells diffuse many times the nearest neighbour separation within one lifetime. On the other hand, when $\chi \gg 1$ then stem cells fail to separate from each other between cell divisions, leading to clustering and, significantly, to a failure of the mean-field description. A continuous transition from smooth to 'wrinkled' clones occurs when $\chi \sim 1$.

To assess the homogenisation time t_h, we note that the mean lateral stem cell separation within the perimeter region is $(1 - f_{\rm S})/(n_l f_{\rm S})$. Drawing upon the literature of non-equilibrium lattice processes (see section 6.B and Ref. [17]), one may show that diffusion of the mean-field density results from asymmetric division as well as from the combined effect of symmetric division and differentiation [17], so that the mean-field diffusion constant is $\lambda(1 - \sqrt{f_{\rm S}}) = \gamma_{\rm SA}$. Therefore, the homogenisation time is estimated as $t_h \sim (1 - f_{\rm S})^2/(\gamma_{\rm SA} n_l^2 f_{\rm S}^2)$. From here we obtain the estimate for the relative magnitude of fluctuations,

$$\lim_{R \gg n_l} \chi \sim \frac{(1 - f_{\rm S})(1 + \sqrt{f_{\rm S}})}{n_l^2 f_{\rm S}^2} \,. \tag{7.7}$$

Thus, when $f_{\rm S} = 1$, then one may see that fluctuations are irrelevant ($\chi = 0$), whereas fluctuations dominate when $f_{\rm S} \ll 1/n_l$ (giving $\chi \gg 1$). The transition between smooth and 'wrinkled' growth occurs when $f_{\rm S} \sim 1/n_l$.

Referring back to the clone fate experiment, this transition clarifies the relationship that exists between a high fraction of terminally differentiated indicator colonies (corresponding to $\gamma_{\rm SA} \sim \lambda$) and the abundance of 'wrinkled' macroscopic colonies.

Numerical simulations of holoclone vs. meroclone growth

To demonstrate the 'wrinkling' transition, we used a simple lattice automata to simulate the growth of colonies with the same approximate mean-field behaviour as embodied by Eq. (7.6). Cells were allowed to evolve according to the same rules as for a confluent sheet (see main text, caption of Fig. 6.1), with the addition of new rules that allowed cells to divide into vacant sites beyond the edge of the clone, e.g. S $\oslash^{(\rm edge)} \rightarrow$ S S. Due to the constraints imposed by the simple (but artificial) lattice description of the colony, these simulations cannot account for the local compressibility of cells near the clone perimeter, so that in effect only the outer-most layer of cells is proliferating, viz. $n_l = 1$. Nevertheless, the simulated clones succeed in capturing many of the qualitative aspects of clone growth, as shown in Fig. 7.5.

In Fig. 7.5(a), where we have chosen a parameter set consistent with $f_{\rm S} \sim 1$ (see figure caption), one may see the late-stage development of a large and circular clone. Although the lattice simulation results in some roughness of the clone perimeter, the large-scale properties of the simulation appear consistent with the hydrodynamics predicted by Eqs. (6.1) and (7.6). In particular, the clone is circular, with a high stem cell fraction driving growth at the clone perimeter. It is striking to note the difference in morphology between the clone perimeter and its centre, which consists of stem cell clusters interspersed with CP (red) and post-mitotic cells (grey).

Turning to Fig. 7.5(b) with lower stem cell fraction on the perimeter, one may see that, over the same period of growth, the clone is smaller in size with a perimeter that is highly irregular, showing 'wrinkles' on a length scale of many cells. Consistent with the mechanism described above, the

wrinkled morphology is associated with the localisation of clone growth to stem-cell rich domains on the perimeter. Yet, as with the smooth clone, the morphology at the clone centre mimicks that of the steady-state *in vivo* system.

To conclude, we have shown that the ability of stem cells to maintain a high density at the clone perimeter is central in determining the clone morphology, and we have related between the proliferative potential of a stem cell, as embodied by the parameter $\gamma_{\mathrm{SA}}/\lambda$, and the transition from smooth to wrinkled clone morphology. These relations reconcile the stem/CP model with observations reported two decades ago in Ref. [24], and have remained unclear throughout this time. In particular, we are now in a position to understand the correlation between a high fraction of terminally-differentiated indicator clones, which implies larger values of $\gamma_{\mathrm{SA}}/\lambda$, and the growth of macroscopic but wrinkled clones. Conversely, it is clear that a low fraction of terminally differentiated clones will correlate to the growth of large and circular clones.

7.2.2 Patterning in clonal cultures

To consolidate the existing phenomenology of holoclone and meroclone growth with the model of stem/CP behaviour in steady-state conditions, we have looked at colonies seeded at clonal density cultured over a period of twelve days, focusing both on the distribution of stem and progenitor cells both at the clone perimeter and at the clone centre.

Experimental results are shown in Fig. 7.6, where a twelve-day colony has been stained for the stem cell marker MCSP as well as the Ki67 proliferation marker. One may see that the clone edges are indeed characterised by high MCSP expression. Away from the clone edge, MCSP-bright cells are clustered as expected. Significantly, Ki67 staining reveals that MCSP-bright cells on the clone perimeter are in cycle, whereas in the clone centre the Ki67 activity is significantly reduced in MCSP-labelled cells that have form strong bonds with their neighbours. Although there is no evidence of a periodic spatial pattern at twelve days, technical difficulties have not allowed us to maintain the colony to later time points. Nonetheless, these preliminary observations, including the appearance of small, quiescent stem cell clusters, are consistent with the stem/CP model.

7.3 Clonal analysis in human epidermal xenografts

The results reported in previous chapters demonstrate that ultimately, quantitative analysis of genetically-labelled clones populations is perhaps the most powerful method for inferring the laws of cell fate in mammalian tissue. Although it is obviously not possible to design controlled *in vivo* lineage tracing experiments in human epidermis, xenografts of human epidermis on mouse offer the potential for quantitative analysis [102].

Let us first consider the fate of a population of genetically labelled CP cells: According to the stem/CP cell model, the stochastic nature of CP cell division would lead to an increasingly broad distribution in clone size at early times, mimicking that seen in murine IFE (chapter 2). At later times, however, the imbalance in the symmetric CP cell division rates (Δ) would lead to the

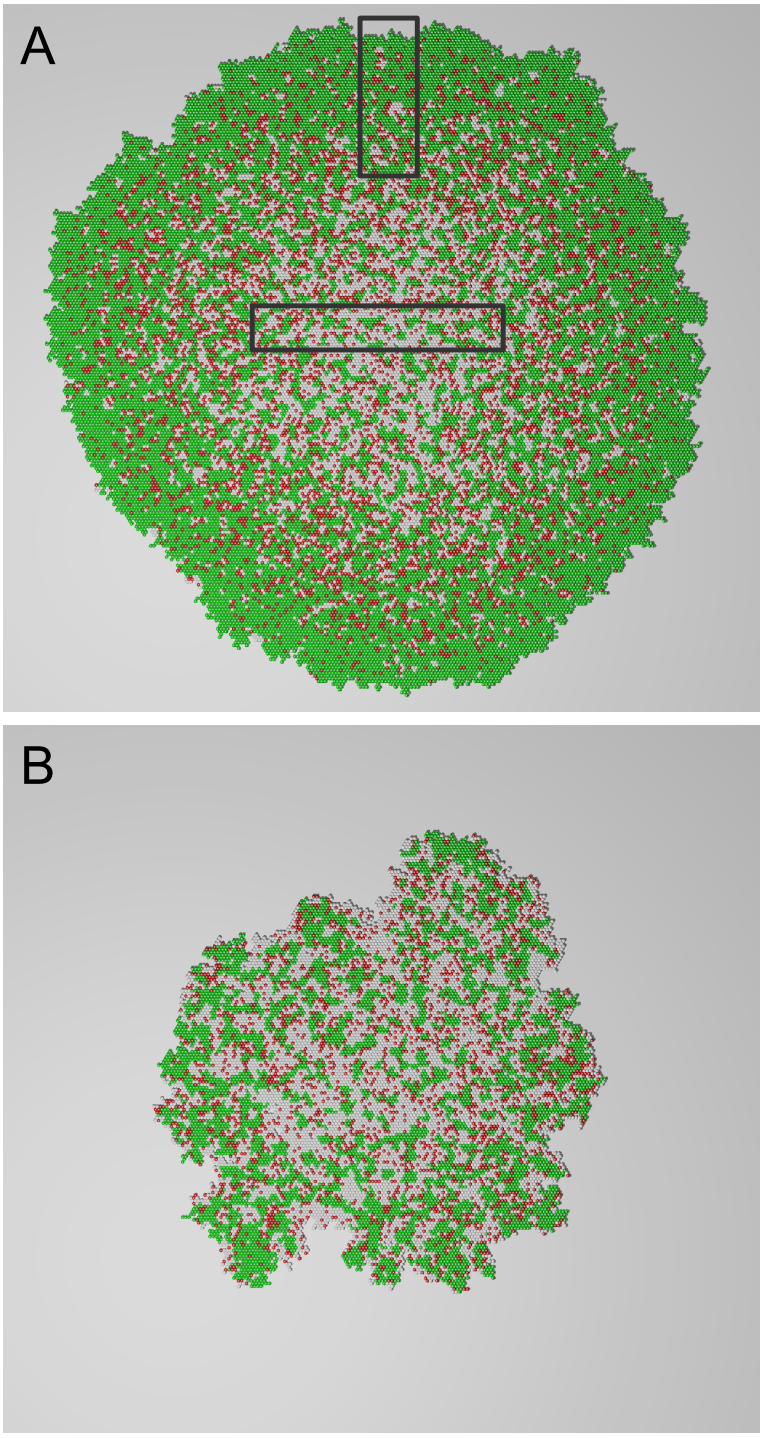


FIG. 7.5: Wrinkled clones result from reduced efficiency of stem cell self-renewal. Cellular automata simulations demonstrate the transition from **a**, smooth and circular clone growth to **b**, irregular and wrinkled development. Starting with a single stem cell, all stem cell divisions on the perimeter give rise to two stem cell daughters in (**a**), whereas in (**b**) 10% of divisions give rise to two CP cells. The two clones are shown at the same time post-seeding. Stem cells are shown in green, CP cells in red and post-mitotic cells in white. In both cases, clone growth proceeds with the same average rate of stem cell division on the clone perimeter. The framed sections in (**a**) compare to the experimental results in Fig. 7.6.

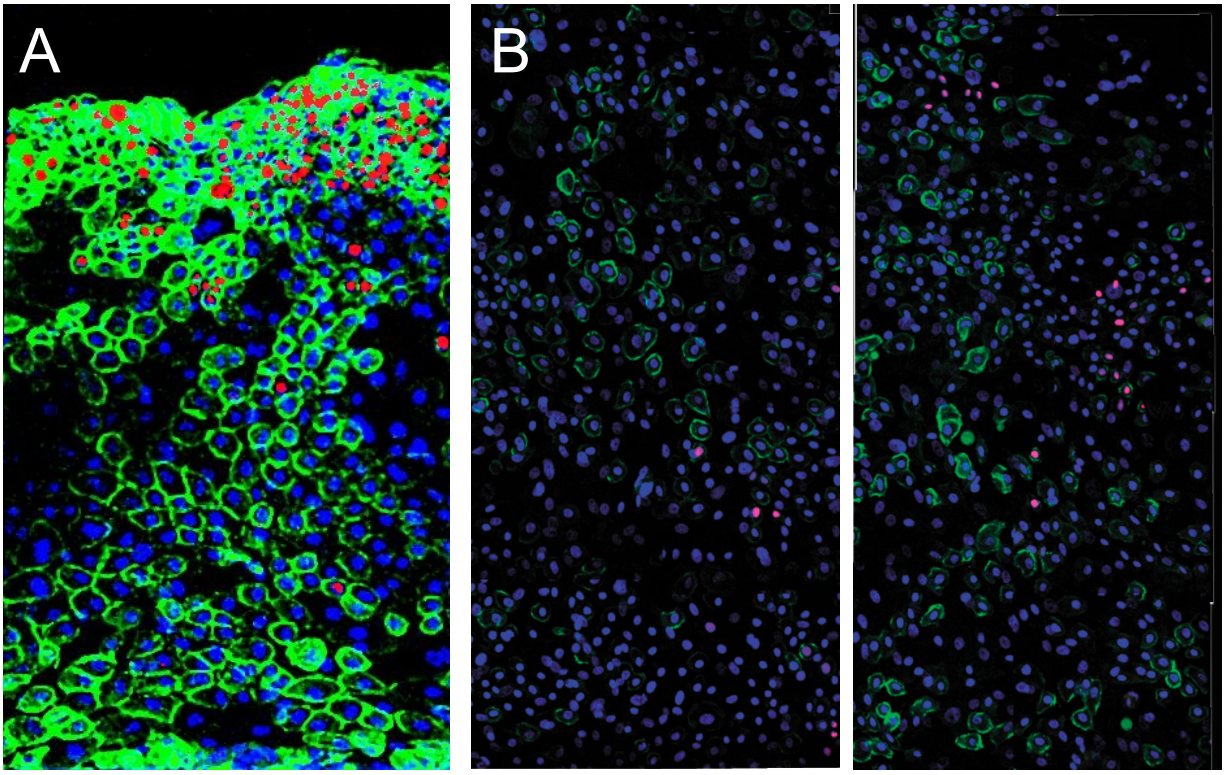


FIG. 7.6: Typical sections of a keratinocyte clone derived from isolated human stem cells *in vitro* at 12 days post-seeding, stained for the stem-cell markers integrin $\alpha 2$ (**a**) and MCSP (**b**), (green), the proliferation marker Ki67 (red), and the nuclear marker DAPI (blue). **a** Typical section showing cells in the vicinity of the clone perimeter. The cell density is high near the clone edge, consisting mainly of integrin-bright cells that show a high level of proliferative activity. By contrast, stem cells appear clustered within the integrin-dull interior, where, consistent with the steady-state theory, proliferation appears on the edges of integrin-bright clusters or in the integrin-dull regions. **b** Two further sections of a keratinocyte clone, showing cells deep within the clone interior. The stem cell marker appears clustered, with Ki67 activity is significantly reduced in MCSP-labelled cells that have form strong bonds with their neighbours.

inevitable extinction of CP cell-derived clones, which are then compensated for by proliferation of the stem cell compartment.

Turning next to a labelled stem cell population, it is clear that, in contrast to the stem/TA hypothesis, the stem cell-derived clones in the stem/CP model do *not* become stationary. Instead, at late times one stem cell must divide symmetrically for each stem cell that differentiates, so that, ultimately, *clone fate adheres to a critical birth-death process.* Thus, drawing upon the analysis in chapter 3, one would expect to observe the hallmark scaling distribution in the number of cells per clone, with the labelled basal layer population consolidating in an ever-decreasing number of clones. However now the characteristic timescale of the birth-death process is $\gamma_{\mathrm{SA}}/a_{\mathrm{S}}$, rather than the symmetric division rate $r\lambda/\rho$ (or $\gamma_{\mathrm{AA}}/\rho$) found in the murine case.

Testing these predictions remains the focus of potential future work.

7.4 Chapter summary: A universal principle for tissue maintenance?

In chapters 6 and 7 we have developed a model of epidermal maintenance that accounts for the wide range of experimental phenomenology in human epidermis, including the observations of stem cell patterning and quiescence, and the variation between smooth and wrinkled clones *in vitro.* The hydrodynamic theory developed in chapter 6 has provided a single framework with which one may address the maintenance of both murine and human epidermis. In both cases, normal adult IFE is largely maintained by a committed progenitor cell population, allowing the majority of stem cells to remain quiescent in homeostatic tissue. The regulatory processes responsible for maintaining stem quiescence allow cells to become active when tissue is disrupted by injury or when cells are transferred to culture.

It is interesting to consider the implications of a model of tissue maintenance by a stochastic committed progenitor compartment, as opposed to the classical, highly regulated stem and TA cell model. Firstly, in the stem/TA model, continual stem cell proliferation is required, exposing stem cells to the risk of oncogenic mutation during DNA replication. The presence of the CP compartment allows stem cell quiescence, avoiding this risk. Secondly, a stochastic CP compartment is far more adaptable to circumstances such as injury than a TA compartment whose behaviour is fixed.

Finally, it is interesting to speculate whether the organising principle applies to other tissues as well as the epidermis: Is homeostasis always maintained by a stochastic progenitor cell compartment, whilst the function of stem cells is tissue repair?

Chapter 8

Conclusions and outlook

8.1 Key lessons

In this work, we have identified a model of tissue maintenance in mouse and human epidermis that identifies the mechanism of steady-state tissue regeneration, defining the roles of stem cells and committed progenitors. The proposed model has a degree of quantitative rigour that is unusual in the field of cell tissue biology, as seen, for example, by the ability of clonal analysis to quantify the division and differentiation rates of cells in the epidermis *in vivo* (chapter 3), and by the ability of the same techniques to predict the effect of a topical drug treatment over the period of a six-month experiment (chapter 4).

Taken in full, the new model challenges many basic concepts of the stem/TA paradigm:

- Stem cells were believed to directly support steady-state tissue maintenance; our results suggest that their role is confined to tissue growth and repair.
- Transit-amplifying cells were believed to age through a limited number of rounds of cell division; our empirical analysis suggests that the fate of progenitor cells is stochastic, with no memory of past events.
- Epidermal tissue was thought to consist of spatially distinct proliferative units containing a well defined number of cells; instead, we have identified that the size distribution of cell families (clones) is non-stationary and is dominated by fluctuations that scale with time.
- It was widely held that stem cells reside in a niche, or micro-environment, that supports their unique function; by analysing the mechanism of collective stem cell patterning in vivo and in vitro, we have found that stem cells are capable of creating their own niche.
- More generally, we have shown that progenitor cells adhere to a set of *simple* rules, which may be end result of rather complex molecular regulation. These rules include the inhibition of stem cell symmetric division by the local cell denisty, and the preferred mutual adhesion of stem cells (see chapter 6).

Thus, in this thesis we have defined a new and predictive "stem/CP" model to replace the traditional qualitative stem/TA hypothesis of epidermal maintenance.

8.2 Outlook: Open questions and new experiments

This thesis has presented a predictive model for cell fate in steady-state epidermal maintenance in mouse and human skin. Using this stem/CP cell model as a platform, one may conceive of three avenues for future investigation:

- *Characterising the molecular circuitry that underlies progenitor cell fate, and the effects of drugs and mutations on steady-state epidermal maintenance:*
  Although the stem/CP cell model makes no reference to the biochemical circuitry underlying cell fate, one may study the role of individual biochemical constituents, or drugs, by altering gene expression (or treating with a drug), and then studying the resulting changes to the experimental phenomenology addressed in this thesis. This approach is exemplified in the study of the effects of all-trans retinoic acid in chapter 4.

  In this context, one specific area of further interest is to characterise the molecular mechanism underlying the *stochastic* nature of CP cell fate. What sets the probabilities for CP cells to give rise to either two differentiated cells, two progenitor cells, or one progenitor and one differentiated cell? In mouse tail-skin these appear to be hard-wired at 8%, 8% and 84% respectively (chapters 2 and 3). Identifying and controlling the origin of this stochastic behaviour has important medical implications.

- *Characterising progenitor cell fate outside of the steady-state:*
  Although the stem/CP model is capable of generating quantitative predictions for clonal analysis in homeostasis, it says little about the behaviour of cells in cases where the tissue is disrupted, for example through injury, cancer growth or during *in vitro* culture growth. Although the absence of steady-state constraints may imply that the analysis of cell fate becomes increasingly complex, it is clear that such conditions are of significant medical and fundamental interest. In particular, being able to relate between the laws of cell fate *in vivo* and *in vitro* may enable more investigations to be carried out in culture.

- *Exploring the universality of the stem/CP model:*
  The stem/CP model predicts that tissue is maintained by a population of CP cells that undergoes division and differentiation according to a stochastic process, with no active contribution from the stem cell compartment. It is interesting to assess the extent to which this model may apply to other tissue types.

References

[1] C. S. Potten and M. Loeffler, Stem cells: attributes, cycles, spirals, pitfalls and uncertainties. Lessons for and from the Crypt, *Development* **110**, 1001 (1990).

[2] H. M. Blau, T. R. Brazelton, and J. M. Weimann, The Evolving Concept of a Stem Cell: Entity or Function?, *Cell* **105**, 829 (2001).

[3] U. Sauer, M. Heinemann, and N. Zamboni, Getting closer to the whole picture, *Science* **316**, 550 (2007).

[4] S. J. Morrison, N. M. Shah, and D. J. Anderson, Regulatory Mechanisms in Stem Cell Biology, *Cell* **88**, 287 (1997).

[5] F. M. Watt and B. L. Hogan, Out of Eden: Stem Cells and Their Niches, *Science* **287**, 1427 (2000).

[6] H. Baharvand, A. Fathi, D. van Hoof, and G. H. Salekdeh, Concise Review: Trends in Stem Cell Proteomics, *Stem Cells* **25**, 1888 (2007).

[7] H. Kitano, Computational systems biology, *Nature* **420**, 206 (2002).

[8] J. Stelling, U. Sauer, Z. Szallasi, F. J. Doyle III, and J. Doyle, Robustness of Cellular Functions, *Cell* **118**, 675 (2004).

[9] D. Noble, The music of life : biology beyond the genome, Oxford University Press, Oxford, UK, 2006.

[10] H. Kitano, Systems biology: A Brief overview, *Science* **295**, 1662 (2002).

[11] M. Ito, Y. Liu, Z. Yang, J. Nguyen, F. Liang, R. J. Morris, and G. Cotsarelis, Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis, *Nat. Med.* **11**, 1351 (2005).

[12] C. Blanpain and E. Fuchs, Epidermal Stem Cells of the Skin, *Annu. Rev. Cell Dev. Biol.* **22**, 339 (2006).

[13] E. Fuchs, Scratching the surface of skin development, *Nature* **445**, 834 (2007).

[14] L. Li and T. Xie, Stem Cell Niche: Structure and Function, *Annu. Rev. Cell Dev. Biol.* **21**, 605 (2005).

[15] E. Clayton, D. P. Doupé, A. M. Klein, D. J. Winton, B. D. Simons, and P. H. Jones, A single type of progenitor cell maintains normal epidermis, *Nature* **446**, 185 (2007).

[16] A. M. Klein, D. P. Doupé, P. H. Jones, and B. D. Simons, Kinetics of cell division in epidermal maintenance, *Phys. Rev. E* **76**, 021910 (2007).

[17] A. M. Klein, D. P. Doupé, P. H. Jones, and B. D. Simons, Mechanism of murine epidermal maintenance: Cell division and the Voter model, *Phys. Rev. E (accepted), arxiv.org/abs/0712.0133* (2008).

[18] A. M. Klein, D. P. Doupé, P. H. Jones, and B. D. Simons, Patterning as a signature of human epidermal stem cell regulation, *To be published* .

[19] P. R. Bergstresser and J. R. Taylor, Epidermal 'turnover time' – a new examination, *British Journal of Dermatology* **96**, 503 (1977).

[20] G. D. Weinstein, J. L. McCullough, and P. Ross, Cell Proliferation in Normal Epidermis, *J Investig Dermatol* **82**, 623 (1984).

[21] J. G. Rheinwald and H. Green, Formation of a keratinizing epithelium in culture by a cloned cell line derived from a teratoma, *Cell* **6**, 317 (1975).

[22] G. G. Gallico, N. E. O'Connor, C. C. Compton, O. Kehinde, and H. Green, Permanent coverage of large burn wounds with autologous cultured human epithelium, *N. Engl. J. Med.* **311**, 448 (1984).

[23] V. Ronfard, J. M. Rives, Y. Neveux, H. Carsin, and Y. Barrandon, Long-term regeneration of human epidermis on third degree burns transplanted with autologous cultured epithelium grown on a fibrin matrix, *Transplantation* **70**, 1588 (2000).

[24] Y. Barrandon and H. Green, Three Clonal Types of Keratinocyte with Different Capacities for Multiplication, *Proc. Nat. Acad. Sci. USA* **84**, 2302 (1987).

[25] C. S. Potten, Cell replacement in epidermis (keratopoiesis) via discrete units of proliferation, *Int Rev Cytol* **69**, 271 (1981).

[26] V. Levy, C. Lindon, B. D. Harfe, and B. A. Morgan, Distinct stem cell populations regenerate the follicle and interfollicular epidermis, *Dev. Cell* **9**, 855 (2005).

[27] I. C. Mackenzie, Relationship between mitosis and the ordered structure of the stratum corneum in mouse epidermis, *Nature* **226**, 653 (1970).

[28] C. S. Potten, The epidermal proliferative unit: the possible role of the central basal cell, *Cell Tissue Kinet.* **7**, 77 (1974).

[29] P. Kaur, Interfollicular Epidermal Stem Cells: Identification, Challenges, Potential, *J. Invest. Derm.* **126**, 1450 (2006).

[30] K. M. Braun, C. Niemann, U. B. Jensen, J. P. Sundberg, V. Silva-Vargas, and F. M. Watt, Manipulation of stem cell proliferation and lineage commitment: visualisation of label-retaining cells in wholemounts of mouse epidermis, *Development* **130**, 5241 (2003).

[31] G. H. Schmidt, M. A. Blount, and B. A. Ponder, Immunochemical demonstration of the clonal organization of chimaeric mouse epidermis, *Development* **100**, 535 (1987).

[32] S. Ghazizadeh and L. B. Taichman, Multiple classes of stem cells in cutaneous epithelium: a lineage analysis of adult mouse skin, *Embo. J.* **20**, 1215 (2001).

[33] T. Kameda, A. Nakata, T. Mizutani, K. Terada, H. Iba, and T. Sugiyama, Analysis of the cellular heterogeneity in the basal layer of mouse ear epidermis: an approach from partial decomposition in vitro and retroviral cell marking in vivo, *Exp Cell Res* **283**, 167 (2003).

[34] S. Ro and B. Rannala, A stop-EGFP transgenic mouse to detect clonal cell lineages generated by mutation, *EMBO Rep.* **5**, 914 (2004).

[35] S. Ro and B. Rannala, Evidence from the stop-EGFP mouse supports a niche-sharing model of epidermal proliferative units, *Exp. Dermatol.* **14**, 838 (2005).

[36] B. Dubertret and N. Rivier, The Renewal of the Epidermis: A Topological Mechanism, *Biophysical Journal* **73**, 38 (1997).

[37] M. C. Gibson, A. B. Patel, R. Nagpal, and N. Perrimon, The emergence of geometric order in proliferating metazoan epithelia, *Nature* **442**, 1038 (2006).

[38] L. Hufnagel, A. A. Teleman, H. Rouault, S. M. Cohen, and B. I. Shraiman, On the mechanism of wing size determination in fly development, *Proc. Nat. Acad. Sci. USA* **104**, 3835 (2007).

[39] N. J. Savill, Mathematical models of hierarchically structured cell populations under equilibrium with application to the epidermis, *Cell Proliferation* **36**, 1 (2003).

[40] J. A. Sherratt and J. D. Murray, Models of epidermal wound healing, *Proc. R. Soc. Lond., Ser. B* **241**, 29 (1990).

[41] P. K. Denman, D. L. S. McElwain, D. G. Harkin, and Z. Upton, Mathematical Modelling of Aerosolised Skin Grafts Incorporating Keratinocyte Clonal Subtypes, *Bulletin of Mathematical Biology* **69**, 157 (2007).

[42] N. J. Savill and J. A. Sherratt, Control of epidermal stem cell clusters by Notch-mediated lateral induction, *Dev. Biology* **258**, 141 (2003).

[43] D. Drasdo, R. Kree, and J. S. McCaskill, Monte-Carlo approach to tissue-cell populations, *Phys. Rev. E* **52**, 6635 (1995).

[44] 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, *Biophysical Journal* **88**, 62 (2005).

[45] M. Block, E. Schöll, and D. Drasdo, Classifying the Expansion Kinetics and Critical Surface Dynamics of Growing Cell Populations, *Phys. Rev. Lett.* **99**, 248101 (2007).

[46] H. Hinrichsen, Non-equilibrium critical phenomena and phase transitions into absorbing states, *Advances in Physics* **49**, 815 (2000).

[47] G. Ódor, Universality classes in nonequilibrium lattice systems, *Rev. Mod. Phys.* **76**, 663 (2004).

[48] L. Alonso and E. Fuchs, Stem cells of the skin epithelium, *Proc. Natl. Acad. Sci. USA* **100 Suppl 1**, 11830 (2003).

[49] L. Lajtha, Stem cell concepts, *Differentiation* **14**, 23 (1979).

[50] K. M. Braun and F. M. Watt, Epidermal label-retaining cells: background and recent applications, *J Investig Dermatol Symp Proc* **9**, 196 (2004).

[51] Unpublished.

[52] G.H. Williams *et al.*, Improved cervical smear assessment using antibodies against proteins that regulate DNA replication, *Proc. Natl. Acad. Sci. USA* **95**, 14932 (1998).

[53] P. Birner *et al.*, Immunohistochemical detection of cell growth fraction in formalin-fixed and paraffin-embedded murine tissue, *Am J Pathol* **158**, 1991 (2001).

[54] T. Das, B. Payer, M. Cayouette, and W. A. Harris, In vivo time-lapse imaging of cell divisions during neurogenesis in the developing zebrafish retina, *Neuron* **37**, 597 (2003).

[55] M. Gho and F. Schweisguth, Frizzled signalling controls orientation of asymmetric sense organ precursor cell divisions in Drosophila, *Nature* **393**, 178 (1998).

[56] T. Lechler and E. Fuchs, Asymmetric cell divisions promote stratification and differentiation of mammalian skin, *Nature* **437**, 275 (2005).

[57] I. H. Smart, Variation in the plane of cell cleavage during the process of stratification in the mouse epidermis, *Br. J. Dermatol.* **82**, 276 (1970).

[58] W. Zhong, J. N. Feder, M. M. Jiang, L. Y. Jan, and Y. N. Jan, Asymmetric localization of a mammalian numb homolog during mouse cortical neurogenesis, *Neuron* **17**, 43 (1996).

[59] I. M. Conboy and T. A. Rando, The regulation of Notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis, *Dev Cell* **3**, 397 (2002).

[60] F. M. Smart and A. R. Venkitaraman, Inhibition of interleukin 7 receptor signaling by antigen receptor assembly, *J Exp Med* **191**, 737 (2000).

[61] S. Temple and M. C. Raff, Clonal analysis of oligodendrocyte development in culture: evidence for a developmental clock that counts cell divisions, *Cell* **44**, 773 (1986).

[62] N. T. J. Bailey, The Elements of Stochastic Processes, J. Wiley & Sons, New York, 1964.

[63] M. Loeffler and H. E. Wichmann, A comprehensive mathematical-model of stem-cell proliferation which reproduces most of the published experimental results, *Cell and tissue kinetics* **13**, 543 (1980).

[64] M. Loeffler and B. Grossmann, A stochastic branching model with formation of subunits applied to the growth of intestinal crypts, *J. Theo. Bio.* **150**, 175 (1991).

[65] R. A. White and N. H. A. Terry, Cell kinetics: Mathematical models and experimental bases, *Math. and Comp. Model.* **32**, 113 (2000).

[66] I. Roeder, Quantitative stem cell biology: computational studies in the hematopoietic system, *Current Opinion in Hematology* **13**, 222 (2006).

[67] D. R. Appleton, N. A. Wright, and P. Dyson, Age distribution of cells in stratified squamous epithelium, *J. Theo. Bio.* **65**, 769 (1977).

[68] C. S. Potten, H. E. Wichmann, and M. Loeffler, Evidence for discrete cell kinetic sub-populations in mouse epidermis based on mathematical-analysis, *Cell and tissue kinetics* **15**, 303 (1982).

[69] H. W. Watson and F. Galton, On the Probability of the Extinction of Families, *Journal of the Anthropological Institute of Great Britain and Ireland* **4**, 138 (1875).

[70] W. H. Press, S. A. Teukolsky, W. Vetterling, and B. P. Flannery, Numerical Recipes in C: The art of Scientific Computing, 2nd ed., Cambridge University Press, Cambridge, UK, 1992.

[71] R. Dover and C. S. Potten, Heterogeneity and cell cycle analyses from time-lapse studies of human keratinocytes in vitro, *J. Cell. Sci.* **89**, 359 (1988).

[72] S. Karlin and H. M. Taylor, A first course in stochastic processes, 2nd ed.,, Academic Press, Inc., San Diego, 1975.

[73] E. R. Fearon, S. R. Hamilton, and B. Vogelstein, Clonal analysis of human colorectal tumors, *Science* **238**, 193 (1987).

[74] E. R. Fearon and B. Vogelstein, A genetic model for colorectal tumorigenesis, *Cell* **61**, 759 (1990).

[75] A. Whittemore and J. B. Keller, Quantitative theories of carcinogenesis, *SIAM Review* **20**, 1 (1978).

[76] D. G. Kendall, Birth-and-death processes, and the theory of carcinogenesis, *Biometrika* **47**, 13 (1960).

[77] B. Houchmandzadeh, Clustering of diffusing organisms, *Phys. Rev. E* **66**, 052902 (2002).

[78] W. R. Young, A. J. Roberts, and G. Stuhne, Reproductive pair correlations and the clustering of organisms, *Nature* **412**, 328 (2001).

[79] T. C. Roosa, F. K. Jugert, H. F. Merk, and D. R. Bickers, Retinoid Metabolism in the Skin, *Pharm. Rev.* **50**, 315 (1998).

[80] J. Lewis and A. Davies, Planar cell polarity in the inner ear: How do hair cells acquire their oriented structure?, *J. Neurobiology* **53**, 190 (2002).

[81] S. J. Gibbs and G. W. Casarett, Spatial Distribution of Cells in Mitotic and DNA-Synthetic Phases of the Cell Cycle in Hamster Cheek Pouch Epithelium, *J. Dent. Res.* **51**, 30 (1970).

[82] I. L. Cameron, D. G. Gosslee, and C. Pilgrim, The Spatial Distribution of Dividing and DNA-Synthesizing Cells in Mouse Epithelium, *J. Cell & Comp. Physiol.* **66**, 431 (1966).

[83] B. I. Shraiman, Mechanical feedback as a possible regulator of tissue growth, *Proc. Nat. Acad. Sci. USA* **102**, 3318 (2005).

[84] T. M. Liggett, Interacting Particle Systems, Springer, New York, 1985.

[85] I. Dornic, H. Chaté, J. Chave, and H. Hinrichsen, Critical Coarsening without Surface Tension: The Universality Class of the Voter Model, *Phys. Rev. Lett.* **87**, 045701 (2001).

[86] O. Al Hammal, H. Chaté, I. Dornic, and M. A. Muñoz, Langevin Description of Critical Phenomena with Two Symmetric Absorbing States, *Phys. Rev. Lett.* **94**, 230601 (2005).

[87] P. L. Krapivisky, Kinetics of monomer-monomer surface catalytic reactions, *Phys. Rev. A* **45**, 1067 (1992).

[88] F. van Wijland, Field theory for reaction-diffusion processes with hard-core particles, *Phys. Rev. E* **63**, 022101 (2001).

[89] S.-C. Park and J.-M. Park, Generating function, path integral representation, and equivalence for stochastic exclusive particle systems, *Phys. Rev. E* **71**, 026113 (2005).

[90] L. Frachebourg and P. L. Krapivisky, Exact results for kinetics of surface catalytic reactions, *Phys. Rev. E* **53**, R3009 (1996).

[91] H. Hinrichsen, Non-equilibrium phase transitions, *Physica A* **369**, 1 (2006).

[92] D. C. Brown and K. C. Gatter, Ki67 protein: the immaculate deception?, *Histopathology* **40**, 2 (2002).

[93] T. Scholzen and J. Gerdes, The Ki-67 protein: From the known and the unknown, *J. Cell. Phys.* **182**, 311 (2000).

[94] P. H. Jones and F. M. Watt, Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression, *Cell* **73**, 713 (1993).

[95] P. H. Jones, S. Harper, and F. M. Watt, Stem cell patterning and fate in human epidermis, *Cell* **80**, 83 (1995).

[96] S. Estrach, J. Legg, and F. M. Watt, Syntenin mediates Delta1-induced cohesiveness of epidermal stem cells in culture, *J Cell Sci* **120**, 2944 (2007).

[97] U. B. Jensen and F. M. Watt, Single-cell expression profiling of human epidermal stem and transit-amplifying cells: Lrig1 is a regulator of stem cell quiescence, *Proc. Natl. Acad. Sci. USA* **103**, 11958 (2006).

[98] J. Legg, U. B. Jensen, S. Broad, I. Leigh, and F. M. Watt, Role of melanoma chondroitin sulphate proteoglycan in patterning stem cells in human interfollicular epidermis, *Development* **130**, 6049 (2003).

[99] S. Lowell and F. M. Watt, Delta regulates keratinocyte spreading and motility independently of differentiation, *Mech Dev* **107**, 133 (2007).

[100] H. Wan, M. G. Stone, C. Simpson, L. E. Reynolds, J. F. Marshall, I. R. Hart, K. M. Hodivala-Dilke, and R. A. Eady, Desmosomal proteins, including desmoglein 3, serve as novel negative markers for epidermal stem cell-containing population of keratinocytes, *J Cell Sci* **116**, 4239 (2003).

[101] U. B. Jensen, S. Lowell, and F. M. Watt, The spatial relationship between stem cells and their progeny in the basal layer of human epidermis: a new view based on whole-mount labelling and lineage analysis, *Development* **126**, 2409 (1999).

[102] S. Ghazizadeh and L. B. Taichmann, Organization of Stem Cells and Their Progeny in Human Epidermis, *J. Invest. Dermatol.* **124**, 367 (2005).

[103] S. Lowell, P. H. Jones, I. Le Roux, J. Dunne, and F. M. Watt, Stimulation of human epidermal differentiation by DeltaNotch signalling at the boundaries of stem-cell clusters, *Curr. Bio.* **10**, 491 (2000).

[104] P. Jones and B. D. Simons, Epidermal homeostasis: do committed progenitors work while stem cells sleep?, *Nat Rev Mol Cell Biol.* (2007).

[105] P. H. Jones, B. D. Simons, and F. M. Watt, Sic Transit Gloria: Farewell to the Epidermal Transit Amplifying Cell?, *Cell Stem Cell* **1**, 371 (2007).

[106] M. I. Koster and D. R. Roop, Asymmetric cell division in skin development: a new look at an old observation, *Dev Cell* **9**, 444 (2005).

[107] J. W. Cahn and J. E. Hilliard, Free Energy of a Nonuniform System. I. Interfacial Free Energy, *J. Chem. Phys.* **28**, 258 (1958).

[108] S. C. Glotzer, E. A. Di Marzio, and M. Muthukumar, Reaction-controlled morphology of phase-separating mixtures, *Phys. Rev. Lett.* **74**, 2034 (1995).

[109] F. L. Ochoa, A generalized reaction diffusion model for spatial structure formed by motile cells, *BioSystems* **17**, 35 (1984).

[110] V. G. Karpov, Negative diffusion and clustering of growing particles, *Phys. Rev. Lett.* **75**, 2702 (1995).

[111] H. Wearing and J. Sherratt, Keratinocyte growth factor signalling: A mathematical model of dermal–epidermal interactions in epidermal wound healing, *Math. Biosci.* **165**, 41 (2000).

[112] P. K. Maini, D. L. S. McElwain, and D. I. Leavesley, Traveling wave model to interpret a wound-healing cell migration assay for human peritoneal mesothelial cells, *Tissue Eng.* **10**, 475 (2004).

[113] P. K. Maini, D. L. S. McElwain, and D. I. Leavesley, Travelling waves in a wound healing assay, *Appl. Math. Lett.* **17**, 575 (2004).

[114] E. L. Ionides, K. S. Fang, R. Rivkah Isseroff, and G. F. Oster, Stochastic models for cell motion and taxis, *Journal of Mathematical Biology* **48**, 23 (January 2004).

[115] N. J. Armstrong, K. J. Paintera, and J. A. Sherratt, A continuum approach to modelling cellcell adhesion, *J. Theor. Bio.* **243**, 98 (2006).

[116] C. M. Elliott and H. Garcke, Diffusional phase transitions in multicomponent systems with a concentration dependent mobility matrix, *Physica D* **109**, 242 (1997).

[117] H. Honda, Geometrical models for cells in tissues, *Int Rev Cytol* **81**, 191 (1983).

[118] H. Honda and H. Yamanaka, A computer simulation of geometrical configurations during cell division, *J Theor Biol* **106**, 423 (1984).

[119] D. A. Clausi and G. W. Brodland, Mechanical evaluation of theories of neurulation using computer simulations, *Development* **118**, 1013 (1993).

[120] H. H. Chen and G. W. Brodland, Cell-Level Finite Element Studies of Viscous Cells in Planar Aggregates, *J Biomech Eng* **122**, 394 (2000).

[121] G. W. Brodland and J. H. Veldhuis, Computer simulations of mitosis and interdependencies between mitosis orientation, cell shape and epithelia reshaping, *Biomechanics* **35**, 673 (2002).

[122] G. Giacomin and J. L. Lebowitz, Exact Macroscopic Description of Phase Segregation in Model Alloys with Long Range Interactions, *Phys. Rev. Lett.* **76**, 1094 (1996).

[123] E. Domany and W. Kinzel, Equivalence of Cellular Automata to Ising Models and Directed Percolation, *Phys. Rev. Lett.* **53**, 311 (1984).

[124] S. Puri and K. Binder, Dynamical scaling in anisotropic phase-separating systems in a gravitational field, *Phys. Rev. B* **46**, 98 (1992).

[125] S. Puri, A. J. Bray, and J. L. Lebowitz, Phase-separation kinetics in a model with order-parameter-dependent mobility, *Phys. Rev. E* **56**, 758 (1997).

[126] Y. Barrandon and H. Green, Cell Migration Is Essential for Sustained Growth of Keratinocyte Colonies: The Roles of Transforming Growth Factor-α and Epidermal Growth Factor, *Cell* **50**, 1131 (1987).