Unravelling stem cell dynamics by lineage tracing

Cédric Blanpain[1,2] and Benjamin D. Simons[3,4,5]

Abstract | During embryonic and postnatal development, the different cells types that form adult tissues must be generated and specified in a precise temporal manner. During adult life, most tissues undergo constant renewal to maintain homeostasis. Lineage-tracing and genetic labelling technologies are beginning to shed light on the mechanisms and dynamics of stem and progenitor cell fate determination during development, tissue maintenance and repair, as well as their dysregulation in tumour formation. Statistical approaches, based on proliferation assays and clonal fate analyses, provide quantitative insights into cell kinetics and fate behaviour. These are powerful techniques to address new questions and paradigms in transgenic mouse models and other model systems.

[1]*Université Libre de Bruxelles, Institut de Recherche Interdisciplinaire en Biologie Humaine et Moléculaire (IRIBHM), 808 route de Lennik, Brussels B-1070, Belgium.*
[2]*Walloon Excellence in Life Science and Biotechnology (WELBIO), Université Libre de Bruxelles, 808 route de Lennik, Brussels B-1070, Belgium.*
[3]*Cavendish Laboratory, Department of Physics, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, UK.*
[4]*Wellcome Trust Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK.*
[5]*Wellcome Trust Medical Research Council Cambridge Stem Cell Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QR, UK.*
Correspondence to C.B., B.D.S.
e-mails: Cedric.Blanpain@ulb.ac.be; bds10@cam.ac.uk
doi:10.1038/nrm3625
Published online 17 July 2013

The coordination of stem cell proliferation and fate specification is central to tissue growth and maintenance. During development, the balance between proliferation and differentiation must be tightly regulated to ensure that precise numbers of lineage-specific precursors are generated in the correct temporal sequence. In the adult, a delicate balance between proliferation and differentiation is essential for homeostasis. Elucidating the mechanisms that regulate this balance will provide fundamental insights into tissue maintenance and repair as well as the pathways that lead to dysregulation of cell fate in disease progression and ageing.

There are several technical challenges associated with the study of stem cell fate behaviour such as, for example, the difficulty to distinguish stem cells from their differentiating progeny. Stem cell-specific markers are scarce and rarely linked to function. Moreover, stem cells operate in a noisy and dynamic environment, as their gene expression levels fluctuate in response to intrinsic factors and/or environmental cues. Transcriptional profiling of cell populations and the use of fixed samples to identify the expression patterns of genes and gene products do not provide information on kinetics and frequently fail to represent cell behaviour at the tissue level. Although live imaging offers the potential to correlate gene expression with fate choice, continuous measurements of stem and progenitor cell fate behaviour over time are usually unfeasible *in vivo*.

Combined with marker-based assays and gene expression profiling, lineage-tracing studies involving measurements of genetically labelled cells and their progeny are beginning to uncover conserved patterns of stem cell fate in tissue maintenance and specification. In some cases, these findings challenge prevailing concepts and raise new questions about molecular regulation of stem cell activity and function. These developments have hinged on the acquisition of high-precision clonal fate data. By highlighting recent progress in this area, the aim of this Review is to discuss lineage-tracing strategies and how their quantitative analysis can be used to develop insights into cell kinetics and fate choice.

We first define some of the key questions concerning stem cell fate behaviour in developing and adult tissues. With this background, we address the basic foundations of lineage-tracing methods and discuss how, in principle, data from these assays can be interpreted through quantitative statistical approaches. We discuss how the application of these methods has provided new insights into the maintenance of two actively renewing tissues: the mouse intestinal epithelium and the interfollicular epidermis. Keeping the focus on epithelial tissues, we consider how lineage-tracing assays can provide insight into stem and progenitor cell fate behaviour in the late stage development of tissues and their dysregulation in tumour progression.

Key questions in stem cell biology

In the course of embryonic and postnatal development, pluripotent stem cells give rise to tissue precursor cells that transit through a differentiation pathway in which their lineage potential becomes increasingly restricted. To ensure the integrity of adult tissue, these tissue-specific progenitors must generate precise numbers of differentiated cells and in the correct temporal sequence.

Different theories have been proposed to describe such behaviour. Some have argued that precursors are specified early in development to generate a stereotypic lineage with a defined cell number and cell type composition[1,2]. Others suggested that lineage specification involves competition between equipotent precursor cells, leading to clones of variable size and composition[3,4].

In adult life, many tissues undergo constant turnover, which requires new cells to replenish cells that were lost. To maintain homeostasis, stem cells must divide asymmetrically so that, on average, one half of the progeny are retained as stem cells while the other half commit to differentiation, either directly or through a series of terminal divisions[5,6]. Because stem cell longevity is considered a defining hallmark of stem cell behaviour, many studies have focused on identifying factors that promote fate asymmetry at the single-cell level so that one daughter cell always retains stem cell identity. However, fate asymmetry may also be achieved at the population level by ensuring that stem cell loss through differentiation is perfectly compensated for by stem cell duplication (FIG. 1a). Through this process of population asymmetric self-renewal, stem cells are continually lost and replaced. Over time, this leads to a reduction in clonal diversity in tissues, as stem cell-derived clones are progressively lost through differentiation, while others expand to maintain the overall number of stem cells. Resolving the pattern of fate asymmetry is key to understanding the factors that regulate stem cell behaviour and tissue maintenance.

Although the regulatory pathways that control stem and progenitor cell fate may be tissue specific, mechanistic strategies may be shared. For example, stem cells are located within a specialized microenvironment called the stem cell niche, which is composed of signals from somatic cells, extracellular matrix proteins, the vasculature, as well as neighbouring cells[6]. Although the niche may promote stem cell competence, its influence on fate is not always clear. Furthermore, evidence of low proliferative activity in various tissues has led many to conclude that quiescence may be a defining characteristic of stem cell behaviour[7,8]. Yet, in other tissues, stem cells undergo rapid proliferation throughout adulthood. Finally, for some tissues, stem cell identity may not be defined by a single signature pattern of gene expression; instead, stem cells may constitute a heterogeneous compartment in which cells transit reversibly between different states of competence, becoming temporarily 'primed' for either quiescence or activity, the latter of which could lead to either proliferation or differentiation[9]. Crucially, although genetic profiling and marker-based assays provide insights into the molecular regulation of activity, understanding the rules governing stem and progenitor cell fate determination requires dynamic information, which can be obtained from lineage-tracing studies.

Lineage tracing: theory

To resolve the kinetics of cell proliferation and fate behaviour, which underpin tissue development and maintenance, several complementary strategies are available. These strategies can be divided into two categories: those that make use of population average measures, providing information on proliferation kinetics and cell lineage specification, and those that recover information at a clonal level, giving insight into the fate behaviour of individual cells.

Proliferation kinetics. To facilitate the study of the proliferative activity of stem and progenitor cells, several labelling techniques have been developed. The proportion of cells in different stages of the cell cycle can be determined by measuring the DNA content using flow cytometry. Conversely, incorporation of nucleotide analogues, such as 5-bromodeoxyuridine (BrdU), during DNA synthesis can reveal the fraction of cycling cells in a tissue as well as their average cell cycle time (BOX 1; FIG. 1b). Moreover, as DNA synthesis usually leads to the symmetric partitioning of nucleotides between daughter cells, the retention of DNA label in pulse-chase experiments can be used to discriminate slow-cycling cells from actively cycling cells, in which the label is rapidly diluted. Indeed, through this process of label retention in daughter cells, the incorporation of nucleotide analogues can be used as a simple, primitive form of clonal marking, at least for a few rounds of division[10,11].

Recently, researchers have developed a pulse-chase strategy on the basis of a transgenic mouse model allowing the expression of a fluorescent protein (that is, GFP fused to the histone 2B (H2B–GFP)) in a doxycycline (DOX)-dependent manner. In the absence of DOX, H2B–GFP is expressed at a very high and homogenous level in proliferative cells (FIG. 1c). Upon DOX administration, H2B–GFP expression is completely repressed; H2B–GFP expressed before induction remains incorporated into chromatin due its high stability and is diluted by half on each division. Thereby, the number of cell divisions that a given cell accomplished can be monitored during the chase period (BOX 1; FIG. 1c,d). Moreover, this approach enables the isolation of live cells on the basis of their slow or rapidly cycling properties as well as fate mapping of slow-cycling cells and their progeny for many more rounds of division than possible with the BrdU assay[12]. As well as providing a quantitative assay to characterize the frequency of slow-cycling cells, this method also allows the dynamic behaviour and fate choice of actively cycling cells to be measured (BOX 1).

Clonal dynamics. In recent years, a more versatile labelling strategy has been developed using transgenic animal models. This approach has been used in several organisms and provides clonal information, allowing the identification of the progeny of individually marked cells[13,14]. In mice, the most popular labelling technique is based on a drug-inducible Cre recombinase together with a reporter system[15]. By placing a drug-inducible Cre recombinase (for example, Cre–oestrogen receptor (ER)) under the control of a gene-specific promoter, its transient activation by drug administration (for example, tamoxifen) leads to the excision of a stop cassette and the permanent expression of a reporter construct in targeted cells and their progeny. The use of different cell-specific promoters and adjustment of the drug dose allow different cell subpopulations to be labelled. Thereby, the behaviour

Equipotent precursor cells
A group of progenitor cells that present the same intrinsic capacity to renew and differentiate. The term equipotency does not imply that all cells will give rise to identical daughter cells, as the cell cycle time may be different between equipotent progenitors, and the choice between renewal and differentiation may be stochastically defined.

Terminal divisions
Cell divisions that lead to the generation of two terminally differentiated cells that will not divide anymore.

Stem cell niche
The particular microenvironment in which stem cells reside. The stem cell niche is thought to regulate stem cell activity and influence fate decisions through the release of extrinsic signals (for example, growth factors, morphogens, nutriments and oxygen).

Pulse-chase
A method that involves the administration of nucleotide analogues for a certain period (pulse), followed by a period during which no nucleotide analogues is administrated (chase). During the chase period, cells that divide will dilute the label equally between the two daughter cells. After a few rounds of cell division (3–4 divisions), the label typically becomes undetectable. By contrast, in non-dividing cells the label remains detectable, and these cells are thus termed label-retaining.

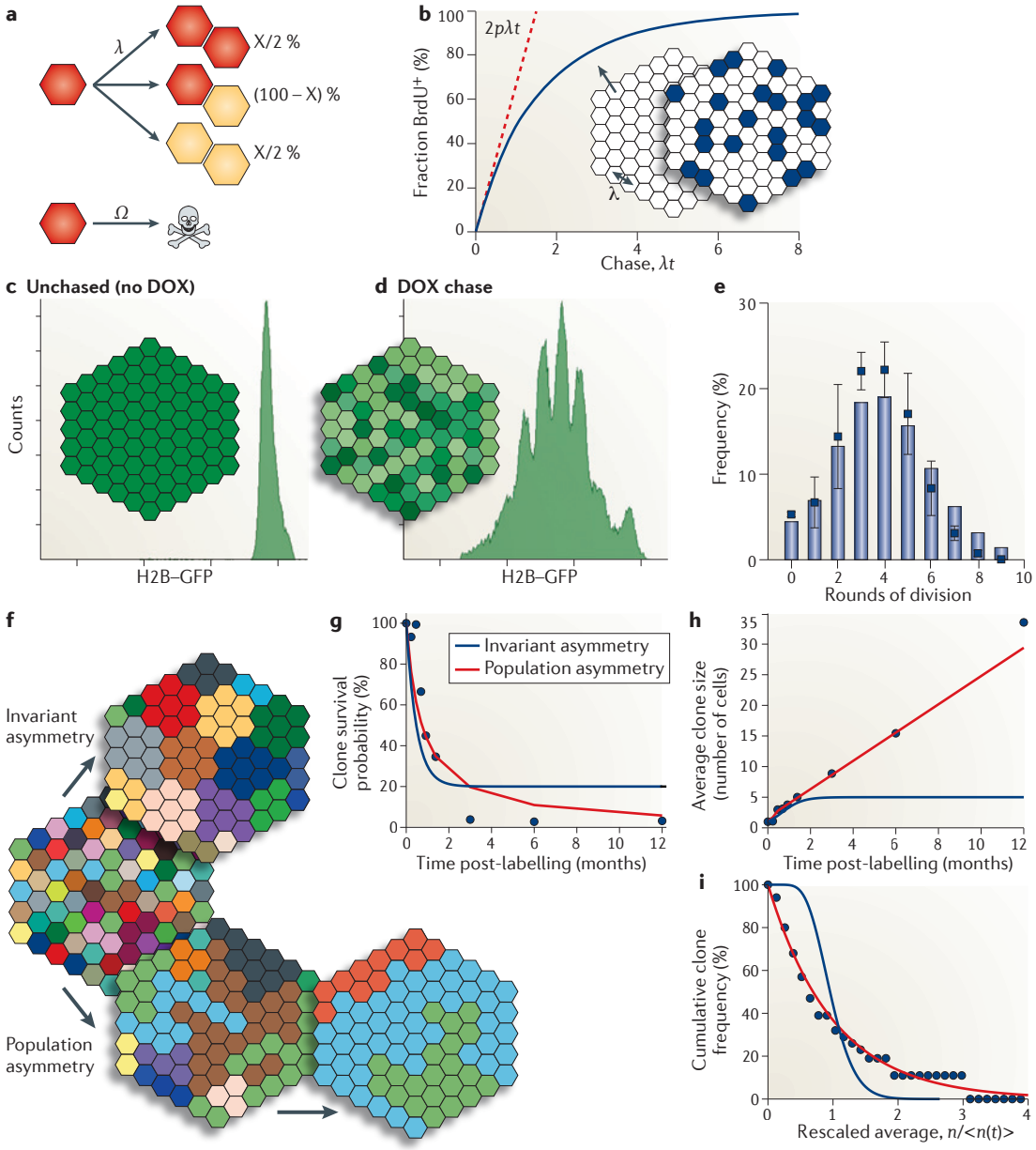


Figure 1 | **Lineage tracing as a quantitative tool.** **a** | Toy model of stem cell self-renewal. Stem cell division at rate λ results, on average, in an asymmetric fate outcome with one stem cell (red) and one differentiated cell (yellow). For $X>0$ (where X denotes the percentage of divisions that lead to symmetric fate), fate is assigned randomly and asymmetry is achieved only at the population level. When differentiated cells are lost at rate Ω, $p = 1/(1+\lambda/\Omega)$ defines the stem cell fraction. **b** | Percentage of labelled cells over time (shown by the blue line) predicted for the model shown in part **a** following continuous 5-bromodeoxyuridine (BrdU) incorporation (the short-time asymptotic is shown in red) (see also BOX 2). **c** | In the histone 2B (H2B)–GFP label-dilution assay, before doxycycline (DOX) administration, fluorescence-activated cell sorting (FACS) analysis shows that all cells express high levels of H2B–GFP (shown in dark green). **d** | Following DOX administration, the level of H2B–GFP expression is reduced by half following cell division, leading to a peak distribution. **e** | Following quantitative analysis, evidence can be found to challenge the model shown in part **a** (the square points represent data from part **d** and the bars (which indicate the proportion of cells that accomplished zero to 10 rounds of cell division after 21 days of chase) from part **a** (see BOX 1 for a description of the model). **f** | Schematic showing clonal evolution of a tissue conforming to the model depicted in part **a** and labelled as a multicolour mosaic. The top panel depicts invariant asymmetry ($X=0$) leading to a mosaic of 'clonal units' (shown as groups of cells with the same colour). The lower panels show the progressive depletion in clonal diversity with population asymmetry ($X>0$). **g**–**i** | Schematics showing the change in surviving clone fraction (part **g**), average clone size (**part h**) and cumulative clone size distribution (**part i**) in the case of invariant asymmetry (blue line) and population asymmetry (red line). When cells exhibit population asymmetry, the surviving clone density progressively diminishes, whereas the average size of survivors increases so that the overall number of labelled cells remains constant. The cumulative clone size distribution, $C_n(t)$, defined as the chance of finding a clone with a size larger than n cells, approaches a hallmark scaling form (see BOX 2). The data and analysis shown in parts **c**–**e** are reproduced, with permission, from REF. 59 © (2012) Macmillan Publishers Ltd. The data shown in parts **g**–**i** are reproduced, with permission, from REF. 56 © (2007) Macmillan Publishers Ltd.

Box 1 | **Proliferation kinetics: population assays**

Three techniques — pulse and pulse-chase, continuous label incorporation and label dilution — provide complementary and quantitative information to analyse proliferation kinetics at the population level. A simplified two-compartment model of a homeostatic tissue with a single equipotent self-renewing stem cell population can be used to illustrate their application. Following division, at rate λ, stem cells give rise to terminally differentiated progeny, which are subsequently lost at rate Ω (see FIG. 1a). As during homeostasis, the generation of differentiated cells must balance their loss, and the progenitor cell fraction, p, is fixed by the condition $p\lambda = (1\text{-}p)\Omega$. This model describes the turnover of basal layer cells in the mouse epidermis. For population average measures, the particular mode of division asymmetry — whether asymmetry is imposed at the level of individual stem cells (termed invariant asymmetry) or at the population level (termed population asymmetry) — is indistinguishable.

Continuous label incorporation and pulse assays

When proliferating cells are continuously exposed to a thymidine analogue such as 5-bromodeoxyuridine (BrdU), they become labelled during S phase. As they progress through mitosis, both daughter cells retain the label, expanding the marked population. At the same time, differentiated cells (labelled and unlabelled) are lost, which gradually leads to all cells becoming labelled over time. If we assume that the timings between consecutive cell divisions and between losses are statistically uncorrelated (that is, a Poisson random process), the percentage of labelled cells varies over time as $(1 - pe^{-2\lambda t} - (1 - p)e^{-\Omega t}) \times 100\%$, with $p = 1/(1+\lambda/\Omega)$ denoting the fraction of cells that are proliferative (see FIG. 1b). Following cells from the initial increase in the labelled cell fraction to the point when almost all cells are labelled, it is possible to infer both the average cell division and loss rates. In cases in which the cell cycle time distribution is more structured (for example, if there is a degree of synchrony in consecutive cell division times), its form may be recovered using a double pulse-labelling assay (such as BrdU and 5-ethynyluridine (EdU)), in which the second pulse assesses the frequency by which labelled cells re-enter the cell cycle.

Label-dilution assay

Whereas assays based on continuous label incorporation inform on cell division kinetics, they do not account for cell-to-cell proliferative heterogeneity. In the histone 2B (H2B)–GFP label-dilution assay, all cells express similar levels of GFP before induction. After induction, the expression level is approximately halved following each round of division. GFP levels can be measured accurately by flow cytometry over several orders of magnitude of fluorescence intensity, allowing as many as 10 rounds of cell division to be monitored (see FIG. 1c,d). Although such assays provide information on the abundance and location (using confocal microscopy) of quiescent cells (see FIG. 3c), the range of GFP expression levels provides quantitative information on proliferation kinetics.

A quantitative description of cell behaviour obtained through H2B–GFP label-dilution assays can be exemplified with the two-compartment model described above[59] (see FIG. 1a). Following division, GFP levels are diluted by a factor of two in both daughter cells, while, on average, one of the daughters commits to differentiation. At the same time, differentiated cells continue to be lost at a rate Ω. According to these dynamics, the chance of finding a surviving cell with a level of GFP that is diluted by a factor $(1/2)^n$ at a time t post-induction is given by:

$$\frac{(\lambda t)^n}{n!}e^{-2\lambda t} + \frac{e^{-\Omega t}}{(2-\Omega/\lambda)^n}\left(1 - \frac{\Gamma[n,(2-\Omega/\lambda)\lambda t]}{(n-1)!}\right)$$

for $n>0$, and $e^{-2\lambda t}$ for $n=0$, in which $\Gamma[n,x]$ denotes the incomplete γ-function (see FIG. 1e).

Finally, although continuous label-incorporation and label-dilution assays provide access to cell kinetics, both methods may impose a damaging effect on cells through, for example, DNA damage and cell death[88]. More importantly, like population-based measures, both approaches are unable to resolve individual cell fate choices. To measure individual cell fates, and to discriminate between different self-renewal strategies, clonal assays should be performed (see BOX 2).

of a single cell-derived clone can be followed or, when the tissue is labelled at the clonal density, cell behaviour can be studied at the population level. By concomitantly expressing multiple fluorescent proteins, different random colour combinations can be induced, leading to mosaic labelling of tissue[16,17].

Clonal density
The density of labelled cells that allows the fate of single labelled cells to be resolved and followed over time.

As with the H2B–GFP dilution assay, inducible genetic labelling based on the Cre recombinase system has been used to trace the transition of cells between different tissue subcompartments at the population level. Although such population average measures may help to identify the size and proliferation dynamics of the stem cell compartment, they do not provide information on the possible heterogeneity in the differentiation potential of the self-renewing population. The real power of inducible genetic-labelling assays lies in their capacity to resolve individual cell fate behaviour. However, obtaining information at the single-cell level may not be straightforward. First, the expression of the promoter may extend over more than one cell compartment, or it may fluctuate within a single subpopulation. Moreover, given that events such as proliferative hierarchy, stochastic fate choice and cell death might not be detected, the fate history of cells in a clone is not uniquely defined by its size and composition alone[18]. However, analysing the properties of a statistical ensemble of clones, across a range of time points, should allow minimizing the impact of such undetected events.

To understand how cell fate behaviour can be inferred from the statistical analysis of clone sizes, the constraints imposed by homeostatic tissues on stem cell self-renewal strategies should be considered. Although stem cells may become temporarily primed for proliferation or differentiation and loss, such bias cannot persist indefinitely under conditions of homeostasis. If it did, 'disadvantaged' cells, tilted towards differentiation, would acquire a strictly limited self-renewal potential, as all their progeny would differentiate and become lost. Therefore, over long timescales, when the cell composition of 'persisting' clones becomes representative of the tissue as a whole, stem cells must function as a single equipotent pool. Although equipotency of precursor cells is not mandatory during development, its existence may be inferred from lineage-tracing studies and, if present, equally exploited by quantitative statistical methods. By taking into consideration a statistical ensemble, in which all possible cell fate histories are proportionately represented, the availability and frequency of fate choices of individual cells can be deduced (BOX 2).

The interpretation of clonal fate data from inducible labelling assays may be complicated by the induction of mixed cell populations. Gene-specific promoters may drive expression predominantly in a particular cell type (although it is rare that they target a single subcompartment) and the labelling might be altered, for example, by cell activity and proximity to the vasculature. Furthermore, Cre-activating drugs can transiently elicit stem cell death and thus influence clonal dynamics[19]. Therefore, it is important to evaluate whether clonal evolution of a marked population is representative of a tissue as a whole and, in particular, whether the frequency and composition of labelled cells are in proportion to the tissue averages.

Lineage tracing in homeostatic tissues

Several case studies exemplify the power and limitations of lineage-tracing approaches for the understanding of stem cell behaviour in tissue maintenance and homeostasis.

Box 2 | **Cell fate choice: clonal assays**

To extract quantitative information from clonal fate studies, it may be necessary to start with a particular model hypothesis and question whether the data (variations in clone size, composition and frequency) can be accommodated by a suitable fit. However, in homeostasis, a more objective approach can be developed in which statistical characterizations of the data itself may be used to place self-renewal within a class of model behaviours. To illustrate how, it is instructive to begin with a simplified model of a homeostatic tissue involving a single self-renewing stem cell population that gives rise to short-lived terminally differentiated progeny (see BOX 1 and FIG. 1a). In this model, clones derived from progenitors initially expand, whereas those derived from differentiated cells become progressively lost. If stem cells follow a pattern of invariant asymmetry (X=0), (defined in FIG. 1a) the tissue separates into a mosaic of 'clonal units', each derived from a single stem cell (see FIG. 1f). When labelled at clonal density, the frequency of surviving clones approaches a constant, whereas their size distribution becomes clustered around the average unit size (see FIG. 1g–i).

If stem cells follow a pattern of population asymmetry (X>0), extinction of clones through chance differentiation is perfectly compensated for by the expansion of 'survivors', so that the overall number of labelled cells remains constant over time (see FIG. 1g,h). Through this process of 'neutral competition', the size distribution of surviving clones converges to a behaviour known as 'scaling'. In this regime, the cumulative clone size distribution, $C_n(t)$, defined as the chance of finding a clone with more than n cells, becomes a function only of $n/\langle n(t)\rangle$, where $\langle n(t)\rangle$ represents the average clone size, that is, $C_n(t)=f(n/\langle n(t)\rangle$[89] (see FIG. 1i). Therefore, although the average size of surviving clones may grow, the chance of finding a clone with a size larger than a multiple of the average remains fixed.

As well as providing a hallmark of stem cell equipotency and population asymmetry, the emergence of scaling represents a striking, generic and parameter-free characterization of the data. Scaling restricts the underlying pattern of self-renewal to a limited class of model behaviours, defined by the scaling function, f, and the average size dependence, $\langle n(t)\rangle$[89]. If, as represented the model in FIG. 1a, population asymmetry relies upon intrinsic (cell-autonomous) regulation, stem cell dynamics are defined by a critical birth–death process for which $\langle n(t)\rangle$ grows linearly with time and $f(x)=exp(-x)$. By contrast, if balance is established by extrinsic cues, then both $\langle n(t)\rangle$ and $f(x)$ depend on spatial organization. In tissues defined by a tubular (one-dimensional) facultative niche, such as the annulus formed by stem cells at the crypt base in the gut, the average size of surviving clones, $\langle n(t)\rangle$, grows with time as a square root power, whereas $f(x)=\exp(-\pi x^2/4)$ (see also FIG. 2d). Similarly, in 'epithelial' tissues defined by a two-dimensional geometry, such as interfollicular epidermis, $\langle n(t)\rangle=t/\ln(t)$, whereas $f(x)=exp(-x)$ (see FIG. 1h,i). For three-dimensional tissues, the scaling dependences alone cannot discriminate between intrinsically and extrinsically regulated population asymmetry.

Although long-term scaling can disclose the pattern of fate, stem cell identity and lineage hierarchy are more challenging to determine, as this requires the analysis of the 'non-universal' short-term behaviour before the onset of scaling. It is necessary to use the range of fate data to discriminate between competitive models. Although scaling behaviour may be detected from as few as 20 clones per time point, the resolution of lineage hierarchy in this transient regime may require a much higher density of data, rising combinatorially with the inclusion of each additional progenitor type. In this case, the task of decoding fate data will be greatly simplified by resolving clones by cell composition and the use of multiple Cre-lines targeting different cell types.

Intestinal epithelium. In mammals, the epithelial lining of the small intestine is organized into crypts and villi (FIG. 2a). Stem cells, which reside at or near the base of the crypt, give rise to transit-amplifying cell progeny that, through a series of terminal divisions, generate the range of absorptive and secretory cell types constituting the lining of the gut. During normal turnover, these terminally differentiated cells become shed into the lumen[20].

Despite long-standing interest, the lineage hierarchy and proliferative potential of intestinal progenitors remains debated[21]. Although most agree that stem cells reside in the base region of the crypt, the identity of the stem cell compartment and the nature of their fate behaviour are controversial. Do stem cells function as a single equipotent pool, or is there an engrained hierarchy (that is, a fixed and predetermined hierarchy) of proliferative potential within this compartment? Are stem cells characterized by a signature gene expression profile? Given their considerable proliferative burden, are tissue-maintaining cells also supported by a minority slow-cycling or quiescent stem cell population?

The earliest studies of lineage tracing of the intestinal epithelium date back to pulse-chase labelling experiments performed by LeBlond and collaborators[22,23]. Following exposure of mice to ^{3}H-thymidine, which labels DNA during cell division, small crypt-based columnar cells (CBCs) were labelled and the fate of their differentiating progeny traced. On the basis of these early studies, it was proposed that intestinal stem cells reside within the CBC population. Later, the significance of this population was challenged by radiation-damage and label-retaining studies, which suggested that cells capable of regeneration were positioned at the forth row from the base of the crypt (commonly termed position '+4')[24,25].

Analysis of the crypt composition in mouse chimaera and studies in which X-linked alleles (*Pgk1a* (phosphoglycerate kinase 1 alpha) and *Pgk1b*) were expressed demonstrated that the mouse intestinal and colonic crypts eventually become monoclonal with time[26]. As an early form of clonal tracing, the mutagen *N*-ethyl-*N*-nitrosourea (ENU) was used as a hereditary labelling strategy to mark a subpopulation of dividing cells in the mouse intestine[27–29]. By noting that labelled crypts undergo a progressive drift towards monoclonality (a process that is termed 'fixation'), it was proposed that either individual crypts are maintained by a single 'master' stem cell positioned at the apex of a proliferative hierarchy or that self-renewal involves neutral competition within an equipotent stem cell pool[30].

Genetic labelling
A method of cell labelling that uses a genetic system (such as a fluorescent reporter gene). The advantage of genetic labelling is its irreversibility, leading to a permanent expression of the reporter gene in the cells initially labelled and all their progeny. Non-genetic labelling, based on, for example, the incorporation of fluorescent dyes in some cells, eventually becomes undetectable as the dyes are diluted.

With the advent of transgenic technology, researchers were able to challenge the results of the ENU assay and, through the use of targeted promoters, search for molecular markers of stem cell identity. In a screen aimed at identifying WNT target genes[31], *Lgr5* (Leu-rich repeat-containing G protein coupled receptor 5) was identified as a gene expressed preferentially in cells intercalated between Paneth cells at the crypt base[32]. By using an inducible genetic labelling in which the Cre–ER recombinase was put under control of the *Lgr5* promoter, it was demonstrated that LGR5-expressing cells (LGR5$^+$ cells) can give rise to all intestinal cell lineages and were capable of long-term maintenance[32]. In a parallel lineage-tracing approach, using the *Bmi1* (a polycomb RING finger oncogene) promoter, which targets rare cells preferentially located at position +4, it was shown that these cells were also multipotent with long-term self-renewal capacity[33]. This raises the possibility that more than one intestinal stem cell population may coexist. Subsequent genetic-labelling studies provided evidence that cells capable of colonizing the entire crypt can express mouse *Tert* (telomerase reverse transcriptase)[34] and *Hopx* (HOP homeobox)[35], two genes that are preferentially induced in cells at position +4. Furthermore, it was shown that HOPX-expressing cells are in a dynamic equilibrium with LGR5$^+$ cells, with both cell lineages giving rise to each other[35]. Finally, it was shown that upon lineage ablation of LGR5$^+$ cells, intestinal homeostasis could be maintained in the absence of LGR5 through BMI1-expressing stem cells[36].

Although these findings are consistent with the existence of different stem cell populations that can be distinguished by distinct marker expression, recent studies suggest that the expression of these genes may not be sharply delineated or exclusive. This raises the possibility that Cre–ER under the control of different promoters may target the same stem cell pool[37–39]. Indeed, stem cells may belong to a subpopulation of cells that co-express all four markers, or expression of these markers may change reversibly over time in response, for example, to extrinsic cues arising from site-specific interactions within the stem cell niche.

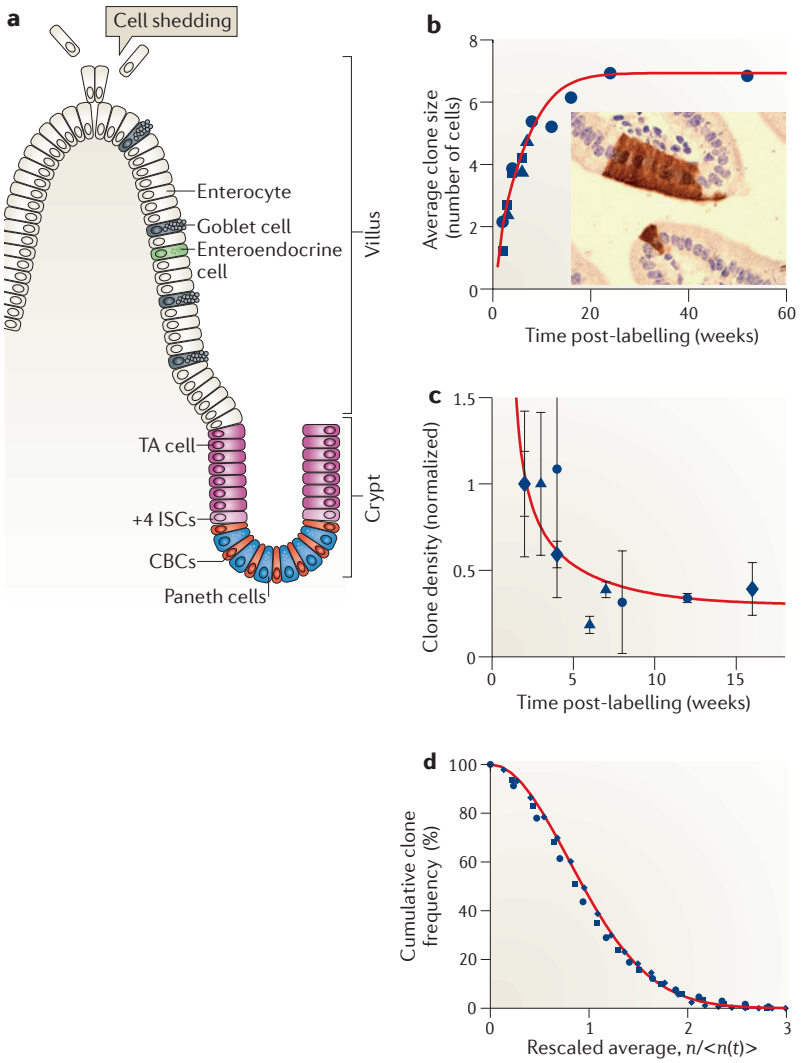
To what extent can quantitative analysis of lineage-tracing data shed light on the question of equipotency and stem cell fate behaviour? Recently, an inducible genetic-labelling approach, in which Cre expression is under the control of a ubiquitous promoter, was developed[40]. Although induction gave rise to clones throughout the crypt region, only clones that anchored to the crypt base survived long term. These clones formed cohesive ribbons of cells emanating from the crypt and terminating at the summit of the villi (FIG. 2b). By using the width of the labelled ribbon as an indication of the size of the underlying stem cell compartment, the size distribution and frequency of surviving clones was recorded over a 1-year time course. Following the transient loss of progenitor-derived clones, the total number of labelled cells remained approximately constant over time. However, this constant cell number was associated with continual clone loss, which was compensated for by a progressive increase in the size of persisting clones, a process that was arrested once individual crypts became monoclonal (FIG. 2b,c). Finally, the acquisition of hallmark scaling behaviour at intermediate times (BOX 2; FIG. 2d), before fixation, ruled out an engrained hierarchy of the proliferative potential in the stem cell compartment and instead pointed to a process of neutral competition in which stem cell loss through differentiation is perfectly compensated for by proliferation of neighbouring cells. This leads to neutral drift dynamics of clones around the crypt circumference.

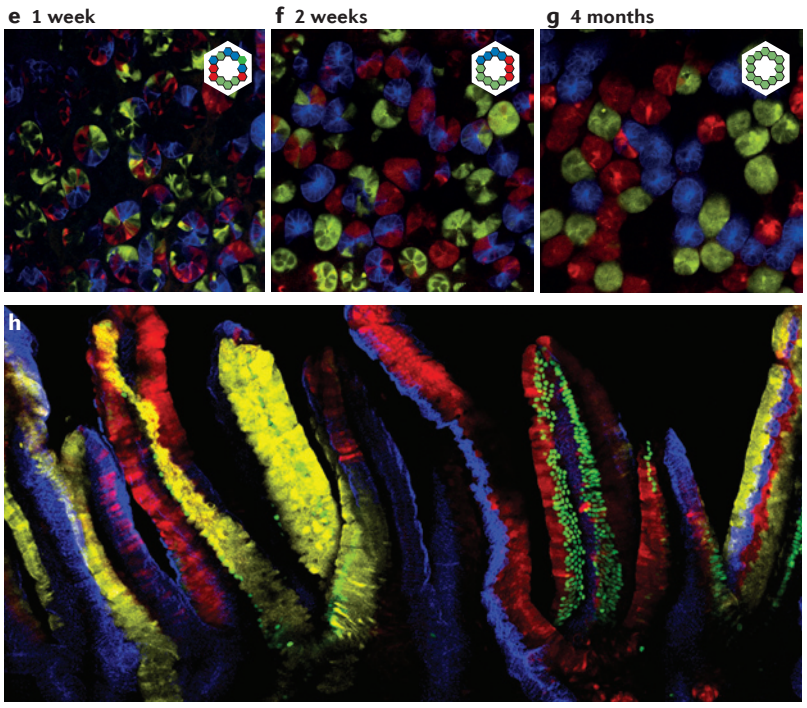
To dissect the molecular signature of this equipotent pool, a multicolour labelling strategy using Cre–ER under the control of the *Lgr5* promoter was used to measure clonal evolution at the cellular level[17]. By revealing the contribution of LGR5-expressing cells in persisting clones, it was possible to identify scaling behaviour of the clone size distribution (BOX 2) and, through the growth of an average size clone, to provide an estimate for the loss and replacement rate of LGR5$^+$ cells. Indeed, these findings were further corroborated through the quantitative analysis of 'domain coarsening' using a multicolour mosaic, which also provided a vivid demonstration of neutral drift dynamics and clonal succession (FIG. 2e—h). On the basis of these findings, it was suggested that stem cell self-renewal is regulated by competition for

Figure 2 | **Lineage tracing of the intestinal epithelium.** **a** | Cellular organization of the epithelial lining of the small intestine. Stem cells, which reside at the base of the intestinal crypt, give rise to transit-amplifying (TA) cell progeny that progressively differentiate into the functional secretory (Paneth, goblet and enteroendocrine cells) and absorptive (enterocytes) cell types of the gut. As cells divide and differentiate, they are conveyed along the walls of the crypt and villi in migration streams involving ribbons of cells until they eventually reach the villus tip where they are shed. In some studies, stem cell function is associated with crypt-based columnar cells (CBCs) that reside throughout the crypt base region, whereas according to other studies stem cells are localized at row +4 (that is, the fourth row from the bottom of the crypt). **b**,**c** | Following inducible genetic labelling of intestinal cells using a transgenic mouse model and an ubiquitous promoter, the average clone size can be determined by measuring the ribbon width on the villi (inset) (part **b**), and the density of surviving clones (part **c**) can be determined (points represent data obtained from mice of different ages, and the red lines show a fit to a simple model of neutral drift stem cell dynamics). **d** | Rescaled cumulative clone size distribution at 2, 3 and 4 weeks post-induction (points) show collapse onto universal scaling function (lines) (see also BOX 2). **e**–**g** | Section through the crypt base following multicolour mosaic labelling of the small intestine at 1 week, 2 weeks and 4 months post-labelling showing hallmark coarsening behaviour of surviving clones (see also FIG. 1f), leading to clonal fixation of crypts (as shown in part **g**, in which cells of the crypt are depicted in the same colour; that is, belonging to a same clone). **h** | Multicolour mosaic labelling of the small intestine at 4 months post-induction showing monoclonal crypts and ribbons of labelled cells on the villi. ISCs, intestinal stem cells. Data in parts **b**–**d** are adapted, with permission, from REF. 40 ©(2010) AAAS. Data in parts **e**–**h** are adapted, with permission, from REF. 17 © (2010) Cell Press.

Scaling behaviour
Behaviour that does not vary under a change of scale. For example, for a population defined by a statistical size distribution, while the average size may change over time, if the chance of finding a member of the population with a size greater than some multiple of the average remains constant over time, the distribution is said to scale.

Neutral drift
A term that was initially used to define the statistical distribution of gene mutations (drift) with no selective advantage (neutral) in a human population. This term can be used to describe a similar phenomenology in other contexts such as the time evolution of the statistical distribution of clone sizes in a lineage-tracing assay.





limited niche access. However, without following the fraction of persisting clones, which is challenging due to mouse-to-mouse and site-to-site variability in the labelling efficiency of the Cre recombinase, it is difficult to determine whether all cells within the LGR5$^+$ compartment have a long-term stem cell potential or, indeed, whether there are non-LGR5-expressing cells that belong to a yet wider stem cell pool. Equally, without detailed information on the short-term dynamics of marked cells, it is difficult to assess whether cells experience a short-term bias in self-renewal potential associated, for example, with positional cues imposed by interactions with the niche.

Finally, although these observations seem to rule out an engrained proliferative hierarchy in the actively cycling CBC population, they cannot exclude the possibility that other cell populations, such as quiescent cells, may have a crucial role in regeneration following injury, as measurements of fate behaviour in homeostasis cannot be used to extrapolate proliferative potential under stress. Indeed, in the mouse intestine, lineage-tracing assays have shown that Delta-like 1 (DLL1)-expressing cells, which are normally committed to secretory cell lineage differentiation, can be recruited back into the stem cell compartment following injury[41]. Finally, an ingenious lineage strategy based on Cre recombination was used recently to label quiescent intestinal cells. This revealed that, in homeostasis, slow-cycling intestinal progenitors are already committed to a Paneth cell lineage, but upon injury repopulate the stem cell niche and revert back into a multipotent stem cell[39]. Such studies illustrate the plasticity of stem and progenitor cell behaviour and emphasize the importance of the microenvironment in directing stem cell potential.

Epidermis. The skin epidermis is a stratified epithelium composed of the interfollicular epidermis (IFE), hair follicles and associated sebaceous and sweat glands[42]. The IFE ensures the barrier function of the skin by forming a waterproof envelope comprising differentiated suprabasal cells. These cells are maintained by the proliferation and differentiation of basal progenitors (FIG. 3a). As cells stratify through the suprabasal cell layers, they express different types of keratin, lose their nuclei and are eventually shed at the skin surface. Although this process continues throughout adulthood, in mice the rate of cell turnover in the epidermis is an order of magnitude smaller than that of the gut.

Similarly to the intestine, the earliest studies of proliferation dynamics in the epidermis were based on nucleotide incorporation in rodents[43–46]. Combined with studies of cellular organization, it was proposed that the IFE was arranged into a mosaic of units (termed epidermal proliferative units (EPUs)) composed of slow-cycling stem cells that support a pool of transit-amplifying cell progeny with a strictly limited proliferative potential[47]. Further evidence in support of proliferative heterogeneity came from tissue culture studies of the human epidermis[48]. Primary cultures of human keratinocytes gave rise to colonies with a distinct proliferative potential. Some cells generated rapidly expanding colonies with undifferentiated morphology

(termed holoclones) that could be maintained by serial passages *in vitro* and were thought to contain epidermal stem cells. Others gave rise to small abortive colonies (paraclones) that undergo terminal differentiation and were thought to represent transit-amplifying cells. In the mid-1990s, it was shown that colony-forming potential correlated with cell adhesion properties and the expression of basal integrins such as β1 integrin[49]. Furthermore, these cells were located at particular locations, such as the rete ridges that form at the bottom of the natural undulations in the palm epidermis[50].

In line with the EPU concept, clonal analysis using retroviral infection[51–53] and ENU-based clonal marking[54,55] revealed the presence of cells that give rise to cell columns spanning from the basal layer to the top of the cornified layer. However, the shape and the size of these EPUs were irregular, indicating a potentially more complex organization than suggested by the stem and transit-amplifying cell paradigm. More than 40 years after the EPU concept had been introduced, the first quantitative clonal analysis of the tail and the ear IFE during homeostasis was undertaken using the ubiquitous *Ah* (also known as CYP1A1)–Cre–ER[56,57]. These studies revealed that following induction, as with the intestine, the number of surviving clones progressively diminished, while their average size increased so that the overall number of labelled cells remained constant (FIG. 1g,h). Such behaviour was difficult to reconcile with the EPU paradigm and seemed to favour a process of population asymmetric self-renewal.

This conclusion was reinforced by the analysis of the clone size distribution, which exhibited a hallmark scaling behaviour (BOX 2; FIG. 1i). From detailed analysis of the short-term dynamics, it was proposed that the tissue was maintained by a single progenitor cell population. The fate of this population was stochastically balanced: most divisions (four out of five) led to an asymmetrical fate outcome, while the remainder resulted in symmetrical duplication or terminal division with equal probability.

As with the crypt, the behaviour of a labelled subfraction of cells cannot rule out the potential significance and function of cells that escape induction. In particular, the unlabelled region may harbour an independent self-renewing population or, as discussed below, it may include a small pool of quiescent cells that may only marginally contribute to normal homeostasis but that may become active following injury. To address potential proliferative heterogeneity in the IFE, another study used two Cre–ER (keratin 14 and involucrin) recombinases that target basal IFE progenitors[58,59]. Following detailed quantitative analysis, the clonal fate data suggested that progenitors induced by the involucrin–Cre–ER assay followed the same stochastic fate behaviour found in the previous studies[56,57,59]. By contrast, clones derived from keratin 14–Cre–ER-targeted cells presented an unexpectedly high survival rate, which further supports a fate biased towards proliferative divisions (FIG. 3b). Once again, through quantitative analysis of short- and long-term data, evidence was found in support of a proliferative hierarchy that involves a second subpopulation of slow-cycling stem cells.

Quantitative analysis of pulse-chase experiments in the epidermis using H2B–GFP-expressing mice provided further evidence in support of proliferative heterogeneity and the inferred cell kinetics (FIG. 1c–e), and associated slow-cycling cells in the mouse tail IFE with domains reminiscent of the rete ridges[59] thought to host epidermal stem cells in human epidermis[50,60,61] (FIG. 3c). Furthermore, transcriptional profiling of the two targeted proliferative cell populations revealed distinct gene expression profiles, with stem cells enriched for markers such as α2β1 (REF. 49) and α6β4 (REFS 62,63), which were previously identified as stem cell markers in the human epidermis.

As well as its ability to undergo cell turnover during homeostasis, the skin epidermis is also capable of extensive repair following wounding. Lineage tracing of different populations of hair follicle stem cells[12,64–67] showed that these cells are rapidly mobilized, migrate and contribute to wound repair. In addition, long-lived slow-cycling IFE stem cells also become active in response to wounding and contribute extensively and stably to repair (FIG. 3e), whereas progenitors seem to make only a minimal contribution[59]. It is, however, possible that the preferential recruitment of certain stem cell types might depend on the severity and the location of the wound. Together, these observations suggest that, in the IFE, progenitors are responsible for the routine maintenance of the tissue, whereas stem cells remain quiescent, ready to mobilize following injury. Curiously, this behaviour contrasts that of the esophagus; lineage-tracing studies show that progenitors are able to transit reversibly to a proliferative state and effect repair following injury[68]. This apparent reversion in fate potential raises the question of whether the proliferation hierarchy identified in the IFE is engrained or whether cells are able to transit reversibly between the stem and progenitor compartments under conditions of normal homeostasis. If such transitions occurred infrequently, they would be difficult to discern in a lineage-tracing assay.

In summary, the ability to reliably interpret clonal fate data rests on the long-term equipotency of the underlying stem cell population, a necessary characteristic for homeostatic turnover. However, leaving aside the later stages of tissue development, in which lineage potential varies over time, even adult tissues (such as hair follicle, muscle or blood) can be influenced by ageing[69–71]. Deciphering how ageing influences proliferative activity, potential and fate behaviour of cells is challenging and requires information on clone size, composition and density at variable induction times as well as chase period.

Lineage tracing in epithelial development

To maintain adult tissues that are actively turned over, homeostasis imposes significant constraints that restrict the range of possible stem cell behaviours. During development, tissue precursors must give rise to the correct number of differentiated progeny with the accurate composition, which requires equally stringent but different regulatory mechanisms. Lineage-tracing studies and clonal analysis are beginning to reveal common strategies

Rete ridges
Defines the bottom of the undulation present in the human skin epidermis, which was thought to contain human epidermal stem cells in certain parts of the body.

of progenitor cell fate behaviour and lineage specification during embryogenesis[4,72,73] and the late stages of tissue patterning. We discuss two examples involving the development of epithelia to demonstrate how genetic tracing experiments can be used to elucidate how the different cell lineages that define these epithelia are specified during development and maintained thereafter during adult homeostasis and regeneration.

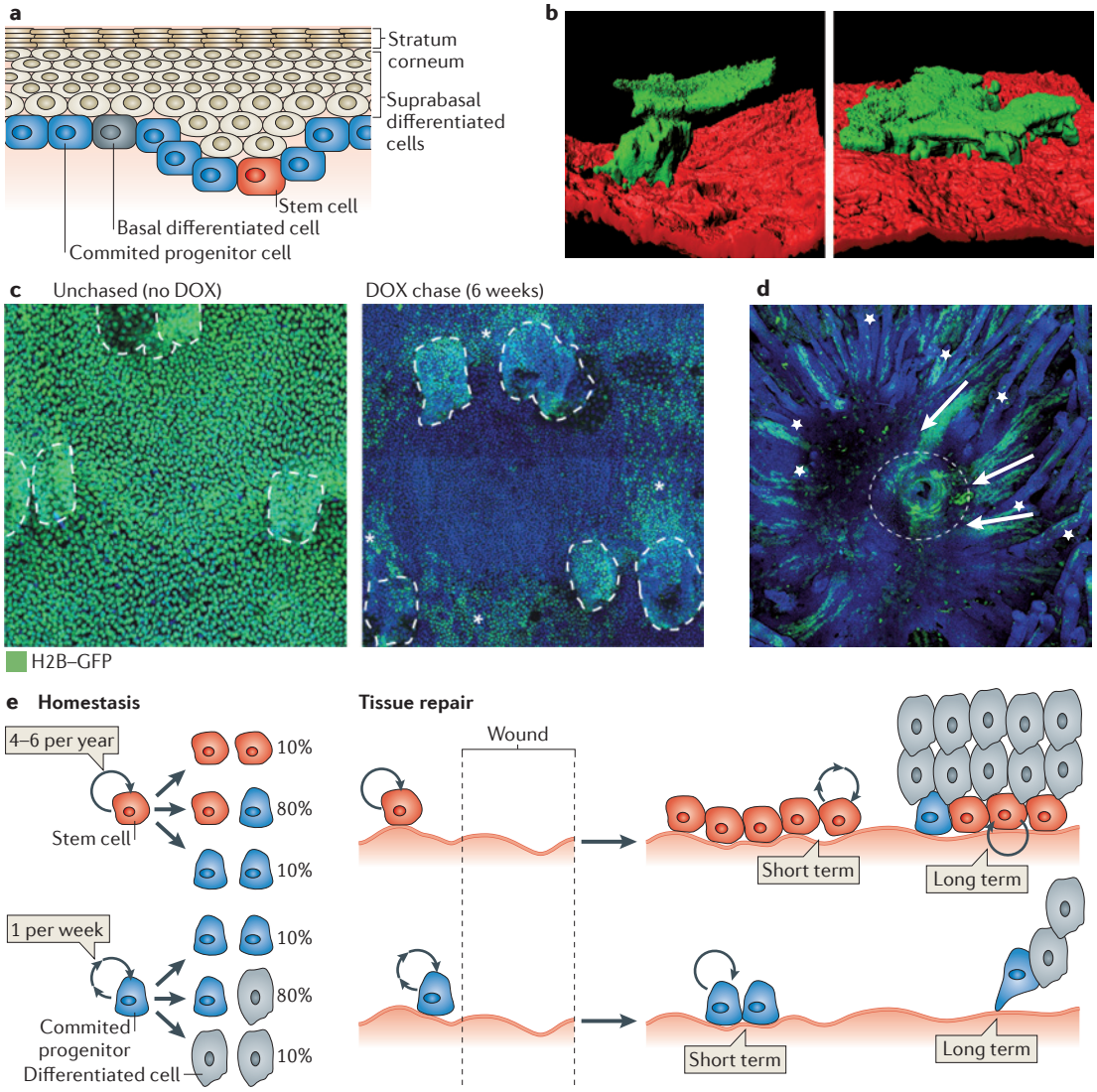


Figure 3 | **Clonal dynamics of the skin epidermis.** **a** | Stratified cellular organization of the mammalian interfollicular epidermis. Proliferative and differentiated cells are found in the basal layer and differentiated cells in the suprabasal cell layers. Following differentiation, cells in the basal layer detach from the basement membrane, stratify and progressively move up through the suprabasal cell layers, and they are shed at the surface of the epidermis. In homeostasis, lineage-tracing studies show that committed progenitors maintain tissue, while a stem cell population remains largely quiescent, ready to mobilize following injury to repair and regenerate tissue. **b** | Single-cell derived clones of epidermal cells (green) attached to the basal lamina (red) 48 weeks after lineage tracing. These studies show heterogeneity in size and shape between different clones. **c** | Studies of the mouse interfollicular epidermis using histone 2B (H2B)–GFP expression reveal the presence of label-retaining slow-cycling cells (denoted by the asterisks) that localize in defined regions of the epidermis. The left panel shows uniform GFP expression before doxycycline (DOX) administration, and the right panel shows the dilution of the label 6 weeks post-DOX administration, revealing the label-retaining slow-cycling cells. Hair follicle regions are indicated by dashed lines. **d** | Lineage tracing using mice expressing keratin 14–Cre–oestrogen receptor (ER) shows the activation and migration of stem cell-derived clones (labelled green) in wound repair following a punch biopsy (marked by the dashed circle). The arrows illustrate the sequence of events from stem cell activation to wound repair. **e** | Model of the cellular hierarchy that sustains epidermal homeostasis and repair. In homeostasis (left panel), the mouse interfollicular epidermis is maintained by committed progenitor cells (blue) following a pattern of balanced stochastic fate, whereby one in five divisions leads to progenitor cell loss through differentiation and replacement (see also FIG. 1a), whereas stem cells (red) remain largely quiescent (dividing only 4–6 times per year). Following injury, stem cells become activated and contribute substantially to repair. The images in parts **c**–**e** are adapted, with permission, from REF. 59 © (2012) Macmillan Publishers Ltd.

The mammary gland and prostate initiate from epithelial buds in the primitive epidermis and the urogenital epithelium, respectively. They progressively expand and invade their underlying mesenchyme to form a dense tubular network that contains inner luminal cells surrounded by outer basal cells (also known as myoepithelial cells)[15] (FIG. 4). Lineage-tracing analysis has been used to define the cellular origin of the basal and luminal cell lineages during postnatal development and adult regeneration in these tissues.

Mammary gland postnatal expansion and regeneration. The assessment of the cell composition following lineage tracing of basal cells or luminal cells demonstrated that the mammary gland initially develops from multipotent embryonic progenitors (FIG. 4b), which are replaced after birth by two distinct types of long-lived unipotent basal and luminal stem cells[74] (FIG. 4c,d). Clonal analysis revealed that only a small fraction of the basal and luminal clones expand considerably during postnatal development and pregnancy (FIG. 4e), suggesting that either unipotent stem cells represent a small fraction of these two cell lineages or that basal and luminal unipotent progenitors represent a pool of equipotent progenitors with a great distribution in their cell cycle time. However, although these basal stem cells are unipotent under physiological conditions, they have the ability to differentiate into both basal and luminal cells upon transplantation[74]. This reveals that the differentiation potential of epithelial cells in transplantation assays does not necessarily reflect their fate under physiological conditions.

Prostate postnatal expansion and regeneration. Similarly to the mammary gland, lineage-tracing studies of androgen-mediated regeneration in the adult prostate suggest the existence of two independent unipotent populations of basal and luminal cells that are self-sustained[75–77]. However, without detailed quantification of clone size and composition, this discovery does not resolve the issues of progenitor heterogeneity and cellular hierarchy.

Using clonal analysis, it has recently been demonstrated that postnatal development of the prostate is also mediated by multipotent basal stem cells that differentiate into basal, luminal and neuroendocrine cells, as well as unipotent basal and luminal progenitors. Although the range of clonal fate data suggests the existence of multiple populations of multipotent and unipotent progenitors (FIG. 4f–h), quantitative analysis of data at just one time point after birth could not rule out an alternative model in which the apparent cellular heterogeneity of basal progenitors could be explained by the stochastic cell fate decision of a single multipotent progenitor[78] (FIG. 4i, j). To address the question of progenitor heterogeneity in this glandular epithelium, further clonal analysis performed on whole mount tissue is required to assess fate outcome in relation to spatial localization and at different times during postnatal development.

Lineage tracing in solid tumours

In cancer, tumours are often found to be heterogeneous in morphology, cell composition, proliferative index, genetic profile and tumour-initiating potential of constituent cells. As with homeostasis, two models have been proposed to explain the basis of such heterogeneity (FIG. 5a). In the first model, all cells are equally capable of contributing to tumour growth through stochastic proliferation and differentiation. In the second model, the tumour is maintained by a minority of self-renewing cells, termed cancer stem cells (CSCs), which give rise to transit-amplifying cell-like progeny[79,80]. Over the past 15 years, a number of studies have demonstrated that a subset of tumour cells present a much greater capacity to reform a secondary tumour upon transplantation into immunodeficient mice[80]. This supports the notion that some tumour cells may be intrinsically more clonogenic than others. However, although such experiments demonstrate the potential of tumour cells in a transplantation assay, they do not assess the actual fate during tumour growth. Lineage tracing using transgenic animal models provides a tool to explore the fate of tumour cells in their native microenvironment.

The established behaviour of normal proliferative cells makes the IFE an attractive system to elucidate mechanisms of tumour initiation and progression. Recently, inducible genetic labelling was used to follow the fate of proliferative cells and their progeny in chemically-induced papilloma, which is a benign skin tumour that presents features of the normal IFE and of more invasive squamous cell carcinoma[81] (FIG. 5b). Curiously, in papilloma, clones derived from most of the labelled tumour cells undergo terminal differentiation, delaminate and are lost within a few weeks of labelling, whereas the surviving clones expand rapidly. This is consistent with the existence of tumour stem cells (FIG. 5c). By using the lineage hierarchy and fate behaviour of normal tissue as a platform, quantitative analysis of clonal fate data revealed the dynamics of papilloma cells. Clonal evolution was found to be consistent with the activation of a quiescent stem cell pool, leading to a slow, aberrant growth of the tissue. By contrast, during malignant progression, most of the tumour cells were actively proliferating, with only 5–15% of the clones presenting signs of squamous terminal differentiation (that is, the formation keratin pearls). In this phase, variability in clonal evolution can be further enhanced by the acquisition of additional mutations, which confer a selective growth advantage on individual clones. Further studies are required to assess whether there is a proliferative hierarchy in the growth potential of progenitors in this phase of invasive cancer.

In a parallel study, a transgenic mouse model was used to analyse clonal evolution within an intestinal adenoma using the *Lgr5*-Cre–ER and a multicolour reporter system[82]. As observed for papilloma cells, clones derived from LGR5-expressing cells contributed disproportionately to tumour growth, which suggests that these cells function as tumour stem cells (FIG. 5d). However, further quantitative studies are required to assess the equipotency of this stem cell pool as well as their potential to maintain long-term benign tumours. Intriguingly, through *in vivo* live imaging of individual breast tumour cells, recent studies have revealed a qualitatively

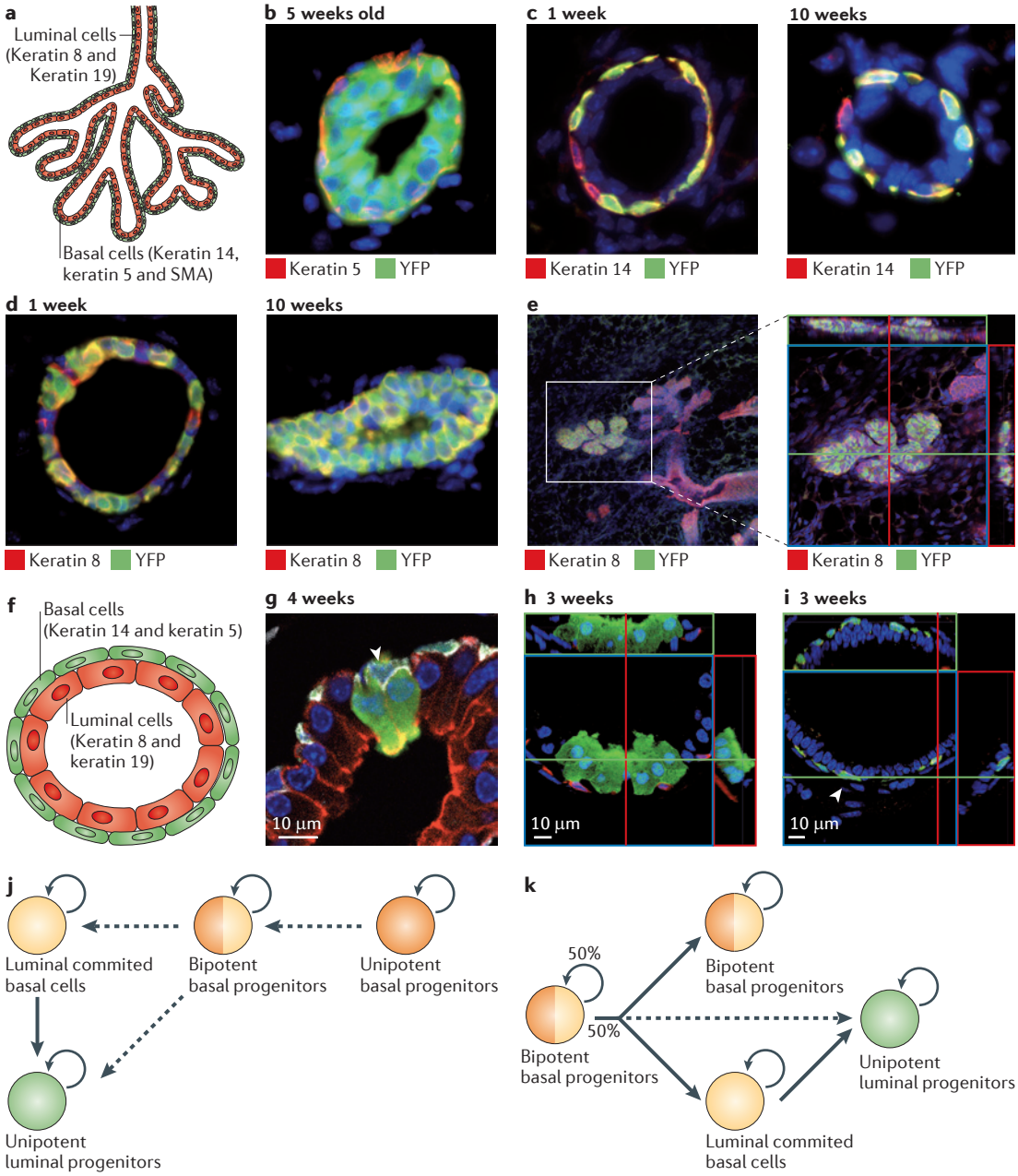


Figure 4 | **Lineage tracing during epithelial development.** **a** | Cellular organization of the mammary gland, with basal cells (also known as myoepithelial cells) shown in green (which express keratin 5, keratin 14 and smooth muscle actin (SMA)), and luminal cells are shown in red (which express keratin 8 and keratin 19). **b** | Lineage tracing of basal cells during embryonic development demonstrates that at the population level, embryonic basal cells give rise to all mammary gland lineages (basal cells and luminal cells are labelled with YFP and shown in green). **c** | Lineage tracing of basal cells during pubertal development demonstrates that these cells only give rise to basal cells (as seen by the green labelling being restricted to the outer layer of the lumen section) at 1 and 10 weeks post-induction of labelling expression. **d** | Lineage tracing of luminal cells during pubertal development demonstrates that luminal cells only give rise to luminal cells (as shown by the green labelling being restricted to the inner layer of the lumen section). **e** | Clonal analysis shows that only 5–10% of luminal cells are capable of massive expansion, giving rise to big luminal clones during pubertal expansion. **f** | Schematic cellular organization of the prostate, with basal cells and luminal cells. **g**–**i** | Clonal analysis of basal cells during postnatal development demonstrates that basal cells can give rise to different cell fate outcomes: bipotent clones (part **g**; the arrow head indicates a basal cell giving rise to basal and luminal cells), luminal clones (part **h**) and basal clones (part **i**; the arrow head indicates a basal cell). **j**,**k** | Two models can potentially explain the clonal fate data obtained during postnatal prostate development. In the first model, basal cells contain different types of progenitors (bipotent, unipotent basal and unipotent luminal) (part **j**), whereas in the second model, basal cells are bipotent progenitors that either undergo renewal or luminal progenitor commitment in a stochastic manner with equal probability (part **k**). Dashed lines indicate hypothetical transitions. Data parts **a**–**e** are adapted, with permission, from REF. 74 © (2011) Macmillan Publishers Ltd. The data shown in parts **f**–**j** are adapted from REF. 78 © (2012) Macmillan Publishers Ltd.

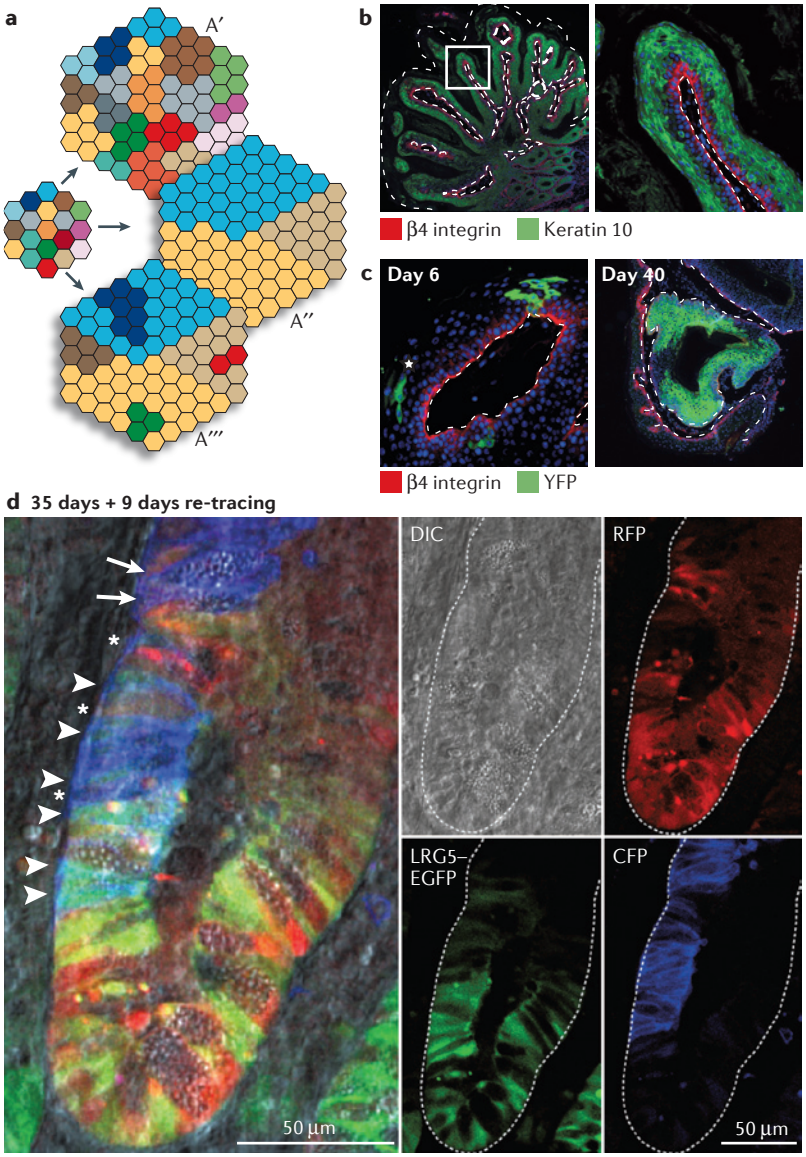


Figure 5 | **Clonal dynamics of tumour cells. a** | Schematic showing the expected clonal fate outcome in a situation in which: all proliferative tumour cells contribute to long-term tumour progression (A′); proliferative tumour cells are organized in a hierarchy in which only a minority of tumour stem cells at the apex have long-term tumour-maintaining potential (A″); and a mixed situation in which a subset of tumour-maintaining cells undergoes neutral competition for clonal dominance (A‴). **b** | Histological organization of a benign skin tumour (papilloma) containing proliferative basal cells (shown in red), which are marked by the expression of β4 integrin, and differentiated tumour cells (green), which are marked by the expression of keratin 10. **c** | Clonal analysis of skin papilloma at 6 days post-induction of labelling expression (YFP) showing clonal heterogeneity: clones of tumour cells are lost by terminal differentiation (denoted by the asterisks) while others expand. Dominant clones fill up most of the tumour after 40 days. Quantitative analysis of lineage-tracing data reveals that papilloma growth is consistent with a normal stem and committed progenitor cell hierarchy (see FIG. 3e, left panel), in which aberrant activation of the stem cell compartment leads to an over-production of tissue. In this case, papilloma growth is supported by a minority population of tumour stem cells that competes for clonal dominance (as in A‴, part **a** of the figure). **d** | Intestinal adenoma re-tracing. *Lgr5* (Leu-rich repeat-containing G protein coupled receptor 5)–Cre–oestrogen receptor (ER)–GFP is expressed in a subset of tumour cells (green) that have been labelled in red at the time of the tumour suppressor gene deletion. Tamoxifen administration to mice with pre-existing tumours induces colour inversion (red to blue) in LGR5-expressing cells and their progeny (blue cells), showing the contribution of LGR5-expressing tumour cells to the tumour growth. Arrow heads indicate $LGR5^{+}$ tumour stem cells, arrows indicate Paneth-like cells and asterisks indicate $LGR5^{-}$ tumour transit amplifying cells. DIC, differential interference contrast image EGFP, enhanced GFP. Images in part **c** are adapted, with permission, from REF. 81 © (2012) Macmillan Publishers Ltd. Images in part **d** are adapted, with permission, from REF. 82 © (2012) AAAS.

similar behaviour of skin and intestinal tumours, with the regression and loss of small clones contrasting the expansion and growth of large dominant clones[83]. This suggests that very different types of tumours may exhibit similar modes of growth.

Conclusions

Proliferation kinetics and lineage-tracing assays based on transgenic animal models can provide crucial quantitative insights into the proliferative potential, lineage hierarchy and fate behaviour of stem and progenitor cells. The application of these techniques to the intestinal epithelium and epidermis emphasizes the importance of stochasticity in the regulation of cell fate choice in tissue development and maintenance, as well as the transition to dysregulated growth in solid tumours. However, in most cases, the underlying molecular mechanisms and signalling pathways that control stochastic fate choice remain elusive. High-resolution lineage-tracing studies using targeted promoters offer the potential to identify cells in various states of priming, and begin to reveal the factors controlling proliferative potential and cell fate choice.

Although lineage tracing provides a functional read-out of cell behaviour over time, studies based on the analysis of clonal fate data in fixed samples offer only statistical insight into cell fate kinetics. In particular, it does not allow the unambiguous reconstruction of individual phylogenies. Although *in vivo* live-imaging assays can identify lineages[4,84], such measures are invasive and challenging to implement and usually can only be used for short-term experiments. The development of single-cell deep-sequencing technologies offers the potential to reconstruct information on different clonal histories in both normal tissue and disease[85–87]. Indeed, as well as identifying the first or key mutation that triggers the development shared by all cells in a given tumour, such data has the potential to reconstruct the pattern and sequence of secondary mutations that may have led to the expansion of a specific clone.

1. Sulston, J. E., Schierenberg, E., White, J. G. & Thomson, J. N. The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Dev. Biol.* **100**, 64–119 (1983).
2. Kao, C. F. & Lee, T. Birth time/order-dependent neuron type specification. *Curr. Opin. Neurobiol.* **20**, 14–21 (2010).
3. Slater, J. L., Landman, K. A., Hughes, B. D., Shen, Q. & Temple, S. Cell lineage tree models of neurogenesis. *J. Theor. Biol.* **256**, 164–179 (2009).
4. He, J. *et al.* How variable clones build an invariant retina. *Neuron* **75**, 786–798 (2012).
5. Morrison, S. J. & Kimble, J. Asymmetric and symmetric stem-cell divisions in development and cancer. *Nature* **441**, 1068–1074 (2006).
6. Morrison, S. J. & Spradling, A. C. Stem cells and niches: mechanisms that promote stem cell maintenance throughout life. *Cell* **132**, 598–611 (2008).
7. Fuchs, E. The tortoise and the hair: slow-cycling cells in the stem cell race. *Cell* **137**, 811–819 (2009).
8. Li, L. & Clevers, H. Coexistence of quiescent and active adult stem cells in mammals. *Science* **327**, 542–545 (2010).
9. Graf, T. & Stadtfeld, M. Heterogeneity of embryonic and adult stem cells. *Cell Stem Cell* **3**, 480–483 (2008).
10. Cotsarelis, G., Cheng, S. Z., Dong, G., Sun, T. T. & Lavker, R. M. Existence of slow-cycling limbal epithelial basal cells that can be preferentially stimulated to proliferate: implications on epithelial stem cells. *Cell* **57**, 201–209 (1989).
11. Cotsarelis, G., Sun, T. T. & Lavker, R. M. Label-retaining cells reside in the bulge area of pilosebaceous unit: implications for follicular stem cells, hair cycle, and skin carcinogenesis. *Cell* **61**, 1329–1337 (1990).
First study to demonstrate that hair follicle stem cells are slow cycling.
12. Tumbar, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303**, 359–363 (2004).
Describes a novel method based on the expression of H2B–GFP to isolate slow-cycling cells and quantify proliferation dynamics *in vivo*.
13. Buckingham, M. E. & Meilhac, S. M. Tracing cells for tracking cell lineage and clonal behavior. *Dev. Cell* **21**, 394–409 (2011).
14. Kretzschmar, K. & Watt, F. M. Lineage tracing. *Cell* **148**, 33–45 (2012).
15. Van keymeulen, A. & Blanpain, C. Tracing epithelial stem cells during development, homeostasis, and repair. *J. Cell Biol.* **197**, 575–584 (2012).
16. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56–62 (2007).
Reports for the first time the use of a multicolour reporter mice to perform lineage-tracing experiments.
17. Snippert, H. J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
18. Schroeder, T. Long-term single-cell imaging of mammalian stem cells. *Nature Methods* **8**, S30–S35 (2011).
19. Zhu, Y., Huang, Y. F., Kek, C. & Bulavin, D. V. Apoptosis differently affects lineage tracing of Lgr5 and Bmi1 intestinal stem cell populations. *Cell Stem Cell* **12**, 298–303 (2013).
20. Barker, N., Bartfeld, S. & Clevers, H. Tissue-resident adult stem cell populations of rapidly self-renewing organs. *Cell Stem Cell* **7**, 656–670 (2010).
21. Barker, N., van Oudenaarden, A. & Clevers, H. Identifying the stem cell of the intestinal crypt: strategies and pitfalls. *Cell Stem Cell* **11**, 452–460 (2012).
22. Cheng, H. & Leblond, C. P. Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine. I. Columnar cell. *Am. J. Anat.* **141**, 461–479 (1974).
Using a labelling approach, this study proposes that columnar basal cells located at the bottom of the crypts correspond to the stem cells of the mouse intestine.
23. Bjerknes, M. & Cheng, H. The stem-cell zone of the small intestinal epithelium. III. Evidence from columnar, enteroendocrine, and mucous cells in the adult mouse. *Am. J. Anat.* **160**, 77–91 (1981).
24. Potten, C. S., Hume, W. J., Reid, P. & Cairns, J. The segregation of DNA in epithelial stem cells. *Cell* **15**, 899–906 (1978).
25. Potten, C. S., Wilson, J. W. & Booth, C. Regulation and significance of apoptosis in the stem cells of the gastrointestinal epithelium. *Stem Cells* **15**, 82–93 (1997).
26. Ponder, B. A. *et al.* Derivation of mouse intestinal crypts from single progenitor cells. *Nature* **313**, 689–691 (1985).
First demonstration that intestinal and colonic crypts become monoclonal with time.
27. Griffiths, D. F., Davies, S. J., Williams, D., Williams, G. T. & Williams, E. D. Demonstration of somatic mutation and colonic crypt clonality by X-linked enzyme histochemistry. *Nature* **333**, 461–463 (1988).
28. Winton, D. J., Blount, M. A. & Ponder, B. A. A clonal marker induced by mutation in mouse intestinal epithelium. *Nature* **333**, 463–466 (1988).
29. Bjerknes, M. & Cheng, H. Clonal analysis of mouse intestinal epithelial progenitors. *Gastroenterology* **116**, 7–14 (1999).
30. Williams, E. D., Lowes, A. P., Williams, D. & Williams, G. T. A stem cell niche theory of intestinal crypt maintenance based on a study of somatic mutation in colonic mucosa. *Am. J. Pathol.* **141**, 773–776 (1992).
31. van de Wetering, M. *et al.* The β-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* **111**, 241–250 (2002).
32. Barker, N. *et al.* Identification of stem cells in small intestine and colon by marker gene *Lgr5*. *Nature* **449**, 1003–1007 (2007).
First study to show, using lineage-tracing experiments, that columnar basal cells correspond to the stem cells of the mouse intestine.
33. Sangiorgi, E. & Capecchi, M. R. Bmi1 is expressed *in vivo* in intestinal stem cells. *Nature Genet.* **40**, 915–920 (2008).
34. Montgomery, R. K. *et al.* Mouse telomerase reverse transcriptase (mTert) expression marks slowly cycling intestinal stem cells. *Proc. Natl Acad. Sci. USA* **108**, 179–184 (2011).
35. Takeda, N. *et al.* Interconversion between intestinal stem cell populations in distinct niches. *Science* **334**, 1420–1424 (2011).
36. Tian, H. *et al.* A reserve stem cell population in small intestine renders *Lgr5*-positive cells dispensable. *Nature* **478**, 255–259 (2011).
Shows, for the first time, the competence of other intestinal cells to acquire stem cell properties upon columnar basal cell depletion.
37. Munoz, J. *et al.* The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent '+4' cell markers. *EMBO J.* **31**, 3079–3091 (2012).
38. Itzkovitz, S. *et al.* Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nature Cell Biol.* **14**, 106–114 (2012).
39. Buczacki, S. J. *et al.* Intestinal label-retaining cells are secretory precursors expressing Lgr5. *Nature* **495**, 65–69 (2013).
40. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010).
41. van Es, J. H. *et al.* Dll1$^+$ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biol.* **14**, 1099–1104 (2012).
42. Blanpain, C. & Fuchs, E. Epidermal homeostasis: a balancing act of stem cells in the skin. *Nature Rev. Mol. Cell Biol.* **10**, 207–217 (2009).
43. Mackenzie, I. C. Relationship between mitosis and the ordered structure of the stratum corneum in mouse epidermis. *Nature* **226**, 653–655 (1970).
44. Potten, C. S. The epidermal proliferative unit: the possible role of the central basal cell. *Cell Tissue Kinet.* **7**, 77–88 (1974).
45. Potten, C. S. & Loeffler, M. Epidermal cell proliferation. I. Changes with time in the proportion of isolated, paired and clustered labelled cells in sheets of murine epidermis. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **53**, 279–285 (1987).
46. Loeffler, M., Potten, C. S. & Wichmann, H. E. Epidermal cell proliferation. II. A comprehensive mathematical model of cell proliferation and migration in the basal layer predicts some unusual properties of epidermal stem cells. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **53**, 286–300 (1987).
47. Potten, C. S. & Morris, R. J. Epithelial stem cells *in vivo*. *J. Cell Sci. Suppl.* **10**, 45–62 (1988).
48. Barrandon, Y. & Green, H. Three clonal types of keratinocyte with different capacities for multiplication. *Proc. Natl Acad. Sci. USA* **84**, 2302–2306 (1987).
First *ex vivo* functional demonstration of the proliferative heterogeneity within the epidermis.
49. Jones, P. H. & Watt, F. M. Separation of human epidermal stem cells from transit amplifying cells on the basis of differences in integrin function and expression. *Cell* **73**, 713–724 (1993).
First study showing that cells with higher proliferation potential can be isolated by flow cytometry from a complex epithelial tissue.
50. Jones, P. H., Harper, S. & Watt, F. M. Stem cell patterning and fate in human epidermis. *Cell* **80**, 83–93 (1995).
51. Mackenzie, I. C. Retroviral transduction of murine epidermal stem cells demonstrates clonal units of epidermal structure. *J. Invest. Dermatol.* **109**, 377–383 (1997).
52. Kolodka, T. M., Garlick, J. A. & Taichman, L. B. Evidence for keratinocyte stem cells *in vitro*: long term engraftment and persistence of transgene expression from retrovirus-transduced keratinocytes. *Proc. Natl Acad. Sci. USA* **95**, 4356–4361 (1998).
53. Ghazizadeh, S. & Taichman, L. B. Multiple classes of stem cells in cutaneous epithelium: a lineage analysis of adult mouse skin. *EMBO J.* **20**, 1215–1222 (2001).
54. Ro, S. & Rannala, B. Evidence from the stop-EGFP mouse supports a niche-sharing model of epidermal proliferative units. *Exp. Dermatol.* **14**, 838–843 (2005).
55. Ro, S. & Rannala, B. A stop-EGFP transgenic mouse to detect clonal cell lineages generated by mutation. *EMBO Rep.* **5**, 914–920 (2004).
56. Clayton, E. *et al.* A single type of progenitor cell maintains normal epidermis. *Nature* **446**, 185–189 (2007).
First quantitative clonal analysis of the interfollicular epidermis.
57. Doupe, D. P., Klein, A. M., Simons, B. D. & Jones, P. H. The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate. *Dev. Cell* **18**, 317–323 (2010).
58. Lapouge, G. *et al.* Identifying the cellular origin of squamous skin tumors. *Proc. Natl Acad. Sci. USA* **108**, 7431–7436 (2011).
59. Mascre, G. *et al.* Distinct contribution of stem and progenitor cells to epidermal maintenance. *Nature* **489**, 257–262 (2012).
60. Lavker, R. M. & Sun, T.-T. Heterogeneity in epidermal basal keratinocytes: morphological and functional correlations. *Science* **215**, 1239–1241 (1982).
61. Lavker, R. M. & Sun, T.-T. Epidermal stem cells. *J. Invest. Dermatol.* **81**, 121s–127s (1983).
62. Tani, H., Morris, R. J. & Kaur, P. Enrichment for murine keratinocyte stem cells based on cell surface phenotype. *Proc. Natl Acad. Sci. USA* **97**, 10960–10965 (2000).
63. Li, A., Simmons, P. J. & Kaur, P. Identification and isolation of candidate human keratinocyte stem cells based on cell surface phenotype. *Proc. Natl Acad. Sci. USA* **95**, 3902–3907 (1998).
64. Ito, M. *et al.* Stem cells in the hair follicle bulge contribute to wound repair but not to homeostasis of the epidermis. *Nature Med.* **11**, 1351–1354 (2005).
65. Levy, V., Lindon, C., Zheng, Y., Harfe, B. D. & Morgan, B. A. Epidermal stem cells arise from the hair follicle after wounding. *FASEB J.* **21**, 1358–1366 (2007).
66. Snippert, H. J. *et al. Lgr6* marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science* **327**, 1385–1389 (2010).
67. Jaks, V. *et al.* Lgr5 marks cycling, yet long-lived, hair follicle stem cells. *Nature Genet.* **40**, 1291–1299 (2008).
68. Doupe, D. P. *et al.* A single progenitor population switches behavior to maintain and repair esophageal epithelium. *Science* **337**, 1091–1093 (2012).
69. Rossi, D. J. *et al.* Cell intrinsic alterations underlie hematopoietic stem cell aging. *Proc. Natl Acad. Sci. USA* **102**, 9194–9199 (2005).
70. Conboy, I. M. *et al.* Rejuvenation of aged progenitor cells by exposure to a young systemic environment. *Nature* **433**, 760–764 (2005).
71. Giangreco, A., Qin, M., Pintar, J. E. & Watt, F. M. Epidermal stem cells are retained *in vivo* throughout skin aging. *Aging Cell* **7**, 250–259 (2008).
72. Gupta, V. & Poss, K. D. Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature* **484**, 479–484 (2012).

73. Tzouanacou, E., Wegener, A., Wymeersch, F. J., Wilson, V. & Nicolas, J. F. Redefining the progression of lineage segregations during mammalian embryogenesis by clonal analysis. *Dev. Cell* **17**, 365–376 (2009).
74. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189–193 (2011).
First lineage-tracing and clonal analysis during mammary gland development and homeostasis.
75. Wang, X. *et al.* A luminal epithelial stem cell that is a cell of origin for prostate cancer. *Nature* **461**, 495–500 (2009).
76. Liu, J. *et al.* Regenerated luminal epithelial cells are derived from preexisting luminal epithelial cells in adult mouse prostate. *Mol. Endocrinol.* **25**, 1849–1857 (2011).
77. Choi, N., Zhang, B., Zhang, L., Ittmann, M. & Xin, L. Adult murine prostate basal and luminal cells are self-sustained lineages that can both serve as targets for prostate cancer initiation. *Cancer Cell* **21**, 253–265 (2012).
78. Ousset, M. *et al.* Multipotent and unipotent progenitors contribute to prostate postnatal development. *Nature Cell Biol.* **14**, 1131–1138 (2012).
79. Reya, T., Morrison, S. J., Clarke, M. F. & Weissman, I. L. Stem cells, cancer, and cancer stem cells. *Nature* **414**, 105–111 (2001).
80. Magee, J. A., Piskounova, E. & Morrison, S. J. Cancer stem cells: impact, heterogeneity, and uncertainty. *Cancer Cell* **21**, 283–296 (2012).
81. Driessens, G., Beck, B., Caauwe, A., Simons, B. D. & Blanpain, C. Defining the mode of tumour growth by clonal analysis. *Nature* **488**, 527–530 (2012).
First quantitative clonal analysis of tumour growth.
82. Schepers, A. G. *et al.* Lineage tracing reveals Lgr5$^+$ stem cell activity in mouse intestinal adenomas. *Science* **337**, 730–735 (2012).
This study, together with reference 81, supports the existence of cancer stem cells during tumour growth in its natural environment.
83. Zomer, A. *et al.* Intravital imaging of cancer stem cell plasticity in mammary tumors. *Stem Cells* **31**, 602–606 (2013).
84. Nakagawa, T., Sharma, M., Nabeshima, Y., Braun, R. E. & Yoshida, S. Functional hierarchy and reversibility within the murine spermatogenic stem cell compartment. *Science* **328**, 62–67 (2010).
85. Bystrykh, L. V., Verovskaya, E., Zwart, E., Broekhuis, M. & de Haan, G. Counting stem cells: methodological constraints. *Nature Methods* **9**, 567–574 (2012).
86. Lu, R., Neff, N. F., Quake, S. R. & Weissman, I. L. Tracking single hematopoietic stem cells *in vivo* using high-throughput sequencing in conjunction with viral genetic barcoding. *Nature Biotechnol.* **29**, 928–933 (2011).
87. Kreso, A. *et al.* Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543–548 (2013).
88. Wilson, A. *et al.* Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell* **135**, 1118–1129 (2008).
89. Klein, A. M. & Simons, B. D. Universal patterns of stem cell fate in cycling adult tissues. *Development* **138**, 3103–3111 (2011).

Acknowledgements
B.D.S. thanks A. Klein for important discussions and contributions, and he gratefully acknowledges the financial support of the Wellcome Trust (grant number 098357/Z/12/Z). C.B. is an investigator of Walloon Excellence in Life Science and Biotechnology (WELBIO), and he is supported by the Belgian Fund for Scientific Research (FNRS) and the European Research Council (ERC).

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
Cédric Blanpain's homepage:
http://blanpainlab.ulb.ac.be/index.htm

ALL LINKS ARE ACTIVE IN THE ONLINE PDF