*Manuscript
Click here to view linked References

Tracing cellular dynamics in tissue development, maintenance and disease

Steffen Rulands[1,2,3,*] and Benjamin D. Simons[1,2,3,*]

[1]Cavendish Laboratory, Department of Physics, J. J. Thomson Avenue, Cambridge CB3 0HE, UK.

[2]The Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Tennis Court Road, Cambridge CB2 1QN, UK.

[3]Wellcome Trust-Medical Research Council Stem Cell Institute, University of Cambridge, UK.

* Corresponding authors

Abstract

The coordination of cell proliferation and differentiation is central to the development and maintenance of tissues, while its dysregulation underlies the transition to diseased states. By combining lineage tracing with transcriptional profiling and marker-based assays, statistical methods are delivering insights into the dynamics of stem cells and their developmental precursors. These studies have provided evidence for molecular heterogeneity and fate priming, and have revealed a role for stochasticity in stem cell fate, refocusing the search for regulatory mechanisms. At the same time, they present a quantitative platform to study the initiation and progression of disease. Here, we review how quantitative lineage tracing strategies are shaping our understanding of the cellular mechanisms of tissue development, maintenance and disease.

Main text

To address the factors that control the development and maintenance of tissues, emphasis has been placed on defining the molecular mechanisms that regulate cell fate choice. Through marker-based assays and gene knockouts, significant progress has been made in resolving elements of the gene regulatory circuitry and signaling pathways that coordinate stem and progenitor cell activity. However, cells function

in a heterogeneous and dynamic environment where gene expression levels adjust in response to promoter activity and environmental cues. Therefore, to define the mechanisms that underpin cell fate, transcriptional information must be integrated with dynamic measures that address functional behaviour.

In the earliest form of lineage tracing, tritiated thymidine incorporation was used as a surrogate clonal marker [1]. Using such methods, C.P. LeBlond, a pioneer of modern stem cell biology, made significant progress in defining stem cell identity in a variety of epithelial tissues [2–4]. Later, lineage tracing strategies based on the incorporation of marker genes by electroporation [5] and lentiviral transfection [6], provided the means to impart hereditary marks on cells. However, it was not until the advent of transgenic animal technology that it became possible to trace the fate of individual cells and their progeny using targeted promoters. These days, intra-vital imaging platforms are providing access to continuous-time lineage data [7–9], while single-cell DNA sequencing methods offer the potential to resolve phylogenies in human tissues [10–12].

Advances in lineage tracing and molecular profiling techniques provide access to information at the cellular and molecular scale. Yet the integration of these measures into new mechanistic insights presents a formidable challenge. In recent years, the development of quantitative statistical methods, based on concepts from statistical physics and mathematics, are gaining traction in biology. These approaches, which place emphasis on “emergent behaviour” and scaling phenomena, provide constraints on the molecular mechanisms that regulate proliferative activity, fate choice and the collective dynamics of stem and progenitor cell populations. Beginning with a review of current lineage tracing strategies, here we provide an overview of how quantitative analysis of clonal data can be used to define mechanisms of cell fate choice in tissue development, maintenance and diseased states.

Lineage tracing strategies

Through advances in transgenics, multiple tracing strategies are available [13,14]. In the **Cre-lox system**, the expression of Cre recombinase is placed under the control of a cell-specific promoter [15]. Following the administration of a drug-inducing agent, the transient expression of Cre leads to the excision of a stop cassette, resulting in the activation of a fluorescent reporter gene. By using targeted promoters, this method allows the controlled labelling of distinct subpopulations at clonal or mosaic density (Figure 1a,b). By coupling the fluorescent reporter to the activation of an oncogene, such methods are increasingly deployed to trace steps in tumour progression, from initiation to invasion [16].

Despite their power, the reliability of genetic labelling assays may be compromised by (the lack of) promoter specificity, toxicity of the Cre enzyme or drug-inducing agent [17], and the faithful identification of clones. In the process of tissue or tumour development, large-scale cell rearrangements may lead to the fragmentation of labelled clones into disconnected cell clusters. Similarly, at non-clonal labelling density, initially separate clones may merge and form cohesive groups of labelled cells of non-clonal origin. Although the impact of clone merger and fragmentation events may be mitigated by multicolour (Brainbow or Confetti) reporter constructs [18,19], such effects cannot be eradicated. However, recent progress in the development of statistical analytical methods provide rigorous quantitative measures that allow lineage relationships to be inferred even from non-clonal density labelling [20].

Although clonal tracing studies provide indirect insight into the fate of marked cells, as a static measure based on the analysis of fixed samples, detailed information on individual fate histories is invariably lost. In the **Mosaic Analysis with Double Marker (MADM) system**, through Cre-induced mitotic recombination, fluorescent reporter genes of distinct colour can be activated in sister cells (Figure 1c,d), allowing definitive lineage information to be recovered [21]. Moreover, by correlating the expression of an oncogene with only one of the fluorescent proteins, the MADM system allows the mutual interaction of mutant and wild-type clones to be studied in situ.

The dispersion of cells in the blood system renders traditional "few-colour" labelling strategies impractical as a lineage tracing system. Here, efforts have been made to increase the spectrum of lineage labels so that clonal information can be recovered. To this end, approaches have been developed based on the incorporation of **genetic "barcodes"** using **lentiviral vectors** [22] as well as the **activation of transposons** [23].

To study proliferation kinetics and resolve slow-cycling cell populations, emphasis has been placed on the **H2B–GFP system** in which the expression of GFP fused to histone 2B is made dependent on the presence of Doxycycline (DOX) [24]: In this approach, H2B-GFP is homogeneously expressed in dividing cells. Upon DOX administration, H2B-GFP production is repressed such that, post-induction, the cellular concentration of H2B-GFP is approximately halved on division (Figure 1e,f). By measuring the intensity of the GFP signal, this pulse-chase strategy allows the distribution of cell divisions to be resolved, both at the population level (by FACS) and spatially (by microscopy). Although such methods have been used to identify quiescent cell populations [25–28], their potential to probe cellular dynamics and fate behaviour of actively cycling progenitors [29,30] remains underexploited.

Following advances in multi-photon microscopy, methods based on **intra-vital imaging** have enabled lineage information and spatial dynamics of epithelial cell populations to be resolved *in vivo* (Figure 1g,h) [7–9]. Technical limitations restrict continuous live-imaging of mouse models to a few days. Although time-lapse imaging over discrete time intervals allows the dynamics of individual clones to be tracked over periods of months or more, ambiguities in the intermediate fate histories prohibit the faithful reconstruction of lineage histories. Nevertheless, with the ability to correlate cell fate choice with position, these approaches have revealed evidence of dynamic heterogeneity, lineage priming and the flexibility of cellular states, providing new insights into the molecular mechanisms regulating stem cell identity and lineage commitment [31–34].

Finally, to trace cell fate behaviour in human tissues, attention has focused on the ability of naturally occurring DNA mutations to serve as a hereditary label (Figure 1i). In one approach, the accumulation of **mitochondrial DNA mutation** (mtDNA) [35,36]

has been used as a surrogate clonal marker [37,38]. In a related approach, the acquisition of **single nucleotide polymorphisms** have been used as a hereditary mark allowing the allele frequency of clonally related cells to be recovered by exome deep sequencing [11,22,39]. Further, advances in single-cell DNA sequencing offer the potential to recover phylogenic information, with applications to both normal and diseased tissues [40]. However, although these approaches provide a valuable window on cell dynamics in human tissues, the neutrality of mtDNA and DNA mutations call into question their reliability.

Clonal distributions as a record of cell fate choice

Although the lineage potential and proliferative capacity of labelled cells is reflected in clone size and composition, for a given clone, such static measures can rarely select from among the multiplicity of potential fate histories the one was actually followed. However, if cell fate behaviour is conserved across a subpopulation, by recovering the distribution of clone size and cell composition from a statistical ensemble, analytical methods based on population dynamics and inference techniques can be used to select the model that best predicts the experimental outcome. By analogy, the balance of a weighted die cannot be discerned from a single throw; but it can be resolved from an analysis of a statistical ensemble of throws. Of course, with the innate complexity of a developing tissue, the challenge of deciphering pattern of cell fate choice from static clonal data may be better compared to recovering the rules of chess from the statistical ensemble of chessboard configurations!

Fortunately, in many cases, the convergence of clone size distributions onto stereotypic behaviours allows rigorous aspects of cell dynamics to be inferred. For example, to achieve homeostasis, the maintenance of cycling adult tissues must rely on the activity of an equipotent stem cell population in which cell proliferation is perfectly balanced by differentiation. Such fate asymmetry may be invariant, achieved at the level of each and every stem cell division. Alternatively, asymmetry may be achieved only at the level of the population where chance stem cell loss through differentiation is compensated by cell duplication so that the overall stem

cell number remains constant (Figure 2a). Through this process of "neutral" cell competition, the chance $C(n,t)$ of a marked cell giving rise to a surviving clone with a size larger than n cells after a time t converges onto a "scaling" form, $C(n,t) = f(n/\langle n(t)\rangle)$, with $\langle n(t)\rangle$ the average clone size and f the scaling function [41], i.e. while the average clone size increases to compensate for the chance differentiation and loss of clones (Figure 2b,c), the chance of finding a clone with a size larger than some multiple of the average remains constant over time (Figure 2d,e). Through the particular form of the scaling function, f, it is possible to determine whether the balance between such stochastic stem cell loss and replacement follows from cell-autonomous regulation or is mediated by extrinsic cues such as neutral competition for limited niche access.

Such approaches have been successful in resolving fate behaviour in multiple epithelial tissue types [42] including interfollicular epidermis [29,43], oesophagus[44], testis [45] and trachea [46], and have explained the nature of stem cell dynamics in the intestinal crypt and the drift towards clonal "fixation"[47] . However, despite the success of these quantitative methods, such long-term scaling dependences may mask short-term bias in survival potential (fate priming) and the sub-lineage dependences of stem cell progeny. To dissect information on transient fate bias, emphasis must be placed on short-term clone size dependences and the combination of targeted promoters, which offer a "stereoscopic" perspective on lineage behaviour, or live-imaging approaches that can resolve the complete lineage history of marked cells [32,34].

Outside homeostasis, the dynamics of clonal evolution is less restricted. As a result, the resolution of cell fate behaviour has proved more challenging. These difficulties are exacerbated by the escalation of clone merger and fragmentation events in growth. Nevertheless, when clonal data can be recovered, quantitative information on lineage potential can often be deciphered. In one example, by targeting radial glial-like (RGL) precursors using the MADM labelling system, quantitative analysis of sister clone size dependence has shown that mouse cortical neurogenesis relies on a remarkably deterministic cell-autonomous programme [48]. Through the correlation of sister clone sizes, these studies revealed that the entry of RGLs into their

neurogenesis is not sporadic, but is tightly-regulated so that, in their proliferative phase, RGLs move sequentially through a defined cascade of symmetrical divisions. Following entry into neurogenesis, RGLs then make a series of invariant asymmetric cell divisions, giving rise to a “quantal” neuronal output before finally entering into gliogenesis.

In a second case study, a multicolour Confetti labelling system has been used to address the lineage and proliferative potential of early Mesp1 expressing cells in the developing mouse heart (Figure 3a) [49,50]. Here, one or few cells were labelled at several time points during gastrulation and their fate was analysed at later times when key stages of cardiac development were completed. In this case, large-scale tissue rearrangements lead to the merger and fragmentation of clones, making lineage analysis problematic (Figure 3b). Nevertheless, using statistical inference methods, clonal information could be inferred from clone fragments during late stage development, from which the early lineage specification of Mesp1 expression precursors was resolved (Figure 3c). As in this case, clonal information is often obtained from sections, which renders the reconstruction of the full clone size distribution unfeasible. However, by taking into account potential cell migratory processes between adjacent sections, further information on cell fate behaviour of precursor cells can be recovered from sectional data alone (Figure 3d). In the present case, quantitative analysis of the sectional clone size distribution (Figure 3e) revealed that cardiac development follows from the early lineage restriction of some 250 Mesp1 precursors that, like RGLs, are specified with largely defined proliferative potential.

Alongside normal developmental processes, genetic labelling approaches have been used to study tumour initiation following oncogenic transformation. In contrast to the coordinated and predictable lineage progression of cell precursors through development, pre-neoplastic transformation and invasive carcinoma follow from a sequence of sporadic transitions that create the hallmark heterogeneity of the tumour environment. On this background, the value of quantitative clonal fate studies may be largely restricted to the earliest phases of transformation, where much of the cell regulatory programme remains intact. Pioneering lineage studies

based on lineage tracing following carcinogen treatment [51] or the clonal activation of oncogenes [52] have been used to quantify the bias in stem and progenitor cell fate of mutant cells over their wild-type neighbours providing a basis to understand the nature of field transformation. For example, the quantitative analysis of clonal evolution following the activation of K-ras or the deletion of Apc have been used to quantify the existence and degree of stem cell fate bias in the mouse intestinal crypt [53,54]. Similarly, quantitative lineage tracing assays have been used to investigate how the stochastic fate behaviour of oesophageal progenitors are perturbed by the activation of a mutation in the Notch signalling pathway [44].

Perspectives

To address the mechanisms that regulate the maintenance of cycling adult tissues, emphasis has been placed on models in which stem cells, defined by signature marker expression, are individually long-lived and progress one-way through a proliferative hierarchy. However, by targeting the functional behaviour of cells, quantitative lineage tracing studies have shown that stem cells are not individually long-lived, but are frequently lost and replaced [55]. Studies based on intra-vital imaging show that stem cells are often not homogeneous but, through signals from the niche, they transit reversibly between states primed for duplication and loss [7,32,56]. Moreover, lineage tracing studies using cell-specific promoters have shown that cells normally committed to differentiation may be recruited back into the stem cell compartment following injury [57]. Together these studies have questioned the molecular basis of stem cell heterogeneity, fate stochasticity and plasticity of cellular states.

Through advances in genetic lineage tracing, DNA sequencing, single-cell expression profiling and genome editing, the ability to probe and manipulate cellular states has never been greater. However, applied in isolation, such techniques offer only limited insight: From the relative abundance of mutant alleles we learn little of the cellular mechanisms that initiate and promote tumour progression. From the transcriptional profile of stem cells we learn little about lineage relationships and functional fate dependences. And from the “rules” of cell fate choice, we learn little of the

molecular circuitry that condition lineage potential and fate decisions. Future progress in developmental and cancer biology will rely on the integration of these complementary techniques, a major challenge that will draw as much on new conceptual and theoretical insights as technological advances.

Acknowledgements

B.D.S. acknowledges the support of the Wellcome Trust (grant number 098357/Z/12/Z) and the core funding to the Gurdon Institute by the Wellcome Trust and Cancer Research UK. S.R. is grateful for the support of the Herchel Smith foundation.

Special interest (•) or outstanding interest (••)

(•) Ref. 8 Ellenbroek and van Rheenen A concise review discussing how intra-vital live-imaging approaches provide a new tool to explore cell fate behaviour in mice under normal conditions of normal homeostasis and in cancer models.

(•) Ref. 11 Martincorena et al. This paper reports on an exome deep sequencing study of physiologically normal skin, illustrating how point mutations can provide a readout of mutant clone dynamics in human tissue.

(•) Ref. 20 Wuidart et al. With an emphasis on prostate and the mammary epithelium, this study combines experimental and statistical methods to provide a rigorous framework for assessing lineage relationships and stem cell fate in different organs and tissues.

(••) Ref. 23 Sun et al. A novel lineage tracing approach based on transposon tagging is introduced to explore the clonal dynamics of haematopoietic stem cells in mice.

(•) Ref. 28 Busch et al. This study shows that, in hematopoiesis, the equilibration between stem cells and their progeny happens on time scales comparable to the life-time of the animal.

(••) Ref. 31 Rompolas et al. Using live-imaging to track interfollicular epidermal stem cells in mice over several generations, this study exposes short-term correlations in the fate of sister cells, which are indicative of coordinated cell activity.

(••) Ref. 32 Ritsma et al. This study shows how the combination of intra-vital imaging, long-term lineage tracing and quantitative analysis can be used to study short-term fate priming of intestinal stem cells in mice.

(•) Ref. 38 Baker et al. This paper shows how sporadic mtDNA mutations in human colon can be used as a lineage label to trace the fate behaviour of intestinal stem cells.

(••) Ref. 39 Simons This study shows how the quantitative analysis of exome sequencing data can be used to study stem cell dynamics in normal tissues, and to identify and define the mutational signature of rare field transformations.

(••) Ref. 48 Gao et al. This paper shows how quantitative analysis of clonal lineage tracing data using the MADM system can provide lineage information on cell precursors in the development of the mouse neocortex.

(••) Ref. 49 Lescroart et al. The study reports on how quantitative statistical analysis of clonal fate data can be used to define the lineage potential and multiplicity of Mesp1 precursors in the developing mouse heart.

References

1. Leblond CP: **The time dimension in histology**. *Am. J. Anat.* 1965, **116**:1–27.

2. Messier B, LeBlond CP: **Cell proliferation and migration as revealed by**

radioautography after injection of thymidine-H3 into male rats and mice. *Am. J. Anat.* 1960, **106**:247–85.

3. LeBlond CP, Clermont Y: **Definition of the stages of the cycle of the seminiferous epithelium in the rat.** *Ann. N. Y. Acad. Sci.* 1952, **55**:548–73.

4. Enesco M, Leblond CP: **Increase in Cell Number as a Factor in the Growth of the Organs and Tissues of the Young Male Rat**. *J. Embryol. Exp. Morphol.* 1962, **10**.

5. Itasaki N, Bel-Vialar S, Krumlauf R: **"Shocking" developments in chick embryology: electroporation and in ovo gene expression.** *Nat. Cell Biol.* 1999, **1**:E203-7.

6. Holt CE, Garlick N, Cornel E: **Lipofection of cDNAs in the embryonic vertebrate central nervous system.** *Neuron* 1990, **4**:203–14.

7. Brown S, Greco V: **Stem cells in the wild: understanding the World of stem cells through intravital imaging.** *Cell Stem Cell* 2014, **15**:683–6.

8. Ellenbroek SIJ, van Rheenen J: **Imaging hallmarks of cancer in living mice.** *Nat. Rev. Cancer* 2014, **14**:406–18.

9. Katie C, Park S, Greco V: **Live imaging of stem cells: answering old questions and raising new ones**. *Curr. Opin. Cell Biol.* 2016, **43**.

10. Navin N, Kendall J, Troge J, Andrews P, Rodgers L, McIndoo J, Cook K, Stepansky A, Levy D, Esposito D, et al.: **Tumour evolution inferred by single-cell sequencing**. *Nature* 2011, **472**:90–94.

11. Martincorena I, Roshan A, Gerstung M, Ellis P, Van Loo P, McLaren S, Wedge DC, Fullam A, Alexandrov LB, Tubio JM, et al.: **High burden and pervasive positive selection of somatic mutations in normal human skin**. *Science (80-.).* 2015, **348**:880–886.

12. Gawad C, Koh W, Quake SR: **Single-cell genome sequencing: current state of**

the science. *Nat. Rev. Genet.* 2016, **17**:175–188.

13. Snippert HJ, Clevers H: **Tracking adult stem cells.** *EMBO Rep.* 2011, **12**:113–22.

14. Kretzschmar K, Watt FM: **Lineage tracing.** *Cell* 2012, **148**:33–45.

15. Sauer B, Henderson N: **Site-specific DNA recombination in mammalian cells by the Cre recombinase of bacteriophage P1.** *Proc. Natl. Acad. Sci. U. S. A.* 1988, **85**:5166–70.

16. Blanpain C: **Tracing the cellular origin of cancer.** *Nat. Cell Biol.* 2013, **15**:126–34.

17. 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* 2013, **12**:298–303.

18. Lichtman JW, Livet J, Sanes JR: **A technicolour approach to the connectome.** *Nat. Rev. Neurosci.* 2008, **9**:417–22.

19. Snippert HJ, van der Flier LG, Sato T, van Es JH, van den Born M, Kroon-Veenboer C, Barker N, Klein AM, van Rheenen J, Simons BD, et al.: **Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells.** *Cell* 2010, **143**:134–44.

20. Wuidart A, Ousset M, Rulands S, Simons BD, Van Keymeulen A, Blanpain C: **Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells**. *Genes Dev.* 2016, **30**:1261–1277.

21. Zong H, Espinosa JS, Su HH, Muzumdar MD, Luo L: **Mosaic analysis with double markers in mice.** *Cell* 2005, **121**:479–92.

22. Lu R, Neff NF, Quake SR, Weissman IL: **Tracking single hematopoietic stem cells in vivo using high-throughput sequencing in conjunction with viral genetic barcoding.** *Nat. Biotechnol.* 2011, **29**:928–33.

23. Sun J, Ramos A, Chapman B, Johnnidis JB, Le L, Ho Y-J, Klein A, Hofmann O,

Camargo FD: **Clonal dynamics of native haematopoiesis**. *Nature* 2014, **514**:322–327.

24. Blanpain C, Lowry WE, Geoghegan A, Polak L, Fuchs E: **Self-renewal, multipotency, and the existence of two cell populations within an epithelial stem cell niche.** *Cell* 2004, **118**:635–48.

25. Blanpain C, Fuchs E: **Epidermal homeostasis: a balancing act of stem cells in the skin.** *Nat. Rev. Mol. Cell Biol.* 2009, **10**:207–17.

26. Wilson A, Laurenti E, Oser G, van der Wath RC, Blanco-Bose W, Jaworski M, Offner S, Dunant CF, Eshkind L, Bockamp E, et al.: **Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair.** *Cell* 2008, **135**:1118–29.

27. Foudi A, Hochedlinger K, Van Buren D, Schindler JW, Jaenisch R, Carey V, Hock H: **Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells.** *Nat. Biotechnol.* 2009, **27**:84–90.

28. Busch K, Klapproth K, Barile M, Flossdorf M, Holland-Letz T, Schlenner SM, Reth M, Höfer T, Rodewald H-R: **Fundamental properties of unperturbed haematopoiesis from stem cells in vivo.** *Nature* 2015, **518**:542–6.

29. Mascré G, Dekoninck S, Drogat B, Youssef KK, Broheé S, Sotiropoulou PA, Simons BD, Blanpain C: **Distinct contribution of stem and progenitor cells to epidermal maintenance.** *Nature* 2012, **489**:257–62.

30. Sada A, Jacob F, Leung E, Wang S, White BS, Shalloway D, Tumbar T: **Defining the cellular lineage hierarchy in the interfollicular epidermis of adult skin**. *Nat. Cell Biol.* 2016, **18**:619–631.

31. Rompolas P, Deschene ER, Zito G, Gonzalez DG, Saotome I, Haberman AM, Greco V: **Live imaging of stem cell and progeny behaviour in physiological hair-follicle regeneration.** *Nature* 2012, **487**:496–9.

32. Ritsma L, Ellenbroek SIJ, Zomer A, Snippert HJ, de Sauvage FJ, Simons BD,

Clevers H, van Rheenen J: **Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging.** *Nature* 2014, **507**:362–5.

33. Hara K, Nakagawa T, Enomoto H, Suzuki M, Yamamoto M, Simons BD, Yoshida S: **Mouse Spermatogenic Stem Cells Continually Interconvert between Equipotent Singly Isolated and Syncytial States.** *Cell Stem Cell* 2014, **14**:658–72.

34. Rompolas P, Mesa KR, Kawaguchi K, Park S, Gonzalez D, Boucher J, Klein AM, Greco V: **Spatiotemporal coordination of stem cell commitment during epidermal homeostasis**. *Science* 2016, **7012**:1–9.

35. Humphries A, Wright NA: **Colonic crypt organization and tumorigenesis.** *Nat. Rev. Cancer* 2008, **8**:415–24.

36. Walther V, Alison MR: **Cell lineage tracing in human epithelial tissues using mitochondrial DNA mutations as clonal markers.** *Wiley Interdiscip. Rev. Dev. Biol.* [date unknown], **5**:103–17.

37. Teixeira VH, Nadarajan P, Graham TA, Pipinikas CP, Brown JM, Falzon M, Nye E, Poulsom R, Lawrence D, Wright NA, et al.: **Stochastic homeostasis in human airway epithelium is achieved by neutral competition of basal cell progenitors.** *Elife* 2013, **2**:e00966.

38. Baker A-M, Cereser B, Melton S, Fletcher AG, Rodriguez-Justo M, Tadrous PJ, Humphries A, Elia G, McDonald SAC, Wright NA, et al.: **Quantification of Crypt and Stem Cell Evolution in the Normal and Neoplastic Human Colon**. *Cell Rep.* 2016, **8**:940–947.

39. Simons BD: **Deep sequencing as a probe of normal stem cell fate and preneoplasia in human epidermis.** *Proc. Natl. Acad. Sci. U. S. A.* 2016, **113**:128–33.

40. Potter NE, Ermini L, Papaemmanuil E, Cazzaniga G, Vijayaraghavan G, Titley I, Ford A, Campbell P, Kearney L, Greaves M: **Single-cell mutational profiling**

and clonal phylogeny in cancer. *Genome Res.* 2013, **23**:2115–25.

41. Klein AM, Simons BD: **Universal patterns of stem cell fate in cycling adult tissues.** *Development* 2011, **138**:3103–11.

42. Simons BD, Clevers H: **Strategies for homeostatic stem cell self-renewal in adult tissues.** *Cell* 2011, **145**:851–62.

43. Clayton E, Doupé DP, Klein AM, Winton DJ, Simons BD, Jones PH: **A single type of progenitor cell maintains normal epidermis.** *Nature* 2007, **446**:185–9.

44. Doupé DP, Alcolea MP, Roshan A, Zhang G, Klein AM, Simons BD, Jones PH: **A single progenitor population switches behavior to maintain and repair esophageal epithelium.** *Science* 2012, **337**:1091–3.

45. Klein AM, Nakagawa T, Ichikawa R, Yoshida S, Simons BD: **Mouse germ line stem cells undergo rapid and stochastic turnover.** *Cell Stem Cell* 2010, **7**:214–24.

46. Watson JK, Rulands S, Wilkinson AC, Wuidart A, Ousset M, Van Keymeulen A, Gottgens B, Blanpain C, Simons BD, Rawlins EL: **Clonal Dynamics Reveal Two Distinct Populations of Basal Cells in Slow-Turnover Airway Epithelium**. *Cell Rep* 2015, **12**:90–101.

47. Lopez-Garcia C, Klein AM, Simons BD, Winton DJ: **Intestinal stem cell replacement follows a pattern of neutral drift.** *Science* 2010, **330**:822–5.

48. Gao P, Postiglione MP, Krieger TG, Hernandez L, Wang C, Han Z, Streicher C, Papusheva E, Insolera R, Chugh K, et al.: **Deterministic progenitor behavior and unitary production of neurons in the neocortex.** *Cell* 2014, **159**:775–88.

49. Lescroart F, Chabab S, Lin X, Rulands S, Paulissen C, Rodolosse A, Auer H, Achouri Y, Dubois C, Bondue A, et al.: **Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development**. *Nat. Cell Biol.* 2014, **16**:829–840.

50. Chabab S, Lescroart F, Rulands S, Mathiah N, Simons BD, Blanpain C: **Uncovering the Number and Clonal Dynamics of Mesp1 Progenitors during Heart Morphogenesis**. *Cell Rep.* 2016, **14**:1–10.

51. Driessens G, Beck B, Caauwe A, Simons BD, Blanpain C: **Defining the mode of tumour growth by clonal analysis.** *Nature* 2012, **488**:527–30.

52. Schepers AG, Snippert HJ, Stange DE, van den Born M, van Es JH, van de Wetering M, Clevers H: **Lineage tracing reveals Lgr5+ stem cell activity in mouse intestinal adenomas.** *Science* 2012, **337**:730–5.

53. Kozar S, Morrissey E, Nicholson AM, van der Heijden M, Zecchini HI, Kemp R, Tavaré S, Vermeulen L, Winton DJ: **Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas.** *Cell Stem Cell* 2013, **13**:626–33.

54. Snippert HJ, Schepers AG, van Es JH, Simons BD, Clevers H: **Biased competition between Lgr5 intestinal stem cells driven by oncogenic mutation induces clonal expansion.** *EMBO Rep.* 2014, **15**:62–9.

55. Stine RR, Matunis EL: **Stem cell competition: finding balance in the niche.** *Trends Cell Biol.* 2013, **23**:357–64.

56. Hara K, Nakagawa T, Enomoto H, Suzuki M, Yamamoto M, Simons BD, Yoshida S: **Mouse spermatogenic stem cells continually interconvert between equipotent singly isolated and syncytial states.** *Cell Stem Cell* 2014, **14**:658–72.

57. Sánchez Alvarado A, Yamanaka S: **Rethinking differentiation: stem cells, regeneration, and plasticity.** *Cell* 2014, **157**:110–9.

58. Doupé DP, Klein AM, Simons BD, Jones PH: **The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate.** *Dev. Cell* 2010, **18**:317–23.

Figure Legends

Figure 1. Current *in vivo* lineage tracing methods. (a) Cre-lox recombination under the control of a cell-specific promoter allows hereditary labelling of targeted cell populations. While lineage history can rarely be inferred for an individual clone at a given time, the fate behaviour of cell populations can often be recovered from the quantitative analysis of a statistical distribution of clones. **(b)** Confocal images of EYFP+ basal cells in clones of mouse ear interfollicular epidermis obtained using AhcreERT R26EYFP/wt mice at 6 months and one-year post induction [58]. Scale bar is 20 μm. **(c)** The Mosaic Analysis with Double Markers (MADM) system relies on Cre-induced mitotic recombination. Fluorescent reporter genes of distinct color can be activated in two sister cells, leading to the expression of one or two markers. Definitive lineage information can be recovered from clones stemming from sister cells labelled in different colors. **(d)** Example of a MADM clone derived from mouse neocortex using Nestin-CreERT2;MADM showing lineage labelled neurons and glia. Sister cells are marked in YFP and RFP (S. Hippenmeyer, unpublished). **(e)** H2B-GFP is stably and homogeneously expressed in cycling cells. Upon DOX administration GFP intensity is halved at each round of cell division such that GFP levels record information about the number of cell divisions a given cell has undergone between DOX administration and analysis. **(f)** H2B-GFP dilution demonstrates regional variability in proliferation in mouse tail interfollicular epidermis [29]. Confocal image of H2B–GFP immunostaining in unchased mice (left) and 6 weeks after DOX administration (right). Slow-cycling cells are preferentially localised at the bottom of epidermal undulations (stars). Dashed lines are hair follicles. Scale bars denote 20 µm. **(g)** Intra-vital imaging provides spatial and temporal lineage information during a specified time window. H Intra-vital imaging of tumour cells in mouse mammary tissue, which were photo-converted from green to red using a violet laser. By providing spatial information, intra-vital imaging is capable of revealing migration patterns of cells [8].

Figure 2. Quantitative lineage tracing studies provide insight into cell fate behaviour in cycling homeostatic tissues. In adult, the fate of dividing stem cells (blue) is constrained by the condition of homeostasis. **(a)** Applied to the turnover of mouse oesophagus, quantitative lineage tracing studies using Ahcre^{ERT} R26$^{EYFP/wt}$ show that clonal evolution is consistent with a model in which progenitor divisions (at rate λ) results in either two differentiating cells (grey, with probability r), two stem cells (with probability r) or one stem cells and one differentiating cells (with probability $1-2r$) [43]. **(b)** According to this model, the average basal layer clone size is predicted to rise linearly over time, $\langle n \rangle \propto t$, to compensate for the ever-diminishing surviving clone fraction. (Points show experimental data and the line shows the linear prediction from the stochastic fate model in (a).) **(c)** The corresponding clone size distribution is predicted to conform to a scaling behaviour in which the chance of finding a clone with a size larger than some multiple of the average remains constant over time and collapses onto an exponential form, $f(x) = \exp(-x)$ (see main text). (Points show data from different time-points post-induction.) **(d)** In mouse spermatogenesis, the stochastic loss of germ line stem cells through differentiation is perfectly compensated by the duplication of neighbours along the one-dimensional seminiferous tubule. In this case, the average clone size is predicted to follow a square root time-dependence, $\langle n \rangle \propto \sqrt{t}$. (Points show data obtained using *GFRα1-CreER*^{T2}*;CAG-CAT-EGFP* mice and the line shows the theoretical prediction.) **(e)** In this case, the cumulative clone size distribution is also predicted to acquire a scaling form with $f(x) = \exp(-\pi x^2/4)$. (Points show experimental data from multiple time points.)

Figure 3. Lineage tracing in the developing mouse heart. (a) Confocal image at low magnification of a mosaically labeled heart in *Mesp1-Cre/Rosa-Confetti* embryo at E12 (Figure courtesy of Ref. [50]). **(b)** Due to cell migration and tissue deformation, labelled clones often fragment into disconnected clusters in developing tissues. Using Confetti labeling, Mesp1 expressing cells were constitutively labelled during

gastrulation. Image shows whole mount of a mouse heart analysed at E12.5 (from [49])**. (c)** In a situation where induction was almost clonal, Lescroart et al. used statistical inference to reconstruct the clonal provenance of fragmented clusters. Using this approach, the time points of lineage specification of Mesp1 precursors to the two heart fields could be inferred. The plot shows the frequency of Mesp1 expressing cells contributing to either the first (FHF) or the second heart field (SHF). **(c)** As in heart, where only the surface touching portion of fragments is accessible, clonal data is often sectional. While the full clone size cannot be reconstructed from sectional data, taking into account processes corresponding to the transfer of cells between layers their fate behavior can nevertheless be partly recovered. **(e)** For symmetrically dividing cells, the size distribution obtains a negative-binomial form (blue line), matching the cumulative clone size distribution (black line) and s.e.m. (gray) of E6.25 (FHF-enriched population) and E7.25 (SHF-enriched population) induced monoclonal-labeled hearts, showing the existence of equipotent progenitors inside these two populations of cardiac progenitors.

Figure 1

Figure 1

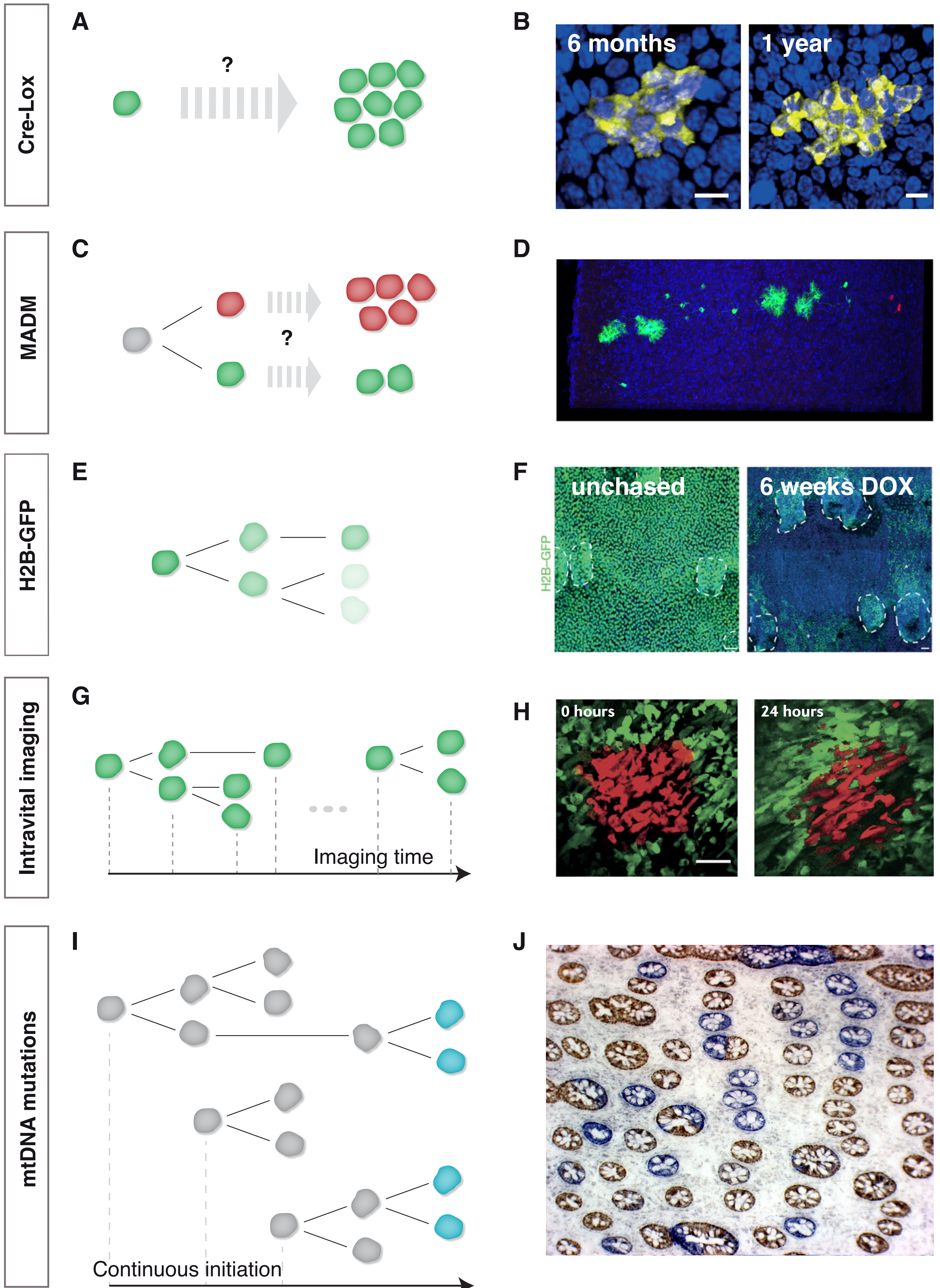


Figure 2

Figure 2

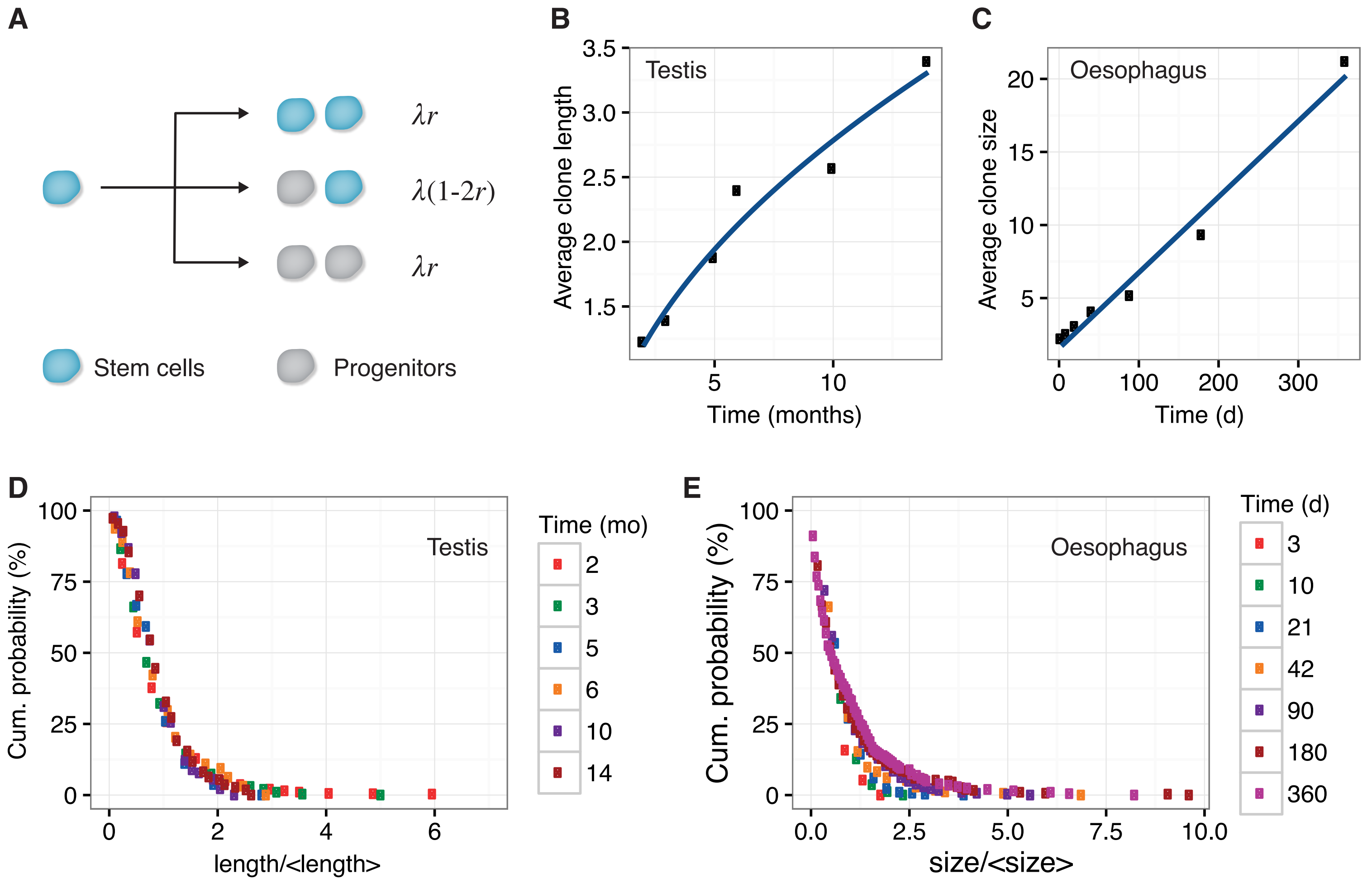
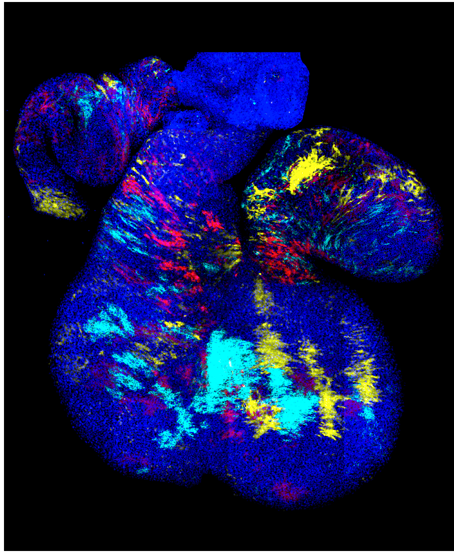


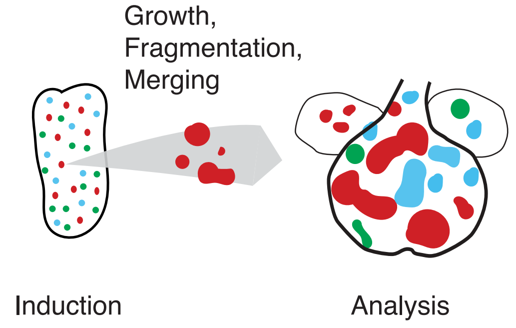
Figure 3

Figure 3

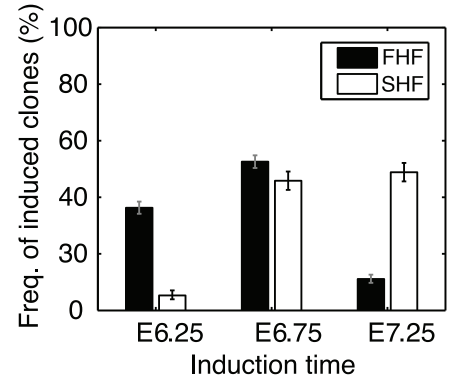
A



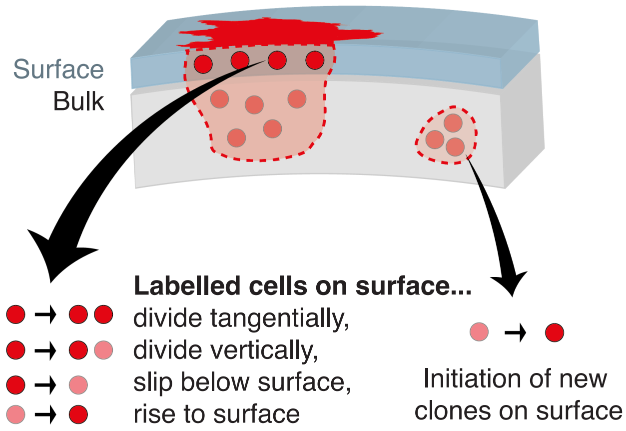
B



C



D



E