LETTERS

A single type of progenitor cell maintains normal epidermis

Elizabeth Clayton[1], David P. Doupé[1], Allon M. Klein[2], Douglas J. Winton[3], Benjamin D. Simons[2] & Philip H. Jones[1]

According to the current model of adult epidermal homeostasis, skin tissue is maintained by two discrete populations of progenitor cells: self-renewing stem cells; and their progeny, known as transit amplifying cells, which differentiate after several rounds of cell division[1–3]. By making use of inducible genetic labelling, we have tracked the fate of a representative sample of progenitor cells in mouse tail epidermis at single-cell resolution *in vivo* at time intervals up to one year. Here we show that clone-size distributions are consistent with a new model of homeostasis involving only one type of progenitor cell. These cells are found to undergo both symmetric and asymmetric division at rates that ensure epidermal homeostasis. The results raise important questions about the potential role of stem cells on tissue maintenance *in vivo*.

The mammalian epidermis is organized into hair follicles interspersed with interfollicular epidermis (IFE), which consists of layers of keratinocytes (Fig. 1a)[4]. In IFE, proliferating epidermal progenitor cells (EPCs) are found in the basal cell layer. On commitment to terminal differentiation, basal cells exit the cell cycle and subsequently migrate into the suprabasal cell layers. Progenitors capable of generating both hair follicles and IFE lie in the hair-follicle bulge, but these cells appear to play no part in maintaining normal interfollicular epidermis[5–9]. Label-retaining studies show that IFE contains slowly cycling basal cells, which have been interpreted as stem cells that support clonal units of transit amplifying (TA) and differentiated cells[10,11], according to the stem/TA cell hypothesis. However, these studies are unable to reveal the dynamics of EPC behaviour during epidermal homeostasis. Previous genetic labelling studies to track the fate of proliferating cells have either required epidermal injury or have yielded too few labelled clones to permit quantitative analysis[12–15].

To track EPC fate in normal epidermis we have used inducible genetic marking to label a sample of cells and their progeny in adult mice. Animals transgenic for the tamoxifen-regulated mutant of *cre* recombinase (Ah*cre*^{ERT}), expressed from the inducible CYP1A1 promoter, were crossed onto the R26$^{EYFP/EYFP}$ reporter strain, in which a conditional allele of enhanced yellow fluorescent protein (EYFP) is targeted to the Rosa26 locus (Supplementary Fig. S1a; refs 16, 17). In the resultant Ah*cre*^{ERT} R26$^{EYFP/wt}$ heterozygotes, EYFP is expressed in a dose-dependent manner following transient expression of *cre* induced by a treatment with βNF and tamoxifen at 6–9 weeks of age (Supplementary Fig. S1). Cohorts of mice were culled for analysis at intervals after a single injection of the inducing drugs. Cells expressing EYFP and their labelled progeny were detected by confocal microscopy of wholemount epidermis[18]. At 2 days post-induction, only singly labelled cells were seen, at a frequency of 1 in 600 cells in the basal layer, indicating that the clusters of cells encountered at later time points are clones, each derived from a single progenitor cell (Fig. 1b and data not shown). Analysis of subsequent cohorts of mice demonstrated clones that remained cohesive and expanded progressively in size (Fig. 1b, Supplementary Fig. S2). We scored clones that contained one or more basal cells; the observed clone-size distribution (that is, the total number of nucleated cells per clone) up to 6 weeks post-labelling, and the basal-layer clone-size distribution up to one year (see Methods) are shown in Fig. 2.

The density of labelled clones containing at least one basal-layer cell in tail epidermis rose from 2 days to a peak at 2 weeks after induction, as EYFP levels accumulated to detectable levels in all labelled cells. Clone numbers then fell to 7 ± 2% (mean ± s.d.) of the peak value by 3 months, and 3 ± 2% at one year; similar results were seen in back skin (data not shown). This decline was accompanied by the appearance of multi-cellular clones containing only suprabasal cells, consistent with clonal loss through differentiation (Supplementary Fig. S4). Analysis of the spatial distribution of IFE clones indicates that labelled clones are not replaced by unlabelled

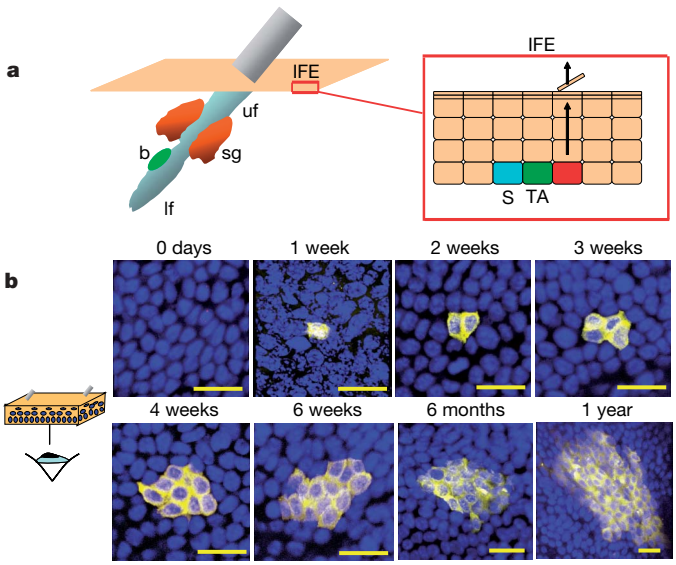

Figure 1 | *In vivo* clonal labelling of epidermal progenitor cells. **a**, Organization of the epidermis. Hair follicles contain stem cells located in the bulge (b, green), with the potential to generate lower hair follicle (lf), sebaceous gland (sg, orange) upper follicle (uf) and interfollicular epidermis (IFE, beige). The schematic shows the organization of keratinocytes in the IFE, as proposed by the stem/TA cell hypothesis. The basal layer comprises stem cells (S, blue), transit amplifying cells (TA, dark green), and post-mitotic basal cells (red), which migrate out of the basal layer as they differentiate (arrows). **b**, Projected *Z*-stack confocal images of IFE wholemounts from Ah*cre*^{ERT} R26$^{EYFP/wt}$ mice viewed from the basal surface at the times shown following induction. Yellow, EYFP; blue, DAPI nuclear stain. Scale bar, 20 μm.

[1]MRC Cancer Cell Unit, Hutchison-MRC Research Centre, Cambridge CB2 0XZ, UK. [2]Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK. [3]Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Robinson Way, Cambridge CB2 0RE, UK.

clones migrating from hair follicles (Supplementary Fig. S5). Moreover, none of the labelled clones can derive from bulge stem cells because this region is not labelled (Supplementary Fig. S1 and Supplementary results).

Before attempting to interpret the clone fate data, it is necessary to assess the extent to which they are influenced by tissue growth or apoptosis. First, the rate of increase in epidermal surface area due to growth was low (estimated at less than 3.5% per month over the time course of the experiment), whereas apoptosis was undetectable in basal-layer cells (see Supplementary results and Supplementary Fig. S6). Furthermore, the number of basal-layer cells per unit area and the proportion of cycling cells (as assessed both by Ki67 and cdc6 immunostaining) showed no significant difference between 2-week and one-year samples. Both techniques of assessing the proportion of cycling cells gave similar results, as did flow cytometry: $22 \pm 3\%$ (mean $\pm$ s.d.) for Ki67; $24 \pm 4\%$ for cdc6; and $22 \pm 1\%$ for flow cytometry (see Supplementary Fig. S7)[19,20]. Finally, there was no significant difference between the proportion of cycling cells in the labelled and unlabelled cell populations, either at 5 days or one year post-induction (see Supplementary results and Supplementary Fig. S2). We therefore conclude that basal-layer cells labelled at induction are typical of the entire basal cell population, and that the observed clonal evolution is representative of the adult system in a state of homeostasis.

According to the stem/TA cell hypothesis, TA cells undergo a limited number of cell divisions followed by differentiation[21]. To test this prediction, we examined clones at 3 weeks, over 90% of which are lost by 12 weeks post-induction. Significantly, clones comprising three or more cells contained both basal and suprabasal cells, indicative of asynchronous terminal differentiation (Fig. 3a). Furthermore, the immunostaining of clones consisting of two basal cells reveals that a single cell division may generate either one cycling and one non-cycling daughter, or two cycling daughters, or two non-cycling daughters (Fig. 3b). This raises the question of whether there is asymmetric cell division within the basal plane as described in the *Drosophila* peripheral nervous system and zebrafish retinal precursors[22,23]. Three-dimensional imaging of wholemount epidermis revealed that only 3% of mitotic spindles lie perpendicular to the basal layer, indicating that, in contrast to embryonic epidermis, the vast majority of EPC divisions generate two basal-layer cells (Supplementary Fig. S8; refs 24, 25). The observation of asymmetric partitioning of numb protein (which marks asymmetric division in neural and myogenic precursors) in clones consisting of two basal cells suggests that planar-orientated asymmetric division also occurs in the epidermis (Fig. 3c)[26,27]. EPC behaviour thus differs substantially from that observed in committed precursors in other systems[28,29].

We next considered the behaviour of the long-lived clones that persist for over 3 months. Within the stem/TA cell hypothesis, the epidermis is organized into epidermal proliferative units comprising about ten basal cells supported by a single self-renewing stem cell[11]. If individual stem cells retain their self-renewal capacity, the stem/TA cell model predicts that the basal-layer clone-size distribution must become time-independent and characteristic of a single epidermal proliferative unit (see Supplementary theory[21]). Such behaviour is in stark contrast to the progressive increase in average clone size observed in the epidermis (Fig. 2).

Faced with this apparent contradiction, one could attempt to revise the stem cell/TA cell model, but staying within the general paradigm. This might include introducing the capacity for stem-cell ageing and/or migration[15]. Alternatively, one could try to exploit the range of experimental data to seek evidence for a new paradigm for epidermal homeostasis. Intriguingly, such evidence is found in the scaling properties of the observed clone-size distribution. Here we argue that the clone fate data are compatible with a model in which IFE is maintained by only one compartment of proliferating cells.

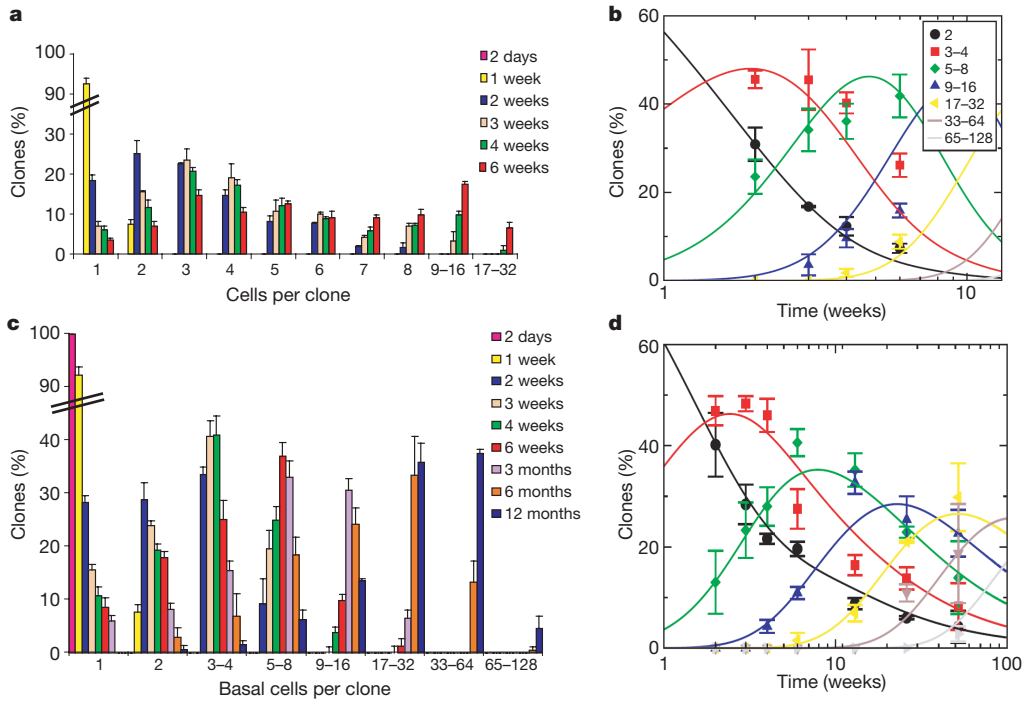

Figure 2 | Clone fate data. a, Distribution of clone size (total cells per clone) as a function of cell number, as measured at 2 days, 1, 2, 3, 4 and 6 weeks post-induction (error bars indicate s.e.m.). **c**, Distribution of basal cells per clone as a function of basal cell number, as measured at 2 days, 1, 2, 3, 4 and 6 weeks, 3, 6 and 12 months post-induction (error bars indicate s.e.m.). **b**, **d**, Distribution of clone size (total cells per clone) (**b**) and basal cells per clone (**d**) as a function of time for different values of cell number (error bars indicate s.e.m.). Here we have aggregated clone sizes in ranges increasing in size in powers of two (see legend within figure). To eliminate possible ambiguities due to labelling efficiency, single cell clones are eliminated from the distribution in **b** and **d**, thereby removing the population of post-mitotic cells labelled at induction. We focus on time points of 2 weeks or more post-induction when EYFP levels have stabilized. Continuous curves show the behaviour of the proposed one-progenitor-cell model with a cell division rate of $\lambda = 1.1$ per week and a symmetric division ratio of $r = 0.08$ (see main text for details).

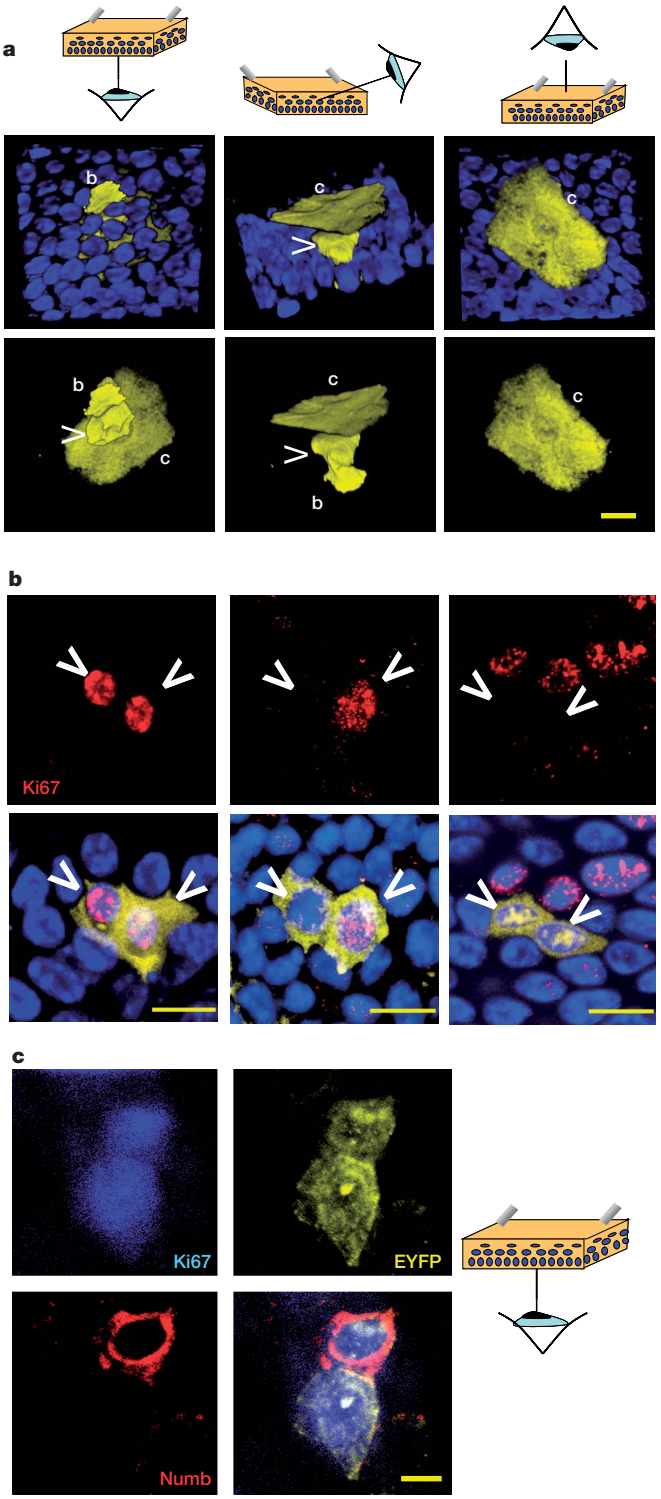

Figure 3 | Asymmetric cell fate in epidermal progenitors. a, Visualization of a three-cell clone exhibiting asynchronous terminal differentiation. Projected *Z*-stack images show one basal cell (labelled b), and two suprabasal cells: a cornified layer cell (labelled c), and a second suprabasal cell indicated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

Whether this model should be considered as an extreme variant of the stem/TA cell hypothesis or a new concept is arguably a matter of semantics, a point we will return to later.

To identify the scaling behaviour, we define the clone-size distribution $P_n(t)$, describing the probability that a labelled progenitor cell develops into a clone with a total of n basal-layer cells at time t after induction. From this we can define the distribution of 'persisting' clones, that is, the distribution of labelled clones containing at least one basal-layer cell:

$$P^{\mathrm{pers}}_{n>0}(t) \equiv \frac{P_n(t)}{1-P_0(t)}$$

With this definition, we show that (Fig. 4a and below), after an initial transient behaviour, the observed clone-size distributions are compatible with the simple scaling form:

$$P^{\mathrm{pers}}_{n>0}(t) = \frac{\tau}{t} f(n\tau/t) \qquad (1)$$

where τ denotes some constant timescale. From this striking observation, we deduce that, at long times, the average number of basal-layer cells within a persisting clone increases linearly with time, a behaviour inconsistent with the existence of long-lived cycling stem cells (see Supplementary theory). More significantly, the scaling indicates that long-time properties of clonal evolution are dictated by only one characteristic timescale τ, consistent with a simple model of clonal fate in which external factors, such as stem-cell ageing or skin injury, do not have a significant impact.

Taken together, all of our experimental observations and the scaling behaviour are consistent with a model of clonal fate involving only one type of EPC and just three adjustable parameters: the overall division rate λ of proliferating (labelled A-type) EPCs; the proportion of cell divisions that are asymmetric, $(1-2r)$; and the rate of transfer Γ of non proliferating (B-type) cells from the basal to the suprabasal layer (see schematic in Fig. 4b). To maintain a steady-state EPC population, the rates of symmetric cell division, A→A + A and A→B + B, must be identical and equal to r. Finally, the observation that the basal-layer cell density remains constant leads to the additional constraint $\Gamma = \lambda\rho/(1-\rho)$, where ρ denotes the proportion of proliferating cells in the basal layer, reducing the number of adjustable parameters to just two.

Defining $P_{n_A,n_B}(t)$ as the probability that a labelled clone involves n_{A} A-type and n_{B} B-type EPCs at time t after induction, its time-evolution is governed by the Master equation :

$$\frac{dP_{n_A,n_B}}{dt} = \lambda\Big\{ r\Big[(n_A-1)P_{n_A-1,n_B} + (n_A+1)P_{n_A+1,n_B-2}\Big] + (1-2r)n_A P_{n_A,n_B-1} - n_A P_{n_A,n_B}\Big\} + \Gamma\Big[(n_B+1)P_{n_A,n_B+1} - n_B P_{n_A,n_B}\Big]$$

subject to the initial condition $P_{n_A,n_B}(0) = \rho\,\delta_{n_A,1}\delta_{n_B,0} + (1-\rho)\,\delta_{n_A,0}\delta_{n_B,1}$. Although an exact analytical solution to this equation is unavailable, at times $t > 1/r\lambda$ the system enters an asymptotic regime where, defining $n = n_{\mathrm{A}} + n_{\mathrm{B}}$, we may show that the basal-layer clone-size distribution for persisting clones acquires the observed scaling form in equation (1) with $f(x) = e^{-x}$ and $\tau = \rho/r\lambda$; that is, the long-time properties of clonal evolution are dictated by the symmetric division rate, $r\lambda$. When combined with the experimentally inferred value for the fraction of proliferating basal-layer cells $\rho = 0.22$, a fit of the data to the asymptotic distribution (Fig. 4a, inset) identifies $r\lambda = 0.088 \pm 0.004$ per week.

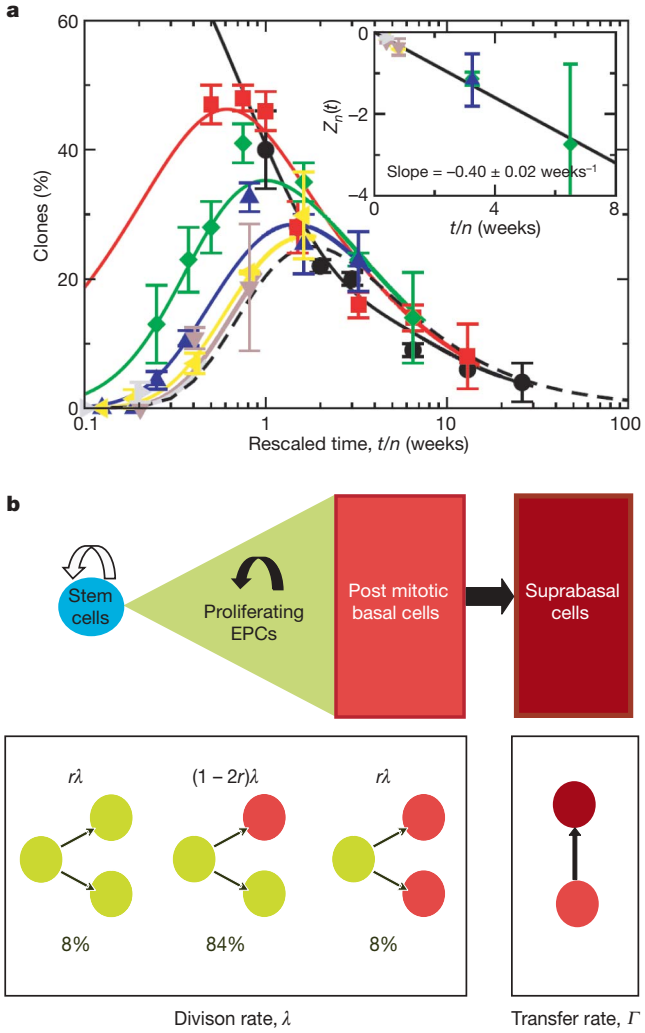

Figure 4 | Scaling and model of epidermal progenitor cell fate. a, The basal-layer clone-size distributions (see Fig. 2d legend for key) are replotted against the rescaled time coordinate, t/n (where n is taken as the upper limit for each distribution; for example, $n = 4$ for the range 3 to 4, and so on). We note that at long time points (>6 weeks), the data sets for different values of n converge onto a single curve (dashed line); that is, the probability of finding a labelled clone with a basal cell number in the range $n/2$ to n at time t post-induction is equal to that of finding a clone with a size in the range n to $2n$ at time $2t$. At shorter timescales, the transient behaviour dominates the distribution leading to a departure from simple scaling. This transient behaviour is very well described by the one progenitor cell compartment model (see Fig. 2 caption and main text). Making use of equation (1), we can identify the universal scaling curve for the grouped data as $G_n(t) \equiv \sum_{m>n/2}^{n} P_m^{\mathrm{pers}}(t) = e^{-n\tau/2t} - e^{-n\tau/t}$, where $\tau = \rho/r\lambda$; that is, $G_n(t/n)$ is independent of n. Therefore, by plotting $Z_n(t) = 1/(2\ln[(1-(1-G_n(t))^{1/2})/2])$ against t/n (inset) at long times (>13 weeks) and large n (>4), the resulting slope may be used to infer $-1/\tau$. **b**, The single progenitor compartment model of epidermal homeostasis. A single population of EPCs (green triangle), with unlimited self renewal potential (filled arrow) maintain the epidermis. Post-mitotic cells in the basal layer (light red) transfer at a steady rate to the suprabasal compartment (dark red). The model proposes no role for stem cells (blue) in the steady state, but a quiescent population may play a role in growth or regeneration following injury (unfilled arrow). The left-hand box shows the proliferative characteristics of EPCs. Cycling cells are shown in green, post-mitotic basal cells in light red. The right-hand box shows the transfer of basal post-mitotic cells to the suprabasal layers (dark red). Taken together, these processes with their respective rates summarize the one-compartment model of homeostasis discussed in the text.

At times $t < 1/r\lambda \approx 11$ weeks, the transient behaviour of the basal-layer clone-size distribution dominates. In this regime, both the basal-layer and total clone-size distributions can be determined from a numerical integration of the corresponding Master equation. Taking $r\lambda$ from the asymptotic data dependence, a one-parameter fit of the basal-layer clone-size distribution to the experimental data obtains a good quantitative agreement over the entire one-year time course for $r = 0.08$ (Fig. 2d), that is, EPC division takes place at a rate of $\lambda = 1.1$ per week (a figure consistent with previous estimates[21]), with 84% of divisions resulting in asymmetric fate, while the cell transfer rate out of the basal layer takes place at a rate of $\Gamma = 0.31$ per week. With the same choice of parameters, the total clone-size distribution also shows a striking quantitative agreement with experiment (Fig. 2b). It should be noted that the slow accumulation of labelled clones over the 2 weeks following induction has no significant effect on the fit of the model (see Supplementary results and Supplementary Fig. S10.)

In conclusion, we have shown that the entire range of clonal fate data reported here is compatible with a model involving a single proliferating cell compartment in mouse tail skin epidermis. Technical limitations prevented us carrying out a similar quantitative analysis in back skin; although the changes in clone diameter observed at this site are in qualitative agreement with a single-compartment model, we cannot exclude alternative models with the available data. It is also important to note that, if present, a small quiescent population of stem cells would be undetectable in our analysis, but would be expected to be highly active in processes such as wound healing[8]. Previous models of epidermal homeostasis hypothesize the existence of a TA cell compartment, which undergoes a limited number of divisions. We show that tail epidermis is maintained by a single population of progenitor cells, which may undergo an unlimited number of divisions.

METHODS

Animals and sample preparation. All animal experiments were conducted as specified by Home Office Project Licence. The generation of Ah*cre*ERT and $R26^{EYFP/EYFP}$ mice has been described previously[16,17]. When $Ah cre^{ERT} R26^{EYFP/wt}$ mice are treated with multiple doses of βNF and tamoxifen, a high level of recombination was seen in the upper hair follicle and IFE (Supplementary Fig. S1). The drug doses were titrated down to produce low-frequency labelling; a single intraperitoneal injection of 80 mg per kg β-naphthoflavone (Sigma-Aldrich) and 1 mg tamoxifen-free base (MP Biomedicals) dissolved in corn oil resulted in EYFP expression in approximately 1 in 600 basal cells of tail IFE and in 1 in 40 basal cells of back IFE, at 2 weeks post-induction[17]. No labelling was detected in the bulge region of the hair follicle and there was no background labelling in untreated $Ah cre^{ERT} R26^{EYFP/wt}$ animals, even at 15 months of age (Supplementary Fig. S1c, d and data not shown). Epidermal wholemounts were prepared as described[18]. For analysis of back epidermis, 60-μm cryosections were used.

Immunostaining and imaging. Immunostaining of wholemounts was performed as described[18]. The following primary antibodies were used; anti-GFP conjugated to AlexaFluor488 or 555 (Molecular Probes), anti-Ki67 (Abcam), anti-numb (Abcam), anti-cleaved caspase 3 (Cell Signalling Technology) and anti-α-tubulin conjugated to FITC (Sigma). Secondary antibodies were from Molecular Probes. Confocal images are presented as *Z*-stack projections; 30–120 optical sections in 0.2–2 μm increments were rendered using Improvision Volocity software.

Analysis of clone size, number and proliferation. Data presented is a typical example of at least two experiments with at least three mice per time point.

Tail epidermis. The patterned organization of tail epidermis enabled definition of a unit area of tail IFE between adjacent rows of hair follicles, which measured 282,000 ± 2,300 μm^2 and contained 4,870 ± 400 (mean ± s.d.) basal-layer cells (Supplementary Fig. S5a). The mean number of labelled clones per unit area of tail IFE was assessed by counting all clones detected by confocal imaging of 50 unit areas in each of three mice at each time point, except at 6 and 12 months when 100 areas were counted. Clone size was determined by *Z*-stack confocal imaging of at least 50 clones containing at least one basal cell in each mouse at each time point. We note that after 6 weeks anucleate cornified layer cells appeared in labelled clones, making it impossible to count total cell numbers; hence, the total number of cells per clone was counted up to 6 weeks, while the total number of basal cells could be scored up to one year. The percentage of

Ki67-positive cells was determined by imaging at least 1,500 basal cells in multiple fields from at least three mice.

Back skin epidermis. 60-μm cryosections were analysed by optical sectioning. To determine clone number, all clones in three 10-mm lengths of epidermis were analysed in each mouse at time points up to 3 months; at 6- and 12-month time points at least 5 cm of epidermis was scored for each mouse. The maximum clone diameter, expressed as number of basal cells, was scored for at least 50 clones in each of three mice at each time point.

Apoptosis. Wholemounts of tail epidermis from 2 days, 1 week, 2 weeks, 3 weeks, 4 weeks and 6 weeks after induction were stained with an anti-cleaved caspase3/AlexaFluor488 conjugate; positive caspase 3 staining was confined to catagen hair follicles. In further experiments, staining for cleaved caspase 3 and apoptotic bodies, visualized by DAPI, was examined in the basal layer in wholemounts from mice at 3 weeks and 6 months post-induction. As a positive control, epidermal wholemounts were irradiated with 160 mJ cm^{-2} ultraviolet (UV)C and analysed after a 16-hour incubation at 37 °C.

Analysis of mitotic spindle orientation. Wholemounts of DAPI-stained tail IFE were analysed by acquiring *Z*-stacks of all nuclei containing condensed chromosomes, as revealed by uniform, intense DAPI staining. Images were rendered in three dimensions, as above, and spindle orientation of all mitotic figures from metaphase to telophase was scored as described[22].

Analysis of *cre* expression. cre^{ERt} messenger RNA levels after induction were analysed by quantitative polymerase chain reaction with reverse transcription (RT–PCR) of RNA prepared from tail epidermis using Trizol (Sigma). Primers used were 5′-CGTACTGACGGTGGGAGAAT and 5′-CCCGGCAAAACAGGTAGTTA, and the product was detected using SyBr Green. GAPDH mRNA was measured with a Taqman probe (Applied Biosystems).

Flow cytometry. A single-cell suspension was prepared from tail epidermis, using a 30 min incubation with Dispase II (Roche), followed by digestion with trypsin EDTA; this method separates the interfollicular epidermis from the dermis and lower hair follicles. After staining with biotin-conjugated anti-β1 integrin antibody (BD Biosciences/Pharmingen) and Alexa488-streptavidin (Molecular Probes), samples were fixed with paraformaldehyde, permeabilized with 0.1% saponin with 100 μg ml^{-1} RNase A and 50 μg ml^{-1} propidium iodide, and analysed on a BD Facscalibur flow cytometer, using propidium iodide channel pulse area/width gating to exclude cell doublets.

Received 29 June 2006; accepted 24 November 2006.
Published online 28 February 2007.

1. Lajtha, L. G. Stem cell concepts. *Differentiation* **14,** 23–34 (1979).
2. Alonso, L. & Fuchs, E. Stem cells of the skin epithelium. *Proc. Natl Acad. Sci. USA* **100** (suppl. 1), 11830–11835 (2003).
3. Braun, K. M. & Watt, F. M. Epidermal label-retaining cells: background and recent applications. *J. Invest. Dermatol. Symp. Proc.* **9,** 196–201 (2004).
4. Gambardella, L. & Barrandon, Y. The multifaceted adult epidermal stem cell. *Curr. Opin. Cell Biol.* **15,** 771–777 (2003).
5. Tumbar, T. *et al.* Defining the epithelial stem cell niche in skin. *Science* **303,** 359–363 (2004).
6. Morris, R. J. *et al.* Capturing and profiling adult hair follicle stem cells. *Nature Biotechnol.* **22,** 411–417 (2004).
7. Levy, V., Lindon, C., Harfe, B. D. & Morgan, B. A. Distinct stem cell populations regenerate the follicle and interfollicular epidermis. *Dev. Cell* **9,** 855–861 (2005).
8. 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).
9. Claudinot, S., Nicolas, M., Oshima, H., Rochat, A. & Barrandon, Y. Long-term renewal of hair follicles from clonogenic multipotent stem cells. *Proc. Natl Acad. Sci. USA* **102,** 14677–14682 (2005).
10. Mackenzie, I. C. Relationship between mitosis and the ordered structure of the stratum corneum in mouse epidermis. *Nature* **226,** 653–655 (1970).
11. Potten, C. S. The epidermal proliferative unit: the possible role of the central basal cell. *Cell Tissue Kinet.* **7,** 77–88 (1974).
12. 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).
13. Kameda, T. *et al.* Analysis of the cellular heterogeneity in the basal layer of mouse ear epidermis: an approach from partial decomposition *in vitro* and retroviral cell marking *in vivo*. *Exp. Cell Res.* **283,** 167–183 (2003).
14. Ro, S. & Rannala, B. A stop-EGFP transgenic mouse to detect clonal cell lineages generated by mutation. *EMBO Rep.* **5,** 914–920 (2004).
15. 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).
16. Srinivas, S. *et al.* Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. *BMC Dev. Biol.* **1,** 4 (2001).
17. Kemp, R. *et al.* Elimination of background recombination: somatic induction of Cre by combined transcriptional regulation and hormone binding affinity. *Nucleic Acids Res.* **32,** e92 (2004).
18. Braun, K. M. *et al.* Manipulation of stem cell proliferation and lineage commitment: visualisation of label-retaining cells in wholemounts of mouse epidermis. *Development* **130,** 5241–5255 (2003).
19. Williams, G. H. *et al.* Improved cervical smear assessment using antibodies against proteins that regulate DNA replication. *Proc. Natl Acad. Sci. USA* **95,** 14932–14937 (1998).
20. Birner, P. *et al.* Immunohistochemical detection of cell growth fraction in formalin-fixed and paraffin-embedded murine tissue. *Am. J. Pathol.* **158,** 1991–1996 (2001).
21. Potten, C. S. Cell replacement in epidermis (keratopoiesis) via discrete units of proliferation. *Int. Rev. Cytol.* **69,** 271–318 (1981).
22. Das, T., Payer, B., Cayouette, M. & Harris, W. A. *In vivo* time-lapse imaging of cell divisions during neurogenesis in the developing zebrafish retina. *Neuron* **37,** 597–609 (2003).
23. Gho, M. & Schweisguth, F. Frizzled signalling controls orientation of asymmetric sense organ precursor cell divisions in *Drosophila*. *Nature* **393,** 178–181 (1998).
24. Lechler, T. & Fuchs, E. Asymmetric cell divisions promote stratification and differentiation of mammalian skin. *Nature* **437,** 275–280 (2005).
25. Smart, I. H. Variation in the plane of cell cleavage during the process of stratification in the mouse epidermis. *Br. J. Dermatol.* **82,** 276–282 (1970).
26. Zhong, W., Feder, J. N., Jiang, M. M., Jan, L. Y. & Jan, Y. N. Asymmetric localization of a mammalian numb homolog during mouse cortical neurogenesis. *Neuron* **17,** 43–53 (1996).
27. Conboy, I. M. & Rando, T. A. The regulation of Notch signaling controls satellite cell activation and cell fate determination in postnatal myogenesis. *Dev. Cell* **3,** 397–409 (2002).
28. Smart, F. M. & Venkitaraman, A. R. Inhibition of interleukin 7 receptor signaling by antigen receptor assembly. *J. Exp. Med.* **191,** 737–742 (2000).
29. Temple, S. & Raff, M. C. Clonal analysis of oligodendrocyte development in culture: evidence for a developmental clock that counts cell divisions. *Cell* **44,** 773–779 (1986).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank Y. Amagase for performing RT–PCR, E. Choolun and R. Walker for technical assistance, S. Penrhyn-Lowe and T. Mills for help with microscopy and R. Laskey, W. Harris, A. Philpott and C. Jones for comments. This work was funded by the Medical Research Council, Association for International Cancer Research and Cancer Research UK.

Author Contributions Experimental work was performed by E.C., D.P.D. and P.H.J., project planning by P.H.J. and D.J.W., biophysical analysis by B.D.S. and A.M.K.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to P.H.J. (phj20@hutchison-mrc.cam.ac.uk).

Supplementary Theory

To consolidate the discussion in the main text, in the following we will expand upon some of the theoretical considerations addressed in this work. In particular, we will elucidate the generic behaviour of the *classical* EPU model and the stem cell/TA cell hypothesis for clonal fate as well as elaborating upon the properties of the single-proliferative compartment model introduced in this work.

Classical EPU model and the predictions of the stem cell hypothesis

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

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

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

distribution, the model may be summarised by the rate equations,

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

where TA_n indexes the n-th generation of the TA cell population, and $N \sim 3-5$ [1] represents the number of rounds of division of the TA cell population. Here λ_S and λ_A (assumed, for simplicity, to be independent of generation) represent the division rates of stem and TA cells respectively, and Γ denotes the post-mitotic cell transfer rate from the basal to the suprabasal layer. Although one can conceive of possible variations, while one assumes that an EPU is maintained by a single long-lived stem cell, the model encapsulated by (1) exemplifies the general properties of any model based on the stem cell/TA cell hypothesis.

To analyse the model, we follow a similar method to that described in the main body of the text: Defining $P_{n_\mathrm{S},n_{\mathrm{A}_1},\cdots,n_{\mathrm{A}_N},n_\mathrm{B}}(t)$ as the probability for finding a given clone with n_S stem cells, $(n_{\mathrm{A}_1} + \cdots + n_{\mathrm{A}_N})$ TA cells, and n_B post-mitotic cells at some time t after induction, its time-evolution is specified by the Master equation

$$\begin{aligned} \frac{dP}{dt} &= n_\mathrm{S}\lambda_S \left[P_{n_{\mathrm{A}_1}-1} - P\right] + \lambda_\mathrm{A} \sum_{k=1}^{N-1} \left[(n_{\mathrm{A}_k}+1)P_{n_{\mathrm{A}_k}+1,n_{\mathrm{A}_{(k+1)}}-1} - n_{\mathrm{A}_k}P\right] \\ &\quad + \lambda_\mathrm{A}\left[(n_{\mathrm{A}_N}+1)P_{n_{\mathrm{A}_N}+1,n_\mathrm{B}-1} - n_{\mathrm{A}_N}P\right] + \Gamma\left[(n_\mathrm{B}+1)P_{n_\mathrm{B}+1} - n_\mathrm{B}P\right], \end{aligned} \tag{2}$$

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

As with the model described in the main text, the time evolution of $P_{\{n\}}(t)$ depends upon the initial conditions. With $P_{\{n\}}(0) = \delta_{n_S,1}(\prod_{k=1}^{N}\delta_{n_{A_k},0})\delta_{n_B,0}$, i.e. with one initially labelled stem cell, the solution to Eq. (2) describes the time-evolution of a persisting clone destined to

form an EPU. At long times, the corresponding probability distribution becomes stationary. Recasting Eq. (2) in terms of the moment-generating function (see, e.g., Ref. [2] for details), one finds a stationary distribution,

$$P_{\{n\}} = \mathcal{P}(n_{\mathrm{A}_1}|\lambda_{\mathrm{S}}/\lambda_{\mathrm{A}}) \left[\prod_{k=2}^{N} \mathcal{Q}(n_{\mathrm{A}_k}|2^{k-1}\lambda_{\mathrm{S}}/\lambda_{\mathrm{A}})\right] \mathcal{Q}(n_{\mathrm{B}}|2^{N}\lambda_{\mathrm{S}}/\Gamma), \tag{3}$$

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

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

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

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

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

On this platform, let us now consider the transient behaviour of a labelled clone population if we assume an equal labelling efficiency of cells in the basal layer. Referring to Eq. (1), one would expect all clones arising from initially labelled TA and B cells to inevitably detach from the basal layer. At the same time, a stem cell labelled at induction will give rise to a persisting clone which converges onto a fully-developed EPU. We conclude that, following a transient period, the size distribution of a randomly-labelled population will *converge onto the steady-state EPU size distribution* (3). This convergence is illustrated in figs. S9a and b where the clone size distributions predicted for the non-equilibrium process (1) are plotted against time for two values of the stem cell density. Notably, the steady-state behaviour differs markedly from that observed in experiment where the size of surviving clones continues to grow, and the distribution $P_{\{n\}}(t)$ does not settle, at least within the observed time period.

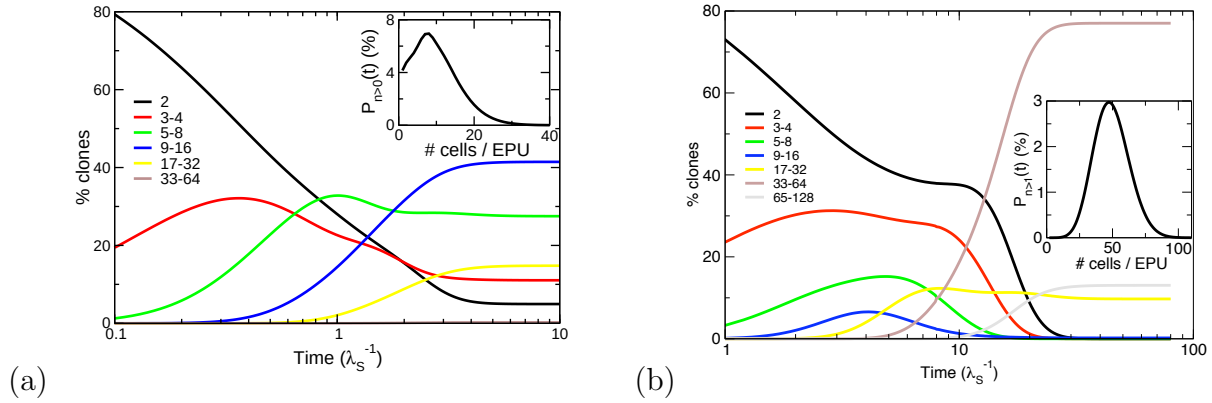

Fig. S9 Predicted clone size distributions for stem/TA cell model. The probability distribution $P_{n>1}^{\text{pers.}}(t)$ for finding a labelled clone with n basal cells is plotted against time (measured in units of the average stem-cell cycle time $1/\lambda_S$) for (a) $\rho_s = 0.1$ (i.e. 10%), and (b) $\rho_s = 0.02$ (i.e. 2%). For consistency with the main text, we have eliminated single cell clones from the statistical ensemble. The distribution is obtained for $N = 3$ generations of TA cells by integrating Eq. (2) starting from an initial condition corresponding to each clone consisting of a single, randomly-labelled cell. To perform the integration, we have used the respective values of stem cell number density, ρ_s, while keeping the overall density of proliferating cells at the experimentally measured value of $\rho_s + \rho_A = 0.22$ (i.e. 22%). The mean-field relations in Eq. (4) constrain the rates λ_A and Γ, leaving only one free parameter, the stem cell cycling rate, λ_S. For $\rho_s = 0.1$, the TA cell division rate is some seven times faster than the stem cell division rate while for $\rho_s = 0.02$, the cycling rates are are comparable (note that larger N will lead to $\lambda_S \ll \lambda_A$). As expected, in the long-time limit, the probability distribution becomes time-independent. *Inset*: The full probability distribution, $P_{n>1}^{\text{pers.}}(t)$, is plotted as a function of n after a time (a) $t = 10/\lambda_S$ and (b) $t = 50/\lambda_S$ when both distributions have converged onto their steady-state.

Single-proliferative compartment model of IFE maintenance

With this background, we turn now to the consideration of the novel single-proliferative compartment model introduced in the main text. When formulated as a non-equilibrium process, the model is described by the set of non-equilibrium rate equations

$$\begin{aligned} \mathrm{A} \xrightarrow{\lambda} & \begin{cases} \mathrm{A}+\mathrm{A} & \text{Prob. } r \\ \mathrm{A}+\mathrm{B} & \text{Prob. } 1-2r \\ \mathrm{B}+\mathrm{B} & \text{Prob. } r \end{cases} \\ \mathrm{B} \xrightarrow{\Gamma} & \mathrm{C}, \end{aligned} \tag{5}$$

where, as usual, A denote EPCs, B represent post-mitotic cells in the basal layer, and C specify the suprabasal layer cells. To maintain the total basal layer cell population, we must impose the condition that $\Gamma = \frac{\rho}{1-\rho}\lambda$, i.e. the net rate at which post-mitotic cells are generated in the basal layer, $\rho\lambda$, is compensated by the rate at which they are removed, $(1-\rho)\Gamma$. By neglecting processes involving the shedding of cells from the surface of the epidermis, the model is restricted to the consideration of the *total* clone size distribution at appropriately short time scales. However, if we focus only on the clone size distribution associated with those cells which occupy the basal layer, the model can be applied up to arbitrary time scales. For this case, the transfer process must be replaced by one in which B type cells are lost from the distribution at a rate Γ, i.e. $\mathrm{B} \xrightarrow{\Gamma} \emptyset$. In either case, the time evolution associated with (5) can be cast in the form of a Master equation. Defining $P_{n_A,n_B}(t)$ as the probability of finding n_A type A cells and n_B type B cells in a given clone after some time t, the basal layer cell probability distribution evolves according to the Master equation

$$\begin{aligned} \partial_t P_{n_A,n_B} = r\lambda\left[(n_A-1)P_{n_A-1,n_B} - n_A P_{n_A,n_B}\right] + r\lambda[(n_A+1)P_{n_A+1,n_B-2} - n_A P_{n_A,n_B}] \\ +(1-2r)\lambda[n_A P_{n_A,n_B-1} - n_A P_{n_A,n_B}] + \Gamma[(n_B+1)P_{n_A,n_B+1} - n_B P_{n_A,n_B}], \end{aligned}$$

with the boundary condition $P_{n_A,n_B}(0) = \rho\delta_{n_A,1}\delta_{n_B,0} + (1-\rho)\delta_{n_A,0}\delta_{n_B,1}$, i.e. if we assume an equal labelling efficiency, at induction, one labels a proliferating cell with probability ρ and post-mitotic cell with probability $1-\rho$. Although the Master equation (and its total cell number generalisation) is not amenable to exact analytic solution, its properties can be inferred from the consideration of the A cell population.

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

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

an example of a Galton-Watson process, long known to statisticians (see, e.g., Ref. [2]). In this case, the probability distribution, which is related to that of the two-component model through the relation $p_{n_\mathrm{A}}(t) = \sum_{n_\mathrm{B}=0}^{\infty} P_{n_\mathrm{A},n_\mathrm{B}}(t)$, can be solved analytically. (Here, we have used a lower case p to discriminate the probability distribution from its two-component counterpart.) For an initial distribution $p_{n_\mathrm{A}}(0) = \delta_{n_\mathrm{A},1}$,

$$p_{n_\mathrm{A}}(t) = \left(1+\frac{1}{r\lambda t}\right)^{-(n_\mathrm{A}+1)} \times \begin{cases} 1 & n_\mathrm{A}=0, \\ \frac{1}{(r\lambda t)^2} & n_\mathrm{A}>0. \end{cases} \tag{6}$$

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

- *IFE is maintained through an ever-decreasing clonal population:* Starting with a single labelled cell, the Galton-Watson process predicts that the persistance probability of the resulting clone (i.e., in this case, the probability that the clone retains at least one proliferating cell), is given by

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

i.e. the persistance probability of a clone decays monotonically asymptoting to the form $1 - p_0(t) \propto 1/t$. Therefore, in contrast to the classical stem cell hypothesis, labelled clones continue to detach from the basal layer indefinitely. At the same time, defining $p^{\text{pers.}}_{n_A>0}(t) = \frac{p_{n_A}(t)}{1-p_0(t)}$ as the size distribution of *persisting* clones, the mean number of basal layer cells in a persisting clone grows steadily as

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

such that the overall cell population remains constant, viz. $\langle n_A \rangle \equiv \sum_{n=0}^{\infty} n_A p_{n_A}(t) = 1$, i.e. the continual extinction of clones is compensated by the steady growth of persisting clones such that the average number of proliferating cells remains constant.

- *Coalescence of clones:* The linear increase in clone size may lead one to worry about the probability of coalesence of neighbouring clones and the attendant contamination of the data. Fortunately, their continual extinction ensures that the fraction of clones cojoined with their neighbours remains small and of same order as that found at induction. (The fraction of clones in contact with their neighbours is estimated by assuming that all clones are randomly and independently distributed, and by noting the observation that clones are oblique in shape, so the distribution in clone area has the same form as $P^{pers.}_{n>0}(t)$.) Indeed, the fact that this fraction is constant is again indicative of the steady-state properties of the IFE maintained throughout the experiment.

- *Larger clones begin to exhibit the stability of the macroscopic system*: If, at some instant, a clone is seen to have, say, N_A proliferating cells then, after a further time t, its size will fluctuate as

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

Thus clones (as defined by the A cell population) will maintain an approximately

stable number of cells providing $t \ll N_A/r\lambda$. For larger clones this time may exceed the lifetime of the animal. At the limit where macroscopic sections of the basal layer are considered, the variations are small. (For example, with a recorded cell density of ca. $17,500\,\mathrm{cells/mm^2}$, a clone with a size of $1\,\mathrm{mm^2}$ will experience a variation in size over 1 year of less than 0.05%.) The increased stability of larger clones also explains the surprising prediction that, given enough time, all clones eventually become extinct (viz. $\lim_{t\to\infty} p_{n>0}(t) = 0$). Calculated explicity, the extinction probability for a clone of size $N_A \gg 1$ scales as $p_0(t) \approx e^{-N_A/r\lambda t}$ approaching unity at long times. However, because this extinction probability is small when $t \ll N_A/r\lambda$, a large enough clone may easily persist beyond the lifetime of the animal.

- *The properties of the proliferating cell population dictate that of the entire clone size distribution*: At times $t \gg 1/r\lambda$, one may show that the full probability distribution for finding $n = n_A + n_B$ cells within a persisting clone scales in proportion to that of $p^{\mathrm{pers.}}_{n_A}(t)$, viz.

$$\lim_{t\gg 1/r\lambda} P^{\mathrm{pers.}}_{n>0}(t) = \frac{\rho}{r\lambda t}\exp\left[-\frac{\rho n}{r\lambda t}\right]. \tag{7}$$

  i.e. the probability distribution acquires the scaling form discussed in the main text. Consequently, one may infer that the mean number of basal layer cells in persisting clones at long times grows linearly with time,

$$\lim_{t\gg 1/r\lambda} \langle n\rangle_{\mathrm{pers.}} = \frac{1}{\rho}(1 + r\lambda t)\,. \tag{8}$$

- *The creation and transfer of differentiated cells dictates the short-time behaviour of the clone size distribution:* The rates λ and Γ at which differentiated cells are created and then transferred into the super-basal region are significantly larger than the rate of symmetric division $r\lambda$, which dictates the long-time behaviour of the clone size distribution. Consequently, at early times ($t \lesssim 1/\Gamma$), the clone size distributions are dominated by these processes, which remain prominent until the population of labelled

differentiated cells associated with each proliferating cell reaches its steady-state value of $(1-\rho)/\rho$. One may therefore infer that, at short times, the mean number of basal layer cells in clones arising from proliferating cells is given by

$$\lim_{t \lesssim 1/\Gamma} \langle n \rangle_{\text{pers.}} = \frac{1}{\rho} - \left(\frac{1}{\rho} - 1 \right) e^{-\Gamma t} \,. \tag{9}$$

Taken together, it is interesting to note that Eqs. (8) and (9) provide a simple means to extract two of the physical parameters that characterise the cell kinetics.

Effect of potential on-going clone accumulation

Our analysis of the clone fate data relies on the simultaneous induction of labelled clones. Yet, at short time scales (of less than two weeks), one observes a small but significant increase in the population of labelled clones. This increase may reflect EYFP accumulation in basal layer cells, or it may signify on-going recombination over this period. It is therefore important to address the potential impact of this effect on the interpretation of the clone fate data.

To this end, we have explored the effect on the predicted labelled clone size distribution of introducing new single cell clones over the first two weeks post-induction. Specifically, one may repeat the theoretical analysis while adjusting the probability distribution $P_{n_A,n_B}(t)$ at each time step to accommodate a small relative increase in the values of the probability $P_{1,0}(t)$. We set the increase in clone population during the first week of experiment to the observed value of 26%, and the total increase by the end of the second week to 35%. The resulting clone size distributions are shown in Fig. S10 alongside the original data and the original model predictions taken from Fig. 2d. The results demonstrate that for clones with two or more cells, the effect on the clone size distributions is small during the first few weeks of the experiment, and becomes negligible for time points beyond six weeks. We conclude that the analysis presented in the main text is not significantly affected by the progressive appearance of labelled clones over the first two weeks of the experiment.

It may at first appear surprising that the change in clone size distributions is so small. To understand why this should be the case, one should recall that, to remove potential ambiguities due to labelling efficiency, single-cell clones have been excluded from the statistical ensemble contributing to the clone size distribution. Therefore, the effect of adding new single-cell clones only becomes apparent upon subsequent division. But with a mean division rate of $\lambda = 1.1$/week, then the earliest labelled cells will have only undergone two

rounds of cell division, on average, before the accumulation is complete. Thus the effect on the clone size distributions remains limited.

At long times, recalling from the main text that the clone size distribution is dominated by the process of symmetric division, then it becomes clear why a 1-2 week uncertainty in initial conditions is negligible: A given progenitor cell undergoes symmetric division, on average, only every six weeks. The uncertainty period is thus far smaller than the interval between symmetric divisions, so that the long-time behaviour remains unaffected.

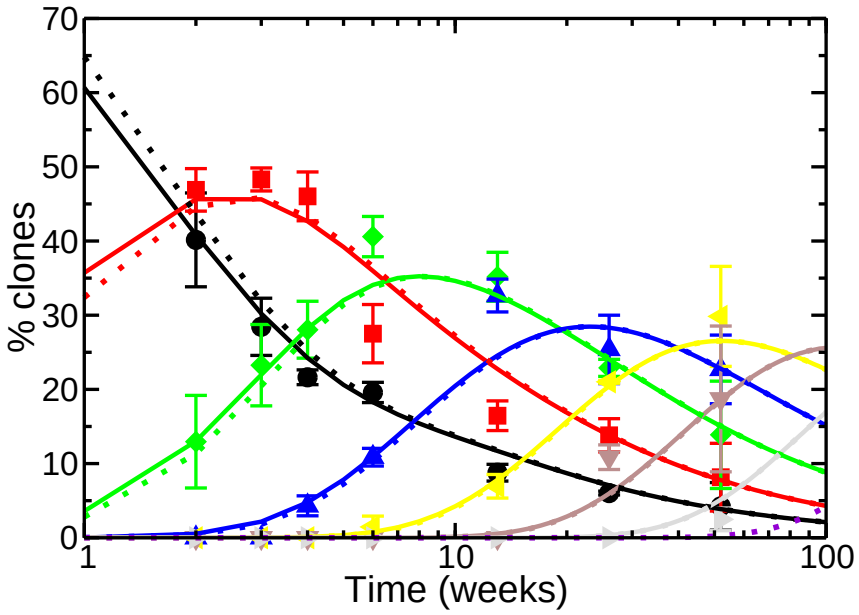

Fig. S10. Clone size distributions allowing for continuing recombination during two weeks post-induction. The solid curves indicate the original model (see fig. 2d, main text). The corrections (dotted lines) were plotted assuming a total increase of 35% in clone density.

Effect of divisions perpendicular to the basal layer

Although there is no direct experimental evidence for perpendicular divisions, in which a progenitor cell gives rise to one daughter cell in the basal layer and one in the supra-basal region, the analysis of spindle orientation (Fig. S8b) suggests that such events might occur with some small probability. It is therefore important to assess their potential impact on the development of clone size distributions.

To accommodate the effect of perpendicular divisions, one may supplement the existing model (process 1 in the main text) by the additional processes, $A \xrightarrow{\mu} A+C$ and $A \xrightarrow{\nu} B+C$, reflecting the potential for divisions that generate a basal and a suprabasal cell. In doing so, one must further modify the rate of transfer from the supra-basal layer, Γ, to ensure that the size of the B cell compartment remains unchanged. Here we have defined μ and ν as the average rates for each of the new processes. We exclude perpendicular division resulting in two daughter progenitor cells because there are no progenitor cells in the supra-basal layers of normal murine epidermis. Then, by incorporating the new processes into the master equation (main text), the corrected clone size distributions are accessible by numerical integration.

The extent of the effect of perpendicular divisions depends on the rate of their occurence (μ, ν). However, referring to Fig. S8(b), it is not clear how these rates relate to the angle of the mitotic spindle. As a result, it is difficult to estimate the actual fraction of divisions which are perpendicular. We have therefore taken a "worst-case" estimate of 10%, corresponding to all divisions lying at an angle beyond 45° from the basal layer resulting in one daughter cell in the supra-basal layer. Using this value, we compare the clone size distribution with and without perpendicular divisions in Fig. S11. The results demonstrate that, when perpendicular divisions occur 10% of the time, their effect is negligible.

It may appear surprising that the effect of perpendicular divisions is small, especially

when compared to the processes of symmetric in-plane division, which occur with only 8% likelihood for each of the two symmetric division channels. It is therefore important to distinguish between the role of symmetric divisions and perpendicular divisions in driving the statistical behaviour of the clone size distributions: The symmetric division processes define a *critical Galton-Watson process* (see ref. Bailey, 1964), which dominates the long-term behaviour of the system. Thus their occurrence, even with 8% likelihood, has a dramatic effect on clone statistics. On the other hand, perpendicular division events do not determine the overall statistical behaviour of the system: they merely accelerate the transfer of basal cells into the suprabasal region, which does not change the *total* population of a clone at all. This merely leads to a small uncertainty in the transfer rate Γ (of order 10%).

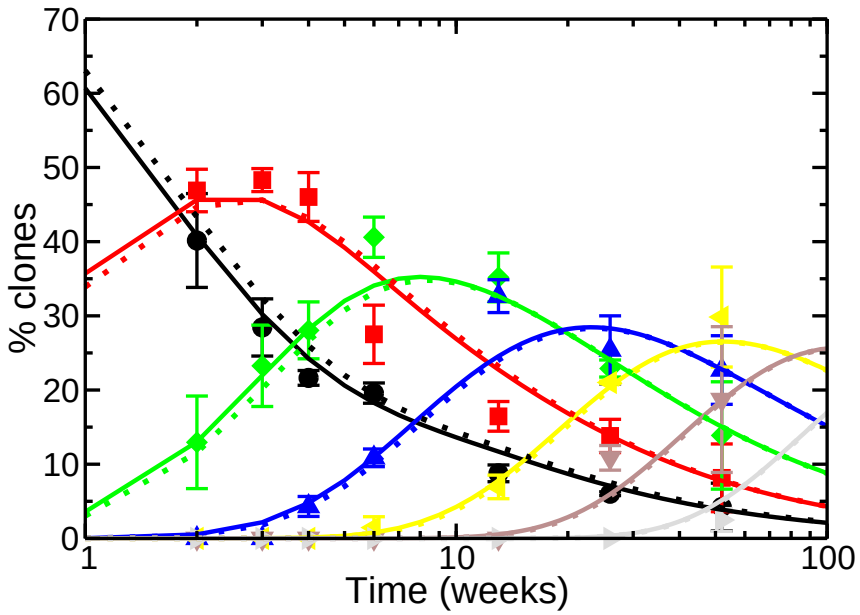

Fig. S11 Clone size distributions including perpendicular division events. The solid curves indicate the original model (see fig. 2d, main text). The corrections (dotted lines) were plotted assuming the (worst-case) estimate where 10% of all divisions result in one daughter cell above the basal layer.

References

[1] Potten, C. S. Cell replacement in epidermis (keratopoiesis) via discrete units of proliferation, *Int. Rev. Cytol.* **69**, 271-318 (1981).

[2] see §8.6 in Bailey, N. T. J. *The Elements of Stochastic Processes*, (John Wiley & Sons, New York. 1964).