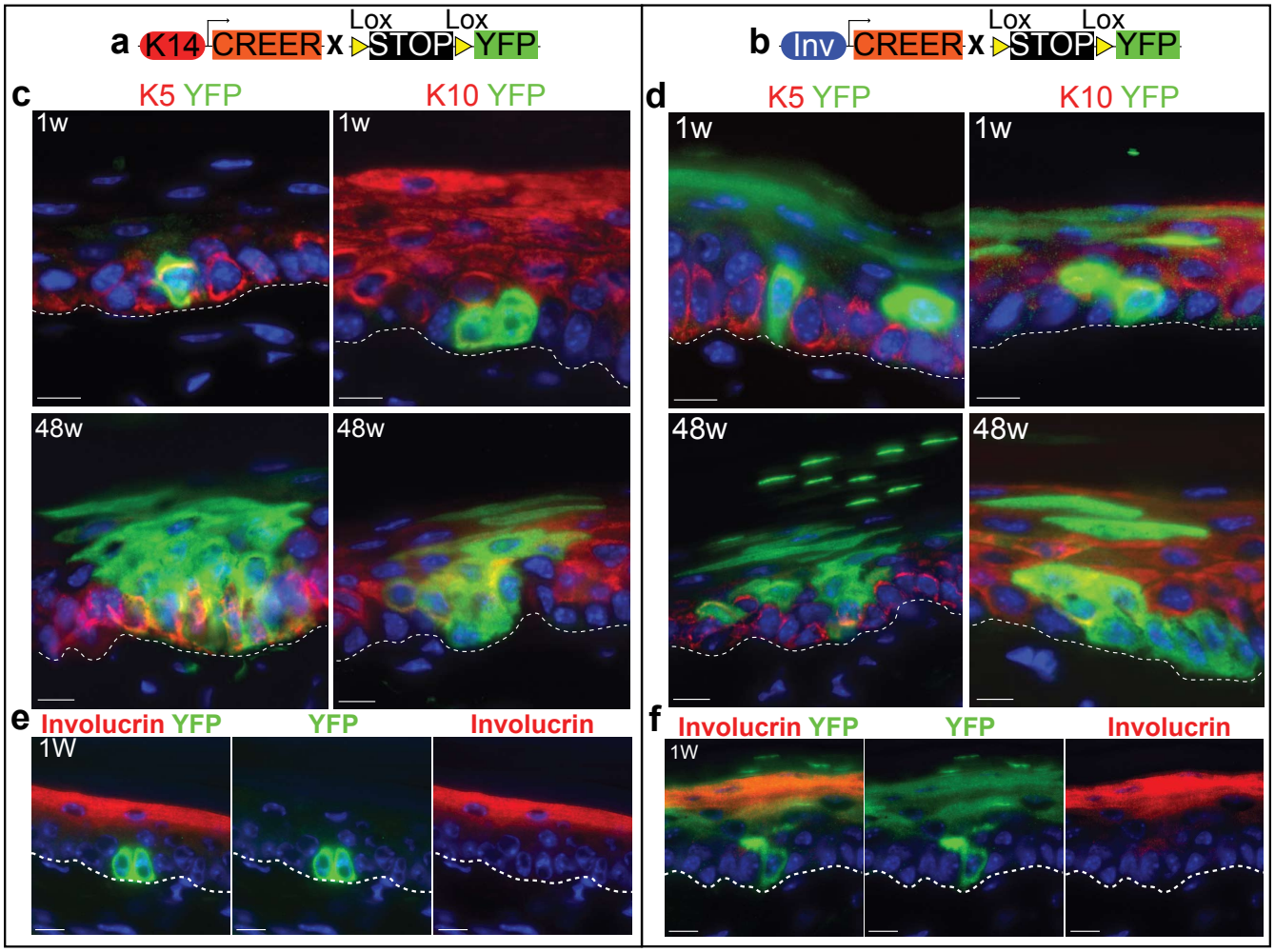
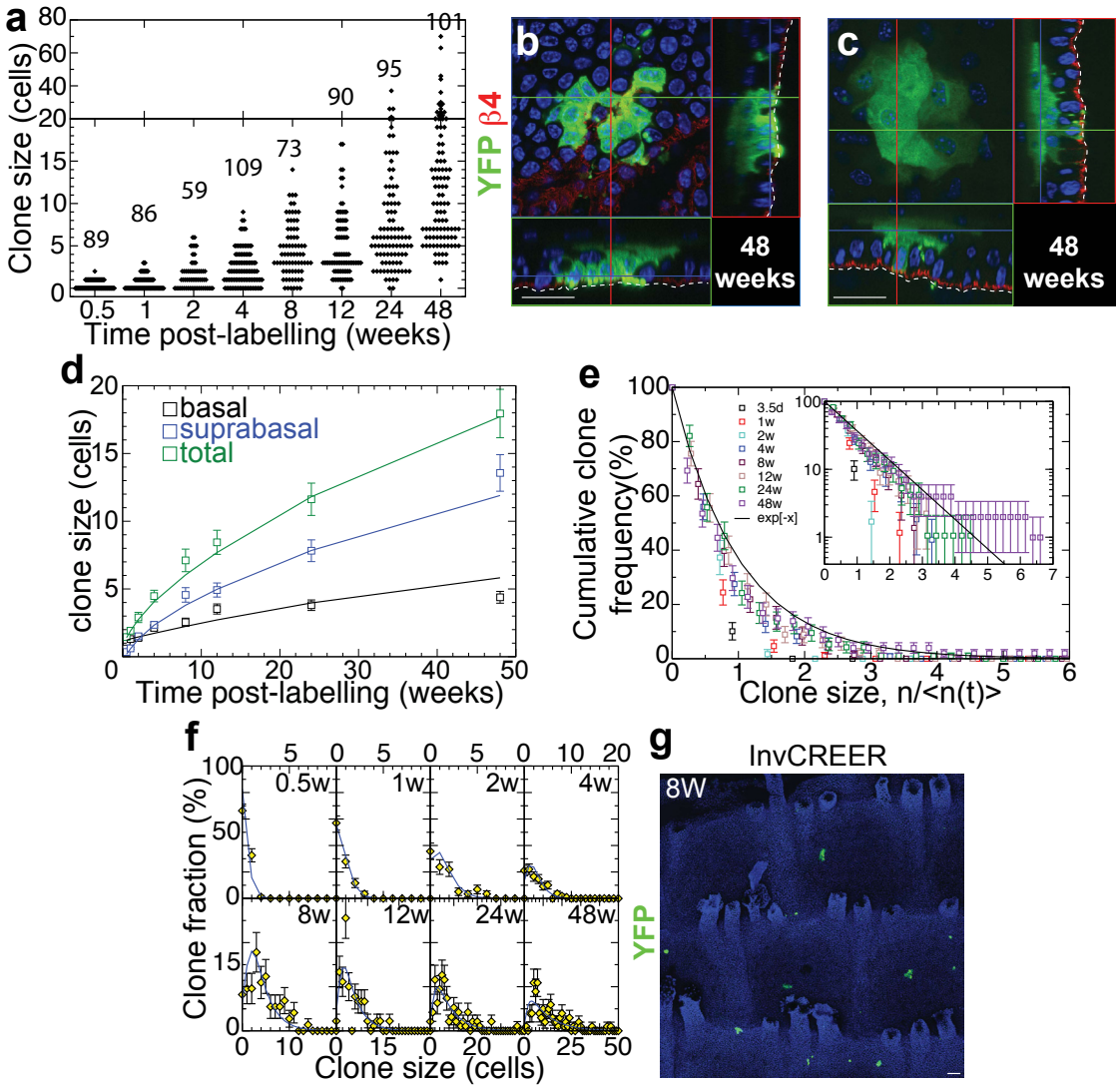
doi:10.1038/nature11393



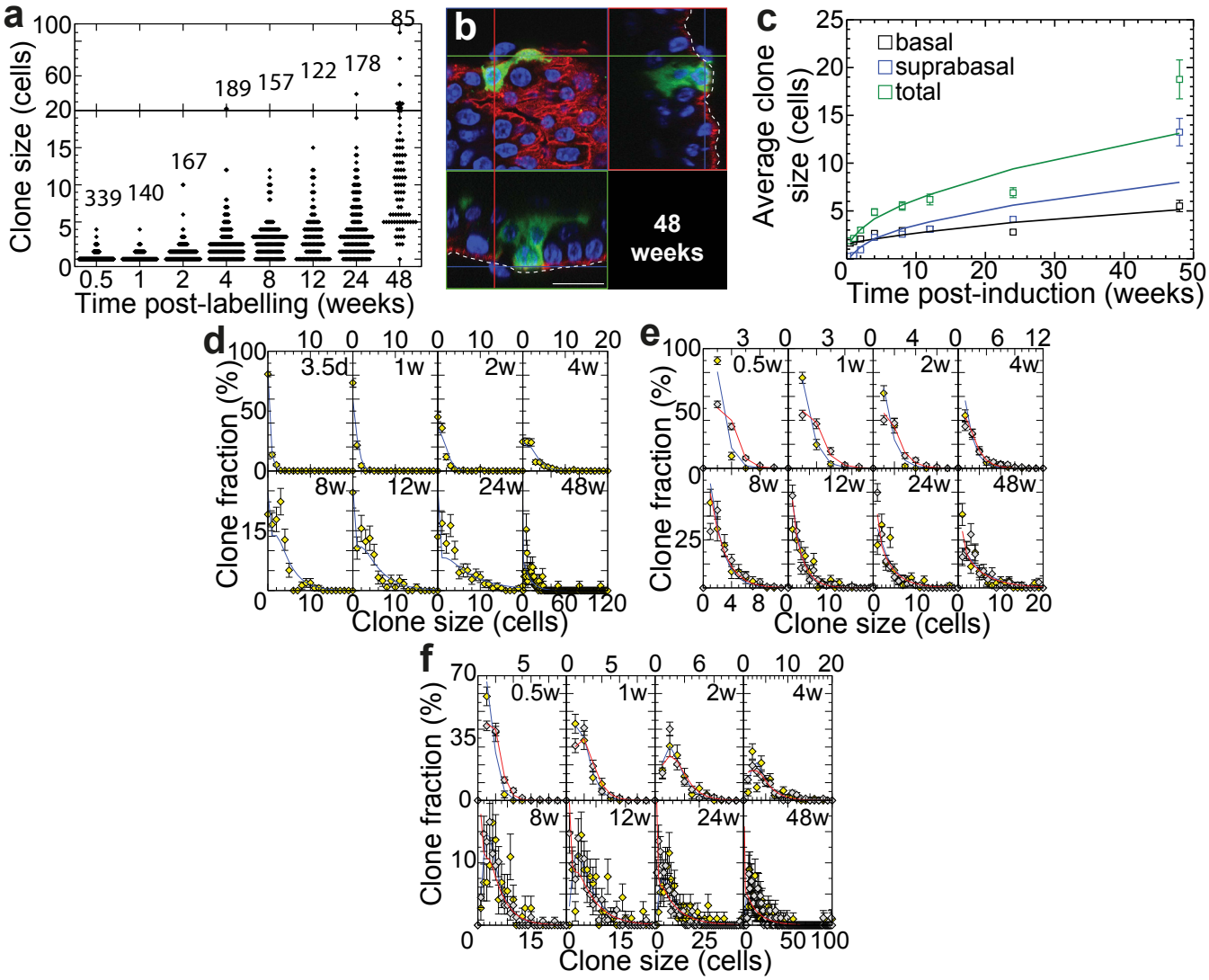

**Supplemental Figure 1. K14CREER and InvCREER target IFE basal cells whereas Involucrin protein is not expressed in the basal layer of the IFE.**

**a, b,** Genetic labelling strategy using (**a**) K14CREER/RosaYFP and (**b**) InvCREER/RosaYFP mice to label basal epidermal progenitors and follow their fate over time. **c, d,** Immunostaining of K5 or K10 and YFP in K14CREER/RosaYFP (**c**) and InvCREER/RosaYFP (**d**) 1 week or 48 weeks after TAM induction, showing the initial targeting of basal cells, some of which give rise to stable columns of basal and differentiated suprabasal cells. **e, f,** Immunostaining of Involucrin and YFP in K14CREER/RosaYFP (**e**) and InvCREER/RosaYFP (**f**) 1 week after TAM administration, showing the presence of Involucrin protein in the suprabasal cells but not in the basal IFE cells targeted by K14CREER and InvCREER/RosaYFP. Dashed lines represent the basal lamina. Hoechst nuclear staining is represented in blue. Scale bars, 10μm



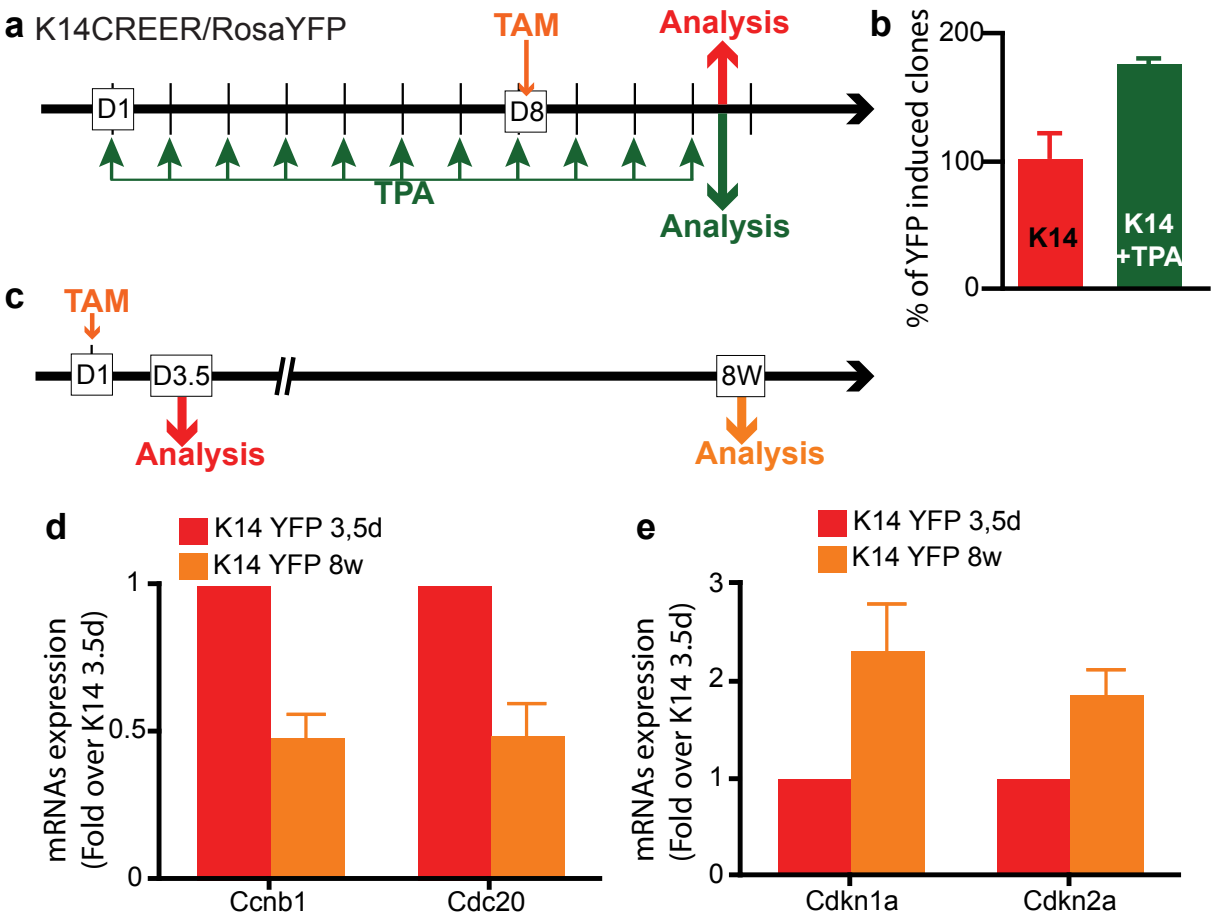

**Supplemental Figure 2. Quantitative analysis of basal clone content and clone size distribution of surviving InvCREER targeted clones.**

**a,** Distribution of InvCREER clone sizes as measured by the suprabasal cell content of surviving clones, imaged by confocal microscopy on whole mount tail epidermis from 3.5 days to 48 weeks following TAM administration. The number of analysed clones is indicated and are identical in panel **f**. **b, c,** Confocal analysis of representative InvCREER-targeted basal clones at 48 weeks post TAM. Dashed lines represent the basal lamina. Scale bars, 10μm. **d,** Time evolution of average clone size showing the basal, suprabasal and total cell content of InvCREER-targeted basal IFE cells. Points show data and lines represent the model prediction. **e,** Graph showing the scaling behaviour of InvCREER clones, as clonal fate data at each time point collapse to a single curve when the clone size is divided by the time. **f,** Frequency distribution of the suprabasal clone size. Symbols represent experimental data and the lines correspond to the model prediction. **g,** Whole mount immunostaining of YFP in InvCREER/RosaYFP mouse 8 weeks after TAM induction, showing the absence of preferential regional location of the InvCREER targeted clones. Error bars show s.e.m. Hoechst nuclear staining is represented in blue. Scale bars, 20μm.



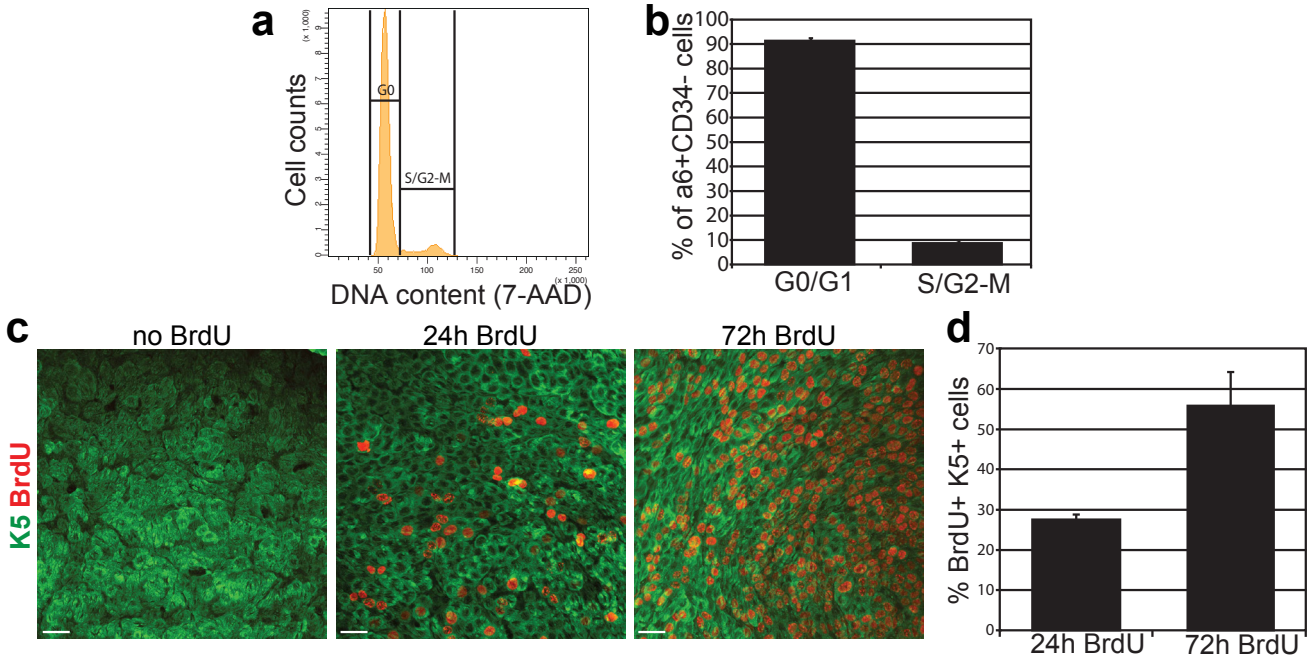

**Supplemental Figure 3. Quantitative analysis of basal clone content and clone size distribution of surviving K14CREER targeted clones.**

**a,** Distribution of K14CREER clone sizes as measured by the suprabasal cell content of surviving clones, imaged by confocal microscopy on whole mount tail epidermis from 3.5 days to 48 weeks following TAM administration. The number of analysed clones are indicated and are identical in panels **c** and **d**. **b,** Confocal analysis of representative K14CREER targeted basal clones at 48 weeks post TAM. Dashed lines represent the basal lamina. Hoechst nuclear staining is represented in blue. **c,** Time evolution of average clone size showing the basal, suprabasal and total cell content of K14CREER targeted basal IFE cells. Points show data and lines represent the prediction of the model. **d,** Frequency distribution of the suprabasal clone size. Symbols represent experimental data and lines correspond to the model prediction. **e-f,** Comparison of the frequency distribution of the basal (**e**) and total clone size (**f**) of K14CREER/RosaYFP (red) and InvCREER/RosaYFP (blue). Symbols represent experimental data and the lines correspond to the model prediction. Error bars show s.e.m. Scale bars, 10μm.

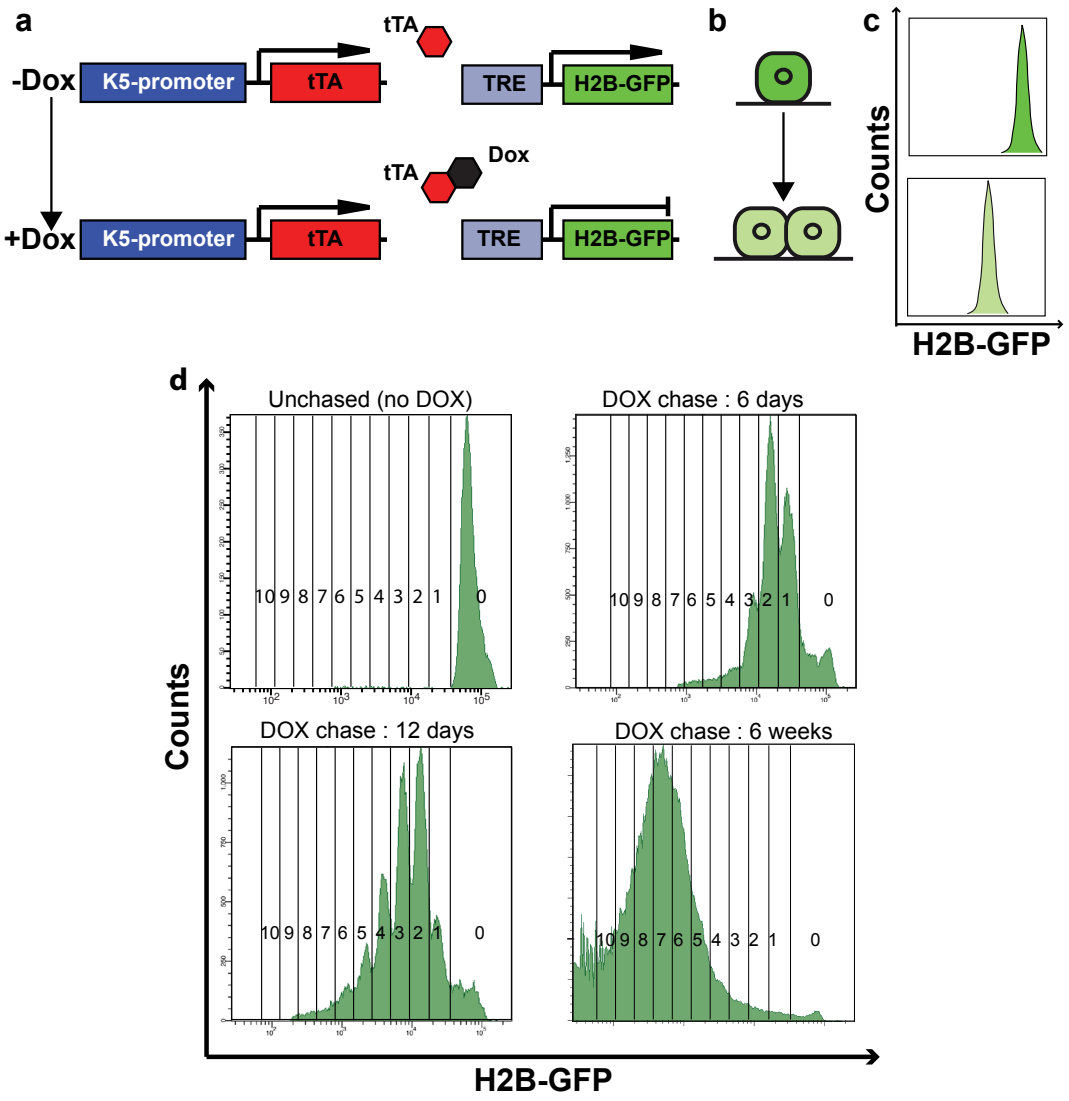



**Supplemental Figure 4. Preferential labelling of dividing cells by low dose of TAM administration in K14CREER/RosaYFP mice.**

**a,** Scheme representing the experimental strategy used to assess the influence of cell proliferation on the K14CREER mediated recombination upon low dose of TAM administration. **b,** Quantification of the density of basal clone induced 3.5 days after 0.2mg of TAM administration to K14CREER/RosaYFP mice treated or not with TPA. **c,** Scheme representing the experimental strategy used to induce and isolate basal α6+CD34- YFP+ cells 3.5 days or 8 weeks after TAM induction in K14CREER/RosaYFP and mice assess by qRT-PCR analysis the expression of cell cycle regulators. **d-e,** qRT-PCR analysis for gene expression of cell cycle regulators in K14CREER/RosaYFP targeted cells after 3.5 days (red) and 8 weeks (orange) of TAM treatment. These results show an upregulation of cell cycle inhibitors (*Cdkn1a, Cdkn2a*) and a decrease of cell cycle positive regulators (*Ccnb1, Cdc20*) in K14CREER/RosaYFP targeted cells after 8 weeks of TAM treatment compared to the one labelled at 3.5 days (n=3). Error bars are s.e.m.

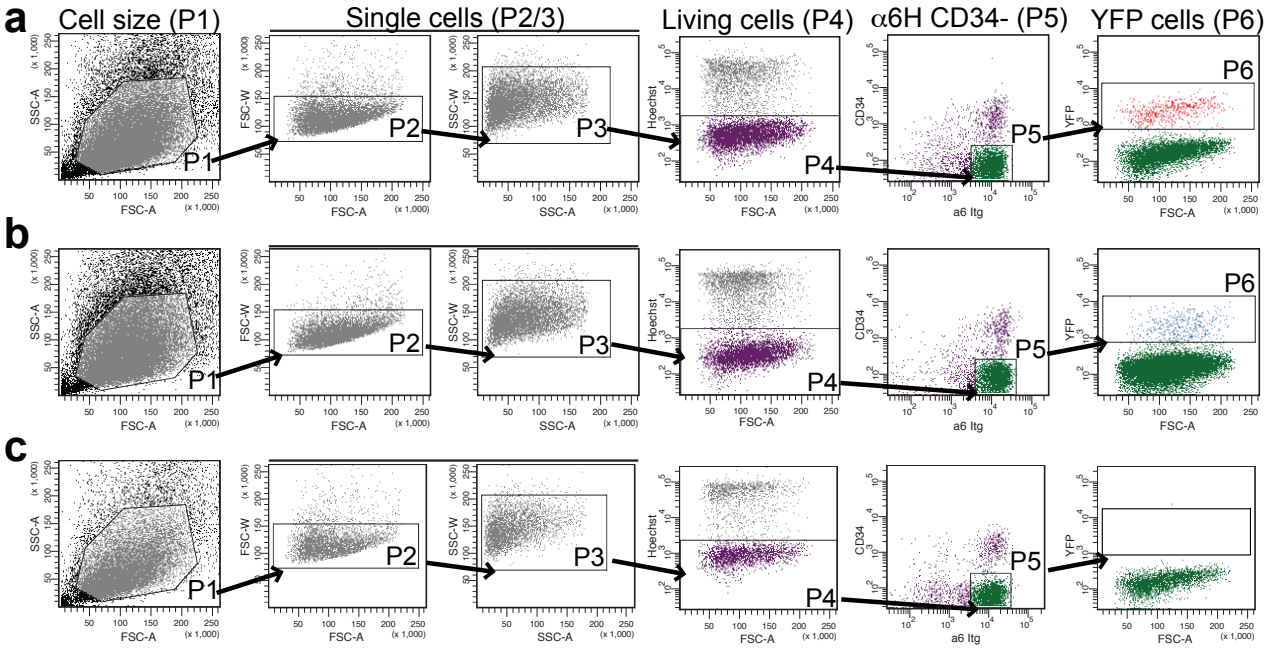



**Supplemental Figure 5. Cell cycle and proliferation kinetics analysis of the tail IFE.**
**a,** DNA content (7AAD) of IFE cells (α6+CD34-) measured by FACS analysis, showing repartition of the cells among the different cell cycle phases. **b,** Quantification of DNA content measured by FACS analysis, showing that about 9% of the basal cells are in S/G2/M stages of the cell cycle. **c,** Maximum intensity projections of K5 and BrdU immunostainings, after 24h and 72h of continuous BrdU administration. **d,** Quantification of BrdU in wholemount epidermis, showing that 28% of basal cells incorporated BrdU after 24h and 56% after 72h (n=2 mice, 400 BrdU+ cells counted at least). Error bars are standard deviations. Scale bars represent 20μm.




**Supplemental Figure 6. Proliferation dynamic analysis of the tail IFE using H2B-GFP label dilution.**

**a-c,** Scheme of the K5tTA/tet(O)-H2BGFP genetic strategy to monitor the rate of cell division based on the quantification of H2B-GFP label dilution. In the absence of doxycyclin (DOX) (**a-c** upper panels), tTA induced uniform and high expression of H2B-GFP in all K5 expressing cells. Upon DOX addition (**a-c**, lower panels), the transcription of H2B-GFP is stopped, and after each division, the H2B-GFP fluorescence is divided by 2, which can be visualized and quantitatively analyzed by FACS. **d**, Examples of H2B-GFP fluorescence peak patterns observed by FACS analysis at different time points following DOX administration to K5tTA/tet(O)-H2BGFP DOX adult mice. Note the presence of discrete peaks of fluorescence that represent cells that accomplished different number of cell divisions. The proportion of cells in each of these peaks of fluorescence is quantified by FACS analysis.

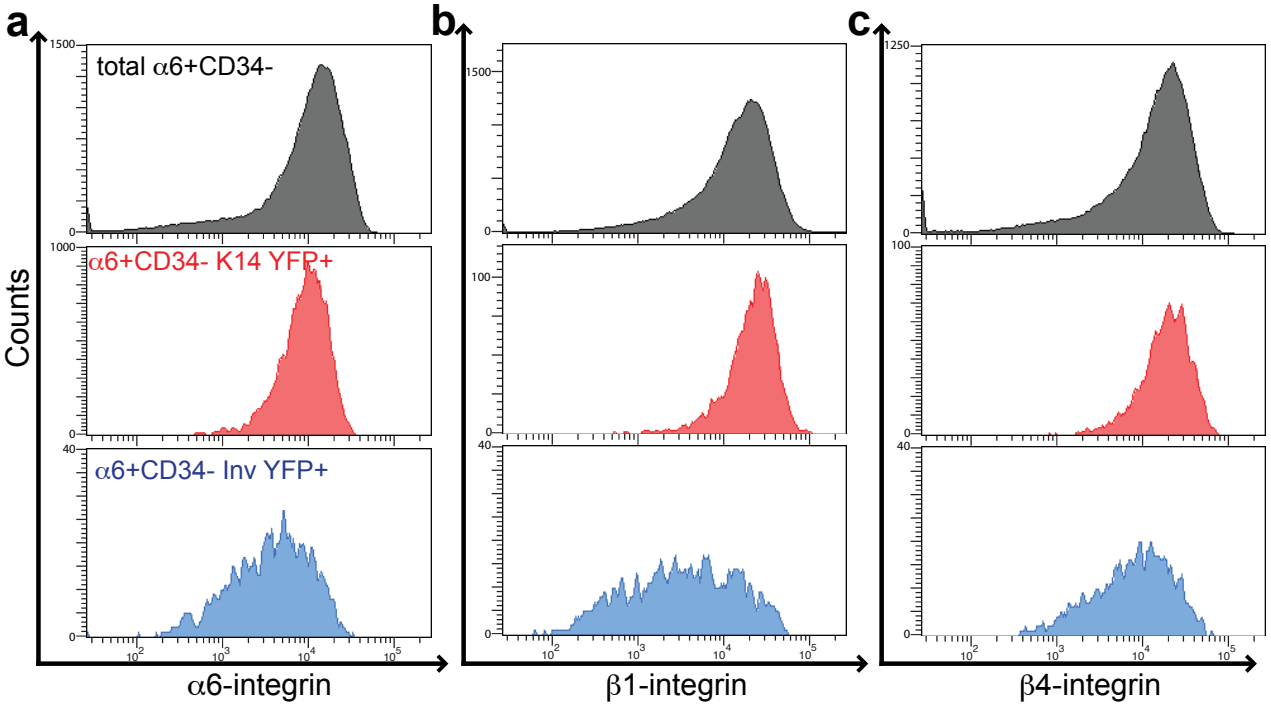



**Supplemental Figure 7. FACS strategy used to isolate basal IFE cells from total epidermis, and from K14 and Inv targeted cells 3.5 days after TAM induction.**

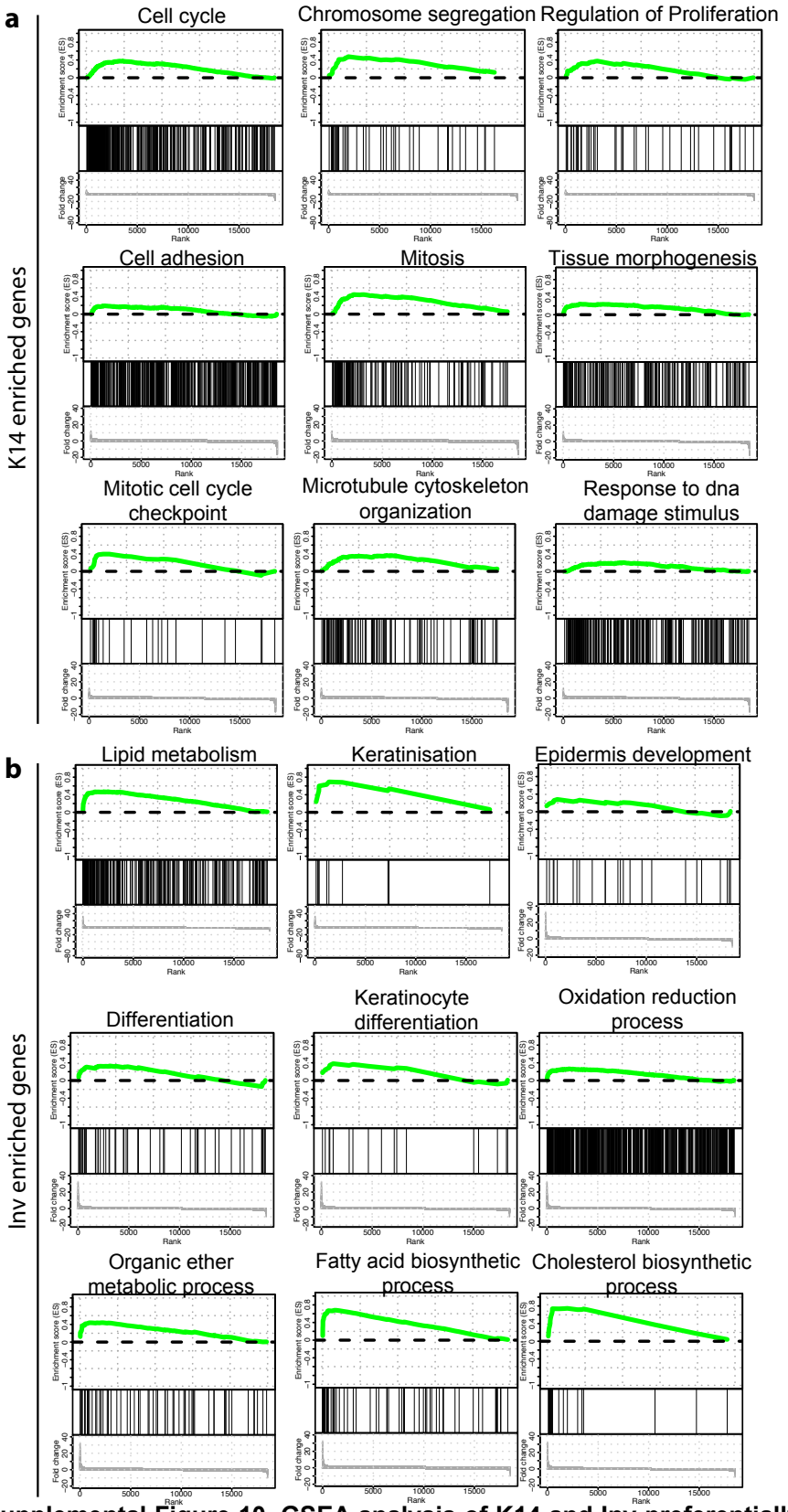
**a-c,** 3.5 days after TAM induction, epidermal cells were dissociated by enzymatic digestion and mechanically separated into single cells, stained with α6 and CD34 antibody. Single living K14 (**a**) and Inv (**b**) labeled epidermal cells or total IFE cells (**c**) were gated by debris exclusion (P1, Forward Scatter/Side Scatter), boublet elimination (P2 and P3), living cells (P4, Hoechst dye exclusion), and basal IFE α6+ CD34- cells (P5) and YFP expression (P6).

| | | | |
|---|---|---|---|
| Cell adhesion | Itga2,Itga3,Itga6,Itgb1,Itgb4,Lama3, Lama5,Lamb1,Lamc1,Lgals1, Col18a1, Cspg4 | Regulation of cell proliferation | Cdkn1a,Cdkn1b,Cdkn1c,Cdkn2a,Cdkn2c, Ereg,Etv4,E2f7,Spry1,Spry2,Tgfbr2, Tgfb1,Tgfb2 |
| Cell cycle | Ccnb1,Ccnd2,Cdc20,Cep55,Cep72, Ereg,E2f1,E2f7,Kifc5b,Kif11,Mki67, Plk1,Wee,Aurka,Bcl2,Brca1,Lig1 | Lipid metabolism | Agpat3,Degs1,Elovl3,Elovl4,Elovl5,Elovl6, Fads6,Fa2h,Lass4,Olah |
| Chromosome segregation | Birc5,Cdca5,Cenpe,Cenpf,Cenph, Top2a | Epidermis Development | Cnfn,Evpl,Grhl3,Notch3,Ovol1,Pou2f3, Pou3f1,Sce1,Sfn |
| Epithelium development | Areg,Bmp1,Bmp2,Bmp4,Celsr1,Dll1, Dll3,Etv4,Frem2,Gli3,Gli2,Ovol2, Wnt3a,Wnt5b,Wnt10a | Keratinization | Krt6b,Sprr1a,Sprr1b,Sprr2d,Sprr2h, Sprr2i,Tgm1,Tgm3 |

**Supplemental Figure 8. Molecular signature of K14CREER and InvCREER targeted basal cells.** Selected genes differentially expressed between K14 and Inv-targeted cells, as determined by microarray analysis. Genes shown are examples of those differentially expressed by more than 1.5 fold in duplicate biological samples. Genes upregulated in K14 SC are in red. Genes upregulated in Inv CP are in blue.

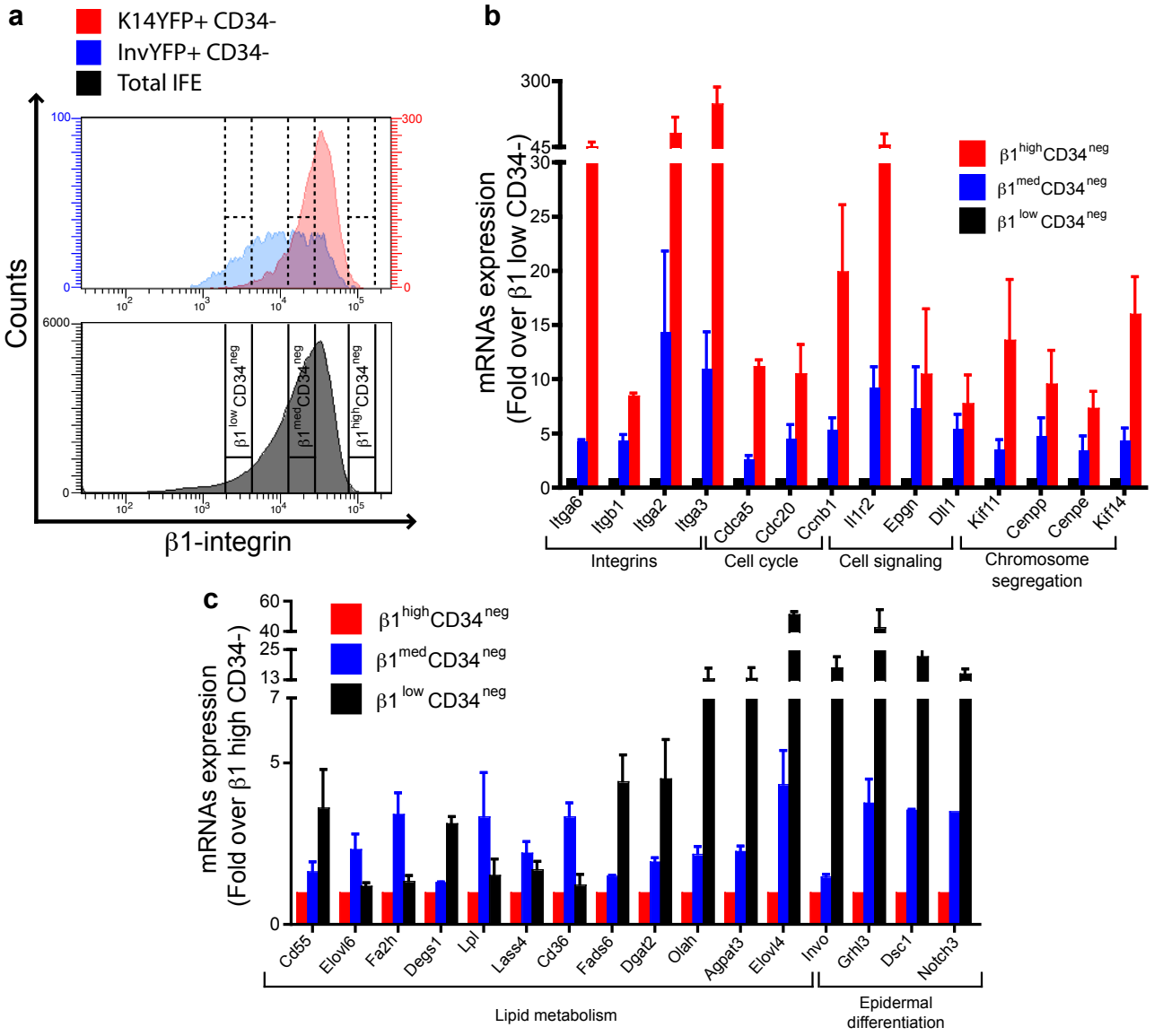



**Supplemental Figure 9. FACS analysis of integrin expression in Inv and K14 targeted cells.** **a, c,** FACS analysis of α6 (**a**), β1 (**b**) and β4 (**c**) integrin expressions in total IFE cells (black), and in InvCREER/RosaYFP (blue) and K14CREER/RosaYFP (red) IFE labeled cells 3.5 days after induction.

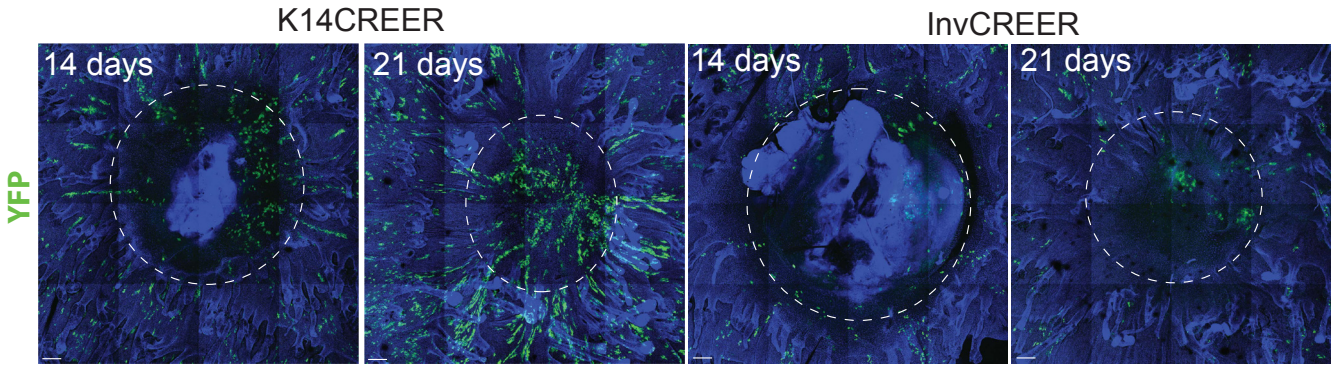



**Supplemental Figure 10. GSEA analysis of K14 and Inv preferentially expressed functional genes categories.**

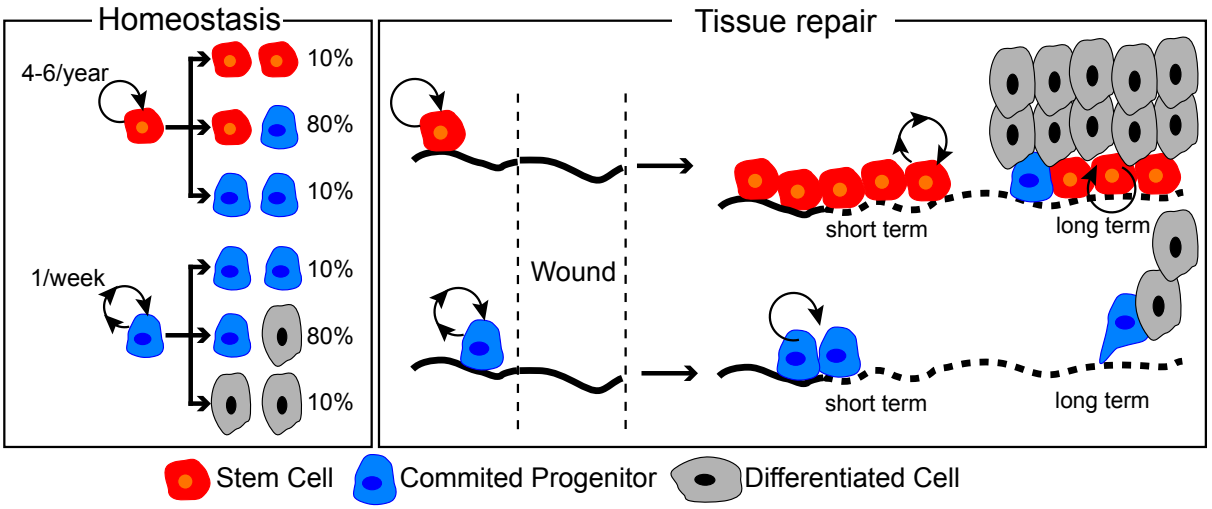
**a,** GSEA showing the distribution of the genes implicated in different functional gene clusters, within the rank order list of all the microarray probe sets of K14CREER basal cells. **b,** GSEA showing the preferential enrichment of the genes implicated in different functional gene clusters, within the rank order list of all the microarray probe sets of InvCREER basal cells.




**Supplemental Figure 11. Molecular characterization of basal cell heterogeneity.**
**a,** FACS analysis of β1 integrin expression level in K14CREER/RosaYFP (red), InvCREER/RosaYFP (blue) labelled cells after 3.5 days of TAM induction and the expression of β1 in total basal epidermal cells from CD1 mice (black). This analysis shows the gating strategy used to gate and sort β1-high, β1-medium and β1-low integrin expressing cells for qRT-PCR analysis. **b,** qRT-PCR analysis of the expression of representative genes, ordered by functional categories, upregulated in the β1-high integrin cell population (n=3). **c,** qRT-PCR analysis of the expression of representative genes, ordered by functional categories, upregulated in the β1-medium and β1-low cell populations (n=3). Error bars are s.e.m.




**Supplemental Figure 12. Different contribution of K14 SC and Inv CP during wound healing.**
Microscopic analysis of YFP immunostaining performed on whole-mounts of wounded tail epidermis 4 weeks after TAM administration to K14CREER/RosaYFP and InvCREER/RosaYFP mice and analyzed after 14 and 21 days following wounding. Dashed lines represent the wounded area. Hoechst nuclear staining is represented in blue. Scale bars, 200μm.




**Supplemental Figure 13. Model of epidermal homeostasis showing the hierarchical organization of SC and CP cells.**
In normal homeostatic conditions, both populations undergo population asymmetric self-renewal with SC dividing more slowly than CP cells. However, on injury, SC become active and proliferate, contributing to long-term repair, while CP cells make a minimal and transient contribution.

# SUPPLEMENTARY TABLE1

| List of genes upregulated in Involucrin compared to K14 (fold change cutoff = 1.5) | | | | | |
|---|---|---|---|---|---|
| **Term** | **Count** | **%** | **PValue** | **Fold Enrichment** | **Benjamini** |
| lipid biosynthetic process | 74 | 6 | 7,9E-29 | 4,6 | 2,0E-25 |
| fatty acid metabolic process | 54 | 4 | 8,5E-24 | 5,2 | 1,1E-20 |
| oxidation reduction | 92 | 8 | 4,4E-15 | 2,4 | 3,7E-12 |
| fatty acid biosynthetic process | 28 | 2 | 1,9E-14 | 6,1 | 1,2E-11 |
| sterol biosynthetic process | 17 | 1 | 7,1E-13 | 9,9 | 3,6E-10 |
| steroid biosynthetic process | 24 | 2 | 3,3E-12 | 5,9 | 1,4E-09 |
| steroid metabolic process | 36 | 3 | 4,1E-12 | 3,9 | 1,5E-09 |
| epidermis development | 31 | 3 | 1,0E-11 | 4,4 | 3,2E-09 |
| ectoderm development | 32 | 3 | 1,0E-11 | 4,2 | 2,9E-09 |
| cholesterol biosynthetic process | 14 | 1 | 4,0E-11 | 10,7 | 1,0E-08 |
| keratinocyte differentiation | 19 | 2 | 4,7E-11 | 6,9 | 1,1E-08 |
| organic acid biosynthetic process | 32 | 3 | 5,2E-11 | 4,0 | 1,1E-08 |
| carboxylic acid biosynthetic process | 32 | 3 | 5,2E-11 | 4,0 | 1,1E-08 |
| epidermal cell differentiation | 19 | 2 | 1,5E-10 | 6,5 | 3,0E-08 |
| sterol metabolic process | 21 | 2 | 6,7E-09 | 4,8 | 1,2E-06 |
| keratinization | 13 | 1 | 1,4E-08 | 8,2 | 2,3E-06 |
| epithelial cell differentiation | 25 | 2 | 9,0E-08 | 3,6 | 1,4E-05 |
| lipid catabolic process | 26 | 2 | 1,2E-07 | 3,4 | 1,8E-05 |
| cholesterol metabolic process | 18 | 1 | 2,6E-07 | 4,5 | 3,7E-05 |
| cellular lipid catabolic process | 14 | 1 | 1,1E-05 | 4,4 | 1,5E-03 |
| coenzyme biosynthetic process | 14 | 1 | 5,2E-05 | 3,8 | 6,6E-03 |
| coenzyme metabolic process | 22 | 2 | 5,7E-05 | 2,7 | 6,9E-03 |
| isoprenoid metabolic process | 12 | 1 | 7,5E-05 | 4,3 | 8,6E-03 |
| cofactor metabolic process | 25 | 2 | 1,0E-04 | 2,4 | 1,1E-02 |
| isoprenoid biosynthetic process | 8 | 1 | 1,5E-04 | 6,4 | 1,6E-02 |
| brown fat cell differentiation | 9 | 1 | 1,6E-04 | 5,4 | 1,6E-02 |
| fatty acid catabolic process | 8 | 1 | 2,1E-04 | 6,1 | 2,0E-02 |
| epithelium development | 31 | 3 | 3,7E-04 | 2,0 | 3,4E-02 |
| phospholipid metabolic process | 22 | 2 | 3,7E-04 | 2,4 | 3,3E-02 |
| organophosphate metabolic process | 23 | 2 | 4,2E-04 | 2,3 | 3,6E-02 |
| vitamin metabolic process | 13 | 1 | 4,6E-04 | 3,3 | 3,9E-02 |
| quinone cofactor metabolic process | 6 | 0 | 5,2E-04 | 8,1 | 4,2E-02 |
| negative regulation of catalytic activity | 16 | 1 | 5,6E-04 | 2,8 | 4,4E-02 |
| cofactor biosynthetic process | 15 | 1 | 5,9E-04 | 2,9 | 4,5E-02 |
| positive thymic T cell selection | 5 | 0 | 6,0E-04 | 11,0 | 4,4E-02 |
| negative regulation of transferase activity | 11 | 1 | 7,0E-04 | 3,6 | 5,0E-02 |
| **List of genes upregulated in K14 compared to Involucrin (fold change cutoff = 1.5)** | | | | | |
| **Term** | **Count** | **%** | **PValue** | **Fold Enrichment** | **Benjamini** |
| cell cycle | 152 | 11 | 1,01E-44 | 3,5 | 3,1E-41 |
| cell cycle process | 109 | 8 | 2,43E-36 | 3,9 | 3,7E-33 |
| cell division | 88 | 7 | 1,28E-33 | 4,4 | 1,3E-30 |
| cell cycle phase | 93 | 7 | 9,10E-32 | 4,0 | 7,0E-29 |
| M phase | 82 | 6 | 1,02E-28 | 4,1 | 6,3E-26 |
| mitotic cell cycle | 74 | 5 | 2,71E-27 | 4,2 | 1,4E-24 |
| M phase of mitotic cell cycle | 62 | 5 | 2,99E-24 | 4,5 | 1,3E-21 |
| nuclear division | 61 | 5 | 5,56E-24 | 4,5 | 2,1E-21 |
| mitosis | 61 | 5 | 5,56E-24 | 4,5 | 2,1E-21 |
| organelle fission | 61 | 5 | 4,53E-23 | 4,3 | 1,5E-20 |
| chromosome segregation | 26 | 2 | 4,31E-13 | 5,7 | 1,3E-10 |

| | | | | | |
|---|---|---|---|---|---|
| cell adhesion | 85 | 6 | 4,33E-11 | 2,1 | 1,2E-08 |
| biological adhesion | 85 | 6 | 4,75E-11 | 2,1 | 1,2E-08 |
| regulation of cell proliferation | 81 | 6 | 1,83E-10 | 2,1 | 4,3E-08 |
| DNA replication | 35 | 3 | 1,56E-09 | 3,2 | 3,4E-07 |
| cytoskeleton organization | 55 | 4 | 4,30E-09 | 2,4 | 8,8E-07 |
| regulation of cell cycle | 42 | 3 | 4,63E-09 | 2,7 | 8,8E-07 |
| tube morphogenesis | 36 | 3 | 1,07E-08 | 2,9 | 1,9E-06 |
| tissue morphogenesis | 44 | 3 | 1,23E-08 | 2,6 | 2,1E-06 |
| morphogenesis of an epithelium | 35 | 3 | 4,99E-08 | 2,8 | 8,0E-06 |
| cell-substrate junction assembly | 10 | 1 | 6,64E-08 | 10,0 | 1,0E-05 |
| DNA metabolic process | 62 | 5 | 7,56E-08 | 2,1 | 1,1E-05 |
| morphogenesis of a branching structure | 28 | 2 | 1,61E-07 | 3,1 | 2,2E-05 |
| negative regulation of cell proliferation | 40 | 3 | 1,63E-07 | 2,5 | 2,2E-05 |
| cell-substrate adhesion | 18 | 1 | 2,54E-07 | 4,4 | 3,2E-05 |
| tube development | 44 | 3 | 2,66E-07 | 2,3 | 3,3E-05 |
| regulation of morphogenesis of a branching structure | 12 | 1 | 4,03E-07 | 6,7 | 4,7E-05 |
| regulation of cell cycle process | 20 | 1 | 5,40E-07 | 3,8 | 6,1E-05 |
| epithelial tube morphogenesis | 25 | 2 | 7,62E-07 | 3,2 | 8,3E-05 |
| regulation of muscle development | 15 | 1 | 8,88E-07 | 4,9 | 9,4E-05 |
| cytokinesis | 12 | 1 | 1,03E-06 | 6,2 | 1,1E-04 |
| gland morphogenesis | 21 | 2 | 1,28E-06 | 3,5 | 1,3E-04 |
| kidney development | 24 | 2 | 1,44E-06 | 3,1 | 1,4E-04 |
| odontogenesis | 15 | 1 | 1,65E-06 | 4,7 | 1,5E-04 |
| branching morphogenesis of a tube | 22 | 2 | 1,76E-06 | 3,3 | 1,6E-04 |
| DNA-dependent DNA replication | 13 | 1 | 2,04E-06 | 5,4 | 1,8E-04 |
| microtubule-based process | 36 | 3 | 2,23E-06 | 2,4 | 1,9E-04 |
| M phase of meiotic cell cycle | 21 | 2 | 2,78E-06 | 3,3 | 2,3E-04 |
| meiosis | 21 | 2 | 2,78E-06 | 3,3 | 2,3E-04 |
| odontogenesis of dentine-containing tooth | 14 | 1 | 3,06E-06 | 4,8 | 2,5E-04 |
| positive regulation of developmental process | 36 | 3 | 3,10E-06 | 2,4 | 2,4E-04 |
| epithelium development | 42 | 3 | 3,54E-06 | 2,2 | 2,7E-04 |
| urogenital system development | 28 | 2 | 3,99E-06 | 2,7 | 3,0E-04 |
| meiotic cell cycle | 21 | 2 | 4,02E-06 | 3,3 | 2,9E-04 |
| regulation of striated muscle tissue development | 14 | 1 | 4,15E-06 | 4,7 | 2,9E-04 |
| enzyme linked receptor protein signaling pathway | 42 | 3 | 4,27E-06 | 2,2 | 3,0E-04 |
| cell fate commitment | 28 | 2 | 4,56E-06 | 2,7 | 3,1E-04 |
| interphase of mitotic cell cycle | 15 | 1 | 5,11E-06 | 4,3 | 3,4E-04 |
| intracellular signaling cascade | 102 | 8 | 5,22E-06 | 1,6 | 3,4E-04 |
| negative regulation of molecular function | 26 | 2 | 5,81E-06 | 2,8 | 3,7E-04 |
| cell junction assembly | 10 | 1 | 6,17E-06 | 6,7 | 3,9E-04 |
| cell-matrix adhesion | 15 | 1 | 6,63E-06 | 4,2 | 4,1E-04 |
| regulation of phosphorylation | 43 | 3 | 8,29E-06 | 2,1 | 5,0E-04 |
| interphase | 15 | 1 | 8,55E-06 | 4,1 | 5,0E-04 |
| regulation of microtubule-based process | 14 | 1 | 9,70E-06 | 4,4 | 5,6E-04 |
| mammary gland morphogenesis | 12 | 1 | 1,03E-05 | 5,1 | 5,8E-04 |
| regulation of organelle organization | 28 | 2 | 1,12E-05 | 2,5 | 6,2E-04 |
| negative regulation of cell differentiation | 31 | 2 | 1,32E-05 | 2,4 | 7,2E-04 |
| vasculature development | 38 | 3 | 1,73E-05 | 2,1 | 9,3E-04 |
| heart development | 35 | 3 | 2,01E-05 | 2,2 | 1,1E-03 |

| | | | | | |
|---|---|---|---|---|---|
| regulation of phosphorus metabolic process | 43 | 3 | 2,08E-05 | 2,0 | 1,1E-03 |
| regulation of phosphate metabolic process | 43 | 3 | 2,08E-05 | 2,0 | 1,1E-03 |
| DNA replication initiation | 7 | 1 | 2,12E-05 | 9,8 | 1,1E-03 |
| blood vessel development | 37 | 3 | 2,40E-05 | 2,1 | 1,2E-03 |
| positive regulation of cell proliferation | 41 | 3 | 2,66E-05 | 2,0 | 1,3E-03 |
| transmembrane receptor protein serine/threonine kinase signaling pathway | 18 | 1 | 2,75E-05 | 3,2 | 1,3E-03 |
| skeletal system development | 41 | 3 | 2,85E-05 | 2,0 | 1,4E-03 |
| regulation of cell adhesion | 20 | 1 | 2,92E-05 | 3,0 | 1,4E-03 |
| positive regulation of cell adhesion | 13 | 1 | 3,13E-05 | 4,2 | 1,4E-03 |
| cell proliferation | 37 | 3 | 3,14E-05 | 2,1 | 1,4E-03 |
| response to DNA damage stimulus | 41 | 3 | 3,35E-05 | 2,0 | 1,5E-03 |
| regulation of mitotic cell cycle | 20 | 1 | 3,97E-05 | 2,9 | 1,8E-03 |
| cell junction organization | 11 | 1 | 4,92E-05 | 4,8 | 2,1E-03 |
| regulation of cytoskeleton organization | 20 | 1 | 6,19E-05 | 2,8 | 2,7E-03 |
| negative regulation of epithelial cell proliferation | 10 | 1 | 6,66E-05 | 5,2 | 2,8E-03 |
| positive regulation of cell-substrate adhesion | 10 | 1 | 6,66E-05 | 5,2 | 2,8E-03 |
| blood vessel morphogenesis | 31 | 2 | 6,82E-05 | 2,2 | 2,9E-03 |
| regulation of transferase activity | 31 | 2 | 7,49E-05 | 2,2 | 3,1E-03 |
| negative regulation of signal transduction | 28 | 2 | 7,59E-05 | 2,3 | 3,1E-03 |
| regulation of cell-substrate adhesion | 12 | 1 | 7,75E-05 | 4,2 | 3,1E-03 |
| negative regulation of catalytic activity | 20 | 1 | 8,22E-05 | 2,8 | 3,3E-03 |
| focal adhesion formation | 6 | 0 | 8,52E-05 | 10,5 | 3,3E-03 |
| negative regulation of cellular component organization | 19 | 1 | 8,55E-05 | 2,9 | 3,3E-03 |
| G1/S transition of mitotic cell cycle | 10 | 1 | 9,20E-05 | 5,0 | 3,5E-03 |
| regulation of kinase activity | 30 | 2 | 9,48E-05 | 2,2 | 3,6E-03 |
| microtubule cytoskeleton organization | 21 | 2 | 1,01E-04 | 2,7 | 3,8E-03 |
| mammary gland development | 18 | 1 | 1,01E-04 | 2,9 | 3,7E-03 |
| transforming growth factor beta receptor signaling pathway | 13 | 1 | 1,02E-04 | 3,8 | 3,7E-03 |
| positive regulation of cell differentiation | 28 | 2 | 1,13E-04 | 2,2 | 4,1E-03 |
| negative regulation of cell communication | 29 | 2 | 1,32E-04 | 2,2 | 4,7E-03 |
| regulation of protein kinase activity | 29 | 2 | 1,32E-04 | 2,2 | 4,7E-03 |
| negative regulation of MAPKKK cascade | 7 | 1 | 1,44E-04 | 7,5 | 5,0E-03 |
| mitotic sister chromatid segregation | 8 | 1 | 1,45E-04 | 6,2 | 5,0E-03 |
| regulation of cell development | 26 | 2 | 1,48E-04 | 2,3 | 5,1E-03 |
| regulation of microtubule cytoskeleton organization | 11 | 1 | 1,49E-04 | 4,3 | 5,1E-03 |
| negative regulation of cell cycle | 12 | 1 | 1,58E-04 | 3,9 | 5,3E-03 |
| angiogenesis | 23 | 2 | 1,72E-04 | 2,4 | 5,7E-03 |
| cell morphogenesis | 41 | 3 | 1,73E-04 | 1,9 | 5,7E-03 |
| cellular component morphogenesis | 45 | 3 | 1,76E-04 | 1,8 | 5,7E-03 |
| positive regulation of ossification | 6 | 0 | 1,80E-04 | 9,3 | 5,8E-03 |
| negative regulation of protein kinase activity | 13 | 1 | 1,91E-04 | 3,6 | 6,1E-03 |
| negative regulation of kinase activity | 13 | 1 | 1,91E-04 | 3,6 | 6,1E-03 |
| sister chromatid segregation | 8 | 1 | 2,16E-04 | 5,9 | 6,8E-03 |

| | | | | | |
|---|---|---|---|---|---|
| regulation of cell death | 64 | 5 | 2,23E-04 | 1,6 | 6,9E-03 |
| developmental growth | 19 | 1 | 2,24E-04 | 2,7 | 6,9E-03 |
| cell cycle checkpoint | 13 | 1 | 2,32E-04 | 3,5 | 7,1E-03 |
| bone mineralization | 7 | 1 | 2,36E-04 | 7,0 | 7,1E-03 |
| growth | 29 | 2 | 2,49E-04 | 2,1 | 7,4E-03 |
| extracellular matrix organization | 19 | 1 | 2,55E-04 | 2,6 | 7,5E-03 |
| microtubule-based movement | 19 | 1 | 2,55E-04 | 2,6 | 7,5E-03 |
| negative regulation of transferase activity | 13 | 1 | 2,81E-04 | 3,4 | 8,2E-03 |
| regulation of ossification | 13 | 1 | 2,81E-04 | 3,4 | 8,2E-03 |
| mammary gland duct morphogenesis | 9 | 1 | 3,19E-04 | 4,8 | 9,2E-03 |
| multicellular organismal metabolic process | 9 | 1 | 3,19E-04 | 4,8 | 9,2E-03 |
| cell projection organization | 41 | 3 | 3,39E-04 | 1,8 | 9,7E-03 |
| gland development | 29 | 2 | 3,51E-04 | 2,1 | 1,0E-02 |
| digestive tract morphogenesis | 10 | 1 | 3,72E-04 | 4,2 | 1,0E-02 |
| negative regulation of macromolecule metabolic process | 58 | 4 | 3,72E-04 | 1,6 | 1,0E-02 |
| regulation of cell division | 11 | 1 | 3,83E-04 | 3,9 | 1,1E-02 |
| regulation of apoptosis | 62 | 5 | 4,10E-04 | 1,6 | 1,1E-02 |
| regulation of epithelial cell proliferation | 14 | 1 | 4,20E-04 | 3,1 | 1,1E-02 |
| neuron differentiation | 48 | 4 | 4,54E-04 | 1,7 | 1,2E-02 |
| chordate embryonic development | 50 | 4 | 4,55E-04 | 1,7 | 1,2E-02 |
| ossification | 19 | 1 | 4,71E-04 | 2,5 | 1,2E-02 |
| digestive system development | 10 | 1 | 4,74E-04 | 4,1 | 1,2E-02 |
| negative regulation of biosynthetic process | 51 | 4 | 5,08E-04 | 1,6 | 1,3E-02 |
| regulation of cell motion | 19 | 1 | 5,29E-04 | 2,5 | 1,4E-02 |
| embryonic development ending in birth or egg hatching | 50 | 4 | 5,66E-04 | 1,6 | 1,4E-02 |
| regulation of programmed cell death | 62 | 5 | 5,66E-04 | 1,6 | 1,4E-02 |
| cell cycle arrest | 13 | 1 | 5,72E-04 | 3,2 | 1,4E-02 |
| ureteric bud development | 11 | 1 | 5,84E-04 | 3,7 | 1,5E-02 |
| chromosome condensation | 8 | 1 | 6,05E-04 | 5,1 | 1,5E-02 |
| regulation of cell migration | 17 | 1 | 7,29E-04 | 2,6 | 1,8E-02 |
| negative regulation of cellular biosynthetic process | 50 | 4 | 7,36E-04 | 1,6 | 1,8E-02 |
| DNA packaging | 18 | 1 | 7,42E-04 | 2,5 | 1,8E-02 |
| positive regulation of cell division | 10 | 1 | 7,47E-04 | 3,9 | 1,8E-02 |
| collagen metabolic process | 8 | 1 | 8,16E-04 | 4,9 | 1,9E-02 |
| extracellular structure organization | 23 | 2 | 8,69E-04 | 2,2 | 2,0E-02 |
| neuron development | 37 | 3 | 8,98E-04 | 1,8 | 2,1E-02 |
| negative regulation of nitrogen compound metabolic process | 47 | 3 | 9,17E-04 | 1,6 | 2,1E-02 |
| protein amino acid phosphorylation | 68 | 5 | 9,18E-04 | 1,5 | 2,1E-02 |
| positive regulation of organelle organization | 12 | 1 | 9,21E-04 | 3,2 | 2,1E-02 |
| multicellular organismal macromolecule metabolic process | 8 | 1 | 1,08E-03 | 4,7 | 2,4E-02 |
| DNA repair | 30 | 2 | 1,12E-03 | 1,9 | 2,5E-02 |
| cell morphogenesis involved in differentiation | 29 | 2 | 1,14E-03 | 1,9 | 2,5E-02 |
| T cell homeostasis | 7 | 1 | 1,14E-03 | 5,4 | 2,5E-02 |
| negative regulation of macromolecule biosynthetic process | 48 | 4 | 1,24E-03 | 1,6 | 2,7E-02 |
| negative regulation of cell death | 32 | 2 | 1,30E-03 | 1,8 | 2,8E-02 |

| | | | | | |
|---|---|---|---|---|---|
| negative regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process | 46 | 3 | 1,33E-03 | 1,6 | 2,9E-02 |
| positive regulation of biomineral formation | 4 | 0 | 1,37E-03 | 14,0 | 2,9E-02 |
| positive regulation of bone mineralization | 4 | 0 | 1,37E-03 | 14,0 | 2,9E-02 |
| in utero embryonic development | 34 | 3 | 1,38E-03 | 1,8 | 2,9E-02 |
| gut development | 10 | 1 | 1,39E-03 | 3,6 | 2,9E-02 |
| negative regulation of protein kinase cascade | 8 | 1 | 1,41E-03 | 4,5 | 3,0E-02 |
| regulation of calcium ion transport into cytosol | 5 | 0 | 1,43E-03 | 8,8 | 3,0E-02 |
| sensory organ development | 33 | 2 | 1,43E-03 | 1,8 | 3,0E-02 |
| cortical cytoskeleton organization | 6 | 0 | 1,45E-03 | 6,5 | 3,0E-02 |
| cell motion | 43 | 3 | 1,56E-03 | 1,6 | 3,2E-02 |
| regulation of protein localization | 16 | 1 | 1,65E-03 | 2,5 | 3,3E-02 |
| bone development | 19 | 1 | 1,70E-03 | 2,3 | 3,4E-02 |
| negative regulation of apoptosis | 31 | 2 | 1,75E-03 | 1,8 | 3,5E-02 |
| gut morphogenesis | 8 | 1 | 1,81E-03 | 4,3 | 3,6E-02 |
| chromosome organization | 46 | 3 | 1,87E-03 | 1,6 | 3,7E-02 |
| cellular response to stress | 46 | 3 | 1,87E-03 | 1,6 | 3,7E-02 |
| exocrine system development | 10 | 1 | 2,02E-03 | 3,4 | 3,9E-02 |
| lung development | 18 | 1 | 2,16E-03 | 2,3 | 4,2E-02 |
| metanephros development | 12 | 1 | 2,35E-03 | 2,9 | 4,5E-02 |
| regulation of intracellular transport | 10 | 1 | 2,41E-03 | 3,3 | 4,6E-02 |
| negative regulation of programmed cell death | 31 | 2 | 2,41E-03 | 1,8 | 4,6E-02 |
| negative regulation of cyclin-dependent protein kinase activity | 5 | 0 | 2,42E-03 | 7,8 | 4,6E-02 |
| regulation of neurogenesis | 20 | 1 | 2,56E-03 | 2,1 | 4,8E-02 |
| respiratory tube development | 18 | 1 | 2,63E-03 | 2,2 | 4,9E-02 |
| nucleoside monophosphate metabolic process | 12 | 1 | 2,70E-03 | 2,8 | 5,0E-02 |
| biomineral formation | 9 | 1 | 2,71E-03 | 3,6 | 5,0E-02 |

# Supplementary Methods: Theory

To explore the range of proliferative heterogeneity and the patterns of fate choice in murine tail interfollicular epidermis (IFE), we have implemented three labeling strategies. Two rely upon inducible genetic labeling using targeted promoters, Inv and K14. The third involves the incorporation and continuous dilution of an H2B-GFP labeling system. To analyze the resulting data, we rely on the application of tested theoretical methodologies. Since aspects of this approach have been reviewed in the recent literature [12-14,21], we present the details of the methodology as Supplementary material reporting only the key findings in the main text.

In section 1, we will explore the lineage tracing data following induction of the InvolCreER:Rosa26-YFP mouse. Here we will lean heavily on the analysis of Clayton et al. [12]. In section 2, we will then analyze the lineage tracing study following induction of the K14CreER:Rosa26-YFP mouse. By drawing on the results of the Inv study, we will provide evidence for a cellular hierarchy in which a second, near-quiescent, stem-like population gives rise to progenitor cells. Once agin, from the clonal fate data, we will define the fate behaviour of this population. Finally, in sections 3 and 4, we will show how BrdU pulse-chase measurements combined with label dilution following induction of the H2B-GFP mouse provides an independent test of the results of the lineage tracing study.

## 1 Lineage tracing using the InvolCreER:Rosa26-YFP mouse

Following TAM administration of the Inv mouse, clones are induced extensively in both the basal and suprabasal cell layers of the IFE. Since suprabasal cells are terminally differentiated, clones derived from the latter will remain as single cells, progressively stratify, and eventually become shed as tissue turns over. By contrast, cells in the basal layer comprise both progenitors, P, and terminally differentiated cells, D. The former may give rise to clones that persist and contribute to the long-term turnover of tissue. To follow the evolution of a defined population, following Ref. [12], we will focus on clones that retain a basal attachment defining a "surviving clone" as one that hosts at least one basal layer cell.

### 1.1 Scaling

Following induction, the population of surviving clones shows a progressive reduction in number (Fig. 1c), while their average size steadily increases (Fig. S2d) such that the overall number of labelled cells remains approximately constant. To look for signatures of the underlying pattern of fate we must turn to the clone size distribution. Beginning with the work of Ref. [12], it has been shown that long-term scaling characteristics of the clonal fate data can discriminate between maintenance strategies involving invariant asymmetrical cell division and population asymmetric self-renewal (for a review of the principles of the methodology, see Ref. [21]). In particular, in an equipotent cell population, if the loss of a progenitor cell is perfectly compensated by the duplication of others, the long-term clone size distribution must asymptote to a scaling form in which the chance of finding a clone larger than some multiple of the average becomes constant. Formally, this scaling behavior is expressed as

$$P_n(t) = f\left(\frac{n}{\langle n(t)\rangle}\right), \tag{1}$$

where $P_n(t)$ describes the chance of finding a surviving clone with a size larger than $n$ (cells). Applied to the present study, we find that the long-term clonal evolution of the basal layer compartment indeed conforms to this scaling paradigm, with the data converging slowly onto a scaling form described the function,

$f(x) = \exp[-x]$ (Fig. S2e). Such behavior is consistent with that found by Jones and colleagues [12,13]from lineage tracing studies of mouse epidermis using a ubiquitous promoter, and suggests that tissue is maintained by a single equipotent progenitor cell pool following a pattern of population asymmetric self-renewal.

## 1.2 Progenitor cell dynamics

With this evidence, we then analyzed the range of clonal fate data (disaggregated into basal and suprabasal compartments) using the cell fate behaviors defined in Ref. [12]: Specifically, we suppose that tissue is maintained by a single progenitor cell population following a pattern of balanced stochastic cell fate in which cell division can lead to one of three possible fate outcomes; two dividing cells, two non-dividing cells, and one dividing and one non-dividing cell. Cells commited to terminal differentiation, D, are then swept from the basal layer into the suprabasal cell layers (denoted by D$^*$), leading to the following dynamics,

$$\mathrm{P} \overset{\lambda}{\mapsto} \begin{cases} \mathrm{P} & \mathrm{P} & \text{Pr. } r(1-\Delta/2) \\ \mathrm{P} & \mathrm{D} & \text{Pr. } 1-2r \\ \mathrm{D} & \mathrm{D} & \text{Pr. } r(1+\Delta/2) \end{cases}, \qquad \mathrm{D} \overset{\Gamma}{\mapsto} \mathrm{D}^* . \tag{2}$$

Here, for reasons that will become clear, we have included an additional factor $\Delta \ll 1$ to accommodate the potential for a (small) imbalance in the rates of proliferation and differentiation. For simplicity, we will assume that the relative balance between proliferation and differentiation is achieved on the basis of internal (cell-intrinsic) regulation. However, previous studies [21]have shown that, in the two-dimensional geometry, pertinent to the epithelia architecture of IFE, mechanisms of balance that rely on extrinsic regulation give rise to a quantitatively similar pattern of clonal evolution.

With this paradigm, the clonal evolution of the basal layer compartment is specified by three independent parameters, the rate of progenitor cell division, $\lambda$, the rate of cell stratification from the basal to the suprabasal cell layer, $\Gamma$, and the relative balance between symmetric and asymmetric fate, controlled by the parameter $r$. With this definition, taking $\Delta = 0$, the ratio of progenitor cells to terminally differentiated basal layer cells is fixed by the rates according to the relation, $\frac{n_\mathrm{P}}{n_\mathrm{D}} = \frac{\Gamma}{\lambda}$, where $n_\mathrm{P}$ and $n_\mathrm{D}$ denote the number density of progenitor cells and differentiated basal layer cells, respectively.

Finally, to capture the clonal dynamics of the suprabasal cell population, we need to introduce a fourth parameter, $\sigma$, describing the rate at which cells in the suprabasal cell layer become undetectable either through shedding, or through the loss of their nuclei, $\mathrm{D}^* \overset{\sigma}{\mapsto} \oslash$. Once again, in homeostasis, since cell loss must be perfectly compensated by the generation of terminally differentiated cells, $\sigma$ is fixed by the relative fraction of basal to suprabasal layer cells according to the relation, $\frac{\sigma}{\lambda} = \frac{n_\mathrm{P}}{n_{\mathrm{D}^*}}$. Here, we suppose that, once transferred to the suprabasal cell layers, cells are equally likely to be shed, independent of their position within these layers. Although such an assumption is surely an over-simplification, further refinements of the model to accommodate details of the stratified outer layers would complicate the analysis without providing significant new insight. We will therefore neglect such effects.

Taken together, the dynamics of the progenitor cell population describe a critical branching or Galton-Watson birth-death process. In this framework, the time-evolution of the surviving clone size distribution is described by the Master equation [12],

$$\begin{aligned} \dot{P}_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}} = \\ \lambda\,[r(1-\Delta/2)P_{n_\mathrm{P}-1,n_\mathrm{D},n_{\mathrm{D}^*}} + r(1+\Delta/2)P_{n_\mathrm{P}+1,n_\mathrm{D}-2,n_{\mathrm{D}^*}} + (1-2r)P_{n_\mathrm{P},n_\mathrm{D}-1,n_{\mathrm{D}^*}} - P_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}] \\ +\Gamma(P_{n_\mathrm{P},n_\mathrm{D}+1,n_{\mathrm{D}^*}-1} - P_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}) + \sigma(P_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}+1} - P_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}})\,. \end{aligned} \tag{3}$$

defining the rate of change of the probability $P_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}$ describing the chance of finding a clone with $n_\mathrm{P}$ basal cells, $n_\mathrm{D}$ differentiated basal layer cells, and $n_{\mathrm{D}^*}$ suprabasal cells. This equation must be solved subject to the initial conduction. In the following, we will assume that the induced terminally differentiated basal layer cells make a negligible contribution to the clonal dynamics since (a) they will turn out to be small in number as compared to the progenitor pool, and (b), they are rapidly swept into the suprabasal cell layers and therefore make no contribution to the surviving clone fraction. We therefore impose the initial condition $P_{n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}(0) = \delta_{n_\mathrm{P},1}\delta_{n_\mathrm{D},0}\delta_{n_{\mathrm{D}^*},0}$, where $\delta_{nm}$ denotes the Krönecker delta symbol.

Here, for simplicity, we have assumed that all elements of the dynamics (division, stratification, shedding) are described by a Poisson random (Markov) process in which the timing between successive events is

uncorrelated. Once again, over several rounds of division, any correlations due to synchrony effects will be rapidly erased from the clonal record, and can therefore be neglected. Although the solution cannot be usefully obtained analytically, as a first order differential equation, the Master equation can be easily integrated numerically using the Euler method, and the results fit to the data. For brevity, we refer to the Ref. [12]for details, and here simply address the findings.

### 1.3 Fit to the data

To fit the model to the experimental data, we can adjust five parameters, the rates $\lambda$, $\Gamma$, and $\sigma$, the fate balance $r$, and the degree of imbalance $\Delta$. However, the clonal fate data already provides some constraints limiting the available phase space. First, the long term persistence of clones (over the one year timecourse) already suggests that any degree of imbalance, $\Delta$, must be small. Under these conditions of homeostasis, for each progenitor cell division, the number of differentiated cells must rise on average by one while the progenitor cell number remains constant. By analyzing the rate of increase of the average total clone size (Fig. S2d), we can deduce a progenitor cell division rate, $\lambda$, a little in excess of once per week. Second, from the analysis of the clonal composition at 3.5 days post-induction, we can see that, of the surviving clones with two labelled cells, some 75% involve one basal and one suprabasal cell showing (a) that the stratification rate, $\Gamma$ is much faster than the division rate $\lambda$ (from which it follows that $n_{\rm D}/n_{\rm P} \ll 1$), and (b) that the vast majority of cell divisions result in asymmetric cell fate. Finally, from the theoretical analysis of the Master equation [12], it has been shown the average basal clone size of surviving clones must converge on the linear time-dependence, $\langle (n_{\rm P}(t) + n_{\rm D}(t)) \rangle \simeq (1 + n_{\rm D}/n_{\rm P})(1 + r\lambda t)$, at times in excess of $1/r\lambda$ post-labelling. From the data, we can then infer that $1/r\lambda \approx 10$ weeks, consistent with an $r$ value of around 0.1.

Using these rough estimates as a guide, when compared with the wide range of experimental data covering the basal, suprabasal and total cell number (Fig. 2a,b and S2a), we find that the best fit of the model is obtained for a progenitor cell division rate of $\lambda = 1.21 \pm 0.05$ per week, a stratification rate of $\Gamma = 4.8 \pm 0.3$ per week, a shedding rate of $\sigma = 0.48 \pm 0.04$ per week, and $r = 0.1 \pm 0.01$. (To obtain the best fit we have further included a small one-day induction time offset, which we attribute to the delayed action of the TAM.) By fine-tuning the data at the longest timepoint, we further estimate the degree of imbalance to be $\Delta \simeq 0.14$, i.e. 9.3% of progenitor cell divisions lead to symmetrical duplication while 10.7% lead to terminal divisions. With these values of the parameters, the resulting clone size distributions can be compared to the theoretical prediction. The results, shown in Figs. 2h,i and S2d,f, reveal an excellent agreement between theory and experiment. Moreover, with these values of the fitting parameters, the ratio $\frac{n_{\rm P}+n_{\rm D}}{n_{\rm D}^*} = \frac{n_{\rm P}}{n_{\rm D}^*}(1 + n_{\rm D}/n_{\rm P}) = \frac{\sigma}{\lambda}(1 + \lambda/\Gamma) \simeq 0.5$, a figure consistent with the measured ratio of basal to suprabasal cells.

From these findings we can infer that the large majority of progenitor cell divisions (ca. 4 in 5) result in asymmetric fate outcome, broadly consistent with the results of Ref. [12]. At the same time, the stratification rate, $\Gamma$, being some four times faster than the proliferation rate, $\lambda$, shows that, on commitment to terminal differentiation, cells rapidly loose their attachment to the basement membrane and stratify, i.e. progenitor cells outnumber basal layer terminally differentiated cells by a ratio of 4 to 1. In summary, analysis of the Inv clone fate data suggests that, as proposed by Clayton *et al.* [12], murine tail IFE is maintained by a single population of committed progenitor (CP) cells following balanced stochastic fate.

## 2 Lineage tracing using the K14CreER:Rosa26-YFP mouse

With these results in hand, we now turn to the findings of the lineage tracing assay using the K14 promoter. In this case, in contrast to the Inv study, the vast majority of clones are derived from basal layer cells. Moreover, despite the mouse-to-mouse variation in labeling efficiency, the survival data does not suggest a significant loss in clone number over the timecourse (Fig. 1c). At the same time, the average basal layer clone size shows a striking behavior in which a near-instantaneous expansion in clone size gives way to a slow progressive increase, reaching a plateau and potentially diminishing slightly at the longest timepoint (Fig. S3c). Indeed, such behavior is reinforced by study of the detailed clone size distribution (Figs. 3a,b and S3a) which records a rapid expansion of clones at short times, with clones involving 4-5 basal layer cells at just 3.5 days post-induction, followed by a slow expansion in which the expansion of the basal layer clone size distribution appears to become "arrested" from 4 weeks post-labeling.

## 2.1 Stem cell dynamics

Although the clonal fate behavior potentially admits a range of possible interpretations of increasing complexity, the analysis of the data is heavily constrained by the findings of the Inv study. Therefore, to address the data, we will introduce a simple hypothesis, which we will then challenge with further experimental investigation. Specifically, we will suppose that the CP cell population in mouse tail is "underpinned" by a second cell compartment – a stem cell (SC) pool – which, in normal homeostasis remains largely quiescent. Following TAM administration, we will suppose that the K14 promoter leads to the induction of cells in both the CP cell pool and the SC compartment. The accelerated expansion of the clone size following induction, which is far more rapid than that exhibited by the Inv mouse, suggests that SCs are induced on entry into cell division. At the same time, the rapid deceleration of the expansion at subsequent timepoints suggests that, following division, SCs re-enter a dormant or quiescent phase while their CP cell progeny go on to proliferate and differentiate. For these clones, the quiescent SC "mother" provides an anchor to the basal layer ensuring long-term clonal persistence, consistent with the experimental data.

We are therefore led to consider the hypothesis for the SC compartment, S,

$$\mathrm{S} \overset{\lambda_\mathrm{S}}{\mapsto} \begin{cases} \mathrm{S} & \mathrm{S} & \text{Pr. } r_\mathrm{S} \\ \mathrm{S} & \mathrm{P} & \text{Pr. } 1-2r_\mathrm{S} \\ \mathrm{P} & \mathrm{P} & \text{Pr. } r_\mathrm{S} \end{cases}, \tag{4}$$

where the CP cell progeny go on to follow the pattern of fate elucidated from the Inv study. Here we have allowed SCs to choose stochastically from one of three possible fates, with frequencies set by fixed probabilities. Moreover, to ensure long-term homeostasis, we suppose that, as with the progenitor pool, the symmetrical fates are approximately balanced. However, once again, we stress that the question of the pattern of fate regulation – internal vs. external - is beyond the resolution of the lineage tracing study, and we choose the former merely for simplicity.

Cast in this form, the model introduces two further parameters into the scheme – the SC division rate, $\lambda_\mathrm{S}$, which, following the discussion above, we anticipate to be very slow as compared to that of the CP cell pool, and the balance between symmetrical and asymmetrical SC division fixed by the parameter $r_\mathrm{S}$. However, it is important to recognize that, if the SC division rate is sufficiently small, the development of clones will be largely controlled by the CP cell progeny of the SCs, and therefore fixed by the dynamics predicted by the Inv study *with the same parameters*.

Once again, the dynamics of the stem/progenitor cell hierarchy can be cast in the form of a Master equation, extended to include the SC compartment,

$$\begin{aligned} \dot{P}_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}} &= \lambda_\mathrm{S}\left[r_\mathrm{S}\left(P_{n_\mathrm{S}-1,n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}} + P_{n_\mathrm{S}+1,n_\mathrm{P}-2,n_\mathrm{D},n_{\mathrm{D}^*}}\right) + (1-2r_\mathrm{S})P_{n_\mathrm{S},n_\mathrm{P}-1,n_\mathrm{D},n_{\mathrm{D}^*}} - P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}\right] \\ &\quad +\lambda(rP_{n_\mathrm{S},n_\mathrm{P}-1,n_\mathrm{D},n_{\mathrm{D}^*}} + rP_{n_\mathrm{S},n_\mathrm{P}+1,n_\mathrm{D}-2,n_{\mathrm{D}^*}} + (1-2r)P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D}-1,n_{\mathrm{D}^*}} - P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}) \\ &\quad +\Gamma(P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D}+1,n_{\mathrm{D}^*}-1} - P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}) + \sigma(P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}+1} - P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}})\,. \end{aligned} \tag{5}$$

Here the probability $P_{n_\mathrm{S},n_\mathrm{P},n_\mathrm{D},n_{\mathrm{D}^*}}(t)$ now includes the number of cells, $n_\mathrm{S}$, that belong to the SC compartment. As before, although this equation is formally intractable analytically, it is easily integrated numerically using the Euler method and the results compared with the experimental data.

## 2.2 Fit to the data

In seeking a fit of the model to the data, we are assisted by two aspects of the cell dynamics. First, as discussed above, the cell fate behavior of the CP cells is fixed by the findings of the Inv study and may be imported directly in the analysis of the K14 data. Second, since SC are largely quiescent, yet enter cycle immediately following induction, the early time behavior – time scales shorter than the typical SC cycle time – is governed solely by the relative labeling efficiency (CP vs. SC) and the known CP cell dynamics. Focussing on this short-term data, we find a best fit of the model to the range of experimental data (basal, suprabasal and total) if we assume that some $50 \pm 5\%$ of cells induced by the K14 promoter belong to the

CP cell compartment, with the remaining cells belonging to the SC pool. Following induction, we estimate that some $80 \pm 10\%$ of SCs undergo an immediate division with $80 \pm 10\%$ resulting in asymmetrical fate outcome (i.e., as with the CP pool, the balance between duplication and differentiation is set by same ratio $r_S = 0.1 \pm 0.05$). However, since the dynamics is dominated by the CP cell compartment, the fits (shown in Figs. 3h,i) accommodate a range of frequencies as evidenced by the large error bar onthe fate fractions. Finally, after division, stem cells fall out of cycle, with the data consistent with a division rate, $\lambda_S$, of less than 3-5 per year, some 10-20 times slower than the CP cell division rate.

### 2.3 Discussion

In summary, we have used the clonal fata for the Inv and K14 mice to infer both the existence of the cellular hierarchy and define the respective division, stratification, and loss rates, the balance of fates, and the relative labelling efficiency of the promoters. Several comments on this fitting procedure are in order. First, the lineage tracing protocol is always vulnerable to the labelling efficiency of the constituitive cell types. We can never rule out the possibility that an unlabelled, and therefore hidden, cell population makes an important contribution to homeostasis (or repair). However, to some extent, we have insulated ourselves against this possibility by ensuring that the cells that are labelled constitute a self-renewing population.

Second, it is important to recognize that the important discriminatory features of the two labelling protocols lie in the shorter-time dynamics. Indeed, from the theoretical curves and a comparison of the data, the clone size distributions of the two labelling strategies are predicted to converge as the clones become dominated by the CP cell contribution (Fig. S3e,f).

Thirdly, we have assumed that the SC form a homogeneous pool in which any may enter cycle with equal probability. However, it seems likely that SC activity may be regulated by extrinsic factors. In this case, it may be that some SCs cycle more rapidly than others. Here we have made a fit to a population average. But, while the overall average cell cycle time remains slow, such heterogeneity could not be ruled out by the data.

Lastly, the clonal fate data provides one window onto the dynamics of the stem and CP cell population. However, to be confident in the findings, it is crucial to find other independent experimental measures that can be used to validate some of the predictions of the model. In the following, we will discuss two such experiments – the incorporation of the thymadine analogue, BrdU, and the H2B-GFP dilute assay.

## 3 Continuous BrdU incorporation

Following continuous BrdU incorportation, the average cell division time of the progenitor cells can be estimated in the following way. If we denote as $f_S$ the fraction of time a progenitor spends in S-phase, the percentage of the progenitor pool that has taken up BrdU after a time $T_{chase}$ of continuous BrdU incorporation is given approximately by

$$\mathrm{BrdU}\% \simeq \min\left[1, f_S + \frac{T_{\mathrm{chase}}}{T_{\mathrm{div.}}}\right] \times 100\%\,,$$

where the first term accounts for the fraction of cells which are in S-phase at the time of the pulse, and the second term accounts for the accumulated fraction which take up BrdU in the chase. Inverting, we can estimate the division time as

$$T_{\mathrm{div.}} \simeq \frac{T_{\mathrm{chase}}}{\mathrm{BrdU}\%/100\% - f_S}\,.$$

From nuclear staining, we can deduce that cells spend ca. 91% of the average cell cycle time in S/G2/M, from which can be place an upper bound on $f_S$.

To apply this estimate, we will assume that CP cells represent the majorty of basal layer cells. From the 24h BrdU chase, we find some $27 \pm 2\%$ of the basal layer population are positive for BrdU while, after 72h, we find $57 \pm 9\%$. From this result, we find an average cell division time of $6 \pm 1$ days, consistent with the results of the clonal fate data.

# 4 Analysis of the H2B-GFP mouse

Prior to Dox administration, cells in H2B-GFP mouse express high levels of GFP. Following Dox administration, the division of progenitor cells leads to the sequential dilution of GFP expression, with roughly equal partitioning of label between cells. To address the quantiative behavior of the data, we will use the predicted stem/CP cell dynamics to infer the statistics of the GFP expression in the basal cell compartment over the timecourse.

## 4.1 Theory

If we denote as $\mathrm{S}_n$ a SC that has undergone $n$ rounds of cell division post-induction, on division, it will transfer either to a SC, $\mathrm{S}_{n+1}$ or CP cell $\mathrm{P}_{n+1}$, with equal probability. Similarly, a CP cell $\mathrm{P}_n$ will transfer to a CP cell $\mathrm{P}_{n+1}$ or a terminally differentiated basal layer cell $\mathrm{D}_{n+1}$, with approximately equal probability. Finally, at a rate $\Gamma$, a terminally differentiated cell will detach and transfer from the basal to the suprabasal cell layer. As a result, from the perspective of the H2B-GFP assay, such a cell will be lost from the ensemble of basal layer cells. Therefore, together, we can represent the fate of SC and CP cells by the cascade,

$$\mathrm{S}_n \overset{2\lambda}{\mapsto} \begin{cases} \mathrm{S}_{n+1} & \text{Pr. } 1/2 \\ \mathrm{P}_{n+1} & \text{Pr. } 1/2 \end{cases}, \qquad \mathrm{P}_n \overset{2\lambda}{\mapsto} \begin{cases} \mathrm{P}_{n+1} & \text{Pr. } 1/2 \\ \mathrm{D}_{n+1} & \text{Pr. } 1/2 \end{cases}, \qquad \mathrm{D}_n \overset{\Gamma}{\mapsto} \oslash\,. \tag{6}$$

Note here that the putative stochasticity of SC and CP cell fate choice is irrelevant, as evidenced by the independence of the processes on the parameters $r$ and $r_{\mathrm{S}}$. We do, however, rely upon the symmetrical partitioning of GFP label between daughters.

For these processes, we can again write down the set of coupled Master equations,

$$\begin{aligned} \dot{s}_n &= \lambda_{\mathrm{S}}\left[s_{n-1}(t)\theta(n-1) - 2s_n(t)\right] \\ \dot{p}_n &= \lambda_{\mathrm{S}} s_{n-1} + \lambda\left[p_{n-1}(t)\theta(n-1) - 2p_n(t)\right] \\ \dot{d}_n &= \lambda\, p_{n-1}(t) - \Gamma d_n(t) \\ \dot{p}_{\mathrm{loss}} &= \Gamma\sum_{n=1}^{\infty} d_n(t)\,, \end{aligned} \tag{7}$$

where $s_n$, $p_n$ and $d_n$ denote the probability that a stem, progenitor and differentiated basal layer cell has undergone $n$ rounds of cell division, and $p_{\mathrm{loss}}$ denotes the loss probability. Here $\theta(n)$ denotes the discrete step function and takes the value of zero for $n < 0$ and unity for $n \geq 0$. These equations must be solved subject to appropriate boundary conditions. Note that the dynamical equations conserve probability such that, as required, $p_{\mathrm{loss}}(t) + \sum_{n=0}^{\infty}[s_n(t) + p_n(t) + d_n(t)] = n_{\mathrm{S}} + n_{\mathrm{P}} = 1$.

In the following, we are interested in the probability that a cell in the basal layer has undergone $n$ rounds of division. For a SC, with the boundary condition $p_n(0) = d_n(0) = p_{\mathrm{loss}}(0) = 0$ and $s_n(0) = \delta_{n,0}$, this probability is given by

$$P_n^{(\mathrm{S})}(t) = s_n(t) + \sum_{m=0}^{n-1}\int_0^t s_{n-m-1}(t')\ \lambda_{\mathrm{S}} dt'\ P_m^{(\mathrm{P})}(t-t')\,,$$

where $P_n^{(\mathrm{P})}(t)$ denotes the corresponding probability that, after $n$ rounds of division, a CP cell has given rise to progeny that still persist in the basal cell layer. The first term simply denotes the "survival" probability of the SC itself, while the second term sums all contributions associated with the probability that the $(n-m)$th SC division at time $0 < t' < t$ results in a CP cell whose progeny persist in the basal cell layer until time $t$, by which time it has undergone a further $m$ rounds of division. As a Poisson-like process, the equation for $s_n(t)$ is easily integrated and obtains the solution, $s_n(t) = \frac{(\lambda_{\mathrm{S}} t)^n}{n!} e^{-2\lambda_{\mathrm{S}} t}$.

To determine $P_n^{(\mathrm{P})}(t)$, we can follow a similar iterative procedure setting

$$P_n^{(\mathrm{P})}(t) = p_n(t) + \int_0^t p_{n-1}(t')\ \lambda dt'\ e^{-\Gamma(t-t')}\,,$$

where $p_n(t)$ must be solved subject to the boundary condition $s_n(0) = d_n(0) = p_{\rm loss}(0) = 0$ and $p_n(0) = \delta_{n,0}$. Once again, the first term represents the chance that a CP cell persists after $n$ rounds of division, and the second term (effectively $d_n(t)$ for these boundary conditions) represents the chance that the $n$th division gives rise to differentiated cell at a time $t'$ which has survived until time $t$. For these boundary conditions, the CP cells also follow a Poission process, with $p_n(t) = \frac{(\lambda t)^n}{n!}e^{-2\lambda t}$, from which integration obtains

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

where $\Gamma[n,x] = \int_x^\infty dt\, t^{n-1}e^{-t}$ denotes the incomplete Gamma function.

Piecing together these results, we thus obtain the total probability that a persisting basal layer cell has derived from $n$ rounds of division of the SC and CP cell population is given by

$$P_n^{\rm tot}(t) = n_{\rm S}\frac{P_n^{(\rm S)}(t)}{\sum_{n=0}^{\rm n_{max}} P_n^{(\rm S)}(t)} + n_{\rm P}\frac{P_n^{(\rm P)}(t)}{\sum_{n=0}^{\rm n_{max}} P_n^{(\rm P)}(t)} ,$$

where $n_{\rm S} = 1 - {\rm n_P}$. For completeness, one may note that, with $\sum_{n=1}^\infty d_n(t) = (1-\Gamma/\lambda)^{-1}(e^{-\Gamma t} - e^{-\lambda t})$, for induced CP cells, $\sum_{n=0}^{\rm n_{max}} P_n^{(\rm P)}(t) = p_{\rm surv.}(t) = 1 - p_{\rm loss}(t) = (1-\lambda/\Gamma)^{-1}(e^{-\lambda t} - (\lambda/\Gamma)e^{-\Gamma t})$. More generally, a useful closed form identity can not be obtained.

Finally, the analysis above supposes that the basal layer cells can be resolved however dilute the label. In practice, there will be a threshold, $n_{\rm max}$ beyond which the GFP label will be lost in the background of the FACS profile. In comparing the experimental data, we will therefore focus on the distribution of cells which have undergone $n \le n_{\rm max}$ rounds of division, i.e. $P_n^{\rm tot} \mapsto \tilde{P}^{\rm tot}_{n\le n_{\rm max}} = \frac{P_n^{\rm tot}}{\sum_{n=0}^{n_{\rm max}} P_n^{\rm tot}}$. Referring to Fig. 3j and S6, we estimate $n_{\rm max} = 9$.

## 4.2 Fit to the data

Referring to the experimental data, making use of the inferred SC and CP cell division rates, $\lambda_{\rm S}$ and $\lambda$, and the stratification rate, $\Gamma$, we find a good agreement between the theoretical prediction and the experimental data (Fig. 3k) over the first 3 weeks if we assume that SC constitute just 5% of the basal layer proliferative cell population. As well as providing further validation of the results of the clonal analysis, this result confirms that the progenitor cell fraction is greatly in excess of the SC population.