Supplementary Theory

In this Supplementary Theory, we provide further details on the modelling approach to study cell fate in the mouse tail epidermis, as well as on the data analysis and statistics used. We start by describing how quantitative analysis of clonal date data can be used to elucidate the lineage hierarchy and progenitor fate behavior under conditions of normal homeostasis. This modeling scheme and its analysis provides a platform to then consider how stem and progenitor cell fates are perturbed in the progression to basal cell carcinoma (BCC) following smoothened activation.

1 Homeostatic renewal of the mouse tail epidermis

To address the lineage hierarchy in mouse tail epidermis, we implement a quantitative clonal lineage tracing strategy inspired by previous studies [1, 2]. Using an inducible genetic labelling strategy based on the CYP1A1 promoter, previous studies by Clayton et al. [1] suggested that the maintenance of tail epidermis relied on the turnover of a single equipotent progenitor population that undergoes balanced stochastic fate. Subsequent clonal analysis based on two distinct promoters, K14-CreER and Involucrin-CreER (Inv-CreER), provided evidence for proliferative heterogeneity, with Inv-CreER targeting a single long-lived progenitor pool while K14-CreER captures a second, long-lived, slow-cycling stem cell pool, consistent with a proliferative hierarchy [2]. Subsequently, further heterogeneity in the spatial skin organization of the tail epidermis has been revealed, showing that the interfollicular epidermis of mouse tail is compartmentalised into a regular hexagonal lattice structure of two non-mixing regions, scale and interscale, that are maintained independently. This raises the question of whether each region conforms to the same self-renewal paradigm, as the previous clonal studies [1, 2] had analyzed the clonal data irrespective of their scale and interscale location.

To assess wether both scale and interscale region contain similar or distinct stem and progenitor cells, we have repeated experiments similar to Ref. [2], but keeping track of the geographical location of clones associated with the K14-CreER and Inv-CreER promoters, which both lead to the labelling of basal progenitors (as well as differentiated cells) in both regions. Using K31 as a scale marker, we segregated the clonal data in three

distinct regions: scale, interscale and border (for clones sited close to the scale-interscale boundary).

1.1 Regional differences in clone size dependence

For a range of chase times following induction, we then assessed the similarity of the clone size distributions between the three regions for each promoter. In order to have a model-independent comparison of the different regions, we use the Kolmogorov-Smirnov test, which is a non-parametric test used to compare two sample distributions. It relies on looking at the maximal distance between two cumulative clone size distribution in two regions. With $P_s^b(x,t)$ representing the cumulative distribution of basal (b) clone sizes x in the scale (s) region at a time t, comprising N clones, and $P_i^b(x,t)$ representing the cumulative distribution of basal clone sizes x in the interscale region at a time t, comprising N' clones, we define the Kolmogorov-Smirnov distance $D_{N,N'}$ as

$$D_{N,N'}(t) = \sup_x |P_i^b(x,t) - P_s^b(x,t)|.$$

The two distributions are considered different with a level of confidence α if $\left(\frac{NN'}{N+N'}\right) D_{N,N'}(t) > c(\alpha)$, where $c(\alpha)$ is the inverse Kolmogorov distribution, from which a P-value can be computed.

Although the clonal distributions were similar soon after induction, strong statistical differences between the scale and interscale distribution were always apparent at later time points. Notably, the distributions of clone sizes in the border region were always intermediate between the scale and interscale distributions. The findings are summarized below, by reporting the P-values at each time point (small P-values are indicative of significant differences). We first test whether scale and interscale distributions are different for the Inv-CreER clones:

Time (weeks)	1	2	4	8	12	24
Basal	-	-	-	-	< 0.005	< 0.05
Suprabasal	-	-	< 0.2	-	< 0.1	< 0.2

and then for the K14-CreER clones:

Time (weeks)	1	2	4	8	12	24
Basal	-	-	< 0.2	-	< 0.2	-
Suprabasal	-	-	< 0.05	-	-	< 0.05

Similarly, we compared measurements of clonal persistence (clonal density) in the scale and interscale region (normalized to their average value at 1 week (1w) post-labelling)

and always found strong differences, both for Inv-CreER ($P < 0.02$ at 12w and $P < 0.1$ at 24w), and K14-CreER clones ($P < 0.1$ at 8w, $P < 0.05$ at 12w and $P < 10^{-7}$ at 24w).

Having established that the clonal distributions in the interscale and scale are distinct, we then considered the cell fate behaviour within these regions separately (Fig. 2 and Extended Data Fig. 2-4).

1.2 Clonal dynamics in interscale during adult homeostasis

1.2.1 Inv-CreER/Rosa-YFP clonal analysis

As a starting point for our analysis, we begin by considering the clonal data obtained from the Inv-CreER mice, focussing on the interscale region. The motivation and basis for our modelling approach have been extensively documented in the literature [1, 2]. As noted in the main text, a key feature of the time course following induction is that the number of surviving clones steadily decreases with time, while the mean size of surviving clones increases approximately linearly. Moreover, the clone size distributions at all time points conform to the scaling behavior where the chance of finding a clone with n cells at time t post-labelling takes the form $P_n(t) = \frac{1}{\langle n(t)\rangle} f(\frac{n}{\langle n(t)\rangle})$, where $\langle n(t)\rangle$ denotes the average clone size, and the function $f(x) = e^{-x}$. This is consistent with the dynamics of critical birth-death process, and implies that clones derived from the Inv-CreER belong to a single equipotent population that undergo stochastic fate choice, which are nearly balanced at the population level.

In the following, we denote as P basal committed progenitors (CPs), which through cell division can give rise to both basal progenitors and suprabasal differentiated cells D, which cannot divide. Here, we have made the assumption that differentiated basal cells migrate rapidly into the suprabasal layer, so that the vast majority of basal cells are either stem cells or committed progenitors. We also define λ_P as the division rate of progenitors and Γ as the shed rate of differentiated cells as they migrate upwards through the suprabasal layer. In this paradigm, progenitor fate is defined by the Markovian process,

$$P \overset{\lambda_P}{\to} \begin{cases} P+P & \text{Pr. } r-\Delta \\ P+D & \text{Pr. } 1-2r \\ D+D & \text{Pr. } r+\Delta \end{cases} \qquad D \overset{\Gamma}{\to} \text{loss}$$

in which the constant fate probabilities are defined intrinsically (i.e. cell-autonomously). The parameter r quantifies the probability of symmetric division/differentiation events. As in Ref. [2], we have included a small positive offset $\Delta \ll r$ to account for a possible small imbalance in progenitor fate towards differentiation and clonal loss.

Formally, defining $P_{n_P,n_D}(t)$ as the probability to find a clone at time t post-induction containing n_P progenitors and n_D differentiated cells is described by the Master equation:

$$\begin{aligned}\frac{dP_{n_P,n_D}}{dt} = \lambda_P\,((r-\Delta)(n_P-1)P_{n_P-1,n_D} + (r+\Delta)(n_P+1)P_{n_P+1,n_D-2} \\ +(1-2r)n_P P_{n_P,n_D-1} - n_P P_{n_P,n_D-1}) \\ +\Gamma((n_D+1)P_{n_P,n_D+1} - n_D P_{n_P,n_D}) \quad (1)\end{aligned}$$

Formally, the solution for the Master equation can not be obtained in closed form [6]. However, since the relative fraction of differentiated cells tracks closely those of the progenitors, the intermediate and long-term behavior of the distribution can be straightforwardly recovered from the progenitor dynamics. In particular, the progenitors conform to a Galton-Watson birth-death type process, which has been studied extensively by statisticians [4]. For $\Delta = 0$, the process is critical, and starting at $t = 0$ with the initial condition $P_{n_P,n_D}(0) = \delta_{n_P,1}\delta_{n_D,0}$, previous studies have shown that the average number of basal progenitors in surviving clones is given by [4, 5, 6]

$$\langle n_P(t)\rangle = 1 + r\lambda_P t \quad (2)$$

This linear increase in average size compensates for the continuous chance loss of clones through differentiation. For a critical process (balanced cell fate), the increase in average basal clone size is simply the inverse of the persisting clone fraction

$$1 - P_0(t) = \sum_{i=1}^{\infty} P_i(t) = \frac{1}{1 + r\lambda_P t} \quad (3)$$

so that the product of the two is constant, and equal to its initial value (Extended Data Fig. 3a). In steady-state, the ratio of total suprabasal to basal cell number is dictated by the ratio of the division and shedding rates: $\frac{n_P^\infty}{n_D^\infty} = \frac{\Gamma}{\lambda_P}$. Therefore, for a perfectly balanced population of equipotent progenitors ($\Delta = 0$), the two key parameters for the growth of the basal and suprabasal compartments are $r\lambda_P$ and $\frac{\Gamma}{\lambda_P}$. Each can be assessed with high precision from the fitting of the mean basal and suprabasal clone sizes. However, to dissect the independent parameter values, for instance r and λ_P, one requires additional information, as we will see later in the text.

When fate becomes slightly imbalanced towards differentiation ($0 < \Delta \ll 1$), the average clone size dependence becomes sub-linear, a manifestation of non-neutral dynamics [2] (Extended Data Fig. 3a). At very long time scales, the effect of Δ will always become predominant, and cause a net decrease in average clone size. However, in the current application, we could never access this regime experimentally.

Fitting procedure

In order to fit the model to the clonal fate data, in the absence of any prior information on the value that the parameters might have, we made use of the classical method of least-squares, which consists in defining a sum $\mathscr{L}$ of residuals

$$\mathscr{L} = \sum_{i=1}^{n} \left(y^i_{\text{obs}} - \widetilde{y}^i_{\text{model}}(\vec{\theta}) \right)^2$$

where y^i_{obs} is a collection of n measured observables, whereas $\widetilde{y}^i_{\text{model}}(\vec{\theta})$ is a collection of predictions from the model using a set of parameters represented by the vector $\vec{\theta}$. The difference, $y^i_{\text{obs}} - \widetilde{y}^i_{\text{model}}(\vec{\theta})$, between a measurement and a prediction is called a residual, and in the case of the Inv-CreER/Rosa-YFP lineage tracing, the parameter vector is $\vec{\theta} = (r, \Delta, \lambda_P, \Gamma)$. We then minimise $\mathscr{L}$ with respect to $\vec{\theta}$ to find the optimal parameters describing the dataset. As the model clone size distribution that we describe throughout the text has no closed-form analytical solutions in the general case [6], we resort to numerical integration of the corresponding Master equations, for various values of the fitting parameters, $\vec{\theta}$.

In each case described below, we start by performing a coarse parameter sweep in order to determine the neighbourhood of the optimal parameters, and we then perform a much finer parameter sweep to obtain the optimal fitting values. For each set of fitting parameters, or estimators, we simulated at least 50000 clones, and performed several replicates to verify that we always converged on the same parameter set.

Once the optimal set of parameters is determined, we complemented the analysis with a confidence interval for each of these parameters, i.e. the confidence that the real parameter values fall within the provided interval. In order to get a confidence interval in a non-parametric fashion, we made use of a bootstrapping method, using a Monte Carlo algorithm for case resampling [9]. This method involves performing the same parameter sweep as the one described before, but for a large number of artificial datasets, which are samples with replacement of the original dataset. For instance, if 100 clones have been counted for a given data point, we randomly pick 100 clones within these (allowing for replacement, so that the same clone can be picked several times), and determine the optimal parameters for that dataset.

We performed this operation for at least 500 iterations, and thereby built an estimate for the probability distribution of our estimators themselves. Once this distribution has been calculated, it is straightforward to deduce a confidence interval for any confidence value of choice. We also plot confidence intervals on each prediction of the model, in order to visualise how the error bars in the fitting parameters extracted influence this prediction. In the remainder of the text, confidence intervals will be stated at the 95% *confidence level*. Throughout the text, we see that the clone size distributions are essentially feature-

less, i.e. well-fitted by a single exponential (something that we verify a posteriori - see paragraph below), so that the mean clone size is the only important parameter. Therefore, we perform our parameter fitting on the joint time evolution of the average basal and suprabasal clone size.

As a technical note, we could in principle include the division rate λ_P in the parameter vector being fitted, and obtain a prediction for it. Nevertheless, from a practical perspective, the data converges very quickly to the scaling form, where the important quantity is $r\lambda_P$. Thus, from the data, we can get precise estimate for $r\lambda_P$, with small confidence intervals. However, the rapidity of the convergence to the scaling form makes it hard to obtain reliable estimates of each independent parameter from the clonal data alone. Nevertheless, we can make use of the extensive proliferation kinetics data already obtained in Ref. [2] via BrdU pulse and H2B-GFP label retention to fix the division rate ($\lambda_P = 1.2 \pm 0.1 \text{week}^{-1}$). Similarly, in the sections on Smoothened activation, we make use of the double labelling BrdU/EdU experiments performed in this study to fix the division rate. During the bootstrapping procedure, we therefore assume that the division rate is normally distribution with the above mean and variance, and pick a value randomly at each iteration, in order to account for the error bars in λ_P on the confidence intervals of the other parameters.

Methods for validating the model

In order to test the validity of our models, as well as the detailed parameters that we extract from the fit to the average clone sizes (which is the only information we use throughout the text, together with division kinetics, to fit all the parameters of the model), we compared the evolution of clonal persistence in time to its predicted values, *once all parameters had been fixed*. Because it is an intuitive parameter to understand, we also compared the evolution of the total labelled cell fraction in time to its predicted value, although we note that this is a related quantity to the clonal persistence and the average clone size (and therefore that this measure does not provide an independent test).

To challenge quantitatively the goodness of the fit, as well as its explanatory power, we verified the model predictions by calculating in both cases the coefficient of determination R^2, which is the simplest indicator of the goodness of a fit. In addition, we calculate the standard error of the fit S, which is an absolute measure of the residuals of the fit, and has been shown to be more adapted for non-linear fitting procedures [10]. For a good fit, R^2 should be as close to 1 as possible, whereas S should be as close to 0 as possible. Specifically, defining n as the number of the points being fitted, $\bar{y} = \frac{1}{n}\sum_i^n y^i_{\text{obs}}$ the average of the observable, $S_{\text{tot}} = \sum_i^n (y^i_{\text{obs}} - \bar{y})^2$, and $S_{\text{res}} = \sum_i^n (y^i_{\text{obs}} - y^i_{\text{model}})^2$, the coefficient of determination is defined as $R^2 = 1 - \frac{S_{\text{res}}}{S_{\text{tot}}}$, while the standard error of the fit is given by $S = \sqrt{\frac{S_{\text{res}}}{n}}$. Moreover, as the above definition of the residuals give a larger relative

importance to large clone size, we also calculate the weighted residuals, where each residual is normalized by the variance σ_i of y^i_{obs} at a given time point: $S^w_{\text{res}} = \sum_i^n \frac{(y^i_{\text{obs}} - y^i_{\text{model}})^2}{\sigma_i^2}$, and the corresponding weighted standard error of the fit $S^w = \sqrt{\frac{S^w_{\text{res}}}{n}}$. As the clone size distribution in the following are exponential, the variance is always very close to the mean, and this definition gives comparatively more weight to the early time points (small average clone sizes).

Throughout this Supplementary Theory, we also verify that the full basal clone size distributions at all six time points are well predicted by the model. It should be noted that although the average of this distribution has been used to fix the parameters of the model (see above), nothing *a priori* dictates that the full experimental distributions should conform to their predicted functional shapes from the theory. This is particularly true of the tumour progression data considered below, where different clones could have widely varying fate and properties, resulting in potentially complex distributions.

To assess the goodness of our fits for the clone size distributions, we again calculate for all time points both the coefficient of determination R^2, as well as the standard error of the fit S. For completeness, we also performed at each time point a Kolmogorov-Smirnoff goodness-of-fit test. In analogy to the two-sample Kolmogorov-Smirnoff test mentioned above, we calculate the one-sample Kolmogorov-Smirnoff statistic D_n between the predicted and experimental cumulative clone size distributions for a number of clones n at a given time point. The null hypothesis in this case is that the experimental distribution follows the predicted value, and is rejected at level α is D_n exceeds a critical value $K_\alpha/\sqrt{n}$, which can be calculated from the cumulative Kolmogorov distribution [11].

In the following, we provide P-values at each time point from the Kolmogorov-Smirnoff test. For instance, if we report a P-value of $P = 0.25$, this means that we cannot see any statistical difference between prediction and the experimental data at the 20% significance level. Larger P-values means that the effect can be said to be not statistically significant with larger statistical significance criterion, and thus higher statistical power. As P-values quantify whether an effect exists, but not the magnitude, or size, of the effect, we combine it in all of the following to the aforementioned coefficients of determination and standard errors of the fit, which complement it to give a better idea of the statistical power of the results.

As a technical note, it should be mentioned that the classical Kolmogorov-Smirnoff test assumes statistical independence between the model and the experimental data. In our case, although we have not directly fitted each individual experimental distribution to obtain the model parameters, we did use the time evolution of the average of the distributions. Therefore, in order to provide the most conservative and rigorous estimates for the P-values, we made use of (unless explicitly specified) the modified tables of critical values of the statistic D_n, testing whether the experimental data follows an exponential

distribution when the mean is estimated from the data (following the classical results of Lilliefors [12], Stephens [13] and others).

Finally, we validate the hierarchical model that we uncovered at homeostasis by showing it can explain quantitatively, with different detailed fate parameters, the experimental data upon SmoM2 activation, both without and with P53 deletion.

Distinguishing imbalance and priming

Importantly, a signature of balanced fate choices is that the total labelled cell fraction (i.e. the product of surviving basal mean clone size and persistence probability) $\Pi(t) = \langle n_p \rangle (1 - P_0)$ remains constant in time with a value dictated by the initial condition (Extended Data Fig. 3a). Nevertheless, we have seen in the main text that $\Pi(t)$ for Inv-CreER labelled progenitors continuously decreases with time, indicative of a bias towards differentiation. This could be explained by two broad classes of hypotheses: either progenitors experience a persistent, if slight, imbalance towards differentiation (requiring another population to restore homeostasis over time), or a single progenitor pool exhibits balance over the long-term but Inv-CreER targets a fraction which is transiently primed towards differentiation (Extended Data Fig. 3a). Here, we explore the detailed predictions of both models.

If progenitors experience a persistent imbalance ($\Delta > 0$), then the predicted labelled cell fraction changes *continuously* with time as

$$\Pi(t) = \Pi(0) e^{-\Delta \lambda_P t}. \tag{4}$$

On the other hand, if a fraction of progenitors is only primed towards a given fate transiently, for a time τ_p with imbalance Δ', the labelled cell fraction changes only during the phase of priming, before settling to a constant value:

$$\Pi(t) = \Pi(0) \begin{cases} e^{-\Delta \lambda_P t} & \text{if } t < \tau_p \\ e^{-\Delta \lambda_P \tau_p} & \text{if } t > \tau_p \end{cases} \tag{5}$$

Extended Data Fig. 3a summarizes the consequence of the three possible models for the evolution of the labelled cell fraction.

Fit of the data

Turning to the data, it is apparent that $\Pi(t)$ continuously decreases for Inv-CreER labelled progenitors, favouring a model of persistent imbalance rather than short-term priming, and indicating the need for a second population to maintain homeostasis in the interscale. However, since the apparent imbalance is rather small, this second contribution need not

be extremely large. In particular, implementing the fitting procedure described above on the parameter vector $\vec{\theta} = (r, \Delta, \Gamma)$, we find the best fits with

$$r = 0.19^{+0.08}_{-0.09}, \qquad \Gamma = 0.15^{+0.04}_{-0.03}, \qquad \Delta = 0.02^{+0.03}_{-0.02}$$

We obtain rather constrained values for both r and Γ, although the value for Δ has larger error bars. This is due to the fact that the impact of Δ on the data is modest and becomes visible only at the later time points. With these parameters, we obtain a good fit of the average clone sizes, both for the basal and suprabasal layers (coefficient of determination of $R^2_b = 0.97$ for the basal layer, $R^2_s = 0.83$ for the suprabasal layer; with respective standard errors of the regression of $S_b = 0.18$ and $S_s = 0.57$ as well as their weighted counterparts $S^w_b = 0.26$ and $S^w_s = 0.46$). One should note that the main source of unexplained variance arises from the very early time point, where the basal cells committed to differentiation are likely to be still in the process of stratifying.

Challenging the model

In order to challenge the model, and in particular the validity of the fitting parameters, we follow three independent strategies:

- On one hand, we verified whether we can predict the full basal clone size distribution at all 6 time points from our fit based on the average time-dependences (Extended Data Fig. 3b). We found that distributions are consistently exponential, and the coefficients of determination for the 1, 2, 4, 8, 12, and 24 week size distributions, respectively are $R^2 = [0.99, 0.97, 0.97, 0.99, 0.97, 0.98]$, indicative of excellent fits. The standard errors of the regression were, respectively $S = [0.05, 0.09, 0.06, 0.02, 0.06, 0.03]$. The P-values evaluating the goodness of the fits were, respectively, $P = [0.39, 0.53, 0.97, 0.999, 0.40, 0.98]$, which means that we could not find any statistically significant difference (even at the 20% significance level) between the experimental distribution and our model.

- Secondly, we made use of the clonal persistence data, or alternatively and equivalently, the labelled cell fraction, which has not been used in any of the model calibrations. As we saw in the main text, our prediction is in good agreement with the model, showing a strong and continuous decrease of the labelled cell fraction over time. We normalised the labelled cell fraction to its experimental value at 2w post-induction (where we are sure that the one-cell clones derived from marked differentiated cells and present at 1w have been shed), and calculated a coefficient of determination for our fit of $R^2 = 0.37$, and a standard error of the regression of $S = 0.16$ (with weighted counterpart of $S^w = 0.17$). Notably, if we exclude the 24w datapoint for which we have few clones and thus higher error bars, the coefficient of determination is $R^2 = 0.84$, and a standard error of the regression of $S = 0.04$.

- Thirdly, we probe the short-term clonal data (namely the 1 week data point) in order to check whether the fate choice probabilities are the same as those extracted from long-term data. In particular, we defined the probability distribution P_{n_b,n_s} of a clone having n_b basal cells and n_s suprabasal cells. Because we only register surviving clones with basal attachment, we only have access to this distribution for $n_b > 0$. Moreover, one-cell clones are ambiguous, as some of them are simply differentiated cells waiting to move suprabasally, so we further restrict the analysis to clones for which $n_b + n_s > 1$. After these two restrictions, we have 48 clones to analyse. Moreover, because we only probe short-term data, we can safely assume that no suprabasal cell loss occurs ($\Gamma = 0$), and that the small long-term imbalance towards differentiation is negligible ($\Delta = 0$). We are then left with only r and λ_P as our fitting parameters, which will enable us to verify *a posteriori* that we had a good estimate for the division rate.

  Following the same type of fitting analysis as before, with least squares and bootstrapping, we found the best fits occurred for

$$r = 0.27^{+0.13}_{-0.16}, \qquad \lambda_P = 0.17^{+0.11}_{-0.09}.$$

  Thus, we obtain a value of r consistent with the previous fits, albeit slightly larger, which might be due to the fact that cell transfer to the suprabasal layers is not instantaneous. Interestingly, we also validate the value for the division rate of the progenitor pool.

1.2.2 K14-CreER/Rosa-YFP clonal analysis

Given the apparent small imbalance towards differentiation of Inv-CreER clones, we turned to the K14-CreER data to check whether K14-CreER marked a second population of cells that can expand to maintain the epithelium. In stark contrast to Inv-CreER, clones labelled from K14-CreER show a net expansion over time, as seen from the fact that the labelled clonal fraction $\Pi_{K14}(t)$ continuously increases over time. Again, this continuous expansion argues against a one-population model with short-term priming (Extended Data Fig. 3a, i.e. where K14-CreER would mark a population primed towards symmetric division and Inv-CreER marks a population primed towards differentiation).

With concrete evidence for proliferative heterogeneity in the interscale region, several classes of models can be distinguished. However, a key distinction that cannot be resolved using the dataset is whether this heterogeneity is intrinsic, associated with discrete progenitor cell types, or whether it is conditioned by extrinsic signals associated with a localized niche, a question that is still unresolved in the field [7, 8]. Within the class of "intrinsic" models, a further distinction is that different progenitor pools could function independently, or that progenitors are arranged in hierarchy with one population giving

rise to the other (stem cell-progenitor model). In the "independent population" model, K14-CreER would target cells with a given imbalance $\Delta_S > 0$, and the labelled cell fraction $\Pi(t)$ would increase exponentially, as discussed in the previous section, to compensate for the loss of progenitors targeted by the Inv-CreER. This seems an unlikely solution, as the Inv-CreER labelled population would then get "weeded out" progressively from the epithelium (with characteristic time $\frac{1}{\lambda_P \Delta_P} \approx 100$ days).

In the stem cell-progenitor model, stem cells S follow a Markovian process in which cell division, at rate λ_s, leads to balanced stochastic fate with defined probabilities,

$$S \to \begin{cases} S+S & \text{Pr. } r_s \\ S+P & \text{Pr. } 1-2r_s \\ P+P & \text{Pr. } r_s \end{cases}$$

while their progenitor cell progeny, P, conform to the dynamics resolved from the Inv-CreER/Rosa-YFP clonal analysis. Defining $P_{n_s,n_P,n_D}(t)$ as the probability to find a clone at time t containing n_S stem cells, n_P progenitors and n_D differentiated cells, the clonal dynamics is defined by the Master equation,

$$\begin{aligned} \frac{dP_{n_S,n_P,n_D}}{dt} = \lambda_S \big(& r_s((n_S-1)P_{n_S-1,n_P,n_D} + (n_S+1)P_{n_S+1,n_P-2,n_D}) \\ & +(1-2r_S)n_S P_{n_S,n_P-1,n_D} - n_S P_{n_S,n_P,n_D}) \\ + \lambda_P \big(& ((r-\Delta)(n_P-1)P_{n_S,n_P-1,n_D} + (r+\Delta)(n_P+1)P_{n_S,n_P+1,n_D-2} \\ & +(1-2r)n_P P_{n_S,n_P,n_D-1} - n_P P_{n_S,n_P,n_D}) \\ & +\Gamma((n_D+1)P_{n_S,n_P,n_D+1} - n_D P_{n_S,n_P,n_D}) \end{aligned} \quad (6)$$

Although an analytic solution to this equation is unavailable, the probability distribution can be recovered straightforwardly by numerical integration. Moreover, according to this dynamics, starting with the initial condition $P_{n_S,n_P,n_D}(0) = \delta_{n_S,1}\delta_{n_P,0}\delta_{n_D,0}$, the labelled cell fraction $\Pi_S(t)$ is predicted to increase over time,

$$\Pi_S(t) = 1 + \frac{\lambda_S}{\lambda_P \Delta_P}(1 - e^{-\lambda_p \Delta_p t}), \quad (7)$$

consistent with the clonal dominance of the stem cell population. This labelled cell fraction plateaus to a value which depends only on the ratio of stem cells and progenitors $\frac{P}{S} = \frac{\lambda_S}{\lambda_P \Delta_P}$, so that $\Pi_S(\infty) = 1 + \frac{P}{S}$. This corresponds to the intuitive idea that each stem cell generates *on average* a given number of progenitors at steady state in the basal layer.

More generally, since K14 is expressed ubiquitously in the basal layer, we expects K14-CreER to label a combination of stem and progenitor cells. Setting f_s as the fraction of

stem cells initially labelled, one can obtain the more general expression for the labelled cell fraction,

$$\Pi_S(t) = f_s(1 + \frac{\lambda_S}{\lambda_P \Delta_P}(1 - e^{-\lambda_p \Delta_p t})) + (1 - f_s)(e^{-\Delta\lambda_P t}). \quad (8)$$

In particlar, this fraction saturates at a value $\Pi_S(\infty) = f_s(1 + \frac{P}{S})$ that is always above unity if the K14-CreER labelling enriches for stem cells as compared to the homeostatic stem cell concentration, which is expected given the findings of the previous study [2]. Although the K14-CreER data allows for a rather large parameter range, the model fit to the data it is heavily constrained by the Inv-CreER analysis, which dictates the kinetics and fate behaviour of the progenitors.

Fit of the data

Implementing the fitting procedure described above on the parameter vector $\vec{\theta} = (r_s, \lambda_S, f_s)$, we obtained the best fitting parameters as

$$r_s = 0.03^{+0.05}_{-0.03}, \qquad \lambda_s = 0.045^{+0.01}_{-0.01}, \qquad f_S = 65\%^{+15\%}_{-15\%}.$$

This suggests that the second, stem-cell, population divides more slowly at a rate of around once every 2 weeks, and that the mode of division is overwhelmingly asymmetric (94%). Moreover, the data also allows us to make an estimate of the fraction of stem cells f_s that is labelled at this clonal dose. Since the size difference between "true" SC-derived K14-CreER clones and Inv-CreER clones is not very large at these time points, the estimation has rather large error bars. Nevertheless, from this analysis, we estimate from the fits that K14-CreER marks around 65% of stem cells and 35% of committed progenitors.

With these parameters, we again obtain a good fit of the average clone sizes, both for the basal and suprabasal compartment (coefficient of determination of $R^2_b = 0.93$ for the basal layer, $R^2_s = 0.93$ for the suprabasal layer, with respective standard errors of the regression of $S_b = 0.42$ and $S_s = 0.48$ as well as their weighted counterparts $S^w_b = 0.23$ and $S^w_s = 0.91$). Finally, with this model, we can estimate the relative fraction of stem cells and progenitors in the basal layer of the interscale at homeostasis as $\frac{n_S}{n_P + n_s} = \frac{2\Delta_P k_p}{k_s} \approx 13\%$, meaning that stem cells constitute a relatively small fraction of the total basal layer pool. As we estimated that 65% of cells labelled with K14-CreER are stem cells, this means that K14-CreER targets preferentially stem cells compared to their actual percentage in the tissue.

Challenging the model

In order to challenge this model, we again test whether the best fitting parameters can be used to predict the the clonal persistence (or labelled cell fraction) for K14-CreER/Rosa-YFP. In particular, we expect "true" SC-derived K14-CreER clones to almost never be lost, so that the bulk of the decrease in persistence is due to the Involucrin-like population (35%).

Indeed, between 2w and 24w, around 40% of the clones are lost, and the persistence curve becomes essentially flat, as predicted by the model. The prediction again falls in line with the measured data (Fig. 2e), although the initial decline is sharper in the data, which could be due to short-term priming effects. Similarly, our model, calibrated by the mean clone size dependence, could very accurately predict the evolution in time of the labelled cell fraction for the K14-CreER/Rosa-YFP data as well (Fig. 2e), with a coefficient of determination of $R^2 = 0.95$ for the fit, and a standard error of the regression of $S = 0.17$ (with weigthed counterpart of $S^w = 0.07$).

Moreover, we also verified whether the theory can predict accurately the full basal clone size distributions (Extended Data Fig. 3c), and found that the coefficients of determination for the 1, 2, 4, 8, 12, and 24w distributions, respectively are $R^2 = [0.999, 0.98, 0.96, 0.98, 0.97, 0.97]$, indicative of excellent fits. The standard errors of the regression are, respectively $S = [0.007, 0.05, 0.08, 0.05, 0.06, 0.05]$, and the P-values evaluating the goodness of the fits are, respectively $P = [0.999, 0.97, 0.08, 0.99, 0.32, 0.65]$, which means that we could not find a statistically significant difference (at the 5% significance level) between the experimental distribution and our model, although the $4w$ data point showed a larger deviation than the others, which would be due to mouse-to-mouse variations.

For the K14-CreER population, it was not possible to infer any information from the short term dynamics as, at 1w post-induction, clones are largely basal, suggesting that they did not have time to transfer yet to the suprabasal layer.

To summarise, the data in the interscale is consistent with a simple model of lineage hierarchy, where a slow-cycling stem cell divides asymmetrically on average to replenish a pool of long-lived progenitors, which are slightly imbalanced towards differentiation. A key signature of this hierarchy in the data is the net and long-term expansion of K14-targeted clones, and the corresponding net and long-term shrinkage of Involucrin-targeted clones. Each of these populations make stochastic fate decisions, resulting in broad, exponential clone size distributions at all time points. Our quantitative model of SCs and CPs in the interscale region remains very much in line with our previous model proposed in Mascre *et al.* [2]. We therefore now turn to the scale data to check whether a similar framework applies.

1.3 Clonal dynamics in scale

Following the strategy of the previous sections, we performed a Kolmogorov-Smirnoff test of the differences between the Inv-CreER and K14-CreER clone size distribution.

In contrast to the interscale, we could find no statistical differences at any time point ($\alpha > 0.2$ at all times, both for basal and suprabasal distribution). Next, we examined how the normalised labelled cell fraction $\Pi(t)$ evolves in time (Extended Data Fig. 4c). In striking contrast to the interscale data, we found that $\Pi(t)$ remains constant at around unity throughout the 24w chase, indicative of Inv-CreER labelling a single equipotent balanced population in the scale. We then turned to the K14-CreER scale data, and saw a statistically identical behaviour (Extended Data Fig. 4c), with the labelled cell fraction remaining constant and hovering around unity. This shows that both K14-CreER and Inv-CreER populations have the same dynamics (Extended Data Fig. 3a), and target a population consisting of a single type of progenitor making perfectly balanced fate choices, as proposed originally by Clayton *et al.* [1], and further supported by Mascre *et al.* [2] and Doupé *et al.* [14]. For the sake of consistency however, we use *a priori* the same model for CPs as interscale, leaving the possibility for an imbalance Δ in scale as well:

$$P \overset{\lambda_P}{\to} \begin{cases} P+P & \text{Pr. } r-\Delta \\ P+D & \text{Pr. } 1-2r \\ D+D & \text{Pr. } r+\Delta \end{cases}$$

Fitting the data

As pointed out in the first section, scale clones are statistically larger on average than interscale clones (Extended Data Fig. 4a). Performing the same type of fitting analysis as before on the average basal and suprabasal mean clone size, we obtained the best fit parameters (Extended Data Fig. 4b):

$$r = 0.22^{+0.10}_{-0.10}, \qquad \Gamma = 0.17^{+0.04}_{-0.04}, \qquad \Delta = 0.00^{+0.02}_{-0.02}.$$

In particular, the values of r and Γ are broadly consistent with the values found in the interscale. Nevertheless, in contrast to interscale, the best fit of the average clone size data is a *perfectly balanced* pool of progenitors ($\Delta \approx 0$). This balance manifests itself as a linear growth of the basal clone size over time (Extended Data Fig. 4a), as predicted by Eq. 2 for a critical birth-death process. This is in striking contrast with the sub-linear growth observed for the Involucrin-targeted clones in interscale (Fig. 2c-e), a signature of subcritical branching processes, i.e. imbalance towards differentiation.

With these parameters, we again obtain a good fit of the average clone sizes, both for the basal and suprabasal layers (coefficient of determination of $R_b^2 = 0.99$ for the basal layer, $R_s^2 = 0.93$ for the suprabasal layer; with respective standard errors of the regression of $S_b = 0.27$ and $S_s = 0.64$ as well as their weighted counterparts $S_b^w = 0.18$ and $S_s^w = 1.5$).

Challenging the model

This lack of measurable imbalance in the mean clone size is especially striking as it confirms, in a completely independent manner, our observation deduced from the persistence data. Indeed, both the K14-CreER and Inv-CreER labelled clonal fraction remains flat from 2w to 24w (Extended Data Fig. 4c), as expected for an equipotent pool of progenitors ($\Delta \approx 0$).

Moreover, we again tested whether the theory also predicted accurately the full basal distribution of clone size at the 6 time points (Extended Data Fig. 4d). As discussed in [3], Inv-CreER labels relatively few cells in the scale region, so that we only looked at the K14-CreER distributions which where more informative and could be compared with the model with better statistical power. As previously, we calculated the respective coefficients of determination as $R^2 = [0.99, 0.97, 0.93, 0.89, 0.98, 0.92]$, showing once again the predictive power of the theory. The standard errors of the regression are, respectively $S = [0.05, 0.08, 0.11, 0.09, 0.04, 0.06]$, and the P-values evaluating the goodness of the fits are, respectively $P = [0.05, 0.99, 0.01, 0.99, 0.99, 0.87]$. One should note that the late time points display no statistically significant difference with the model, although the early time point (1w and 4w) do. This could be due to the presence of basal differentiated cells, which leads us to under-estimate clone sizes at early times points.

The scale data provides an ideal control to test the statistical validity of our approach. We therefore compared the experimental clone size distribution in scale successively to our prediction for the interscale Inv-CreER parameter set (single progenitor pool with slight imbalance) and for the interscale K14-CreER parameter set (hierarchy of stem and progenitor pools). Using P-values as before to assess the goodness of our fits, we first test whether the scale dataset was statistically different to the interscale Inv-CreER prediction, which yields respectively, $P = [0.001, 0.88, < 10^{-5}, 0.33, 0.008, 3.10^{-5}]$, showing very significant differences in the majority of cases, and emphasizing the superiority of the "imbalanced" parameter set to predict the interscale Inv-CreER data. Further, we tested whether the scale dataset was statistically different to the interscale K14-CreER prediction, which yields respectively, $P = [0.01, 0.99, 0.002, 0.99, 0.37, 0.03]$, showing more statistically significant differences than the scale parameter set.

1.4 Discussion of the homeostatic system

In summary, our analysis shows that different regions of the tail epidermis are maintained by two qualitatively different lineage hierarchies. Although both contain long-lived committed progenitors targeted by Inv-CreER, the progenitors in scale act as a single balanced self-renewing population, whereas the progenitors in interscale are slightly tilted towards symmetric differentiation, leading to a net clonal loss compensated by the activity of a small population of slowly-cycling stem cells. In each case, we have extracted from the fit

of the basal and suprabasal mean clone size the probabilities underlying the fate choices of each population of cells.

Relevance of the scale-interscale regionalization

The existence of two distinct differentiation programs in scale vs interscale has been long-known in the tail epidermis [15]. The interscale suprabasal cells undergo orthokeratic differentiation, with the formation of a granular cell layer and loss of nuclei in cornified layers, whereas scale suprabasal cells undergo parakeratotic differentiation, with nuclei retention and without a granular cell layer. Although it had been reported that scale and interscale are maintained by different cell populations [3], our study shows additionally that the two regions display a different hierarchy of homeostatic maintenance, and differential responses to oncogenic activation. Interestingly, although the scale/interscale distinction is thought not to exist in human epidermis and other murine epidermal regions, a recent study has shown that a similar and unsuspected scale/interscale regionalization exists in murine backskin [16]. Interestingly, Fig. 5f of Ref. [16] seems to show in tail skin interscale that the labelled cell fraction of the Dlx1 population rises steadily in time, whereas the labelled cell fraction of the Slc1a3 population decreases steadily in time. This would be characteristic of a hierarchy of stem cells (traced by Dlx1) and committed progenitors (traced by Slc1a3) in interscale, which agrees with our results using K14 and Involucrin. Within a year, their Dlx1 cell fraction increases by roughly a factor 2.5, whereas their Slc1a3 cell fraction is nearly entirely lost, which is also close quantitatively to our values. By contrast to interscale, their scale tracing displays a roughly constant labelled cell fraction from 2w to 3 months, both for the Slc1a3-CreER and Dlx1-CreER, again in perfect agreement with our framework. One should note that they see a subsequent increase in labelled cell fraction in both cases from 3 months to 1 year, which is more puzzling, but could be due to the method of quantification, relying on projected area rather than clonal counting.

As a whole, this data is consistent with ours and argues for a single equipotent progenitor population in scale, and a hierarchy of stem cells and committed progenitors in interscale. In the future, it would thus be interesting to examine quantitatively the generality of our findings in other epidermal contexts

Relationship to in vivo live-imaging data

It should be noted that, although the exact parameters could be modified if one considered more complex models, including for example short-term fate priming or spatial correlation in cell fate behavior, such models would still need to contain the core principles described above. Interestingly, a recent study [17], investigating murine epidermal homeostasis using *in vivo* live-imaging, found that paw and ear are maintained by an equipotent population of progenitors with equal probability to divide or differentiate. This is in line

with the model originally proposed by Clayton *et al* [1], as well as our findings in scale. Importantly, Ref. [17] confirmed that there are no fate correlations between generations, a key prediction inferred from static lineage tracing assays. They do find a correlation between siblings as well as a short refractory period for division. However, as mentioned above, such short-term correlations would not influence our findings, but simply the short-term data (less than a week), before the data converges towards the scaling form predicted by the stochastic fate model. One should note that the second cellular population we have uncovered in interscale is rather small in fraction, and thus that a large amount of live-imaging data would be acquired in order to statistically uncover it.

Finally, an important question that remains un-answered, and must still be addressed from a molecular perspective, is whether fate choice and differentiation are determined intrinsically upon division, or are specified later and extrinsically by neighboring cells or a microenvironmental niche.

In the next section, given the spatial compartmentalisation and hierarchical organisation that we have uncovered from the analysis of the homeostatic system, we ask the natural question of how the tail epidermis would react to a perturbation that drives the system out of homeostasis. In particular, we consider the effects of clonal activation of the constitutively active form of Smoothened, a potent oncogene that leads to the formation of basal cell carcinoma, being careful to distinguish clones in the scale and interscale region.

2 Oncogenic activation using SmoM2 over-expression

As noted in the main text, a fascinating feature of hedgehog over-expression is that it only leads to basal cell carcinoma in the interscale region from K14-CreER, whereas scale clones remain smaller and never reach an invasive stage. It is therefore tempting to examine its relationship with the regionalisation of the lineage hierarchies that we have uncovered in the homeostatic data.

Firstly, as stated in the main text, the regional differences upon Smo activation become even more obvious than in homeostasis, with BCC present only in the interscale. Nevertheless, for the sake of consistency, calculation of the Kolmogorov-Smirnoff distance between the clone size distributions in scale vs interscale for K14-CreER clones gives the following P-values:

Time (weeks)	1	2	4	8	12	24
Basal	–	0.001	< 0.001	< 0.005	< 0.005	< 0.05
Suprabasal	-	-	–	–	< 0.005	< 0.01

which shows consistently strong statistical differences between scale and interscale. In the case of the Involucrin SmoM2, it was not meaningful to calculate Kolmogorov-Smirnoff distance significance values between scale and interscale, as there are only very few scale

clones. Of note, we have defined here, and throughout the rest of the text, a suprabasal cell as any cell not possessing an attachment to the basement membrane. This comprised in the case of oncogenic activation both of normally differentiating cells, as well as basal-like cells displaying differentiation defects.

Secondly, within interscale, as stated in the manuscript, only K14-CreER clones progress to BCC, whereas Inv-CreER clones remain frozen at the hyperplasia and dysplasia stages. To check whether the full size distribution of clones is statistically different, we again calculated the Kolmogorov-Smirnoff distance between the clone size distributions in interscale for K14-CreER clones vs for Inv-CreER clones, which gives the following P-values:

Time (weeks)	1	2	4	8	12	24
Basal	–	< 0.001	–	< 0.0005	< 0.001	< 0.001
Suprabasal	-	-	–	0.001	0.06	0.002

again indicative of consistently strong statistical differences after a few weeks following induction. Importantly, at the later time points, which are the most informative to examine differences between K14-CreER and Inv-CreER clones, we found that Inv-CreER clones closely follow a single exponential distribution, whereas the distribution of K14-CreER clones cannot be fitted by a single exponential, and displays two characteristic sizes (Fig. 4h). Such behavior hints that Inv-CreER clones may still behave as a single equipotent population upon Smo overactivation, whereas K14-CreER clones represent a mixture of at least two populations. Additionally, we find that the small K14-CreER clones follow closely the distribution of Inv-CreER clones (Fig. 3h and Fig. 4h), indicative of the fact that K14 labels again a mixture of clones, overlapping partially with Inv-CreER clones. Therefore, we see strong similarities in the phenomenology of the data in interscale upon Smo overactivation compared to the homeostatic hierarchical organisation. We can then use the homeostatic model we have uncovered in the previous sections to heavily constrain the possible models of oncogenic growth upon Smo overactivation. We thus follow the same outline, concentrating successively on the interscale and the scale, and contrasting the Inv-CreER and K14-CreER datasets for each.

2.1 Clonal dynamics in interscale

2.1.1 Involucrin SmoM2 data

Starting with the Involucrin SmoM2 clonal assay in the interscale, in contrast with the homeostatic data, the parameters upon Smo activation are much less constrained. In particular, one may expect the development of a larger imbalance, Δ, towards either positive or negative values. More importantly, the parameters do not need to be time-invariant but could adjust in concert with the growth of the lesion. Therefore, in order to test whether parameters such as the division rate evolve in time, we first analysed the EdU/BrdU double pulse experiments.

Division rate in Involucrin clones: double pulse experiment

For the 8w data point, a striking feature of the data is the absence of correlation between the fraction of basal double labelled cells and the size of the clone. This was true when looking at either the fraction of EdU or BrdU positive cells as a function of clone size (Extended Data Fig. 5d). Quantitatively, after 6 days of continuous BrdU incorporation, a linear regression ($y = ax + b$) of the fraction of double positives vs basal clone size gave a weak coefficient of determination ($R^2 = 0.07$ for the double-pulse fraction, $R^2 = 0.08$ for the BrdU fraction) showing the lack of predictive power of the clone size. Further, the low value of the coefficient ($a = -4.0\ .10^{-3}$ both for double-pulse and BrdU fraction) confirms this. Similarly, after 2 days of continuous BrdU incorporation, $R^2 = 0.04$ for the double-pulse fraction, $R^2 = 0.12$ for the BrdU fraction, with ($a = +1.0\ .10^{-3}$ both the double-pulse fraction and $a = +3.0\ .10^{-3}$ for the BrdU fraction).

We therefore grouped all of the clones together to compute how the average fraction of double-labelled cells increases over time, and fit this increase with a gamma-distribution (widely use to model statistical events such as cell division time probability [18]), from which we extract a division time of $T = 7.2 \pm 0.6$ day at 8w post-induction (Fig. 3c). To confirm this finding, we also examined the double-pulse data at 4w post-induction, and found the same independence of cell kinetics upon clone size. At 6d of continuous BrdU incorporation, we found $R^2 = 0.008$ for the double-labelling fraction, $R^2 = 0.02$ for the BrdU fraction, with a similarly low value of a, as before.

Nevertheless, when comparing the double-labelling fraction distribution between 4w and 8w post-Smo activation, we could find dramatic differences, with the 4w fraction being much larger than the 8w fraction (from T tests: $P < 10^{-4}$, $P < 10^{-5}$ and $P < 10^{-5}$ after respectively 2d, 4d and 6d of continuous BrdU incorporation). Using the same fitting procedure, the best fit value of the division rate at 4w post-induction was $T = 3.6d \pm 0.5$, twice as fast than at 8w. In order to probe whether this decrease in division rate persisted to later time points, we performed the same EdU/BrdU analysis at the 12w post-induction time point. Interestingly, we found a best fit values for the division rate at 12w of $T = 9.8d \pm 0.3$ (Fig. 3c), again significantly slower than at 8w, validating our hypothesis of continuous growth attenuation (Fig. 3e).

In order to model the clonal data, we therefore made an empirical fit of the decrease in division rate with an exponential decay with a long-term offset (Fig. 3e), and took this curve as an input for the stochastic simulations of the clonal data. It should be noted that a very surprising result of the EdU/BrdU double-labelling assay is that, for most of the time course, the division rate in the tumour is lower than the division rate of the surrounding tissue. This already indicates that Involucrin-targeted basal cells must be quite heavily tilted towards symmetric division in order to explain their substantial size difference compared to homeostatic Involucrin clones.

Apoptosis rate in Involucrin clones

As noted in the main text, we noticed that starting at 8w post-induction, apoptosis occurs in Involucrin SmoM2 clones (Fig. 3d and Extended Data Fig. 5e-i). At the time of staining, the average percentage of caspase3 positive cells is $f_a = 1.4\%$. We next asked whether apoptosis occurred randomly across clones, or if it clustered in a given clone at a given time. We resorted to a numerical analysis, first simulating an exponential distribution of basal clone sizes, with the average value matching the experimental one (Extended Data Fig. 5i). We then assumed apoptosis was a stochastic Markovian process, with a given rate Γ_a, which is therefore the only free parameter in this simulation. In order to fix Γ, we adjust it via a classical least-square method, described above, by fitting the fraction of clones with zero apoptotic cells to its experimental value of 21%. We then compared the cumulative distributions of the (non-zero) fractions of caspase3 positive cells in the experimental and theoretical cases (Extended Data Fig. 5i). If apoptosis occurred randomly, and not cooperatively, in a given clone, one should get theoretically a Poisson distribution within the variable clone size distribution. Cooperative apoptosis would manifest into fatter tails for the experimental distribution compared to the theoretical one.

From this analysis, it was apparent that the two distributions were markedly different ($P < 0.05$ using Kolmogorov-Smirnoff distance), and that there was a clear effect of cooperativity in the experimental distributions, showing that apoptosis occurred preferentially in some clones. This is important, as stochastic apoptosis merely rescales the effective growth rate (being the product of the division rate and imbalance towards division). On the other hand, collective apoptosis, if it targets a given clone persistently, removes clones independently of their size, and therefore affects only the persistence of clones, but not the average clone size of persisting clones.

Moreover, we also performed caspase3 stainings at 4w post-induction, and noticed, as stated in the main text, that no apoptotic cell could be observed. On the other hand, performing caspase3 stainings at 12w post-induction, we found an increase of the fraction of apoptotic cells as compared to the 8w time point, with $f_a = 4\%$.

Estimating absolute apoptotic rates from caspase3 staining is rather challenging, as one needs to have real-time information on how long caspase3 is expressed before cell removal [19]. Moreover, it is difficult to infer from the data whether basal apoptosis involves true progenitors, or whether it targets basal differentiated cells, which reside for a small amount of time in the basal layers. Given the data above, we thus assumed that the apoptotic rate is zero up to 6w post-induction, and grows linearly thereafter. Guided by the fact that the basal mean clone size plateaus for later time points in Involucrin SmoM2 clones, we thus set $\Gamma(t) \approx 0.005(t-42)/42 d^{-1}$, which triples from 8w to 12w post-induction, as seen in the above data. We note, however, that this apoptotic rate is very small compared to the division rate for the bulk of the dynamics. In particular, it has little discernable influence of the mean clone size up to 12w post-induction, and only

acts on the 24w time point, where the division rate becomes low enough that apoptosis becomes important.

Fitting of the data

With this in mind, we ran the same type of fitting procedure as before on the parameter vector $\vec{\theta} = (r, \Delta, \Gamma)$, using the mean basal and suprabasal clone size evolution in time (Fig. 3f). We redefine Δ according to the following dynamics:

$$P \overset{\lambda_P}{\to} \begin{cases} P+P & \text{Pr. } r+\Delta \\ P+D & \text{Pr. } 1-2r \\ D+D & \text{Pr. } r-\Delta \end{cases}$$

so that $\Delta > 0$ upon Smo activation. Anticipating the result, it is worth noting that the ratio of basal to suprabasal cell numbers remains roughly unchanged throughout the course of the 24w chase. As the division time continuously changes (Fig. 3f), this requires the shed rate Γ to adjust too. So we rather assumed, for the sake of simplicity, that Γ has the same functional dependence, and we fit the ratio Γ/λ_P. Finally, we noticed that at late datapoints, there are variable small numbers of small clones which are purely basal, without any suprabasal cells. Presumably, this corresponds to a different growth regime of terminally differentiated clones, but, as we do not wish to add additional parameters, and in order not to contaminate the data with these small clones, we considered in the analysis clones with a least one suprabasal cell $n_s > 0$ (this holds for the entire Smoothened analysis, both for Involucrin and K14 clones, both in data and the corresponding simulations).

We then obtained the best-fit parameters:

$$r = 0.28^{+0.06}_{-0.06}, \qquad \Delta = 0.115^{+0.01}_{-0.005}, \qquad \Gamma/\lambda_P = 1.0^{+0.1}_{-0.1}$$

As expected, the parameter Δ can be determined with very high precision, as it is the one that has the most drastic influence on the basal clone size. On the other hand, r can be determined with less confidence, as its influence is weaker. Surprisingly, the data can therefore be accurately fit using fate probabilities that remain constant in time. With these optimal fitting parameters, we obtained coefficients of determination of $R_b^2 = 0.99$ for the basal clone size and $R_s = 0.98$ for the suprabasal clone size (with respective standard errors of the regression of $S_b = 1.8$ and $S_s = 2.3$, as well as their weighted counterparts $S_b^w = 0.78$ and $S_s^w = 1.15$). Again, these parameters provide an excellent fit (Fig. 3f). Therefore, even though there are significant unknowns on how the division rate evolves at very long times, as well as on the quantitative evolution of the apoptosis rate, the present framework can be thought of as a coarse-grained result of this complex dynamics.

Challenging the model

Next, we assessed whether this simple theory could predict at least the main features of the clone size distributions (Fig. 3h). From a theoretical perspective, the clone size distributions for an equipotent population of imbalanced progenitor cells are expected to collapse onto a single exponential. For each time point (1, 2, 4, 8, 12 and 24 weeks post-induction), we calculated the coefficients of determination for the fit to the basal clone size distribution: $R^2 = [0.35, 0.66, 0.85, 0.99, 0.98, 0.99]$. The standard errors of the regression are, respectively $S = [0.27, 0.13, 0.09, 0.04, 0.03, 0.04]$, and the P-values evaluating the goodness of the fits from the Kolmogorov-Smirnoff test are, respectively $P = [0.09, 0.07, 0.007, 0.2, 0.38, 0.23]$. The P-values evaluating the goodness of the fits, from Kolmogorov-Smirnoff test with the Stephens correction for exponentiality testing with unknown mean, are, respectively $P = [0.008, 0.006, < 0.001, 0.07, 0.15, 0.06]$.

Therefore, one should note that the late time points display no statistically significant difference with the model at the 95% confidence level, although the short-term data does. As with homeostasis, this could be due to the presence of basal differentiated cells, which makes us under-estimate the clone size at early times points. This is also, in part, due to our simplified assumption of Markovian dynamics of cell division, which means that cells do not have a refractory period after a rond of division, i.e. they can re-enters cycle arbitrarily fast. This produces long-tails in the size distribution at early times points which are unrealistic. However, it should still be noted that, at the later time points, our model gives a faithful fit for the clone size distribution, converging towards a single exponential, an additional validation of our modelling approach for the Involucrin SmoM2 data. This provides further evidence that the Involucrin SmoM2 cells behave as a remarkably uniform and equipotent population throughout the 24w time course of the experiments following Smoothened activation.

Finally, we examined the behavior of "clone in clones" using short-term EdU incorporation as a surrogate clonal marker, as described in the main text (Extended Data Fig. 5j). An interesting feature of the model is that the clonal dynamics are well-fit by assuming a constant imbalance towards symmetric division. We therefore questioned whether this corresponds to the real fate outcomes at 24w. We concentrated on 2-cell "clones" of EdU, and scored whether both cells are basal (probability p_{b-b}), suprabasal (probability p_{s-s}), or whether one cell is basal and one suprabasal (probability p_{b-s}). This corresponds in our model, respectively, to the probabilities $r + \Delta$, $r - \Delta$ and $1 - 2r$.

Turning to the data, as stated in the main text, we obtained a probability for symmetric renewal of $p_{b-b} = 0.49 \pm 0.09$, a probability of asymmetric division of $p_{b-s} = 0.37 \pm 0.07$, and a probability of symmetric differentiation of $p_{s-s} = 0.14 \pm 0.04$ (Extended Data Fig. 5j). The amount of asymmetric divisions and symmetric division is the same, within error bars, as predicted in the models. The experimental value for the symmetric division probability is much larger than the symmetric differentiation probability, confirming a large

imbalance towards symmetrical division in the basal compartment. One should note that our model underestimates the probability of symmetric division. This could be due to the fact that we have underestimated apoptosis, which would require an increased imbalance at the latest time point to compensate. Overall, the data therefore validates our idea that fate imbalance is strongly entrenched upon Smo activation. All Inv-CreER labelled progenitors behave as a surprisingly homogeneous equipotent and imbalanced population throughout the 24w course of the experiment although, with time, division rates decrease, which prevents an exponential uncontrolled growth.

2.1.2 K14 SmoM2 data

We now contrast the previous findings with the clonal dynamics in K14 SmoM2 mice. As we saw in the main text, the main feature of this dataset is the presence of numerous BCC in the interscale region with K14-CreER, and statistically bigger clones than with Inv-CreER. We followed the same strategy as before, starting with the double-pulse labelling experiments in order to get a better handle on the range of optimal parameters. In contrast to the Involucrin data, a striking feature of the data is the clear positive correlation between clone size and double-labelling fraction, indicative of the fact that the larger clones proliferate faster than the small clones (Extended Fig. 7d). Given the results of the homeostatic K14-CreER lineage tracing experiment, we assessed whether a similar population hierarchy that exists during homeostasis also influences tumor initiation, and whether CPs and SCs behave differently upon Smo overactivation.

Division rate in K14 clones

In order to test this hypothesis, we took the simplest approach, and divided the double pulse labelling data into two groups: clones with a basal cell number larger (respectively smaller) than the median, and compared the double-labelling kinetics of both groups. In contrast to the Inv-CreER data, the distribution of double pulse fraction was statistically different for large and small clones for nearly all time points considered (Extended Data Fig. 7d). The results of Student's T-Test for samples of unequal variance at 8w post-induction was $P < 0.001$ at 4d of continuous BrdU, $P < 0.001$ at 6d of continuous BrdU, and $P < 0.05$ at 8d of continuous BrdU (we couldn't detect statistically significant differences at 2d, as the double pulse fraction is very low for both in that condition). Overall, performing the same fitting procedure as before, the division time was $T = 6d \pm 0.9$ (Fig. 4c, e).

Moreover, we also performed the same statistical test to compare the double pulse fraction in small K14-CreER labelled clones with the Inv-CreER labelled clones at the same time point, and could not find any statistically significant difference (Fig. 4c). This confirms that we are targeting two populations in the interscale with K14-CreER,

and suggests strongly that one of these populations is identical to the one induced with Inv-CreER.

We then turned to the double pulse data at 4w post-induction, and saw the same trend, albeit with a smaller difference. Both division rates were significantly faster than at 8w, and the fraction of double pulse labelling was higher for larger clones ($P < 0.05$ at 2d and 4d of continuous BrdU incorporation). The same fitting procedure gave a division time of large K14 clones of $T = 3.0d \pm 0.3$ (Fig. 4e and Extended Data Fig. 7g). Again, the proliferation kinetics of small K14 clones were statistically indistinguishable from Involucrin clones at the same time point.

Finally, as for the Involucrin SmoM2 data, we performed the same EdU/BrdU analysis at the 12w post-induction time point. We found a best fit values for the division rate at 12w of $T = 9.0d \pm 0.4$, again significantly slower than at 8w, showing that the large K14 clones continuously slowed down like Involucrin SmoM2 clones, although their division rate remained larger throughout the course of the experiment (Fig. 3e and 4e). Since we found that the outcome of the model at the 24w time point was sensitive to the offset, we also performed an EdU/BrdU double pulse experiment with a 6d BrdU chase at 20w post-induction, and found a division rate of $T = 15.0d \pm 3$ (Fig. 4e and Extended Data Fig. 7g).

Therefore, the cell kinetics experiments support the hypothesis that K14-CreER clones proliferate at a sustainably higher rate, which could explain their increase in size compared to Inv-CreER clones. As for the Inv-CreER population, we parametrized the decrease in division rate of the K14-CreER population by an exponential decay with an offset, and used it as an input in the stochastic simulations.

Apoptosis rate in K14 clones

As noted in the main text, at 8w post-induction, the overall fraction of caspase3-positive basal cells is 4-fold smaller for K14-CreER labelled clones ($f_a = 0.4\%$) than for the Inv-CreER labelled clones (Fig. 4d and Extended Data Fig. 7e-h). Moreover, although Involucrin clones showed no dependence of clone size on the fraction of caspase3 positive cells, an anti-correlation was observed in K14 clones (Extended Data Fig. 7e-h). Because of the overall low number of apoptotic cells, it was not possible to check statistically whether two populations co-exist in the small clones, one being similar to Involucrin. Similar to the Involucrin case, detailed modelling showed that apoptosis was clustered in some specific K14 clones (Extended Data Fig. 7e-h). We therefore assumed that the rate of apoptosis in basal cells of K14 clones followed the same linear trend as that found in the Involucrin clones, but with the 3-fold reduced value. In summary, in addition to an enhanced division rate, large K14 clones display a decreased apoptosis rate. This contributes to make the effective growth rate larger for K14-CreER targeted clones.

Fitting of the data

As with the K14 homeostasis analysis, we performed the fitting procedure by assuming that we label a fraction f_s of stem cells, and a fraction $1-f_s$ of progenitors. As we use the same promoters and tamoxifen doses, we assume that this fraction f_s is equal to its homeostatic value determined above ($f_s = 65\%$). One key difference with the homeostasis data, however, is that the the constraint of homeostasis is relaxed, and therefore that there can be a strictly positive imbalance $\Delta_s > 0$ towards symmetric division for stem cells,

$$S \to \begin{cases} S+S & \text{Pr. } r_s+\Delta_s \\ S+P & \text{Pr. } 1-2r_s \\ P+P & \text{Pr. } r_s-\Delta_s \end{cases}$$

The probability $P_{n_s,n_P,n_D}(t)$ to have a clone at time t post-induction containing n_S stem cells, n_P progenitors and n_D differentiated cells then follows from the Master equation:

$$\begin{aligned} \frac{dP_{n_S,n_P,n_D}}{dt} = \lambda_S\big((r_s+\Delta_s)(n_S-1)P_{n_S-1,n_P,n_D} + (r_s-\Delta_s)(n_S+1)P_{n_S+1,n_P-2,n_D} \\ +(1-2r_S)n_S P_{n_S,n_P-1,n_D} - n_S P_{n_S,n_P,n_D}\big) \\ +\lambda_P\big((r-\Delta)(n_P-1)P_{n_S,n_P-1,n_D} + (r+\Delta)(n_P+1)P_{n_S,n_P+1,n_D-2} \\ +(1-2r)n_P P_{n_S,n_P,n_D-1} - n_P P_{n_S,n_P,n_D}\big) \\ +\Gamma\big((n_D+1)P_{n_S,n_P,n_D+1} - n_D P_{n_S,n_P,n_D}\big) \end{aligned} \quad (9)$$

The ensemble of fitting parameters are then given by $\vec{\theta} = (r_s, \Delta_s, r_P, \Delta_P)$.

To analyze the data, we first considered the simplest possibility that the properties of progenitors (r_P and Δ_P) are exactly the same as the ones in Involucrin SmoM2 clones. One should note that the 24w mean suprabasal clone size increases drastically compared to the previous values at 8w and 12w post-induction. Presumably, this is due to the fact that BCC invade the dermis, allowing more suprabasal cells to accumulate. However, as we wish to avoid over-fitting the data and that this corresponds to a qualitatively different growth regime, we exclude the 24w suprabasal clone size datapoint from the fitting, to compare different models more objectively.

With the above ansatz, we could never obtain a satisfactory joint fit for both the mean basal and suprabasal clone size, even for the optimal parameters, which were

$$r_s = 0.15^{+0.2}_{-0.15}, \qquad \Delta_s = 0.0^{+0.01}_{-0.0},$$

yielding coefficients of determination of $R_b^2 = 0.79$ for the basal clone size, and $R_s = 0.91$ for the suprabasal clone size (with respective standard errors of the regression of $S_b = 29$

and $S_s = 8.2$, as well as their weighted counterparts $S_b^w = 0.7$ and $S_s^w = 2.5$). Moreover, one should note that the only way to avoid getting very large clone sizes is to have balanced fate for stem cells Δ_s, and even that overestimates the mean basal clone size at all time points. This stems from the fact that the difference in clone size between K14-CreER and Inv-CreEr is not large enough to accommodate imbalance in both SCs and CPs in K14-CreEr-targeted clones. Moreover, if SCs really were balanced, stem cells, with division rate λ_s would become diluted as clone size increases ($n_s/n_P \to 0$), which means large clones should be composed overwhelmingly of progenitors cells. However, such behavior would predict that we should not see any difference in the proliferation kinetics between large K14 clones and Involucrin clones, which is inconsistent with the data. Thus, this analysis rules against the possibility of transformed cancer stem cells producing imbalanced progenitors and suprabasal cells similar to the Involucrin SmoM2 data: In such a case, the clone sizes would be much larger than observed, even for balanced stem cell outcomes.

There are several alternative models that could be advanced to fit the observed dataset, all of which require that the properties of progenitors and differentiated cells in K14 clones are different to the progenitors and differentiated cells in Involucrin clones. The simplest possibility is that progenitors derived from cancer stem cells retain similar, nearly balanced fate properties as in homeostasis, fixing the parameters r_P and Δ_P to the values deduced in the first section: $r_P = 0.19$, $\Delta_P = 0.02$. Obviously, these parameters could be slightly different in K14-CreER/Rosa-SmoM2 clones to the homeostatic case. However, the key features detailed below would not be affected. We assume for the sake of simplicity that both stem cells and progenitors divide at the same rate, which we measured above. The parameter vector in this model is then simply $\vec{\theta} = (r_s, \Delta_s)$, and fitting the evolution of mean basal clone size (Fig. 4f), we find the optimal fitting parameters (Fig. 4g) as

$$r_s = 0.25^{+0.1}_{-0.1}, \qquad \Delta_s = 0.085^{+0.005}_{-0.015}.$$

As expected, we get precise estimates with small error bars for the values of the imbalance Δ_s, which is the crucial parameter determining clones sizes, whereas the fraction r_s changes little of the dynamics in the case of large imbalance, and consequently has large error bars. With these parameters, we can get an good fit of the data (Fig. 4f and Extended Data Fig. 7k) $R_b^2 = 0.99$ for the mean basal clone size and $R_s^2 = 0.98$ for the suprabasal clone size (with respective standard errors of the regression of $S_b = 3.8$ and $S_s = 3.17$, as well as their weighted counterparts $S_b^w = 0.42$ and $S_s^w = 2.67$), showing the superiority of this model (which additionally is consistent with the proliferation kinetics experiments by construction). We thus note that we can still fit rather accurately the clone size evolution with r_s and Δ_s as the only free parameters. The resulting proportion of stem cells to progenitors in BCC is expected to be around 1 to 4 with this parameter set.

Challenging the model

Again, we wished to see whether this simple theory could predict at least the main features of the clone size distributions (Fig. 4h). Strikingly, we noted that whereas the distribution of Involucrin SmoM2 clones was consistently well-fit by a single-exponential decay, the K14 SmoM2 clone distribution adopted a more complex shape, characterized by long tails, especially prominent at the 8w, 12w and 24w post-induction data point. Moreover, the distribution of small K14 clones followed closely the distribution of Involucrin SmoM2 clones, indicating that the two populations uncovered at homeostasis were still maintained upon oncogenic activation. This mirrors closely the prediction of our model, showing two distinct slopes in the clone size distributions.

For each time point (1, 2, 4, 8, 12 and 24 weeks), the coefficients of determination for the fit to the basal size distribution were, respectively $R^2 = [0.77, 0.92, 0.96, 0.97, 0.97, 0.97]$. The standard errors of the regression were, respectively $S = [0.29, 0.07, 0.06, 0.12, 0.08, 0.07]$, and the P-values evaluating the goodness of the fits were, respectively $P = [< 0.001, < 0.001, 0.46, < 0.001, 0.06, 0.1]$. Similar to the Involucrin SmoM2 data, at the earliest time point, the model deviates from the experimental values. This is likely due to our simplified assumption of Markovian cellular division, which means that cells can re-enters cycle arbitrarily fast, and produces artificially long tails of the clone size distribution at early times points. However, we find that our model captures the essential features of the data in the latest time point, resulting in low standard errors of the regression, and model predictions not significantly different from the data at the 95% confidence levels. Obviously there are deviations to the predicted curve, which could arise from our simplified assumption of fate probabilities remaining constant in time, throughout the 24w time-course. Nevertheless, with these limitations in mind, it is still interesting to note that our simplified model can recapitulate the key features of the data (Fig. 4h).

In particular, we statistically tested a key feature of the K14 SmoM2 data (Fig. 4h) by checking whether the distribution could be fitted by the one-population model used for the Involucrin SmoM2. Determining again the best fit parameters in a one-population model from the mean clone size time course, we therefore checked whether we could fit the full basal clone size distribution. Crucially, the P-values evaluating the goodness of the fits are, respectively, $P = [< 0.01, 0.02, < 0.005, 0.26, < 0.001, < 0.001]$. This confirms that the clone size distributions of the K14 SmoM2 data cannot be fit by a single exponential, and that the hierarchical model has a higher predictive power, especially for the latest time points of BCC progression, which are the most relevant.

Furthermore, a key prediction of the model is that a large fraction of basal cells in K14 clones consist of stem cells, which divide to give rise to basal progenitors. Therefore, we expect short-term fate outcomes to be heavily tilted towards basal-basal outcome in K14 clones. We therefore turned to the same retracing experiment as in the Involucrin case, which use EdU as a clonal marker to follow "clones in clones". We used precisely the same setting of short EdU pulse followed by a 3-day chase in 24w post-induction mice. We first concentrated on fate outcomes in BCCs, as we were certain that these originated

from "true" K14 stem cells (Extended Data Fig. 7i bottom). Strikingly, concentrating as before on two-cell clones, the large majority consisted of two basal cells, in stark contrast to retracing in Involucrin clones.

The probabilities of obtaining the combination basal-basal (p_{b-b}), basal-suprabasal (p_{s-b}) and suprabasal-suprabasal (p_{s-s}) were respectively:

$$p_{b-b} = 0.77 \pm 0.07, \qquad p_{b-s} = 0.21 \pm 0.02, \qquad p_{s-s} = 0.02 \pm 0.002.$$

These data are consistent with our prediction of a lineage hierarchy maintaining K14 clones, with stem cells dividing to give rise to progenitors. Next, we examined dysplasic and hyperplasic clones (Extended Data Fig. 7i top), and found their fate outcome to be intermediate between those of BCCs and those of Involucrin SmoM2:

$$p_{b-b} = 0.58 \pm 0.08, \qquad p_{b-s} = 0.36 \pm 0.05, \qquad p_{s-s} = 0.06 \pm 0.01.$$

Again, this is in line with our hypothesis of medium sized-clones in K14 consisting of both "true" K14 stem cell clones, and Involucrin-like progenitor clones.

Persistence of K14 vs Involucrin clones

Yet another important difference between K14 and Involucrin clones upon Smo activation is their differential persistence (Extended Data Fig. 5b and Extended Data Fig. 7b). Although K14 clones are maintained across the 24w time course, Involucrin clones are slowly lost with time, reaching 40% persistence after 24w post-induction. In our theory with stochastic apoptosis, we expect the persistence of imbalanced clones to plateau to a constant value [5] after a transitory period of time, mirroring the K14 data. A possible explanation for the loss of Involucrin clones could be the collective apoptosis we have uncovered in the previous section, which would remove Involucrin clones at a rate 4-times higher than K14 clones, or a greater rate of symmetric differentiation that our model currently predicted.

2.2 Scale

We have seen in the main text that SmoM2 expressing scale clones never give rise to BCC, and are statistically smaller, both for K14 and Involucrin.

Nevertheless, when examining the persistence of scale clones (Extended Data Fig. 6c), we noticed that it is also much lower than interscale clones (Extended Data Fig. 5b and Extended Data Fig. 7b). Strikingly, K14 and Involucrin had statistically indistinguishable persistence (Extended Data Fig. 6c), arguing again that we mark the same population, in analogy to the homeostatic behavior.

We then compared the mean basal clone size of K14 and Involucrin clones in the scale. Although K14 clones are consistently larger (Extended Data Fig. 6a,b), this difference is fully explained by very short-term differences in cell fate associated, for example, with fate priming. Indeed, when we normalised basal clone size by its value at 1w, K14 and Involucrin data collapse on each other, arguing again that the long term behaviour of K14 and Involucrin is identical in the scale (Extended Data Fig. 6d).

In analogy to the homeostasis section, we therefore plotted Π, the total labelled cell fraction as a function of time for scale (Extended Data Fig. 6a-b). As expected from the aforementioned normalization, the labelled cell fractions were very similar for K14 and Involucrin clones. Moreover, we noticed that, whereas it increases monotonously in interscale (both for Involucrin and K14), it remains roughly constant after 8w for scale, after an initial phase of small expansion. This suggests that, in contrast to interscale tumors (Extended Data Fig. 5c and Extended Data Fig. 7c), which display net growth over the entire course of the 24w experiment, scale tumors have limited growth potential, and display near-balanced fate over the long-term, after a phase of moderate net growth in the first weeks. Crucially, K14 and Involucrin in scale display the same long-term kinetics, consistent with the one-population model uncovered at homeostasis. Nevertheless, we found evidence for a phase of short-term priming, which might be longer in cells targeted by K14CreER, explaining the short-term growth advantage of K14-targeted clones.

2.3 Discussion of the SmoM2 over-expression data

Our results show that oncogenic activation has profoundly different effects depending on the location within tail epidermis and the cell-of-origin in which it is targeted. In Involucrin-targeted clones, SmoM2 over-expression causes a prolonged bias towards symmetric division compared to symmetric differentiation, resulting in net clonal expansion. This net expansion would be exponential if the effective division rate did not continuously decay, under the combined effect of enhanced apoptosis and progressively decreasing division rate.

One of the most striking and surprising features of the data is that the Involucrin-targeted cohort behaves extremely uniformly: the decay in division rate appears uniform in time, and does not depend on the size of the clone considered. Moreover, clone size distributions are well-fit by single-exponentials throughout the 24w time course, again indicative of Involucrin-targeted clones behaving long-term as a homogeneous population, even upon oncogenic activation. Although we have presented in the sections above a best fit for the detailed fate probabilities ($P \rightarrow PP, PD, DD$), one should note that the only key parameter for the growth characteristic of the clones is $\lambda_P \Delta_P$, i.e. the product of the effective division rate and the imbalance probability towards symmetric renewal. Therefore, as we measure λ_P independently, we can get a rather precise prediction for the value of Δ_P throughout the time course. On the other hand, the parameter r_P,

which quantifies the percentage of asymmetric divisions, is largely unimportant for the tumour dynamics, in contrast to homeostatic dynamics. This is because r_P creates a linear contribution to growth (neutral drift, $\langle n_P \rangle \propto \lambda_P r_P t$), whereas $\Delta_P > 0$ creates an exponential contribution to growth (biased drift, $\langle n_P \rangle \propto e^{\lambda_P \Delta_P t}$). This explains why we get a large confidence interval when extracting a value of r_P from the fits.

The fact that Involucrin-targeted clones function as an homogeneous equipotent population over half a year raises the interesting question of subsequent accumulation of driver and passenger mutations. Indeed, several studies in the past decades have suggested the key importance of successive acquisition of mutations for many types of cancer expansion and metastasis [20, 21, 22, 23]. As this would cause departure from equipotency, this is likely to be only a small effect for the Involucrin-targeted clones. We conjecture that this might be due to the limited long-term proliferative potential of these cells, which would make the appearance of additional mutations unlikely.

In contrast to the Involucrin tracing, we have shown that K14-targeted clones have a large survival advantage over the long term upon SmoM2 over-activation. Because of a larger resistance of growth arrest, as well as the hierarchical fate organization of the tissue, K14-targeted clones can grow to much larger size and reach an invasive stage. Again, here we have extracted from our fits the detailed fate probabilities of stem cells upon division ($S \to SS, SP, PP$), supplemented with the fate probabilities of progenitors. As above, the key parameters for growth are the parameter combinates $\lambda_S \Delta_S$ and $\lambda_P \Delta_P$, whereas the percentage of asymmetric divisions, measured by r_S, is largely unimportant for tumour growth. We therefore again provide a precise confidence interval for Δ_S, as opposed to r_S, which prevents us from being able to make conclusive claims as to whether the imbalance towards symmetric renewal comes from a decreased probability of asymmetric divisions ($S \to SP$) or decreased probability of symmetric differentiation ($S \to PP$). This question could only be assessed specifically by performing a "clone-in-clones" analysis at various time points of tumour progression.

Although the same two-population hierarchical model deduced from homeostasis could explain accurately the K14-targeted SmoM2 data, it has been reported that secondary mutations can play a key role later on for basal cell carcinoma development, as well as its resistance to treatment [24]. Again, extensive modelling effort has been concentrated on understanding the evolution of resistance to treatment, both theoretically [20], and applied in several types of cancer, such as chronic lymphocytic leukemia [25], or solid tumors in the skin, pancreas and colon [26]. An extension of our current work would therefore be either to perform long-term retracing in larger basal cell carcinoma, or examine the effect of drug therapy, in order to see whether similar evolutionary behavior would apply.

3 P53 deletion causes CPs to become competent for BCC progression

As stated in the main text, we found that P53 displays lower stabilization in K14-targeted SmoM2 clones (Extended Data Fig. 8a), which progress to BCC, compared to Involucrin-targeted SmoM2 clones. Given the importance of P53 in the regulation of cell cycle and apoptosis, this raised the question of whether an additional mutation, such as P53 deletion, could allow CPs to progress to the BCC stage. Indeed, we crossed our Inv-CreER/Rosa-SmoM2 and K14CreER/Rosa-SmoM2 mice with p53fl/fl mice, and found that Involucrin-targeted clones became competent to form BCC. This finding underlines the role of P53 deletion in BCC progression, and also raised the key issue of how P53 deletion mediates this response. We also noted a marked increase in clone sizes for both conditions. Interestingly, comparing the distributions of basal clone sizes at 8w and 12w post-induction with or without P53 deletion, we found that the increase in clone size was not driven by the uncontrolled growth of a small population, but rather by a net global size increase throughout the population. This was true both for the K14- and Involucrin-targeted clones (Extended Data Fig. 8c-d). This indicates that the effect of P53 deletion impacts the majority of clones.

Given the consistent differences in division and apoptotic rates described in the sections above between K14 SmoM2 and Involucrin SmoM2 clones, we therefore performed similar caspase3 stainings (Fig. 5e and Extended Data Fig. 8f,g) and EdU/BrdU double pulse labelling experiments (Fig. 5f) as before, to verify whether P53 deletion affected any of these rates. As stated in the main text, although we find that BCC formed from Involucrin-targeted P53fl/fl clones have lowered levels of apoptosis, mirroring our findings for K14-targeted SmoM2 clones, the overall levels of apoptosis in small clones are largely unaffected by P53 deletion. In particular, we found that apoptosis in Involucrin Smo/P53 clones is still markedly higher than in K14 Smo/P53. Coupled with our findings that clone sizes become uniformly larger upon P53 deletion (Extended Data Fig. 8c-d), this indicates that apoptosis is likely to play a secondary role in tumor growth and BCC progression, although its down-regulation could enhance the observed survival of BCC.

However, a pulse of EdU followed by 6 day continuous BrdU incorporation at 12w post-induction revealed a significant increase in the fraction of double labelled cells due to P53 deletion (Fig. 5f), both for Involucrin and K14-targeted clones (respectively $P = 0.037$ and $P = 0.0045$ from a Mann-Whiney U non-parametric test). Interestingly, the distributions of double labelled fractions in Involucrin and K14 targeted clones became statistically indistinguishable upon P53 deletion ($P = 0.87$, Mann-Whitney U test), in contrast to what we have described above without P53 deletion. Fitting with the same procedure as before yielded respectively a 20% and 30% increase in division rate upon P53 deletion in K14 and Involucrin targeted clones at 12w post-induction.

A natural hypothesis is therefore that P53 deletion allows CPs to escape growth arrest, and to reach division rates close to the ones of K14-targeted clones. This allows Involucrin-targeted clones to grow to much larger sizes, which could help to reach a critical size for BCC progression and invasion into the dermis. To verify this hypothesis quantitatively, we again numerically simulated Eq. 6-9, using the exact same parameters we had extracted for K14 and Involucrin SmoM2 clones, but simply increasing the division rates by respectively 20% and 30% uniformly throughout the 12w time course. Obviously, this is an approximation, as the division rate could vary in a more complicated manner and/or p53 could affect fate outcomes for CPs and SCs. However, we found that such an increase in division rates was sufficient to explain the bulk of the increase in basal clone size, thus yielding a good prediction for the evolution of the basal clones sizes throughout the 12w time-course (see Fig. 5g for prediction, together with a 95% confidence interval, given the confidence interval on the model parameters determined from the SmoM2 data). For Involucrin-targeted clones (respectively K14-targeted clones), we found coefficients of determination and standard error of the fit $R_s^2 = 0.98/S_s = 6.2$ (respectively $R_s^2 = 0.99/S_s = 5$) for the basal clone size, indicative of a high predictive power of the model in both cases.

Interestingly, even though P53 deletion increased significantly both clones sizes, we found that SCs derived clones remain larger than CPs derived clones (Fig. 5i and Extended Data Fig. 8c,d), as well as more persistent (Extended Data Fig. 8e), which underlines that the hierarchical organisation of the skin and the nature of the cell of origin still plays a role in the dynamics of oncogenic growth upon P53 deletion (Extended Data Fig. 9).

References

[1] Clayton, E., Doupe, D. P., Klein, A. M., Winton, D. J., Simons, B. D., and Jones, P. H. (2007). A single type of progenitor cell maintains normal epidermis. Nature, 446(7132), 185-189.

[2] Mascre, G., *et al.* (2012). Distinct contribution of stem and progenitor cells to epidermal maintenance. Nature, 489(7415), 257-262.

[3] Gomez, C., Chua, W., Miremadi, A., Quist, S., Headon, D. J., and Watt, F. M. (2013). The interfollicular epidermis of adult mouse tail comprises two distinct cell lineages that are differentially regulated by Wnt, Edaradd, and Lrig1. Stem Cell Reports, 1(1), 19-27.

[4] Bailey, N. T. J. The Elements of Stochastic Processes, J. Wiley and Sons, New York, 1964, Chap. 8.6.

[5] Klein, A.M., Doupe, D.P., Jones, P.H. and Simons, B.D., (2007). Kinetics of cell division in epidermal maintenance. Physical Review E, 76(2), p.021910.

[6] Antal, T., and Krapivsky, P. L. (2011). Exact solution of a two-type branching process: models of tumor progression. Journal of Statistical Mechanics: Theory and Experiment, 2011(08), P08018.

[7] Lim, X., *et al.* (2013). Interfollicular epidermal stem cells self-renew via autocrine Wnt signaling. Science 342, no. 6163: 1226-1230.

[8] Clevers, H., Loh, K.M. and Nusse, R. (2014). An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. Science 346, no. 6205: 1248012.

[9] Manly, B.F.J. (2006). Randomization, bootstrap and Monte Carlo methods in biology. Vol. 70. CRC Press.

[10] Brown, A.M. (2001). A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. Computer methods and programs in biomedicine 65, no. 3: 191-200.

[11] Massey, F.J. (1951). The Kolmogorov-Smirnov test for goodness of fit. Journal of the American statistical Association 46, no. 253: 68-78.

[12] Lilliefors, H.W. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. Journal of the American Statistical Association 64, no. 325: 387-389.

[13] Stephens, M.A. (1974). EDF statistics for goodness of fit and some comparisons. Journal of the American statistical Association 69, no. 347: 730-737.

[14] Doup, D. P., Klein, A. M., Simons, B. D., and Jones, P. H. (2010). The ordered architecture of murine ear epidermis is maintained by progenitor cells with random fate. Developmental cell, 18(2), 317-323.

[15] Schweizer, J., and Marks, F. (1977). A developmental study of the distribution and frequency of Langerhans cells in relation to formation of patterning in mouse tail epidermis. Journal of Investigative Dermatology, 69(2), 198-204. Chicago

[16] Sada, A., Jacob, F., Leung, E., Wang, S., White, B. S., Shalloway, D., and Tumbar, T. (2016). Defining the cellular lineage hierarchy in the interfollicular epidermis of adult skin. Nature cell biology.

[17] Rompolas, P., Mesa, K. R., Kawaguchi, K., Park, S., Gonzalez, D., Brown, S., ... and Greco, V. (2016). Spatiotemporal coordination of stem cell commitment during epidermal homeostasis. Science, aaf7012.

[18] Till, J.E., McCulloch, E.A., and Siminovitch, L. (1964). A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. Proceedings of the National Academy of Sciences of the United States of America, 51(1), 29.

[19] Benjamin, C.W., Hiebsch, R.R. and Jones, D.A. (1998). Caspase activation in MCF7 cells responding to etoposide treatment. Molecular pharmacology, 53(3), pp.446-450.

[20] Durrett, R., and Moseley, S. (2010) Evolution of resistance and progression to disease during clonal expansion of cancer. Theoretical population biology 77, no. 1: 42-48.

[21] Nowak, M.A., Komarova, N.L., Sengupta, A., Jallepalli, P.V., Shih, I-M. Vogelstein, B. and Lengauer, C. (2002). The role of chromosomal instability in tumor initiation. Proceedings of the National Academy of Sciences 99, no. 25: 16226-16231.

[22] Bozic, I., Antal T., Ohtsuki, H., Carter, H., Kim, D., Chen, S., Karchin, R., Kinzler, K.W., Vogelstein, B. and Nowak, M.A. (2010). Accumulation of driver and passenger mutations during tumor progression. Proceedings of the National Academy of Sciences 107, no. 43: 18545-18550.

[23] Lengauer, C., Kinzler, K.W. and Vogelstein, B. (1998). Genetic instabilities in human cancers. Nature 396, no. 6712: 643-649.

[24] Pricl, S., Cortelazzi, B., Col, V.D., Marson, D., Laurini, E., Fermeglia, M., Licitra, L., Pilotti, S., Bossi, P. and Perrone, F. (2015). Smoothened (SMO) receptor mutations dictate resistance to vismodegib in basal cell carcinoma. Molecular oncology 9, no. 2: 389-397.

[25] Komarova, N.L., Burger, J.A. and Wodarz, D. (2014). Evolution of ibrutinib resistance in chronic lymphocytic leukemia (CLL). Proceedings of the National Academy of Sciences 111, no. 38: 13906-13911.

[26] Bozic, I., *et al.* (2013) Evolutionary dynamics of cancer in response to targeted combination therapy. Elife 2: e00747.