# Stem cell lineage survival as a noisy competition for niche access

**Bernat Corominas-Murtra[a,1,2], Colinda L.G.J. Scheele[b,1], Kasumi Kishi[a], Saskia I.J. Ellenbroek[b], Benjamin D. Simons[c,d,e], Jacco van Rheenen[b,2], and Edouard Hannezo[a,2]**

[a]Institute for Science and Technology Austria. Am Campus 1, A-3400 Klosterneuburg, Austria; [b]Division of Molecular Pathology, Oncode Institute, The Netherlands Cancer Institute, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands.; [c]The Wellcome Trust/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge CB2 1QN, UK; [d]Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, UK; [e]The Wellcome Trust/Medical Research Council Stem Cell Institute, University of Cambridge, Cambridge CB2 1QN, UK

**Understanding to what extent stem cell potential is a cell-intrinsic property, or an emergent behavior coming from global tissue dynamics and geometry, is a key outstanding question of systems and stem cell biology. Here, we propose a theory of stem cell dynamics as a stochastic competition for access to a spatially-localized niche, giving rise to a stochastic conveyor-belt model. Cell divisions produce a steady cellular stream which advects cells away from the niche, while random rearrangements enable cells away from the niche to be favourably repositioned. Importantly, even when assuming that all cells in a tissue are molecularly equivalent, we predict a common ("universal") functional dependence of the long-term clonal survival probability on distance from the niche, as well as the emergence of a well-defined number of functional stem cells, dependent only on the rate of random movements vs. mitosis-driven advection. We test the predictions of this theory on datasets on pubertal mammary gland tips, embryonic kidney tips as well homeostatic intestinal crypt. Importantly, we find good agreement for the predicted functional dependency of the competition as a function of position, and thus functional stem cell number in each organ. This argues for a key role of positional fluctuations in dictating stem cell number and dynamics, and we discuss the applicability of this theory to other settings.**

Stem cell dynamics, biophysical modelling, stochastic processes, mammary morphogenesis, intestinal renewal

Many biological tissues are renewed via small numbers of stem cells, which divide to produce a steady stream of differentiated cells and balance homeostatic cell loss. Although novel experimental approaches in the past decade have produced key insights into the number, identity, and (often stochastic) dynamics of stem cells in multiple organs, an outstanding question remains as to whether stem cell potential is a cell-intrinsic, "inherited" property, or rather an extrinsic, context-dependent state emerging from the collective dynamics of a tissue and cues from local "niches", or microenvironments (1–8). Although recent experiments have provided evidence for the latter in settings such as the growing mammary gland (9), adult interfollicular epidermis (10, 11), spermatogenesis (12) or the intestinal epithelium (13), a more global theoretical framework allowing to quantitatively interpret these findings is still lacking.

The case of the intestinal crypt serves as a paradigmatic example of the dynamics of tissue renewal, and is one of the fastest in mammals (13). The intestinal crypt consists of a small invagination in the intestine where the epithelial cells populating the intestinal walls are constantly produced. The very bottom of the crypt hosts a small number of proliferative, Lgr5+ stem cells, (14) that divide and push the cells located above them to the transit amplification (TA) region, where cells lose self-renewal potential. Cells are eventually shed in the villus a few days later, constituting a permanent "conveyor-belt" dynamics. Lineage tracing approaches, which irreversibly label a cell and its progeny (3), have been used to ask which cell type will give rise to lineages that renew the whole tissue and have revealed that all Lgr5+ cells can stochastically compete in an equipotent manner on the long term (15–18), but still display positional-dependent short term biases for survival (13). Interestingly, similar conclusions have been reached in pubertal mammary gland development (9), where branching morphogenesis occurs through the proliferation of the cells in the terminal end buds of the ducts (19), the region where the mammary stem cells (MaSCs) reside (9, 20). In both cases, intravital imaging revealed random cellular motions enabling cells to move against the cellular flow/drift defined by the conveyor belt dynamics. Moreover, in the intestine, tissue damage, or genetic ablation of all Lgr5+ stem cells, caused Lgr5- cells to recolonize the crypts and re-express Lgr5+ to function as stem cells (13), arguing for extensive reversibility and flexibility in the system (21). In addition, Lgr5- and Lgr5+ cells of the fetal gut were also shown to nearly equally contribute to intestinal morphogenesis (22). Altogether, this supports proposals that the definition of stem cell potential should evolve to emphasize, instead of molecular markers, the functional ability of cells to renew over the long-term (23, 24).

## Significance Statement

What defines the number and dynamics of the stem cells that generate and renew biological tissues? Although several molecular markers have been described to predict stem cell potential, we propose a complementary approach that mathematically describes "stemness" as an emergent property arising from a stochastic competition for space. We predict from that competition the robust emergence of a region made of functional stem cells, as well as give simple predictions on lineage survival probability. We test our results with data obtained from intravital live-imaging experiments in mammary gland development, existing data from kidney development and from the self-renewal of the crypt, to show that our framework can predict the number of functional stem cells and lineage survival probability.

B.C.-M., B.D.S, J.v.R and E.H. designed research; B.C.-M., C.L.G.J.S., K.K., S. I.J. E., and E.H. performed research and analyzed data; B.C-M and E.H. wrote the paper with input from C.L.G.J.S., B.D.S and J.v.R.

The authors declare no conflicts of interest.

[1]B.C.-M. and C.L.G.J.S. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: bernat.corominas-murtra@ist.ac.at, j.v.rheenen@nki.nl, edouard.hannezo@ist.ac.at

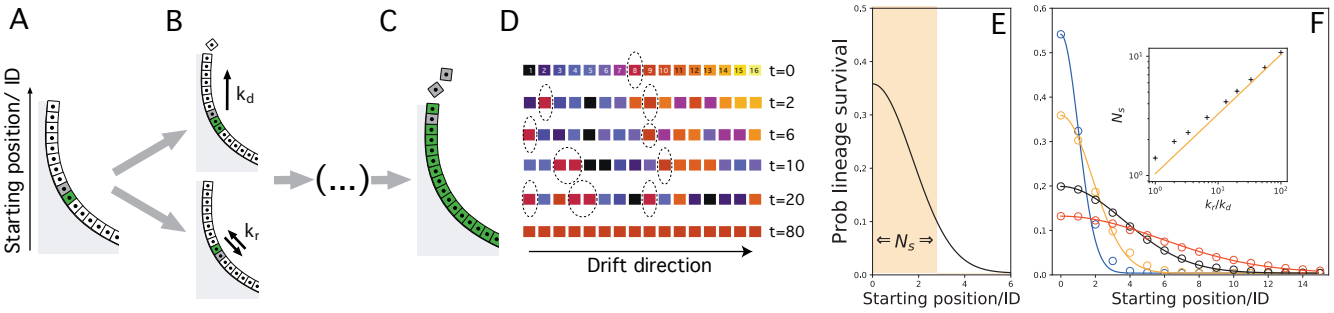


**Fig. 1.** Stochastic conveyor belt as a paradigm for stem cell renewal. A/ A cell in the epithelial wall of the crypt can B/ duplicate at rate $k_d$ pushing the upper cells up, creating a conveyor-belt mechanism; or switch its position randomly at rate $k_r$, introducing a stochastic or noisy ingredient in the dynamics. C/ At longer time scales, the lineage of a single starting cell colonizes the whole system. D/ Example of SCB dynamics. At $t = 0$ we have $N = 16$ lineages in the system, depicted with different colors and at starting positions $1, \ldots, 16$ respectively. In time, lineages are progressively eliminated, but stochastic cell rearrangements makes it possible for a lineage far from the origin (starting position $n = 8$ in red and highlighted with a dashed circle) to win the competition. E/ Probability that a given lineage colonizes the entire system as a function of initial position of its mother cell, decaying as a Gaussian of width $\sqrt{k_r/k_d}$, see text for details. The width of this distributions defines a functional stem cell region ($N_s$ cells, highlighted in orange, plotted for $k_r/k_d = 3$). F/ Numerical simulations of the 1-dimensional SCB dynamics. We compute the long term survival probability $p(c_n)$ as a function of initial starting position $n = 0, 1, 2, \ldots$, with respect to the base of the system for several values of $k_r/k_d$ (1, 3.3, 13.3 and 33 in resp. blue, orange, black and red). Dots show the outcome of the simulations and lines show the analytical prediction $p(c_n) \sim \exp\{-\frac{k_d}{2k_r} n^2\}$, as shown in equation (4). Inset shows plot of best fit for the variance of the numerical distributions (black crosses) against the analytical model prediction $\sim \sqrt{k_r/k_d}$ (orange solid line).

However, this new definition raises a number of outstanding conceptual problems: What then defines the number of functional stem cells in a tissue? How can short-term biases be reconciled with long-term equipotency? Is there a sharp distinction between stem and non-stem cells, or is there instead a continuum of stem cell potential together with flexible transition between states? Qualitatively, it is clear that fluctuations and positional exchanges are needed to prevent a single cell in the most favourable position to be the unique "functional" stem cell (defined as cells whose lineage colonizes a tissue compartment on the long-term). Incorporating these features in a dynamical model of stem cell growth and replacement, able to make predictions e.g., on the probability of lineage perpetuation, would represent an important step towards the understanding of how stem cells operate in the process of tissue growth and renewal.

In this paper we develop a reaction-diffusion formalism for stem cell renewal in the presence of noise and local niches, taking into account local tissue geometry as well as cell division and random cell movements (Fig. 1a–c). Importantly, within this purely extrinsic and dynamical approach, which does not need to posit any intrinsic "stem cell identity", a well-defined number of functional stem cells emerges, which only depends on the geometry and a balance between the noisiness of cell movements and division rates advecting cells away from niche regions. This model also predicts that stem cell potential should decay continuously as a function of distance from the niche, with a "universal" Gaussian functional dependence. We test this prediction against published live-imaging datasets for the homeostatic intestinal crypt (13) and during the branching of embryonic kidney explants (25), and find a good quantitative agreement for the full survival probability of cells depending on their initial position relative to the niche. Furthermore, we use our theoretical results to extract the amplitude of the random positional fluctuations in the developing mammary gland using static lineage tracking experiments (9). This enables us to predict the number of functional stem cells for this system, finding values consistent with previously reported estimates.

## Dynamics of tissue renewal and development

To develop the model, we first consider the simplest situation of a one-dimensional column of cells, with a rigid boundary condition at the base (mimicking, for instance, the bottom of the crypt), so that each cell division produces a pushing force upwards transmitted to the cells above (or in the case of growing mammary gland or kidney, driving ductal elongation). This model is motivated by its simplicity, as it is able to qualitatively derive the essential traits of the complex dynamics studied here. As we shall see, further refinements, aimed at making predictions for real systems, consider more realistic geometries. From this simple dynamics, we define the number of functional stem cells as the typical number of cells that have a non-negligible probability to produce long-term progenies (without "losing" the competition against other cells). If the dynamics was fully devoid of noise (a simple conveyor belt) and all cell divisions were symmetric, then one of the bottom-most cells would always win the competition. In the case of a 1-dimensional array of cells, this problem is trivial. If one considers a cylindric geometry, there would be a single row of functional stem cells, which is the limiting case of the model described in Ref. (16) of symmetric and stochastic 1-dimensional, neutral competition along a ring of equipotent cells. However, live-imaging studies shows that, in multiple settings including mammary gland (9), kidney morphogenesis (25, 26) and intestinal crypts (13), there is widespread rearrangement of cells through stochastic cell movements. Intuitively, such rearrangements are expected to increase the number of "functional" stem cells, as re-arrangements allow cells away from the niche to relocate to favourable positions, and would thus provide a biophysical mechanism for setting the number of stem cells assumed in models such such as that developed in Ref. (16).

The simplest abstraction of the system is a 1-dimensional column of $N$ cells. Each cell divides at constant rate $k_d$. In 1D, we assume a rigid boundary at the bottom so that cell proliferation generates a net flow of cells along the positive axis, i.e. advection away from the niche. In addition, the position of the cells can fluctuate stochastically at rate $k_r$ (either via

local cell-cell rearrangements, or more global movements of cells relative to the niche, see sections S1A and S4 of the SI for details), allowing cells far away from the niche to reposition despite the overall flow.

At $t = 0$, each cell is characterised by its starting position $n$ (distance from the niche), and will give rise in time to a lineage denoted $c_n$, which can span the entire tissue. However, as soon as a cell reaches the position $N$, it disappears from the system, resulting after a sufficiently large time period in a single surviving lineage. This competitive dynamics can be metaphorically understood as a conveyor belt with random fluctuations in the cell positions, sketched in Fig. 1a–c. This is why we call it *Stochastic Conveyor Belt* (SCB) dynamics, and use it to model tissue renewal (e.g. intestinal crypt homeostasis) or organ growth (e.g. kidney and mammary gland morphogenesis). The only difference between these two general cases is a change of reference frame (see section S1 and Fig. S1 of the SI). In Fig. 1d, we show an example of a typical run of the simulated SCB dynamics in 1 dimension, until monoclonality is achieved (see also Video S1-3 and section S5A of the SI for details).

To make quantitative predictions from the dynamics outlined above, we start by following the prevalence of a single lineage. Here, the action of the other lineages can be imposed as an average drift force that depends on the position of each cell of the lineage we follow. The equation accounting for the time evolution of the prevalence of lineage $c_n$, to be referred to as $\rho_n(z,t)$, in the continuum limit is:

$$\frac{\partial \rho_n}{\partial t} = -k_d \frac{\partial}{\partial z}(z\rho_n) + \frac{k_r}{2}\frac{\partial^2 \rho_n}{\partial z^2} + k_d \rho_n \quad . \qquad [1]$$

We refer to this reaction-diffusion equation (27, 28) as the SCB equations (see SI, section S1B for details). The first term on the right hand-side is a drift term, accounting for the average push up movement at position $z$ due to random cellular proliferation at rate $k_d$ at lower levels, $\sim k_d z$. The second term is a diffusive term (29, 30) accounting for the random reallocations of cells, occurring at rate $k_r$. The third term is a proliferative term, accounting for the exponential proliferation of each cell of the lineage under study, at rate $k_d$.

Considering initial conditions $t_0 = 0$, $\rho_n(z,0)$ a Gaussian centered on $n$ with $\sigma^2 = 1/2$ (a density representing a single cell at position $n$) and natural boundary conditions, the solution of equation (1) can be approximated by (see SI, section S1B, for details):

$$\rho_n(z,t) \approx \sqrt{\frac{k_d}{2\pi k_r}} \exp\left\{ -\frac{k_d}{2k_r}\left(\frac{z - n e^{k_d t}}{e^{k_d t}}\right)^2 \right\} \quad . \qquad [2]$$

Next, we sought to relate this lineage prevalence to the experimentally relevant quantity of long-term lineage survival, in other words, *how likely is it for a cell starting at a given position n to take over the entire crypt?*. Although lineage fixation is a concept that only makes sense in the discrete lineage, we observed that lineage prevalence converges asymptotically towards a simple scaling form $\rho_n(\infty)$:

$$\rho_n(\infty) \equiv \lim_{t\to\infty} \rho_n(z,t) \quad , \qquad [3]$$

which is a constant that does not depend on position $z$ or time $t$, but only on the starting position of the lineage. This argues that, on the long-term, lineages starting at different positions $n$ and $n'$ have well-defined relative prevalence, leading to the natural assumption that the long term lineage survival probability of lineage $c_n$ is proportional to this asymptotic lineage prevalence. This means that the probability of lineage survival, $p(c_n)$, can be expressed as:

$$p(c_n) \approx \frac{\rho_n(\infty)}{\sum_j \rho_j(\infty)} \propto \exp\left\{ -\frac{k_d}{2k_r} n^2 \right\} \quad . \qquad [4]$$

The above equation, which is a central result of the study, defines the probability that a cell starting at position $n$ will "win the competition" and colonize the whole one-dimensional system (see SI, section S1 for details).

In spite of the approximations outlined above, stochastic numerical simulations of the model system show excellent agreement with equation (4)(Fig. 1e,f). We also note that although we have assumed here that positional rearrangements occur between two cells, more complex sources of positional noise $k_r$ can be considered (which can be mechanistically dependent or independent on $k_d$), and lead to the same qualitative results. These include, for instance, post-mitotic dispersal, as seen during the branching morphogenesis of the kidney uteric bud (26) and where daughter cells can travel long distances outside the epithelium post-division, or correlated "tectonic" movements of the epithelium, where cells could collectively reposition relative to the niche, as proposed during mammary or gut morphogenesis (9, 22) (see SI section S4 for details).

## Functional stem cell numbers and dynamics in the stochastic conveyor belt

The prediction for the probability of long-term lineage survival under the SCB dynamics is surprisingly simple, decaying as a Gaussian distribution as a function of position away the niche, with a length scale that is simply the amplitude of the stochastic fluctuations divided by the proliferation rate, $\sim \sqrt{k_r/k_d}$ (see Eq. (4)). Intuitively, cells close to the origin have the highest chance to win and survive, whereas this probability drops abruptly for cells starting the competition further away, i.e. around $N_s$ cell diameters away from the base, with:

$$N_s^{1D} = 1 + \sqrt{\frac{k_r}{k_d}} \quad . \qquad [5]$$

Note that the first term satisfies the boundary condition that, in the case $k_r = 0$, the system has a single functional stem cell (located at the base) in 1 dimension. Eq. (5) thus implies that multiple rows of cells possess long-term self-renewal potential (as assessed for example, in a lineage tracing assay), emerging through their collective dynamics, and with a number that depends only on the ratio of the division to rearrangement rates (resp. $k_d$ and $k_r$). Although Eq. (5) is the outcome of a 1-dimensional approximation, we show that it holds and can be generalized in more complex geometries (see section S3 of the SI). In particular, in a cylindrical 2-dimensional geometry, we show the functional stem cell number would simply be the same number $N_s^{1D}$ of cell rows (arising from the stochastic conveyor belt dynamics) multiplied by the number of cells per row (fixed by the geometry of the tissue). Moreover, the above result can be generalized, giving an estimate of $N_s$ for general geometries (see Eq. (26) of the SI, where we give the general expression for $N_s$ in arbitrary organ geometries). This general result will be at the basis of the forthcoming

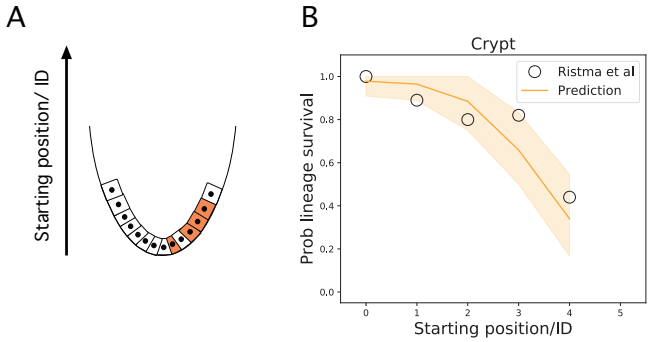


**Fig. 2.** A/ Schema of the self-renewal of the crypt epithelia, showing the origin of the coordinate system at the bottom of the system. B/ The probability that a given lineage remains within the system as a function of the starting position after a time lapse against the predictions of the conveyor belt dynamics for the crypt. Data corresponding to the probability that a lineage remains in the system for the small intestinal crypt, reported in (13) depending on its starting position. The orange line represents the prediction of the stochastic conveyor belt dynamics, fitting well the data for $k_r/k_d \approx 1$. Shaded areas represent the confidence interval (1S.D.) of the prediction.

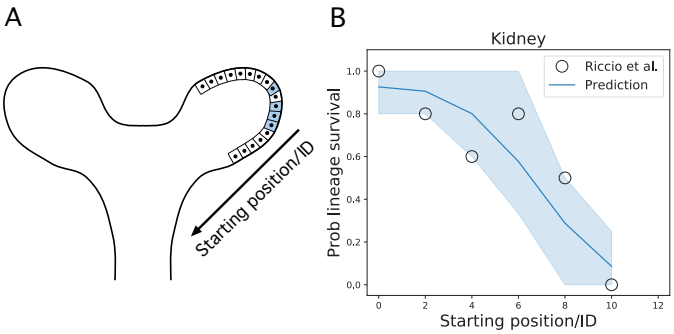


**Fig. 3.** A/ Schema of the kidney tip during development. The conveyor-belt dynamics holds, the only difference is the reference frame: Whereas in the stem cell replacement model of the intestinal crypt the reference frame is the bottom of the gland, in the kidney and mammary gland, the reference frame is taken from the newly created ducts. B/ The probability that a given clonal remains within the system as a function of the starting position of the mother cell after a given time against the predictions of the conveyor belt model dynamics. Black circles represent real data points, obtained by counting the amount of cells of a given lineage remaining in the system (from Ref. (25)). We observe that the distribution is much broader, fitting well to the theory for a ratio $k_r/k_d \approx 16$ in kidney, over an order of magnitude larger than in intestinal crypt. Shaded area represents the confidence interval (1S.D.) of the prediction.

sections, when dealing to more realistic geometries to explore the dynamics of the organs under study. Importantly, our framework generalizes the work of Ref. (16), as we do not fix the stem cell number $N_s$ explicitly, which rather emerges from an interplay between geometry and SCB dynamics, together with the competitive dynamics being qualitatively different in the flow direction (see section S1 of the SI).

We now turn to experimental data to test whether the proposed dynamics can help predict the number of functional stem cells in several organs, as well as the evolution of the survival probability with starting position of a clone. Although the division rate $k_d$ is well-known in most systems considered, the stochastic movement rate $k_r$ is harder to estimate, and can potentially vary widely, from rather small in intestinal crypts (13), to large in mammary and kidney tips, with extensive clonal fragmentation and random cell movements (9, 25).

**Predictions on clonal dynamics and survival.** Intravital live-imaging provides an ideal platform to test the model, as it provides both knowledge of the starting position of a given cell as well as its clonal time evolution (whereas classical lineage tracing relies on clonal ensembles obtained from fixed samples). In small intestinal crypts, different Lgr5+ cells have been predicted to have very different lineage survival potential on the short-term, depending on their position within the stem cell niche, resulting in an effective number of stem cells smaller than the number of Lgr5+ cells (13, 31). We thus reanalyzed quantitatively this dataset by plotting the survival probability of a clone as a function of its starting position $n$ (Fig. 2) after a given time period assumed to be large enough for equation (4) to hold. We then compared this to a 2-dimensional stochastic simulation of the model (see SI section 5 for details). Importantly, we found a good qualitative and quantitative agreement between model and data, with the survival probability decaying smoothly with the starting position (Fig. 2b). The only parameter here was $k_r/k_d \approx 1$, which fits well with short-term live imaging experiments and the idea of cell division promoting rearrangements (13).

To back these simulations with an analytical prediction on stem cell numbers, the details of tissue geometry must be taken into account (with the number of cells per row $i$ needing to be estimated, while the number of rows participating in the competition arising as an emergent property from the 1-dimensional model). A good approximation is based on that fact that the crypt can be abstracted as a hemispherical monolayer with radius $R$ (measured in units of cell diameter) coupled to a cylindrical region (Fig. S1, S3, S4 and section S3 of the SI for details), so that one can get the number of stem cells, $N_s^{2D}$, as:

$$N_s^{2D} \approx 2\pi R^2 \left[1 - \cos\left\{\frac{1}{R}\left(1 + \sqrt{\frac{\pi}{2}\frac{k_r}{k_d}}\right)\right\}\right] \quad . \qquad [6]$$

With $k_r/k_d \approx 1$ as above, and estimating $R \approx 2$ for the radius, our simple theory then predicts that the number of functional stem cells should be $N^{2D} \approx 11$, which agrees well with measurements of (13), as well as inferred numbers from continuous clonal labelling experiments (31). This is expected, as our model reduces to the 1-dimensional ring model of Ref. (16) for low $k_r/k_d$.

We then sought to test the model further using a published dataset on embryonic kidney branching in explants (25). This has been recently noted to be a highly stochastic process, with neighbouring cells at the start of the tracing ending up either surviving long-term in tips or being expelled to ducts. Moreover, Ref. (25) observed extensive random cell intercalations, in addition to the previously described mitotic dispersal (26), where cells extrude from the epithelium post-division and reinsert at a distance of $d_c$ cell diameters away. Importantly, these processes can still be captured as an effective diffusion coefficient $k_r$ in our framework (see section S4 of the SI for details). Specifically, knowing that the fluctuations may occur at each duplication, and that they imply a displacement up to $d_c \approx 2-4$ cell lengths, we can estimate that $k_r/k_d \approx d_c^2$ at the minimum (i.e. discounting other fluctuations). Note that the conveyor belt dynamics applies exactly for tip elongation as in crypt: The only difference is that the reference frame

from which the dynamics is observed changes (see section S1 and figure S1 of the SI for details).

The above observation argues again that noise will play a key role in kidney tip cell dynamics. Strikingly, extracting from Ref. (25) the probability of survival as a function of distance from the edge of a tip, we found that the 2-dimensional simulations of our model provided again an excellent prediction for the full probability distribution (Fig. 3a,b), with cells much further away (compared to the intestinal crypt) having a non-negligible probability to go back and contribute. Again, the only fit parameter was the ratio $k_r/k_d = 16$, which agrees well with our estimate of the noise arising from mitotic dispersal. Taking into account the full 2-dimensional geometry as above, and estimating in this case a tip radius of $R = 3-5$ cells, this predicts $N_s \approx 90 \pm 10$, which could be tested in clonal lineage tracing experiments.

These two examples show that the same model of SCB dynamics and its prediction of the master curve for the survival probability of clones can be used in different organs to understand their stem cell dynamics, and shows that ratios of relocation to advection $k_r/k_d$ can be widely different even in systems with similar division rates $k_d$.

**Number of functional stem cells in the developing mammary gland.** Next, we sought to test the suitability of the SCB dynamics to model stem cell dynamics of mammary gland morphogenesis, where extensive cell movements have been reported within tips via intravital live-imaging (9), with rapid rearrangements occurring on time scales of a few hours (Fig. 4 and Fig. S6A). In this case, however, tips cannot be followed for long-enough for survival probabilities to be directed measured as in Figs. 2 and 3 for intestine and kidney, respectively. However, extensive clonal dispersion has been observed in quantitative clonal lineage tracing experiments during pubertal growth (9, 32), and we therefore sought to infer the value of noise from these experiments (Fig. S6B)

Turning back to published lineage-tracing datasets, where single mammary stem cells are labelled at the beginning of puberty (3 weeks of age) and traced until either 5 weeks or 8 weeks of age, clones in tips displayed extensive fragmentation, which is expected to be directly related to the ratio $k_r/k_d$ (Fig. 4c,d and Fig. S6B-D). We thus ran as above 2-dimensional simulations of our SCB dynamics (see section S5 of the SI for details), using measured values of the tip width and length to set the geometry. As a metric for clonal dispersion, we then computationally measured for each labelled cell the distance to its closest clonal neighbour: for a fully cohesive clone, all cells should be touching and the distance to the closest neighbour should be always one cell diameter. Increasing the value of $k_r/k_d$ robustly increased the closest neighbour distance. We then performed the same measurements in the experimental data set, both for the 5 weeks and 8 weeks time points (Fig. 4c,d), and also for luminal and basal cell types separately, given the dominant unipotency of these cell populations in pubertal development (9, 32–34). We found highly consistent results in all four cases (average closest distance of around 1.85 cell diameter) which allowed us to infer a ratio of (see section S5 and figure S6 of the SI for details):

$$k_r/k_d \approx 2-5 \quad , \qquad [7]$$

in mammary gland, emphasizing the importance of considering stochasticity in the conveyor belt picture. Indeed, we found

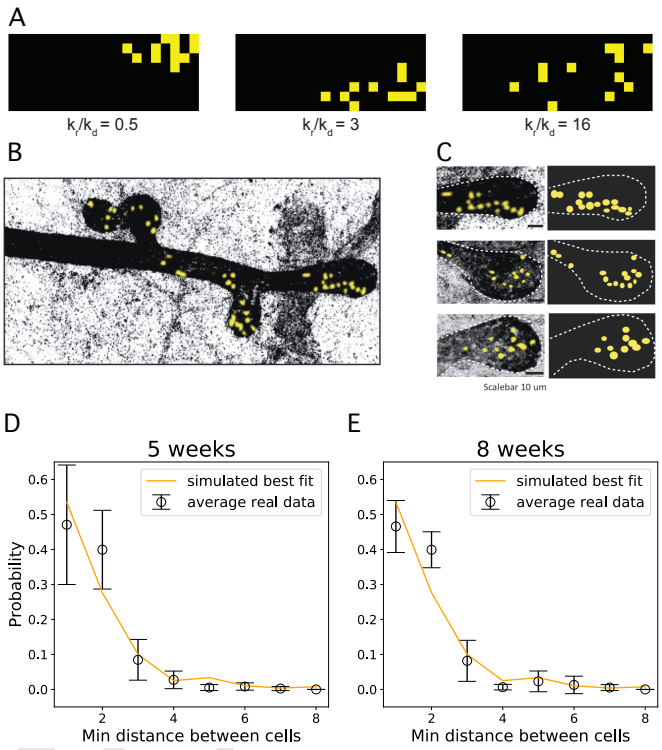


**Fig. 4.** A/ Inferring the relation $k_r/k_d$ from the clone dispersion using a simulation of the stochastic conveyor belt dynamics in 2 dimensions. The distribution of distances of the closest neighbours is highly sensitive to the relation $k_r/k_d$. Here we show numerical simulations of fragmentation under increasing (left to right) values of $k_r/k_d$. B/ Growing tips of a developing mammary gland together with sparse lineage-tracing experiments, where a single lineage (yellow here, induced in 3w-old animals) can be observed. Clonal dispersion due to random cell rearrangements is observed. C/ Close-up of three different mammary tips (left) and corresponding reconstructions to extract relative cellular positions (right). The geometry of the end buds can be approximated by a hemispherical structure connected to a cylindrical one whose radius can be inferred to be around $2-5$ cell diameters. D-E/ Probability distributions of nearest distances between clonally-related cells in tips (resp. from 5-week and 8-week old mice). Black dots represent experimental data (basal and luminal cells have been treated together for this analysis, as they do not show different behaviour at the level of the dynamics). Orange lines are from 2-dimensional numerical simulations of the SCB model (see SI text for details) showing a good fit from $k_r/k_d \approx 3$ for both time points. Error bars represent mean and SD.

that, with this fitting parameter, the model reproduced well the probability distribution of closest distances, both at the 5 weeks and 8 weeks time points (Fig. 4c,d).

In addition to this value, we must again pay attention to the geometry of the mammary tip, with basal cells forming a 2-dimensional monolayer (similar to the previous cases) while luminal cells form multiple layers in 3 dimensions within the tip. Assuming that the intercalation between cells occurs mainly at the same layer, the system of luminal cells in the tip of the mammary gland can be abstracted as $R-1$ successive hemispherical 2-dimensional layers. Let us emphasize the dependence of $N^{2D}$, as defined by equation (6), on $R$, writing $N_s^{2D} \equiv N_s^{2D}(R)$. In that case, the amount of luminal stem cells can be inferred as:

$$N_s^{3D} = \sum_{R'<R} N_s^{2D}(R') \quad . \qquad [8]$$

Taking the fitted range of $k_r/k_d \in (2,5)$, together with an estimation of the radius of $R = 5 \pm 2$, Eq. 8 then predicts that

a number of luminal stem cells per tip of $N_s^{3D} = 170 \pm 110$, in good quantitative agreement with experimental estimates from lineage tracing of $N_s^{\text{exp}} = 172 \pm 102$ (mean±s.d.) (9). For basal cells, using the same parameters for a 2-dimensional monolayer, Eq. 6 predicts that $N_s^{2D} = 37 \pm 11$, against empirical observations reporting an amount of basal stem cells of at least 15 (32), and $N_s^{\text{exp}} = 93 \pm 76$ (mean±s.d.) (9). Although the prediction thus falls in the correct range, the under-estimation of basal stem cell number may be due to the highly anisotropic geometry of basal stem cells.

## Discussion

The main objective of this study was to provide new insights to the question of whether stem cell function is a cell-intrinsic, inherited, property, or rather an extrinsic, context-dependent notion emerging from the collective dynamics of a tissue (2, 3, 7, 8). To that end, we took a complementary standpoint to the one based on the classification of molecular markers and their potential functional role, adopting a purely dynamical/geometrical descriptions of niches. Combining the two would be a logical extension for future work. We analyzed stem cell lineage survival as a purely dynamical process of competition for finite niche space, taking into account the presence of stochastic cell rearrangements, cell proliferation and tissue geometry. This gives rise to a complex reaction-diffusion process that can be abstracted as a "stochastic-conveyor belt". We show that survival probability as a function of starting position away from the most favourable position adopts a simple universal Gaussian shape, so that a well-defined number of functional stem cells (i.e. cells which have a non-negligible probability of surviving long-term) arises in the theory, set by tissue geometry and the ratio between random reallocation and cell proliferation rates, $k_r/k_d$. We applied this theory to recent live-imaging data tracing stem cell survival as a function of position in the homeostatic intestinal crypt and kidney morphogenesis, and find good quantitative agreement. We also use the model to infer values of $k_r/k_d$ from fixed lineage tracing experiments in mammary gland morphogenesis, and show that this inference allows us to predict the typical number of stem cells in this system. Interestingly, the ratio of noise to advection $k_r/k_d$ appeared to be an order of magnitude larger in kidney development as compared to the intestinal crypt (with mammary gland being intermediate), which explained well the widely different number of functional stem cells observed in each.

Although we have sketched here the simplest source of noise in cellular movements (random exchange of position in cell neighbors), our results are highly robust to different types of microscopic mechanisms, and should thus be seen as representative of a general class of models for stem cell dynamics with advection and noise. In mammary gland and kidney morphogenesis, direct cell-cell rearrangements are observed (9, 25), while kidney also displays mitotic dispersal (26), where noise arises from the randomness of cell re-insertion in the layer after division. Furthermore, on short-time scales, directed cellular movements have been observed in kidney tip morphogenesis, with Ret and Etv4 mutant clones being statistically overtaken by wild-type cells, leading to the proposal that Ret/Etv4 were involved in directional movement towards tips (25). However, tips maintain heterogeneity in Ret expression through branching, arguing that cells must shuttle between high-Ret and low-Ret states (25), which would effectively contribute to movement stochasticity on long time scales. Finally, "tectonic" movements, which collectively reposition cells towards/away from niches, can also be captured in the model (Fig. S5). These are particularly relevant in developmental settings, such as gut morphogenesis, where the global shape of the epithelium changes, displacing collectively cells from villus to crypt regions (22), or upon tip-splitting during branching morphogenesis (9). Active migration, as observed in adult intestinal homeostasis (35) could also contribute to such collective random repositioning events. In the future, it would be interesting to further understand quantitatively random cell re-arrangements $k_r$, and how they could be modulated by parameters such as tissue density, aspect ratio, active cell migration or division rates (see section S3C of the SI for details). Mechanical models of cell motility upon rheological transitions (36–38), or of re-arrangements and junctional remodelling upon cell divisions (39, 40) in densely packed tissues could also help to understand quantitatively what sets $k_r$ in each system.

The proposed framework can, in principle, be applied to any tissue dynamics in which niche signals and/or cellular proliferation is localized, leading to directional flows (41, 42). On the other hand, substantial extension of the model would be necessary in the context of an "open niche" such as spermatogenesis (43) or skin homeostasis (11), where renewing cells form a 2-dimensional layer of neutrally competing progenitors, thus with little in-plane cellular flows. Finally, the theory could be extended to cases of non-neutral growth. Live-imaging of skin tumor growth for instance is consistent with very low values of $k_r/k_d$ (44), as little to none clonal dispersion is observed, which would tend to favor deterministic growth in our model. Nevertheless, this does not occur as tumor cells trigger higher proliferation rates of normal cells (44), resulting in complex geometrical changes and encapsulation of the malignant clone. Incorporating these types of complex signalling and geometric feedbacks between multiple cell populations (45–47) in our model would thus have particular relevance to understand the dynamics of tumor initiation (48, 49). Our approach must be taken as part of a more general enterprise, namely understanding the role of both intrinsic cues, and complex collective dynamics, in defining the functional stem cells.

## Materials and Methods

Additional information on the theoretical, computational and experimental methods used can be found in the SI Materials and Methods. In brief, all mice for mammary gland experiments were females from a mixed background, housed under standard laboratory conditions. All experiments were performed in accordance with the Animal Welfare Committee of the Royal Netherlands Academy of Arts and Sciences, The Netherlands. Clonal dispersion in the developing mammary tips was measured in whole mount glands from R26-CreERt2;R26-Confetti mice, traced from 3 weeks of age and sacrificed at mid-puberty (5 weeks) or at the end of puberty (8 weeks). The length and the width of the tips were measured, and the coordinates of each labelled confetti cell in the tip were determined, to calculate the distance between each cell and their closest neighbor. Raw data used to generate Fig. 4 can be found in Supplementary Data 1.

## Acknowledgements

We thank all members of the Hannezo, Simons and Van Rheenen groups for stimulating discussions. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement number No 648804 to J.v.R. and 851288 to E.H.). It has also received funding from the CancerGenomics.nl (Netherlands Organisation for Scientific Research) program (to J.v.R.), the Doctor Josef Steiner Foundation (to J.v.R). B.D.S acknowledges funding from the Royal Society E.P. Abraham Research Professorship (RP/R1/180165) and the Wellcome Trust (098357/Z/12/Z).

1. Wang J, Xu L, Wang E, Huang S (2010) The potential landscape of genetic circuits imposes the arrow of time in stem cell differentiation. *Biophysical journal* 99(1):29–39.
2. Simons, B. D. ans Clevers H (2011) Strategies for homeostatic stem cell self-renewal in adult tissues. *Cell* 145:851–862.
3. Blanpain C, Simons BD (2013) Unravelling stem cell dynamics by lineage tracing. *Nature reviews Molecular cell biology* 14(8):489.
4. Watt FM, Huck WT (2013) Role of the extracellular matrix in regulating stem cell fate. *Nature reviews Molecular cell biology* 14(8):467.
5. Blanpain C, Fuchs E (2014) Plasticity of epithelial stem cells in tissue regeneration. *Science* 344(6189):1242281.
6. Plaks V, Kong N, Werb Z (2015) The cancer stem cell niche: how essential is the niche in regulating stemness of tumor cells? *Cell stem cell* 16(3):225–238.
7. Moris N, Pina C, Arias AM (2016) Transition states and cell fate decisions in epigenetic landscapes. *Nature Reviews Genetics* 17(11):693.
8. Yang H, Adam RC, Ge Y, Hua ZL, Fuchs E (2017) Epithelial-mesenchymal micro-niches govern stem cell lineage choices. *Cell* 169(3):483–496.
9. Scheele C, et al. (2017) Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* 542(7641):313–317.
10. Rompolas P, Mesa KR, Greco V (2013) Spatial organization within a niche as a determinant of stem-cell fate. *Nature* 502(7472):513.
11. Rompolas P, et al. (2016) Spatiotemporal coordination of stem cell commitment during epidermal homeostasis. *Science* 352(6292):1471–1474.
12. Kitadate Y, et al. (2019) Competition for mitogens regulates spermatogenic stem cell homeostasis in an open niche. *Cell Stem Cell* 24(1):79–92.
13. Ritsma, L. ea (2014) Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging. *Nature* 507:362–365.
14. Barker Nea (2007) Identification of stem cells in small intestine and colon by marker gene lgr5. *Nature* 449:1003–1007.
15. Snippert, H. J. ea (2010) Intestinal crypt homeostasis results from neutral competition between symmetrically dividing lgr5 stem cells. *Cell* 143:134–144.
16. López-García C, Klein AM, Simons BD, Winton DJ (2010) Intestinal stem cell replacement follows a pattern of neutral drift. *Science* 330:822–825.
17. Snippert HJ, Clevers H (2011) Tracking adult stem cells. *EMBO Rep* 12:113–122.
18. Klein AM, Simons BD (2011) Universal patterns of stem cell fate in cycling adult tissues. *Development* 138(15):3103–3111.
19. Hannezo E, et al. (2017) A unifying theory of branching morphogenesis. *Cell* 171(1):242–255.
20. Visvader JE, Stingl J (2014) Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev.* 28:1143–1158.
21. van Es, J. H. ea (2012) Dll1+ secretory progenitor cells revert to stem cells upon crypt damage. *Nature Cell Biol.* 14:1099–1104.
22. Guiu J, et al. (2019) Tracing the origin of adult intestinal stem cells. *Nature* 570(7759):107.
23. Krieger T, Simons BD (2015) Dynamic stem cell heterogeneity. *Development* 142(8):1396–1406.
24. Post Y, Clevers H (2019) Defining adult stem cell function at its simplest: The ability to replace lost cells through mitosis. *Cell Stem Cell* 25(2):174–183.
25. Riccio P. ea (2016) Ret and etv4 promote directed movements of progenitor cells during renal branching morphogenesis. *PLoS Biol* 14(2):e1002382.
26. Packard A, et al. (2013) Luminal mitosis drives epithelial cell dispersal within the branching ureteric bud. *Developmental cell* 27(3):319–330.
27. Britton N (1986) *Reaction-Diffusion Equations and Their Applications to Biology*. (Academic Press).
28. Grindord P (1996) *The Theory and Applications of Reaction-Diffusion Equations*. (Oxford University Press).
29. Gardiner CW (1983) *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*. (Springer-Verlag:Berlin).
30. Van Kampen NG (2007) *Stochastic Processes in Physics and Chemistry*. (North Holland), 3 edition.
31. Kozar S, et al. (2013) Continuous clonal labeling reveals small numbers of functional stem cells in intestinal crypts and adenomas. *Cell stem cell* 13(5):626–633.
32. Davis FM, et al. (2016) Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nature communications* 7:13053.
33. Lilja AM, et al. (2018) Clonal analysis of notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. *Nature cell biology* 20(6):677–687.
34. Wuidart A, et al. (2018) Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nature cell biology* 20(6):666–676.
35. Krndija D, et al. (2019) Active cell migration is critical for steady-state epithelial turnover in the gut. *Science* 365(6454):705–710.
36. Bi D, Lopez JH, Schwarz JM, Manning ML (2014) Energy barriers and cell migration in densely packed tissues. *Soft Matter* 10(12):1885–1890.
37. Garcia S, et al. (2015) Physics of active jamming during collective cellular motion in a monolayer. *Proceedings of the National Academy of Sciences* 112(50):15314–15319.
38. Popovic M, Druelle V, Dye NA, Jülicher F, Wyart M (2020) Inferring the plasticity of epithelial tissues from their geometry. *ArXiv ePrint* (arXiv:2002.05133 [physics.bio-ph]).
39. Firmino J, Rocancourt D, Saadaoui M, Moreau C, Gros J (2016) Cell division drives epithelial cell rearrangements during gastrulation in chick. *Developmental cell* 36(3):249–261.
40. Pinheiro D, et al. (2017) Transmission of cytokinesis forces via e-cadherin dilution and actomyosin flows. *Nature* 545(7652):103–107.
41. Moad M, et al. (2017) Multipotent basal stem cells, maintained in localized proximal niches, support directed longranging epithelial flows in human prostates. *Cell Reports* 20:1609–1622.
42. Han S, et al. (2019) Defining the identity and dynamics of adult gastric isthmus stem cells. *Cell Stem Cell* 25(3):342–356.e7.
43. Jörg DJ, Kitadate Y, Yoshida S, Simons BD (2019) Competition for stem cell fate determinants as a mechanism for tissue homeostasis. *ArXiv e-print: arXiv:1901.03903v2 [q-bio.CB]*.
44. Brown S, Pineda C, Xin T, et al. (2017) Correction of aberrant growth preserves tissue homeostasis. *Nature* 548:334–337.
45. Zhou X, et al. (2018) Circuit design features of a stable two-cell system. *Cell* 172(4):744–757.
46. Nabhan AN, Brownfield DG, Harbury PB, Krasnow MA, Desai TJ (2018) Single-cell wnt signaling niches maintain stemness of alveolar type 2 cells. *Science* 359(6380):1118–1123.
47. Hannezo E, Heisenberg CP (2019) Mechanochemical feedback loops in development and disease. *Cell* 178(1):12–25.
48. Sánchez-Danés A, et al. (2016) Defining the clonal dynamics leading to mouse skin tumour initiation. *Nature* 536(7616):298.
49. Muzzopappa M, Murcia L, Milán M (2017) Feedback amplification loop drives malignant growth in epithelial tissues. *Proceedings of the National Academy of Sciences* 114(35):E7291–E7300.