HHS Public Access

Author manuscript

Nat Phys. Author manuscript; available in PMC 2022 May 01.

Published in final edited form as:

Nat Phys. 2021 December ; 17(12): 1391–1395. doi:10.1038/s41567-021-01383-0.

Clonal dominance in excitable cell networks

Jasmin Imran Alsous[1,7], **Jan Rozman**[2,3,7], **Robert A. Marmion**[4], **Andrej Košmrlj**[5], **Stanislav Y. Shvartsman**[1,4,6,*]

[1]Flatiron Institute, Simons Foundation, New York, NY 10010, USA

[2]Jožef Stefan Institute, Ljubljana 1000, Slovenia

[3]Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana 1000, Slovenia

[4]The Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ 08544, USA

[5]Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ 08544, USA

[6]Department of Molecular Biology, Princeton University, Princeton, NJ 08544, USA

[7]These authors contributed equally

Abstract

Clonal dominance arises when the descendants (clones) of one or a few founder cells contribute disproportionally to the final structure during collective growth [1–8]. In contexts such as bacterial growth, tumorigenesis, and stem cell reprogramming [2–4], this phenomenon is often attributed to pre-existing propensities for dominance, while in stem cell homeostasis, neutral drift dynamics are invoked [5,6]. The mechanistic origin of clonal dominance during development, where it is increasingly documented [1,6–8], is less understood. Here, we investigate this phenomenon in the *Drosophila melanogaster* follicle epithelium, a system in which the joint growth dynamics of cell lineage trees can be reconstructed. We demonstrate that clonal dominance can emerge spontaneously, in the absence of pre-existing biases, as a collective property of evolving excitable networks through coupling of divisions among connected cells. Similar mechanisms have been identified in forest fires and evolving opinion networks [9–11]; we show that the spatial coupling of excitable units explains a critical feature of the development of the organism, with implications for tissue organization and dynamics [1,12,13].

*Corresponding author (stas@princeton.edu).

Author contributions

J.I.A., S.Y.S., and A.K. designed the research. J.I.A. performed the experiments and analyzed the data. S.Y.S. and A.K. developed the theory. J.R. implemented and analyzed all theoretical models and simulations. R.A.M. designed the flies. J.I.A., J.R., and S.Y.S. wrote the manuscript with input from all co-authors.

Competing interests statement

The authors declare no competing interests.

Code and Data availability

All custom codes and experimental data for cluster sizes in egg chambers are available at: https://github.com/rozmanj/CD_ExcitableNetworks.

Drosophila oocytes develop within egg chambers, which comprise a germline cyst enveloped by a single-layered follicle epithelium that originates from two follicle stem cells that reside in the stem cell niche [14,15]. In the youngest egg chamber (‘stage 1’), the epithelium contains ~50 follicle cells; however, as the germline expands, the follicle cells increase in number by ~20-fold, reaching ~600-1,000 cells in ‘stage 6’ egg chambers, while undergoing minimal cell death or cell rearrangements (Fig. 1a, b) [14–17]. Thereafter, the follicle cells switch to endoreplication and become ‘post-mitotic’ [17]. Excluding the pairs of polar cells, follicle cell divisions almost all occur with incomplete cytokinesis, resulting in daughter cells that remain connected through stabilized intercellular bridges called ring canals (Fig. 1b, c; Extended Data Fig. 1) [18–20]. Ring canals therefore encode the history of cell divisions, forming lineage trees of connected cells, henceforth clusters, which tile the epithelium like a jigsaw puzzle (Fig. 1c). By tracing ring canal connections [18–20], one can reconstruct the distribution of cluster sizes and their joint growth dynamics [21].

Using images of egg chambers with fluorescently labeled ring canals and cell membranes, we analyzed cluster size distributions in entire epithelia during follicle cell proliferation (Extended Data Fig. 1; Supplementary Table 1; Supplementary Video 1; Methods). Although the epithelium originates from two follicle stem cells, each of which contributes about half of the follicle cells in ‘stage 1’ egg chambers [22], with a 53%+/−28% reported contribution to the final tissue [15], the epithelium of ‘stage 1’ chambers is tiled by around a dozen clusters of 2-8 connected cells and ~10 unconnected cells, rather than just two clusters of interconnected cells (Extended Data Fig. 1; Supplementary Table 2). This finding is consistent with the fact that in the stem cell niche, follicle cell divisions are more likely to be complete, or ring canals are more likely to break as follicle cells envelop the germline [18].

Starting from these initial conditions, we found that as the epithelium proliferates, cluster sizes diverge, and one or two dominant clusters emerge. In post-mitotic egg chambers, the largest cluster contains ~30-40% of all cells in the tissue, with no obvious pattern or orientation with respect to egg chamber axes (Fig. 2a, b; Supplementary Table 3), consistent with findings from transgene expression experiments [19,20] and with the fact that follicle cells covering the anterior or posterior of the germline cyst are equally likely to arise from either stem cell [22]. This observed fractional coverage of dominant clusters is reminiscent of the dominant clones that emerge during zebrafish morphogenesis that cover ~30-60% of the ventricular surface [1]. Note that while ~3% of follicle cell divisions are complete and fragment a cluster into two (Extended Data Fig. 1; Methods), such occurrences can only decrease the fractional coverage of the largest clusters. As such, the experimental data constitute a lower bound on how large the starting clusters that become dominant can grow. Furthermore, as the epithelium starts with approximately two dozen clusters and singles, the emergence of dominant clusters of the sizes reported here is unlikely to arise through cluster fragmentation alone.

To characterize the joint divergence in the cluster sizes during development, we used the Gini coefficient, a metric commonly used to describe the extent of inequality in wealth distribution within a population; here this metric is used to describe the inequality in cell distribution among the clusters in the epithelium of a given egg chamber. The Gini

coefficient (G) is a normalized sum of unique pairwise differences in wealth (x_i) in a population of n individuals [23]:

$$G = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} |x_i - x_j|}{2n \sum_{i=1}^{n} x_i}. \quad (1)$$

As such, increasing values of G indicate increasing inequality, with $G = 0$ indicating equal sharing of resources among constituents, and $G = 1$ representing the case where one entity in the population harbors all resources. In the biological system, x_i describes the number of cells in the i^{th} cluster and n describes the number of clusters. We found that while the distribution of cluster sizes in the youngest egg chambers yields a Gini coefficient of ~0.3, the Gini coefficient increases to ~0.8 in post-mitotic egg chambers with ~600-1,000 cells (Fig. 2c). Qualitatively similar results were obtained using other diversity measures (Extended Data Fig. 2; Methods) [24].

To rationalize our findings, we analyzed several theoretical models of collective cell growth. For all models presented below, the following applied: The cluster size distributions in ‘stage 1’ egg chambers were used as initial conditions (Supplementary Table 2; Methods), and simulations terminated when the total number of cells reached 1,000, thus matching the approximate end of mitosis [17]. All models account for the empirically derived probability, $p_c = 0.026$, that a cell division is complete and does not result in a ring canal (Extended Data Fig. 1; Methods). Since the distribution of ring canals was previously shown to form a network intermediate between linear and maximally branched [18], each ring canal connection between the dividing cell and its linked neighbors was given an equal probability of being assigned to either daughter. All models considered are topological, accounting for the network of connected cells and its fragmentation by complete divisions, but not the spatial constraints of the tissue.

To determine whether the observed extent of cluster divergence can arise due to stochastic effects arising from the multiplicative nature of growth, we implemented a simple ‘Independent model’ of cell divisions in which, starting from experimental initial conditions, a random cell is chosen to divide at each time step, with equal probability for all cells across clusters. As larger clusters contain more cells, they are more likely to host new divisions, thus leading to a divergence in cluster sizes (Gibrat’s law [25]). This formulation evokes preferential attachment models of network growth [26], with a key difference being that it is larger clusters, rather than the more connected nodes, that are likely to receive new vertices. We found that by the time the total number of cells in the simulation reaches 1,000, the largest emergent clusters make up ~13% of the total cells, with average final Gini coefficients of ~0.62 - compared to the significantly larger experimental values of ~36% and ~0.81, respectively (Extended Data Fig. 3; Methods). These results also suggest that fragmentation by complete divisions alone is insufficient to account for the emergence of dominant clusters, and that, as we demonstrate below, to reproduce the observed extent of divergence in cluster sizes and emergence of dominant clusters, for each cell that divides, multiple divisions must occur in the corresponding cluster.

The correlated nature of follicle cell divisions has been described before: Expression of the mitotic marker *Cyclin B* (*CycB*) was shown to be synchronized in small domains of ~5-10 cells [17] (Extended Data Fig. 4; Supplementary Videos 2 and 3), and staining patterns of *Phosphohistone H3* were found to exhibit a significant difference between the experimentally observed values and the theoretically expected number of adjacent mitotic cells if entry into mitosis were random [18]. Such intercellular coordination in cell division could arise from sibling cells jointly entering mitosis because their cell cycles were synchronized through the prior division, namely, through a 'Cell Cycle Timer' scheme [18]. However, such a model does not capture the statistics of cluster sizes and the dynamics of their evolution, yielding average largest final clusters representing ~11% of the tissue, with a maximum final Gini coefficient of ~0.55 - compared to the experimental values of ~36% and ~0.81, respectively (Extended Data Fig. 3; Methods).

Synchronization of divisions also raises the possibility that coordination can arise through intercellular exchange of mitosis promoting factors. During zebrafish cardiogenesis, cells undergoing transient membrane fusions, proposed to mediate exchange of cytoplasmic contents, were highly proliferative [27]; in *Drosophila*, ring canals have a diameter of ~250 nm, and are thus large enough to permit intercellular diffusion of cytoplasmic contents and equilibration of intercellular protein levels [18–20]. We therefore drew on the vast literature on excitable systems and the cell cycle to test whether the emergence of clonal dominance can arise through coupling of mitosis between connected cells [9,28,29]. The cell cycle of eukaryotic cells is commonly divided into a temporal sequence of cell growth (G1), DNA replication (S), a second gap phase (G2), and mitosis (M). Cyclin-dependent kinases (*Cdks*) and their activating cyclins (*Cyc*) induce phosphorylation events that drive the cell in or out of the S or M phases [30]; in *Drosophila,* the G2-to-M transition is regulated by activating *CycB/* or *CycA/Cdk1* complexes, where sufficiently high *CycB* levels can raise *Cdk1* activity above the threshold required for mitosis [31]. The result of mitosis is two daughter cells that must again complete the cell cycle before dividing. The analogy to excitable systems rests on the following similarities: Excitable units (cells) have a unique equilibrium rest state that is stable with respect to small perturbations, however, a perturbation exceeding some threshold (e.g., increase in *CycB* levels) can trigger a large excursion in the state space (mitosis), resulting in the system being first excitable, then refractory, and eventually excitable again. When units are coupled, local excitations can spread and induce the neighbors' transition from a resting to an excited state, thus enabling signal propagation.

To test how well this picture of coupled and excitable cells reproduces the statistics of cluster size distribution and the dynamics of their emergence, we analyzed a simple model of excitable systems that has few parameters and captures key attributes of the biological system: the 'Forest Fire' model [10,32,33]. Here, excitable units correspond to cells, ring canals to edges, and individual clusters to isolated graphs. A cell can be in one of three states: Tree (*T*), during which it can divide but currently is not (G2/M), Fire (*F*) which corresponds to mitosis, or Refractory (R), during which the cell is incapable of dividing following a previous division (G1/S) (Fig. 3a). A unit can transition from T to F at each simulation time step with probability p_1, modeling a spontaneous entry into mitosis. The F state then transitions in the next time step to the refractory state (R); from there, it can again

return to the excitable state (T) in any subsequent time-step, with probability p_r. To represent the ring canal-mediated coupling of divisions between neighbors, the T-to-F transition can also be induced, with probability p_t, by a neighbor in state F. A key modification to the Forest Fire model is the inclusion of cell divisions: When a unit in state F transitions to the refractory state, it now acquires a new and linked neighboring cell that is also in the refractory state, representing the second daughter cell that arises from the division, while a complete division, occurring with probability p_c, results in unconnected daughter cells. The model therefore describes changes in the state of cells in the network and in the structure of the network.

Our results show that the Forest Fire model predicts the robust emergence of experimentally comparable dominant clusters for a significant portion of the parameter space (Fig. 3b and c; Extended Data Fig. 5; Supplementary Video 4), captures the observed dynamics of cluster size divergence (Fig. 3d and e; Extended Data Figs. 2 and 6), and reproduces the (standardized) moments of the cluster size distributions (Extended Data Fig. 7), provided that two conditions are met: Firstly, p_r must be higher than p_i, so that numerous cells in state T are present in the tissue, allowing for the propagation of divisions, which can only be induced in directly linked cells in state T (the size of the largest cluster at a given p_t depends largely on the ratio of p_r and p_i, but not their individual values). Secondly, to match the experimental findings, coupling must be of at least intermediary strength ($p_t \geq 0.65$). If coupling is too high (≥ 0.9), the dominant cluster sizes can exceed what is observed experimentally; it is therefore necessary for p_r to be larger than p_i by only ~1-2 orders of magnitude to permit a sufficiently high fraction of refractory cells that hinder the propagation of divisions (Supplementary Video 5).

For both the Independent and Cell Cycle Timer models, neither of which accounts for coupling between cells, we also analyzed an extended differential growth model, in which clonal dominance could arise due to heritable advantages that predispose some founder cells to progress through the cell cycle more rapidly (Methods) [3,4]. To reproduce our experimental findings, such a growth advantage would have to be significant (e.g., cell divisions occurring at least ~4 times faster), and present in at most a fifth of the starting clusters (Extended Data Figs. 8 and 9). Given the stem cells' equivalent contribution to the epithelium [15,22], and the reported randomness in cell division orientation as well as shape and position of syncytia [19,20], genetically- or positionally-encoded advantages are unlikely to play a primary role in the emergence of dominance.

To conclude, the experimental and theoretical results above demonstrate that clonal dominance can arise spontaneously through coupling of cell division, and that this phenomenon can be interpreted within the framework of spatially distributed excitable dynamical system. Note that various abstractions made in the implementation of the Forest Fire model, such as divisions occurring in one time-step, loss of ring canals occurring only through complete cell divisions (Methods), and coupling existing only between directly connected cells (Supplementary Video 6), render the estimation of model parameters from experiments not directly interpretable. Furthermore, given its topological nature, the model does not account for factors arising from the spatial positions of cells. For example, growth of a dominant cluster could increase local cell density, thus potentially impeding further

proliferation within that cluster [34]. Notably, follicle cells appear to be of uniform size despite their disparate contributions to the final tissue; similar findings were reported in the *Drosophila* imaginal disc, where cells overexpressing *CycD* and *Cdk4* exhibited faster cell divisions and contributed disproportionately more to the tissue but displayed no detectable alterations in cell cycle phasing or cell size [35]. Nonetheless, despite simplifications and its phenomenological character, the Forest Fire model of growing excitable networks is consistent with the correlated nature of divisions and provides quantitative agreement with the statistics of cluster sizes and the dynamics of their evolution.

The results presented here have general implications for dynamics of multicellular development. First, a main outcome of this model is that it provides evidence, at least indirectly, of the excitable character of the cell cycle and gene regulatory processes [28, 36]. Second, dynamic proliferative behaviors that lead to clonal dominance can confer advantages on the developing system. For example, clonal dominance is proposed to play a key role in shaping the vertebrate organ [1] and to drive skin expansion during zebrafish development [37]; in the follicle epithelium, it may facilitate tissue-scale transitions, such as the switch to endoreplication (Extended Data Fig. 1) [17]. Third, our previous work has shown that ring canals connecting germline cells within a convex enclosure result in entropically-constrained tree packing configurations [28]; future work will investigate the role that such topological links play in tiling the surface of a tissue.

Methods

Experiments.

Briefly, *Drosophila melanogaster* flies were raised under standard conditions at 25°C and dissected using an established protocol [38]. Egg chambers were extracted by manual dissection, and images were obtained through imaging of immunostained and/or through genetically fluorescently labeled samples. Experimental data for analysis of cluster size distributions were acquired by confocal microscopy of fixed egg chambers, for which entire epithelia were analyzed. Cluster size distributions for each egg chamber were determined by analyzing z-stacks of entire egg chambers (3D images) in Bitplane's Imaris [39]. FIJI and Adobe Premiere Pro were used for video annotation.

Fly stocks.

Strains used in this study are listed in Supplementary Table 1.

Immunofluorescence and antibodies.

Ovaries from well-fed adult flies were dissected and fixed in 4% paraformaldehyde at room temperature for 20 minutes and stained with the following primary antibodies from the Developmental Studies Hybridoma Bank (DSHB): mouse anti-Hindsight (a27B8 1G9, 1:300), rat anti-E-cadherin (DCAD2, 1:500), and mouse anti-CycB (1:500). The following secondary antibodies were used (1:300 dilution): Alexa-Fluor goat anti-mouse 488nm, goat anti-rat 647nm, and goat anti-mouse 647nm. DAPI (1:500 dilution) was used to label nuclei. Samples were mounted in a 50-50 mixture of Aqua-Poly/Mount (Polyscience) and an optical clearing medium, RapiClear 1.47 (Sunjin Lab Co.).

Protocol for live imaging.

We used a modified version of an established protocol for live imaging of dissected egg chambers [40]. Briefly, ovaries from ~3 flies were dissected in Schneider's media (ThermoFisher #21720-001). Separated ovarioles were transferred to a MatTek dish (MatTek #P35G-1.5-10-C) containing 200 μL of Schneider's media supplemented with Fetal bovine serum (Sigma #F4135), and insulin (Sigma #I 0516). The lid was kept on the dish to prevent drying during imaging.

Microscopy.

Imaging was performed either on a Nikon A1 confocal microscope, using a 100x or 60x/1.3 NA oil objective, or a Zeiss LSM 710 point scanning confocal microscope with a 25x/0.8 or 40x/1.2 Apochromat water objective lens. To image fixed samples, z-stacks (step sizes 350nm or 500nm) were acquired using the 405nm diode laser, 561nm diode-pumped solid-state (DPSS) laser, a 638nm diode laser, and a 488nm argon-gas laser line. Time series were obtained using 1.5-2 μm step sizes when ring canal markers were imaged and 3 μm otherwise; time intervals were set anywhere between 1.5-4 minutes.

Reconstructing cluster sizes.

We used microscopy and image processing of egg chambers with labeled ring canals and cell membranes to reconstruct the sizes of clusters of cells connected through ring canals. Because live-cell imaging of egg chambers is limited to several hours whereas the cells proliferate over ~50 hours, the evolution of cluster size distributions was reconstructed from fixed samples of egg chambers in which entire epithelia were analyzed. We therefore acquired three-dimensional images of egg chambers at various stages of development spanning those ~50 hours, during which epithelial cell number increases by ~20-fold through mitosis that occurs with incomplete cytokinesis ~97% of the time (Fig. 1a–c; Extended Data Fig. 1a, b, f; Supplementary Video 1). The youngest egg chambers analyzed were 'stage 1' egg chambers with ~50 cells, in which the germline cluster is fully enveloped by the epithelium; the oldest egg chambers analyzed were ~'stage 6 and 7' egg chambers, with ~600-1,000 cells, before any significant changes in follicle cell morphology occurred in the latter stages. Since sibling cells remain connected through ring canals that persist during most of egg chamber development [15,18], cells belonging to a cluster were identified by tracing which cells they were connected to through ring canals (Fig. 1b, c). While ring canals appear to lie closer to the apical surface of the epithelium, their position along the cell membrane can vary (Extended Data Fig. 1c), thus necessitating the acquisition of closely spaced z-stacks for reconstructing cluster sizes.

For each egg chamber, Bitplane's Imaris was used to identify individual cells using either the annotation, filaments, or the spots modules (Extended Data Fig. 1d) [39]; the number of cells in each cluster was obtained from the corresponding Statistics tab. Cells unconnected to any others are referred to as singles; these include the pairs of polar cells (Extended Data Fig. 1e) [18]. Note that since the follicle epithelium derives from two somatic stem cells, all cells in the epithelium, regardless of which cluster they belong to, are descendants of one of these two stem cells [15,22].

To reconstruct cell connections from live egg chambers (Supplementary Video 6), we generated triply labeled flies expressing fluorescent membrane (*Resille-GFP*), ring canal (*Pav-RFP*) and nuclear (*PCNA-GFP*) markers. Adjacent cells with a ring canal along a shared membrane at any of the optical z-slices were considered connected and belonging to a single cluster. Live egg chambers could be only partially imaged in the z-direction; as such, the cell lineage tree in Supplementary Video 6 is a partial one.

Probability of ring canal formation.

By reconstructing the number and sizes of clusters across stages from different egg chambers, we found that the number of clusters increases with the total number of cells in the epithelium (Extended Data Fig. 1f). This observation can be explained either by some cell divisions being complete and not resulting in a ring canal, or by ring canals being lost. Assuming the former, the probability of a complete division was determined by fitting a linear function to the plot of the total number of clusters against the total number of cells, as in [18]; the slope of the fitted function gives the probability of a complete division, $p_c = 2.6\% \pm 0.3\%$ - a value somewhat lower than the ~10% reported estimate [18]. This empirically determined probability of complete cytokinesis is accounted for in all theoretical models. Since ‘stage 1’ egg chambers already contain around two dozen clusters and singles combined, fragmentation by itself is unlikely to play a key role in the emergence of dominant clusters.

Measures of cluster size divergence.

In addition to the Gini coefficient, several other diversity indices were used to quantify the extent of cluster size divergence [24]: The Shannon index is given by $Sh = -\sum_{i=1}^{n} \frac{x_i}{N} \log\left(\frac{x_i}{N}\right)$, where n is the total number of clusters, N is the number of cells, and x_i is the number of cells in the i^{th} clusters. The Shannon index has a maximum value of $Sh_{\max} = \log n$ when cells are evenly distributed between all clusters, with less even distributions corresponding to lower values of Sh. From the Shannon index, we obtain the evenness (Shannon’s equitability) $J_E = \frac{Sh}{Sh_{\max}}$, given as the ratio between the Shannon index and the maximum possible Shannon index ($Sh_{\max}$) for a given number of clusters. The Theil index $T = \frac{1}{n}\sum_{i=1}^{n} \frac{x_i}{\bar{x}} \log\left(\frac{x_i}{\bar{x}}\right)$, where $\bar{x}$ is the average number of cells in a cluster, can also easily be shown to equal the difference between the maximum and the actual Shannon index, $T = Sh_{\max} - Sh$. The Simpson’s index (with replacement) $S_r = \sum_{i=1}^{n} \left(\frac{x_i}{N}\right)^2$ gives the probability of two randomly selected cells belonging to the same cluster. The Hoover index $H = \frac{\sum_{i=1}^{n} |x_i - \bar{x}|}{2\sum_{i=1}^{n} x_i}$ gives the fraction of cells that would have to be redistributed to achieve a perfectly even distribution of cells between all cluster. The Berger-Parker index is equivalent to the fractional size of the largest cluster as shown in Fig. 3d (see Extended Data Fig. 2 for a summary of these diversity measures in experiments and simulations).

Simulation implementation.

The simulations were implemented using C++ programming language as discussed below; the data analysis and generation of plots was performed in Wolfram Mathematica.

Individual simulation runs for both the Independent, Cell Cycle Timer, and Forest Fire models were initialized using experimental data taken from the 11 ‘stage 1’ chambers, with initial cluster sizes given in Supplementary Table 2 (data from EC1 is used for Fig. 3b and Supplementary Video 4; the non-dividing nature of polar cells is not included in the simulations). When initializing the simulations, cells are linked in a linear fashion without branching. Simulations were run until the total number of cells reached 1,000, and all three models include the possibility of complete divisions, using the experimentally determined probability $p_c = 0.026$. Divisions are implemented as described in the main text: After a cell divides, each ring canal connecting the mother cell to its linked neighbors is given an equal probability of being assigned to either of the daughters; the daughter cells are themselves linked by a ring canal with probability $1 - p_c$, or remain unliked with probability p_c, thus fragmenting the cluster. When averaging over simulation runs, 200 simulations were run for each of the 11 initial conditions, for a total of 2,200. Final theoretical results for dominant cluster sizes were compared to the experimental data in Supplementary Table 3.

Independent model.

In this model, at each time step, one cell is randomly selected to divide, with equal probability for all cells. Results for the Independent model are shown in Extended Data Fig. 3.

Cell Cycle Timer model.

In this model, the times of the first division for each of the starting cells are taken from a normal distribution with mean t_0 and standard deviation σ_0. At each step of the simulation, the cell that is set to divide soonest is identified and made to divide; the time to the next division of each of the resulting daughter cells is then taken from the same normal distribution as before (each daughter cell has a separate division time). Without loss of generality, the mean division time is set to $t_0 = 9.6$ h, matching a reported experimental value [15], while the standard deviation σ_0 was varied over several orders of magnitude, up to $\sigma_0 \approx t_0$. If the normal distribution yielded a nonpositive division time during the simulation, a new time interval was selected again from the same distribution until a positive value was obtained. Results for the Cell Cycle Timer model are found in Extended Data Fig. 3.

Forest Fire Model.

In this model, all cells started the simulation in state T. At each time-step, each cell that start the time-step in state T can transition to F spontaneously with probability p_i, and additionally with probability p_t for each linked neighbor that starts the time-step in state F, representing the induction of divisions by a neighboring dividing cell. Each cell that starts the step in state R can transition back to T, with probability p_r. Afterwards, all cells that start the step in state F sequentially divide, producing two cells in state R.

The analyzed parameter sets are all combinations of p_t between 0 and 1, in increments of 0.05, p_i values $10^{0.75-i}$, $10^{0.5-i}$, $10^{0.25-i}$, and 10^{-i} for integer values of i between 1 and 5, inclusive, as well as $p_i = 1$, and p_r values $5 \cdot 10^{0.25-i}$, $5 \cdot 10^{-i}$, $5 \cdot 10^{-0.25-i}$, and $5 \cdot 10^{-0.5-i}$, for i between 1 and 5, inclusive, as well as $p_r = 1$. Extended Data Fig. 5 shows the average size of the largest cluster at the end of a simulation as a fraction of the total final cell number, for a large part of the studied parameter space.

Error calculations in the Forest Fire model.

To make a quantitative comparison between theoretical and experimental results for the various measures of diversity $d^{(\alpha)}$ (Berger-Parker, Gini, Shannon, Evenness, Theil, Simpson, and Hoover), we calculate the chi-square coefficient between the average values of diversity indices in simulations and the experimental data points for egg chambers with more than 61 cells (the size of the largest egg chamber used as an initial condition). The $\chi^2(\alpha)$ coefficient of diversity index α is defined as follows:

$$\chi^2_{(\alpha)} = \sum_i \frac{\left(d^{(\alpha)}_{\text{experiment}}(N_i) - d^{(\alpha)}_{\text{theory}}(N_i)\right)^2}{\left(s^{(\alpha)}_{\text{theory}}(N_i)\right)^2}$$

where the index i runs over all experimental egg chambers with more than 61 cells, N_i is the number of cells in the i^{th} experimental egg chamber, $d^{(\alpha)}_{\text{experiment}}(N_i)$ is the value of diversity index α for the i^{th} experimental egg chamber, $d^{(\alpha)}_{\text{theory}}(N_i)$ and $s^{(\alpha)}_{\text{theory}}(N_i)$ are the average value and the standard deviation, respectively, of the diversity index α in the Forest Fire model at N_i cells for a given set of parameters, averaged over 2,200 simulations; α is one of the seven diversity indices used (Extended Data Fig. 2): Berger-Parker index (equivalent to largest cluster fraction), Gini coefficient, Shannon index, Evenness, Theil index, Simpson's index, and Hoover index. To estimate the overall agreement between theory and experiment, we compute the average, $\langle \chi^2 \rangle$, between chi-square coefficients for all seven diversity indices. See Extended Data Figs. 6 for $\langle \chi^2 \rangle$ plots for the studied parameter space; the parameter set used in Fig. 3b, d, and e and in Extended Data Figs. 2 and 7 corresponds to the lowest $\langle \chi^2 \rangle$ at $p_i = 0.0001$, the value used in Fig. 3c (note that a relatively low value of p_i was chosen so that the part of the diagram where $p_r \gg p_i$ can be shown, as p_r cannot exceed 1).

Caveats to point out in the Forest Fire simulations are that first, 2,200 points are not obtained at each N_i, as Forest Fire simulations can and do add multiple cells at a single time step. As a result, averages and standard deviations computed at different parameter sets and cell numbers are not based on an equivalent number of points. Second, if at least two simulation results are not obtained at some N_i, an infinite χ^2 coefficient is assumed (in general, this is only relevant when both p_i and p_r are very high, producing rapid growth that can result in at least one of the experimental N_i values being overshot by all simulation runs). Lastly, as the largest egg chamber used as an initial condition has 61 cells (EC9), the theoretical values in Fig. 3d, e and Extended Data Figs. 2 and 7 are only shown for cell numbers greater than

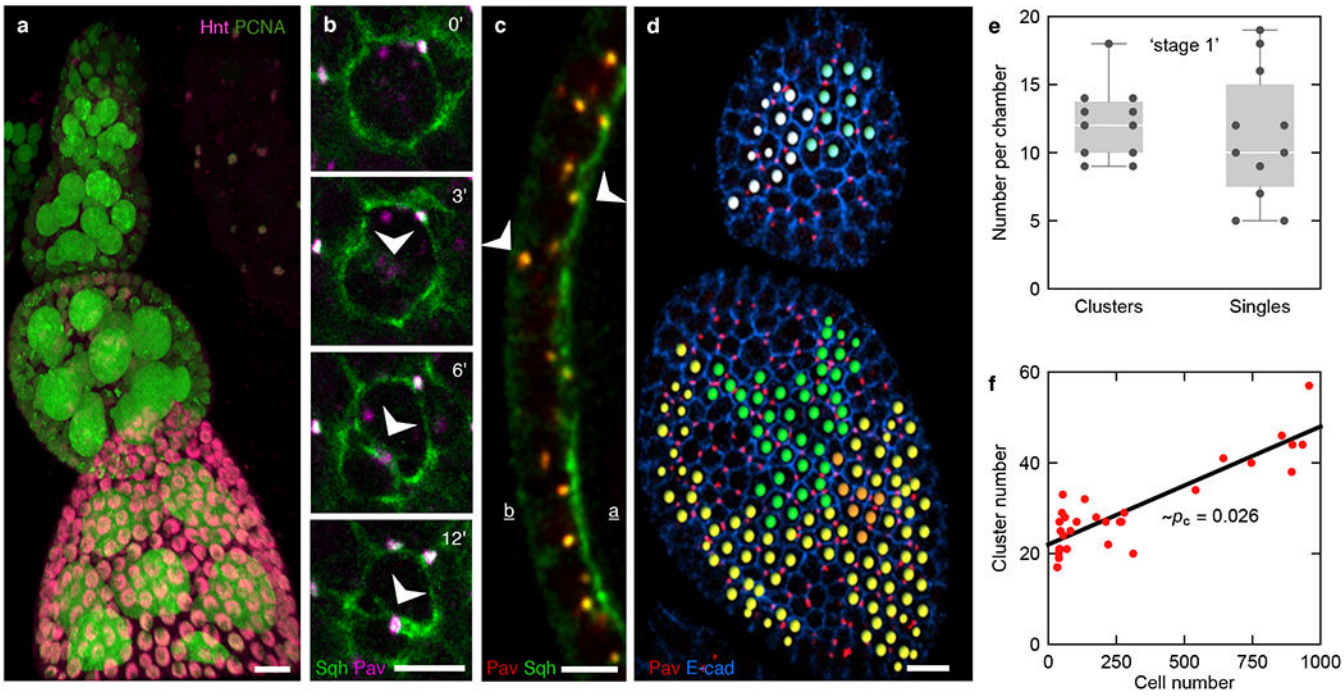
61. Furthermore, the theoretical values are only drawn for up to 1,000 cells, even if some simulations overshoot that number due to multiple cells being added in the final time step.

Differential growth patterns.

To analyze the potential contribution of differential growth patterns due to hereditary biases of some of the founder cells to the emergence of dominant clusters, we extended the Independent and Cell Cycle Timer models to now include a fast- and a slow-dividing population of cells. Each starting cluster was assigned probabilities p_f and $p_s = 1 - p_f$ of containing only fast- or only slow-dividing cells, respectively, and the trait of being fast- or slow-dividing was set to be hereditary. In the extended Independent model, the probabilities of a fast- or slow-dividing cell being chosen to divide were weighted, through w_f and w_s, respectively; in the extended Cell Cycle Timer model, division times for fast-and slow-dividing cells were taken from normal distributions with means and standard deviations (t_f, σ_f) and (t_s, σ_s), respectively.

In both cases, we found that to reproduce the experimental findings, the starting population of fast-dividing cells had to be relatively small, approximately $p_f \leq 20\%$, and the extent of differential growth had to be significant, with either approximately $w_f / w_s \geq 3$ or $t_f / t_s \leq 0.25$ (Extended Data Figs. 8 and 9). Therefore, under the assumptions here, small hereditary biases would not be sufficient to explain the observed extend of clonal dominance. Furthermore, should the origin of clonal dominance in the epithelium arise from hereditary biases, it is unlikely that two populations of cells with different division rates originate from the two stem cells, as the starting fraction of fast-dividing cells must be quite small to reproduce the experimental findings, and the two stem cells have been shown to contribute relatively evenly to ‘stage 1’ egg chambers [22]. Lastly, experimental studies have demonstrated the random configuration and orientation of syncytia on egg chamber’s surface [19,20], suggesting that positional biases are similarly unlikely.

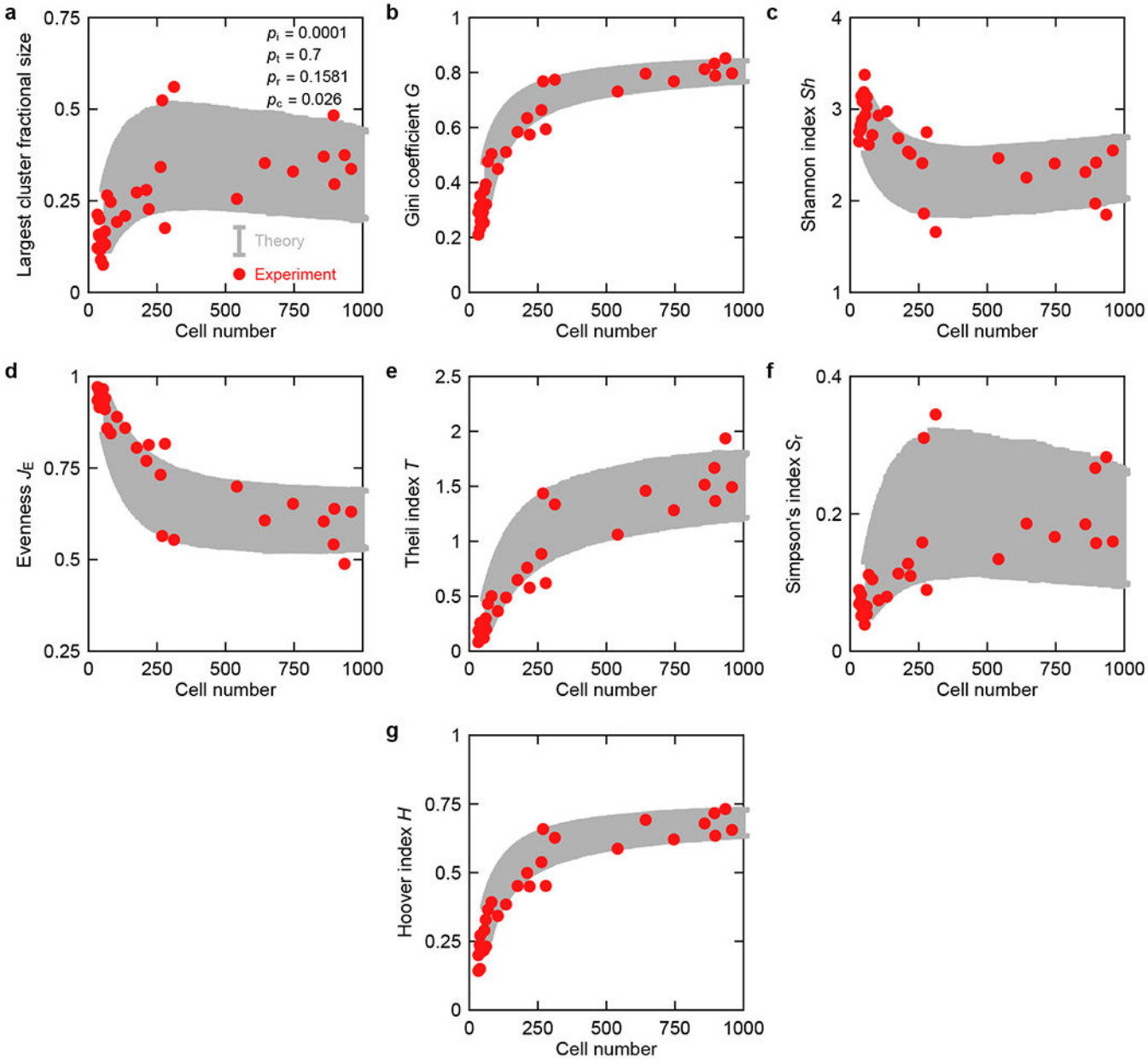
Extended Data



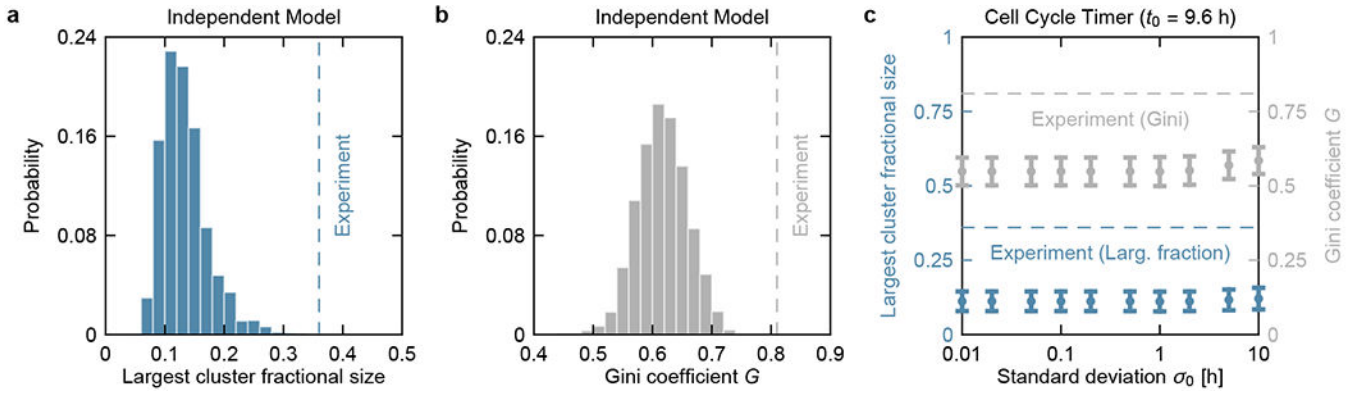
Extended Data Fig. 1. Incomplete cytokinesis leaves daughter cells connected through stable ring canals.

a, Egg chambers expressing the nuclear *Proliferating Cell Nuclear Antigen* (*PCNA*, green) and stained for *Hindsight* (*Hnt*) (magenta), which labels endocycling follicle cells that have exited mitosis [17]. *Drosophila* follicle cells increase in number through mitosis before transitioning to endocycling, during which the follicle cells duplicate their DNA without dividing; follicle cells that have exited the mitotic cell cycle are 'post-mitotic'. The endocycle is thought to start at stage 6, a point at which the egg chambers will have started to visibly elongate [17, 41]. Scale bar = 10 μm. **b**, Formation of a ring canal (arrowhead) in dividing epithelial cells in an egg chamber expressing fluorescently labeled *Spaghetti squash* (*Sqh,* green) and *Pavarotti* (*Pav,* magenta). Time in minutes; scale bar = 5 μm. **c**, Cross-sectional view of an epithelium with fluorescently labeled *Sqh* (green) and *Pav* (red), showing the variable positions (arrowheads) of the ring canals along the membrane (a is apical, b is basal), which necessitates acquisition and analysis of 3D images for identification of intercellular connections and cluster sizes. Scale bar = 5 μm. **d**, Confocal images of the surfaces of two egg chambers expressing fluorescently labeled *Pav* (red) and labeled with anti-*E-cadherin* (*E-cad*, blue) overlaid with reconstructed clusters: Spots of a given color denote cells belonging to a cluster: two (white, blue) and three (green, yellow, and orange) clusters of interconnected cells are shown in the younger and older egg chamber, respectively. Cluster size determination and visualization were performed in Bitplane's Imaris [40]. Scale bar = 10 μm. **e**, Box plots of the number of clusters and singles (cells unconnected to others through ring canals) in 'stage 1' egg chambers (n = 11); the whiskers encapsulate the entire data, the bottom and top sides of the box indicate the first and third quartile of the data, respectively, and the white line shows the median. **f**, Plot of the

number of clusters as a function of total cell number, used to extract the probability that a division does not lead to a ring canal p_c=0.026 ± 0.003. Line shows a linear fit.

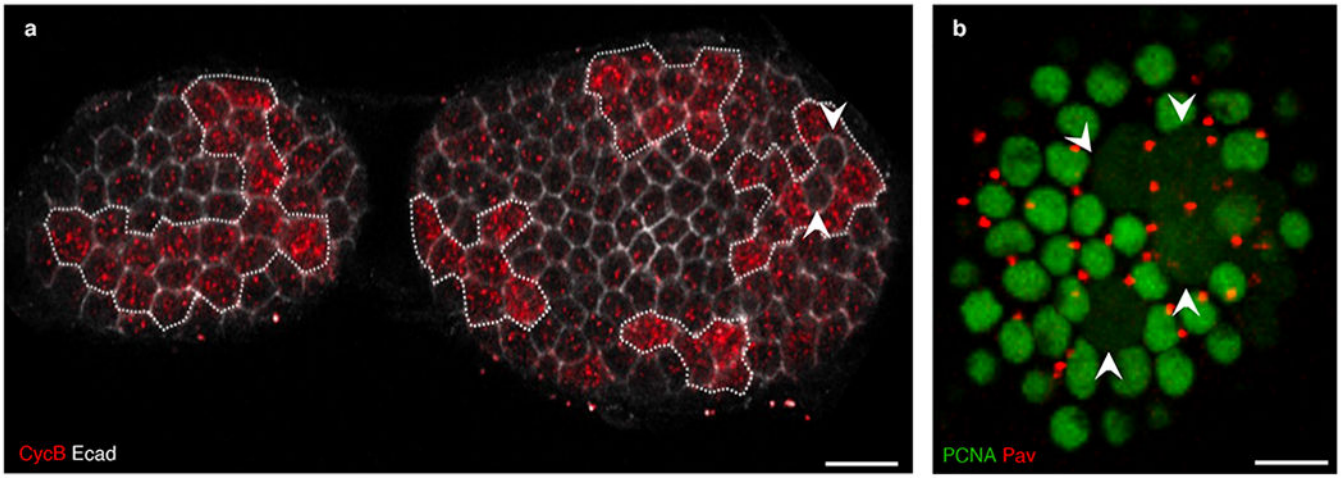


Extended Data Fig. 2. Comparison of diversity measures used to quantify cluster size divergence.
a, Largest cluster fraction (i.e. Berger-Parker index), **b**, Gini coefficient, **c**, Shannon index, **d**, Evenness, **e**, Theil index, **f**, Simpson's index (with replacement), and **g**, Hoover index for experimental egg chambers (red points), along with theoretical predictions given by the Forest Fire model simulations (gray; parameters p_i = 0.0001, p_t = 0.7, p_r = 0.158114, p_c = 0.026) averaged over 2,200 simulation runs. Gray error bars indicate the standard deviation in simulations.



Extended Data Fig. 3. Largest cluster fractions and Gini coefficients as obtained from the Independent and Cell Cycle Timer models of uncoupled cell divisions.

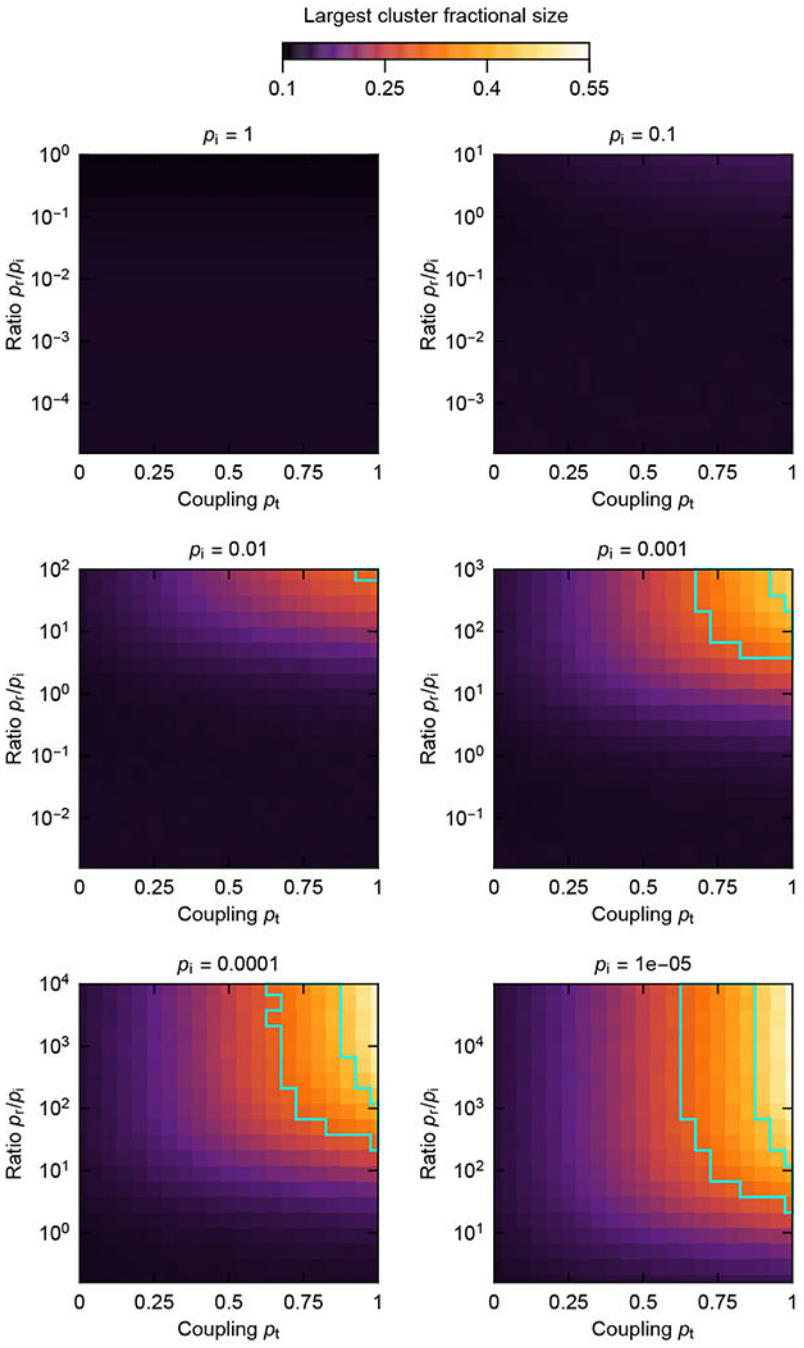
a, Distribution of final largest cluster sizes as fractions of all cells in the Independent model, based on 2,200 simulation runs. Dashed lines indicate the experimental values obtained by averaging over all egg chambers with >600 cells (Largest cluster size fraction = 0.36, s.d. = 0.06; $n = 7$). **b**, Distribution of final Gini coefficients in the Independent model, again based on 2,200 simulation runs. Dashed lines indicate the experimental values obtained by averaging over all egg chambers with >600 cells ($G = 0.81$, s.d. = 0.03; $n = 7$; see Supplementary Table 3 for experimental data in **a** and **b**). **c**, Final largest cluster size as a fraction of all cells (blue) and the Gini coefficient (gray) as obtained from the Cell Cycle Timer model with the mean cell division time $t_0 = 9.6$ h [15], at different values of standard deviation σ_0 of cell division time, averaged over 2,200 simulation runs. Error bars indicate standard deviation in simulations; dashed lines indicate experimental values averaged over all egg chambers with >600, as in **a** and **b**.



Extended Data Fig. 4. The correlated nature of epithelial cell divisions.

a, Epithelial surface of two egg chambers stained for the mitotic marker *Cyclin B* (*CycB*), showing domains (~5-15 cells) of coordinated expression. Arrowheads pointing to two adjacent dividing cells in one of those domains. **b**, Surface of a follicle epithelium expressing fluorescently labeled *Pav* (red), and the nuclear Proliferating Cell Nuclear

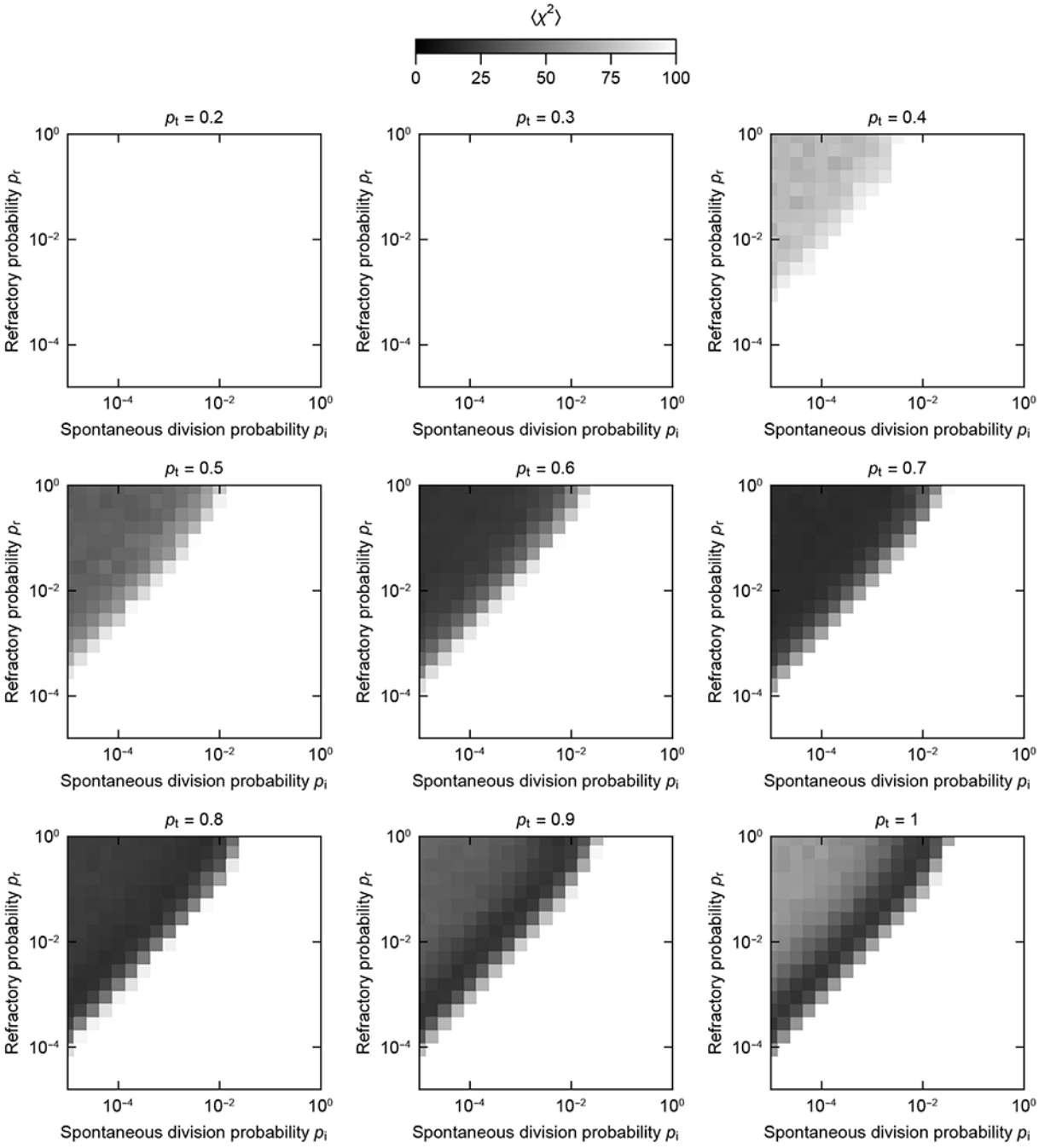
Antigen (*PCNA*); arrowheads point to dividing connected cells. Scale bars in **a** and **b** = 10 μm.



Extended Data Fig. 5. Largest cluster sizes in the Forest Fire model.

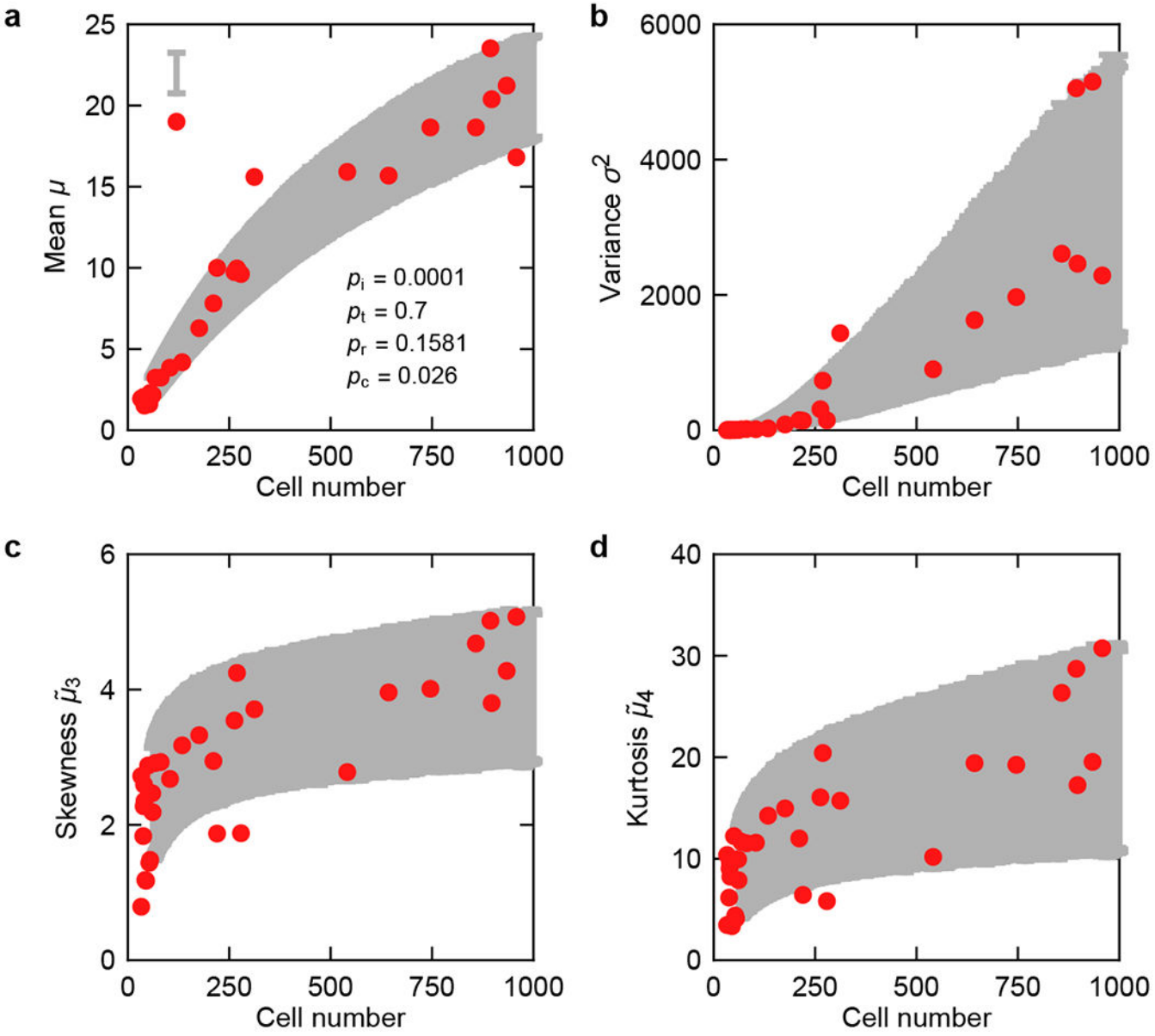
Final size fractions of the largest cluster as a function of p_t and p_r/p_i for different orders of magnitude of p_i and averaged over 2,200 simulations at each parameter set. Regardless of the value of p_i, similar average sizes occur at the same value of p_t and a given p_r/p_i ratio. Cyan-bordered region indicates where experimental and theoretical values are less than an

experimental standard deviation apart (experimental value average over all chambers with >600 cells: largest cluster size fraction = 0.36, s.d. = 0.06, $n = 7$; Supplementary Table 3).



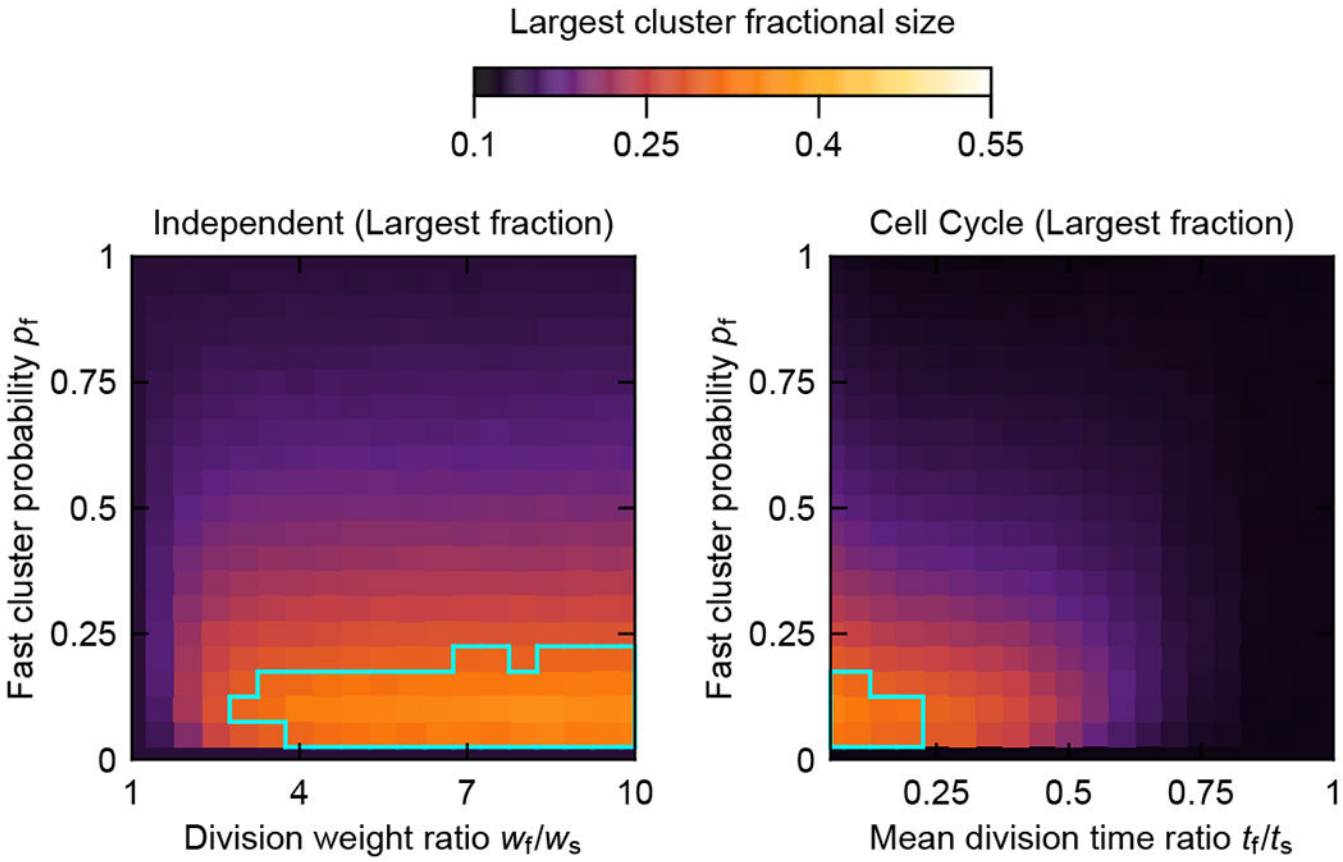
Extended Data Fig. 6. Analysis of the Forest Fire model parameter space

Plots of the average chi-square coefficient ($\langle \chi^2 \rangle$; see Methods) of the diversity indices between the Forest Fire model simulations and the experimental data for all egg chambers with >61 cells, for different values of p_t as a function of p_i and p_r. Average chi-square coefficient values exceeding 100 are shown in white.



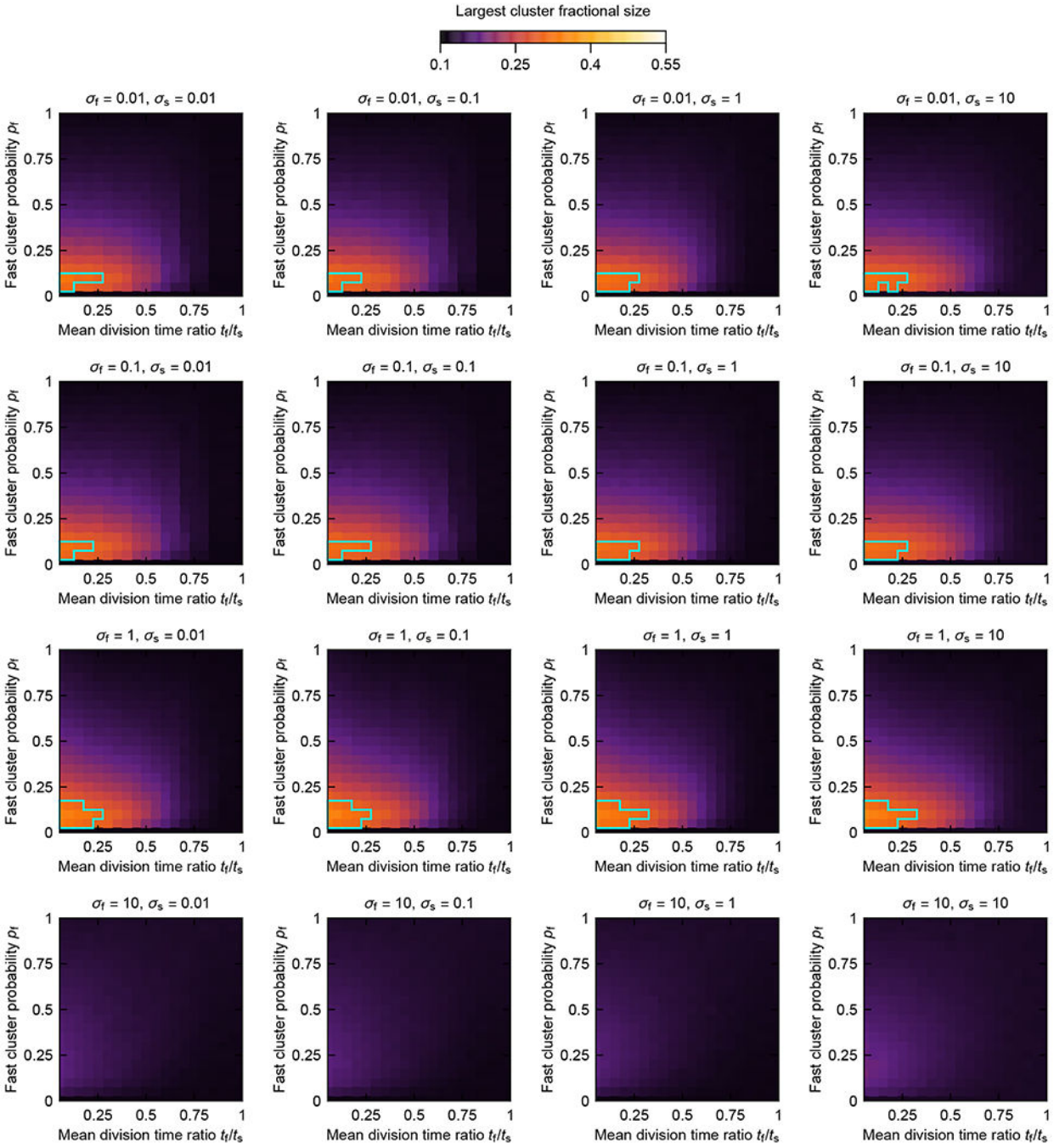
Extended Data Fig. 7. Comparison of the (standardized) moments of the distribution of cluster sizes.

a, Mean [$\mu = \langle x_i \rangle$], **b**, Variance [$\sigma^2 = \langle (x_i - \mu)^2 \rangle$], **c**, Skewness [$\tilde{\mu}_3 = \left\langle (x_i - \mu)^3 / \sigma^3 \right\rangle$], and **d**, Kurtosis [$\tilde{\mu}_4 = \left\langle (x_i - \mu)^4 / \sigma^4 \right\rangle$], where x_i is the number of cells in the i^{th} cluster and $\langle \cdot \rangle$ indicates the average over all clusters in a chamber, for experimental egg chambers (red points) and theoretical predictions given by the Forest Fire model simulations (gray; parameters $p_i = 0.0001$, $p_t = 0.7$, $p_r = 0.158114$, $p_c = 0.026$), averaged over 2,200 simulation runs. Gray error bars indicate standard deviation in simulations.



Extended Data Fig. 8. Average largest cluster sizes obtained from the Independent and the Cell Cycle Timer models with differential growth.

a, Final size of the largest cluster as a fraction of total cell number for the Independent model with differential growth, averaged over 2,200 simulation runs, as a function of the ratio of division probability weights for the fast and slow-growing cell populations, w_f/w_s, and the probability of each starting cluster being assigned to the fast-growing population, p_f. Cyan-bordered region here and in **b** shows where theoretical and experimental values of the final largest cluster fraction are less than an experimental standard deviation apart (experimental value average over all chambers with >600 cells: largest cluster size fraction = 0.36, s.d. = 0.06, $n = 7$; Supplementary Table 3). **b**, Final size of the largest cluster as a fraction of total cell number for the Cell Cycle Timer differential growth model for t_s = 9.6 h and $\sigma_s = \sigma_f$ = 0.5 h, averaged over 2,200 simulation runs, as a function of the mean division time ratio between the fast and slow cell populations, t_f/t_s, and the probability of each starting cluster being assigned to the fast population, p_f.



Extended Data Fig. 9. Analysis of the Cell Cycle Timer differential growth model parameter space.

Final largest cluster size as fraction of total cell number different σ_f and σ_s, for t_s = 9.6 h, averaged over 2,200 simulation runs, as a function of the mean division time ration between the fast and slow cell populations, t_f/t_s, and the probability of each starting cluster being assigned to the fast population, p_f. Cyan-bordered region indicates where experimental and theoretical values are less than an experimental standard deviation apart (experimental value average over all chambers with >600 cells: largest cluster size fraction = 0.36, s.d. = 0.06, n = 7; Supplementary Table 3).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Gary Laevsky for expert help with imaging and Enrico Coen for advice on how to improve the manuscript and broaden its scope. We thank Eric Wieschaus, Trudi Schüpbach, Alexander Berezhkovskii, Matej Krajnc, Sayantan Dutta, and Corina Tarnita for helpful discussions, Daniel Hellwig for assistance with movie making, Jonathan Jackson for manuscript editing, and the Martin Lab (Biology, MIT) for lab stocks and microscopy. This work was supported by the NIH R01GM134204-02 (S.Y.S.), and the Slovenian Research Agency (research core funding no. P1-0055) (J.R.).

References

[1]. Gupta V, & Poss KD Clonally dominant cardiomyocytes direct heart morphogenesis. Nature 484, 479–484 (2012). [PubMed: 22538609]

[2]. Hall BA, Piterman N, Hajnal A, & Fisher J Emergent stem cell homeostasis in the *C. elegans* germline is revealed by hybrid modeling. Biophys. J 109, 428–438 (2015). [PubMed: 26200879]

[3]. Luria SE & Delbrück M Mutations of bacteria from virus sensitivity to virus resistance. Genetics 48, 491–511 (1943).

[4]. Shakiba N et al. Cell competition during reprogramming gives rise to dominant clones. Science 346, 6438, eaan0925 (2019).

[5]. Lopez-Garcia C, Klein AM, Simons BD & Winton DJ Intestinal stem cell replacement follows a pattern of neutral drift. Science 330, 6005, 822–825 (2010). [PubMed: 20929733]

[6]. Snippert HJ et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. Cell 143, 134–144 (2010). [PubMed: 20887898]

[7]. Kamimoto K et al. Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling. eLife 5, e15034 (2016). [PubMed: 27431614]

[8]. Sereti K-I et al. Analysis of cardiomyocyte clonal expansion during mouse heart development and injury. Nat. Commun 9, 754 (2018). [PubMed: 29467410]

[9]. Mikhailov AS Kinematics of Wave Patterns in Excitable Media. In: Holden AV, Markus M, Othmer HG Nonlinear Wave Processes in Excitable Media. NATO ASI Series (Series B: Physics, Volume 244. Springer, Boston, MA, 1991).

[10]. Bak P, Chen K & Tang C A forest-fire model and some thoughts on turbulence. Phys. Lett. A 5-6, 297–300 (1990).

[11]. Pei S, Tang S & Zheng Z Detecting the influence of spreading in social networks with excitable sensor networks. PLoS One 10, e0124848 (2015). [PubMed: 25950181]

[12]. Clayton E et al. A single type of progenitor cell maintains normal epidermis. Nature 446, 185–189 (2007). [PubMed: 17330052]

[13]. Imran Alsous J, Villoutreix P, Stoop N, Shvartsman SY & Dunkel J Entropic effects in cell lineage tree packings. Nat. Phys 14, 1016–1021 (2018). [PubMed: 30881478]

[14]. King RC Ovarian development in Drosophila melanogaster (Academic Press, New York, 1970).

[15]. Margolis J & Spradling AC Identification and behavior of epithelial stem cells in the *Drosophila* ovary. Development 121, 3797–3807 (1995). [PubMed: 8582289]

[16]. King RC & Vanoucek EG Oogenesis in adult *Drosophila melanogaster*. X. Studies on the behavior of the follicle cells. Growth 24, 333–338 (1960). [PubMed: 13830957]

[17]. Deng W-M, Althauser C & Ruohola-Baker H Notch-Delta signaling induces a transition from mitotic cell cycle to endocycles in *Drosophila* follicle cells. Development 128, 4737–4746 (2001). [PubMed: 11731454]

[18]. Airoldi SJ, McLean PF, Shimada Y & Cooley L Intercellular protein movement in syncytial *Drosophila* follicle cells. J. Cell Sci 124, 4077–4086 (2011). [PubMed: 22135360]

[19]. McLean PF & Cooley L Protein equilibration through somatic ring canals in *Drosophila*. Science 340, 1445–1447 (2013). [PubMed: 23704373]

[20]. McLean PF & Cooley L Bridging the divide. Fly 8, 12–18 (2013).

[21]. Buckingham ME & Meilhac SM Tracing cells for tracking cell lineage and clonal behavior. Dev. Cell 21, 394–409 (2011). [PubMed: 21920310]

[22]. Nystul T & Spradling AC Regulation of epithelial stem cell replacement and follicle formation in the *Drosophila* ovary. Genetics 184, 503–515 (2010). [PubMed: 19948890]

[23]. Dorfman R A formula for the Gini coefficient. Rev. Econ. Stat 61, 146–149 (1979).

[24]. Xu S, Böttcher L & Chou T Diversity in biology: definitions, quantification and models. Phys. Biol 17, 031001 (2020). [PubMed: 31899899]

[25]. Sutton J Gibrat's Legacy. Journal of Economic Literature, Volume XXXV, 40–59 (1997).

[26]. Barabasi A-L & Albert R Emergence of scaling in random networks. Science 286, 509–512 (1999). [PubMed: 10521342]

[27]. Sawamiphak S, Kontarakis Z, Filosa A, Reischauer S & Stainier DYR Transient cardiomyocyte fusion regulates cardiac development in zebrafish. Nat. Commun 8, 1525 (2017). [PubMed: 29142194]

[28]. Kromer J, Khaledi-Nasab A, Schimansky-Geier L & Neiman AB (2017). Emergent stochastic oscillations and signal detection in tree networks of excitable elements. Sci. Rep 7, 3956 (2017). [PubMed: 28638071]

[29]. Tyson JJ Modeling the cell division cycle: cdc2 and cyclin interactions. Proc. Natl. Acad. Sci. U.S.A 99, 7328–7332 (1991).

[30]. Hochegger H, Takeda S & Hunt T Cyclin-dependent kinases and cell-cycle transitions: does one fit all? Nature 9, 910–916 (2008).

[31]. Shcherbata HR, Althauser C, Findley SD & Ruohola-Baker H The mitotic-to-endocycle switch in *Drosophila* follicle cells is executed by Notch-dependent regulation of G1/S, G2/M and M/G1 cell-cycle transitions. Development 131, 3169–81 (2004). [PubMed: 15175253]

[32]. Drossel B & Schwabl F. Self-organized critical forest-fire model. Phys. Rev. Lett 69, 1629–1632 (1992). [PubMed: 10046273]

[33]. Kromer J, Khaledi-Nasab A, Schimansky-Geier L & Neiman AB Emergent stochastic oscillations and signal detection in tree networks of excitable elements. Sci. Rep 7, 3956 (2017). [PubMed: 28638071]

[34]. Puliafito A et al. Collective and single cell behavior in epithelial contact inhibition. Proc. Natl. Acad. Sci. U. S. A 109, 739–744 (2012). [PubMed: 22228306]

[35]. Datar SA, Jacobs HW, de la Cruz AFA, Lehner CF & Edgar BA The *Drosophila* Cyclin D–Cdk4 complex promotes cellular growth. EMBO J. 19, 4543–4554 (2000). [PubMed: 10970848]

[36]. Batchelor E, Loewer A, Mock C & Lahav G Stimulus-dependent dynamics of p53 in single cells. Mol. Syst. Biol 7, 488 (2011). [PubMed: 21556066]

[37]. Roan H-Y, Tseng T-Z & Chen C-H Whole-body clonal mapping identifies giant dominant clones in zebrafish skin epidermis. Development doi.10.1242/dev.199669 (2021).

[38]. Wong LC & Schedl P Dissection of Drosophila ovaries. J. Vis. Exp 1, 52 (2006).

[39]. Bitplane Inc. Imaris Version 8.3.1, imarisxt. Zurich, Switzerland.

[40]. Prasad M, Jang ACC, Starz-Gaiano M, Melani M, & Montell DJ A protocol for culturing *Drosophila melanogaster* stage 9 egg chambers for live imaging. Nat. Protoc 2, 2467–2473 (2007). [PubMed: 17947988]

[41]. Jia D, Xu Q, Xie Q, Mio W & Deng W-M Automatic stage identification of *Drosophila* egg chamber based on DAPI images. Sci. Reports 6, 18850 (2016).

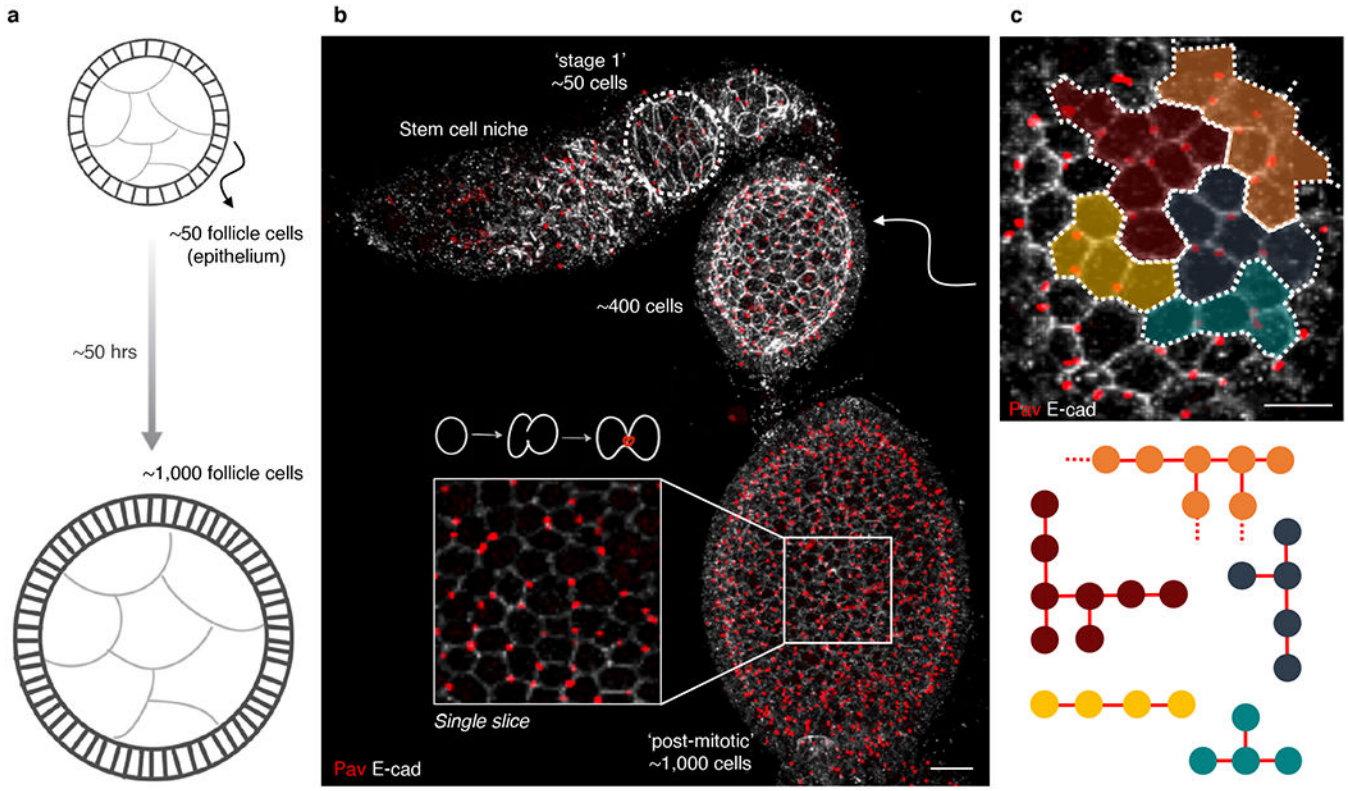


Fig. 1: Ring canals encode the history of cell divisions, permitting reconstruction of cell lineage trees.

a, Schematic of a cross-section of an egg chamber, illustrating the increase in follicle cell number in the epithelium by ~20-fold through cell divisions. The epithelium remains a single-layered tissue as the underlying germline cluster expands in volume. **b**, Projection of a string of mitotic egg chambers arranged from youngest (*top*: a 'stage 1' egg chamber is outlined, with ~50 cells total) to oldest (*bottom*: ~1,000 cells), with the stem cell niche indicated. Ring canals (*Pavarotti*, *Pav*) and cell membranes (*E-cadherin*, *E-cad*) are labeled. Inset shows a single optical plane, highlighting ring canal connections that form between dividing follicle cells. Schematic shows how a cell division with incomplete cytokinesis results in two daughter cells connected through a ring canal, thus forming clonal clusters of interconnected cells that encode the history of cell divisions. **c**, Reconstructions of adjacent clusters of interconnected cells on a section of the surface of the ~400-cell egg chamber in **b** (arrow), with the corresponding color-coded cell lineage tree representations shown below; dashes signify that the tree extends beyond what is visible in this 2D image. Scale bar in **b**, **c** = 10 μm.

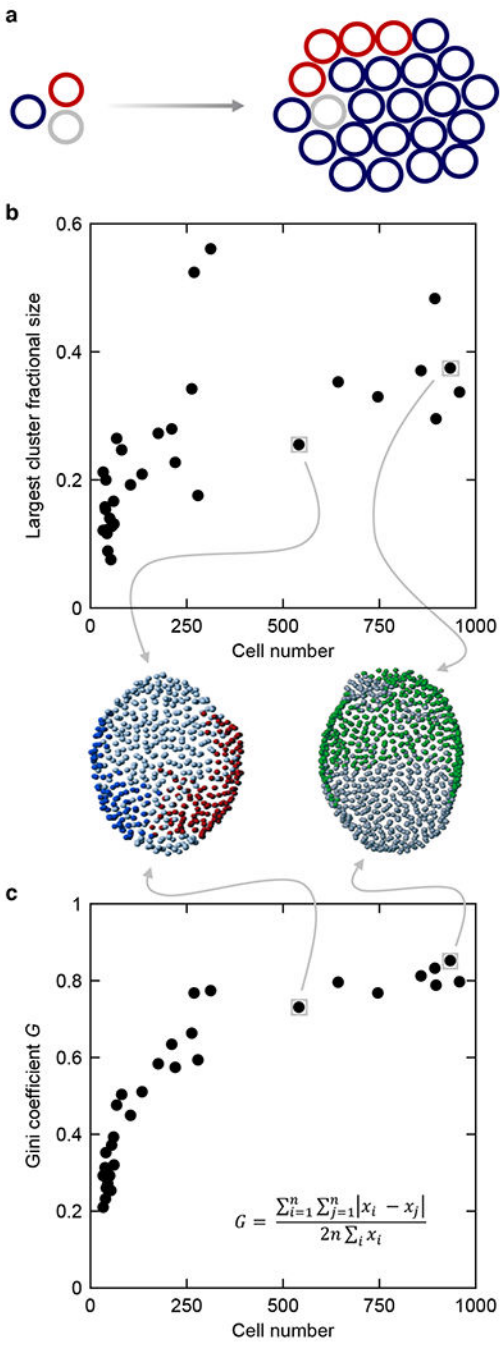


Fig. 2: As cells proliferate, cluster sizes diverge, and dominant clusters of connected cells emerge.

a, Schematic illustrating the various contributions of founder cells to a proliferating tissue, leading to emergence of a dominant clone (blue). **b**, Plot of the fractional coverage of the largest cluster and **c**, the Gini coefficient (G) as a function of total follicle cell number in the epithelium of a given egg chamber. The Gini coefficient is used to characterize the joint divergence in the sizes of clusters (Eq. 1). Cluster size statistics were obtained by analyzing the complete epithelium of fluorescently labeled egg chambers across developmental stages (Fig. 1b and c; Methods). Insets show two egg chambers with their largest clusters of

connected cells reconstructed, where cells are depicted as spheres (*Left*: largest two clusters are shown in red and blue; *Right*: largest cluster is shown in green).

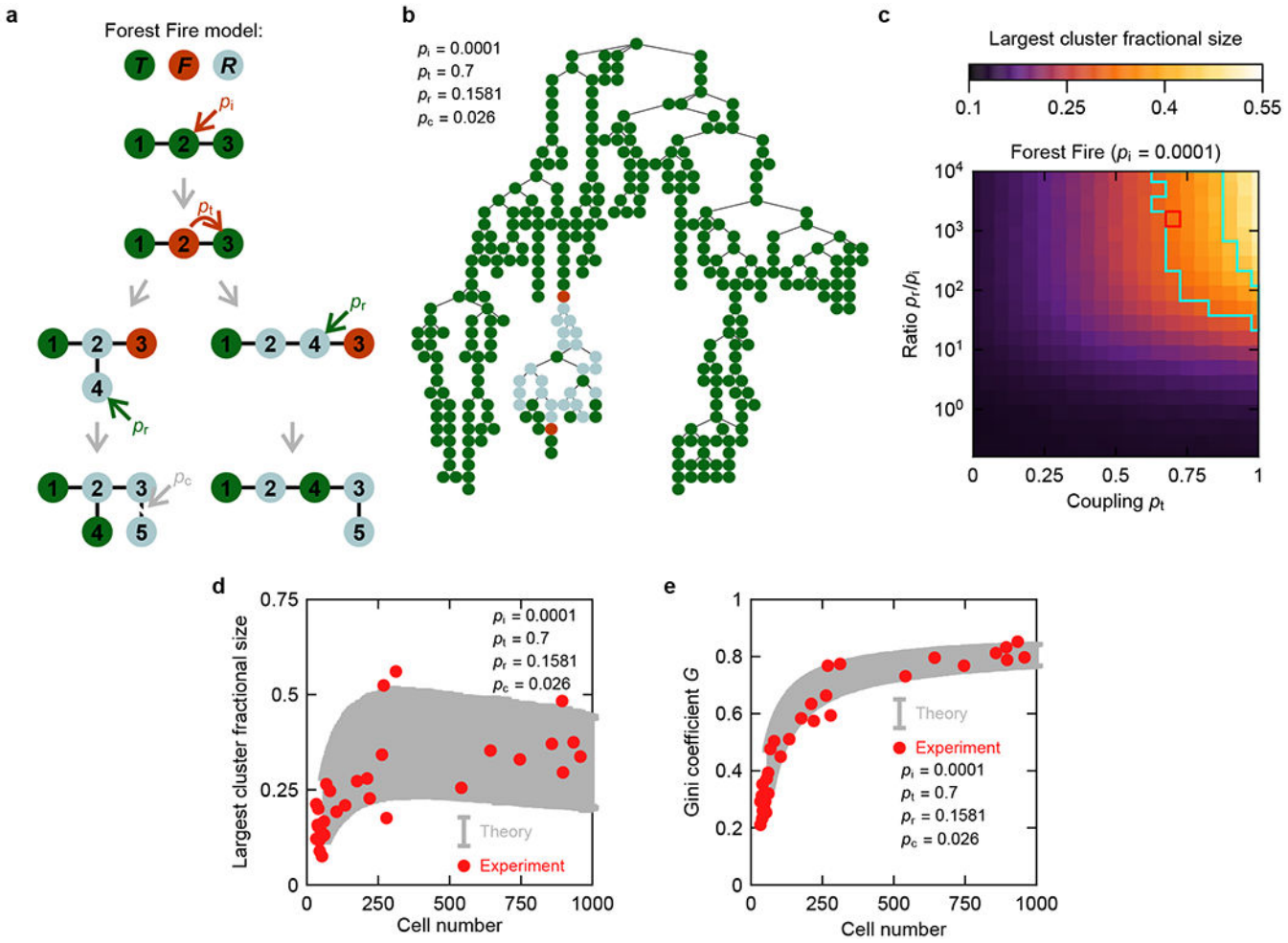


Fig. 3: A Forest Fire model captures the observed statistics of cluster sizes and dynamics of their evolution.

a, Schematic of the main events in the expanded Forest Fire model and associated probabilities. Each cluster is represented as a separate graph, individual cells as nodes on the graph, and ring canals as edges. Each cell can at any timepoint be in one of three states: Tree (*T*, green; a cell capable of dividing), Fire (*F*, red; mitosis), or Refractory (*R*, gray; a cell currently incapable of dividing following mitosis). The four parameters of the model are p_i, p_t, p_r, and p_c, denoting the probabilities of a spontaneous division, division induction, refraction, and a complete division, respectively. A cell in state *F* that transitions to state *R* acquires a new neighbor, also in state *R*, representing the second cell that emerges from the division. **b**, Example of a dominant cluster comprising 334 cells, obtained following the final division in a simulation run for the parameters shown. Note the region of refractory (gray) cells following propagation of cell divisions through adjacent cells, with two currently dividing cells on the boundary (red). **c**, Final size of the largest cluster as a fraction of total cell number obtained from the Forest Fire model as a function of p_t and p_r/p_i, at p_i = 0.0001, averaged over 2,200 simulations at each parameter set; cyan-bordered region indicates where theoretical values are within one experimental standard deviation of the experimental value, averaged over chambers with >600 cells (Largest cluster size fraction = 0.36, s.d. = 0.06;

n = 7; Supplementary Table 3); Extended Data Fig. 5 shows a more extensive parameter sweep. Red square indicates the parameter set that yields the best overall agreement with experiments across all considered diversity indices at p_i = 0.0001 (Methods). This parameter set is used in panels **b**, **d**, and **e**. **d**, Predicted dynamics of the largest cluster fractional size and **e**, the Gini coefficient, respectively, as a function of total cell number, averaged over 2,200 stochastic trajectories (gray; error bars indicate standard deviation of the values across different simulation runs) for the same parameters as **b**, overlaid with experimental values (red).