nature physics

LETTERS

https://doi.org/10.1038/s41567-021-01383-0

Check for updates

Clonal dominance in excitable cell networks

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

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, whereas 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].

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 ~20-fold in number, 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 and Extended Data Fig. 1)[18–20]. Ring canals therefore encode the history of cell divisions, forming lineage trees of connected cells, henceforth called 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 labelled ring canals and cell membranes, we analysed cluster size distributions in entire epithelia during follicle cell proliferation (Extended Data Fig. 1, Supplementary Table 1, Supplementary Video 1 and 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 connected cells (Extended Data Fig. 1 and 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 and Supplementary Table 3); this is 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, which cover ~30–60% of the ventricular surface[1]. Note that whereas ~3% of follicle cell divisions are complete and fragment a cluster into two (Extended Data Fig. 1 and 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]:

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

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 harbours all resources. In the biological system, x_i describes the number of cells in the ith cluster and n describes the number of clusters. We found that whereas 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 and Methods)[24].

[1]Flatiron Institute, Simons Foundation, New York, NY, USA. [2]Jožef Stefan Institute, Ljubljana, Slovenia. [3]Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, Slovenia. [4]The Lewis-Sigler Institute for Integrative Genomics, Princeton University, Princeton, NJ, USA. [5]Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, NJ, USA. [6]Department of Molecular Biology, Princeton University, Princeton, NJ, USA. [7]These authors contributed equally: Jasmin Imran Alsous, Jan Rozman. ✉e-mail: stas@princeton.edu

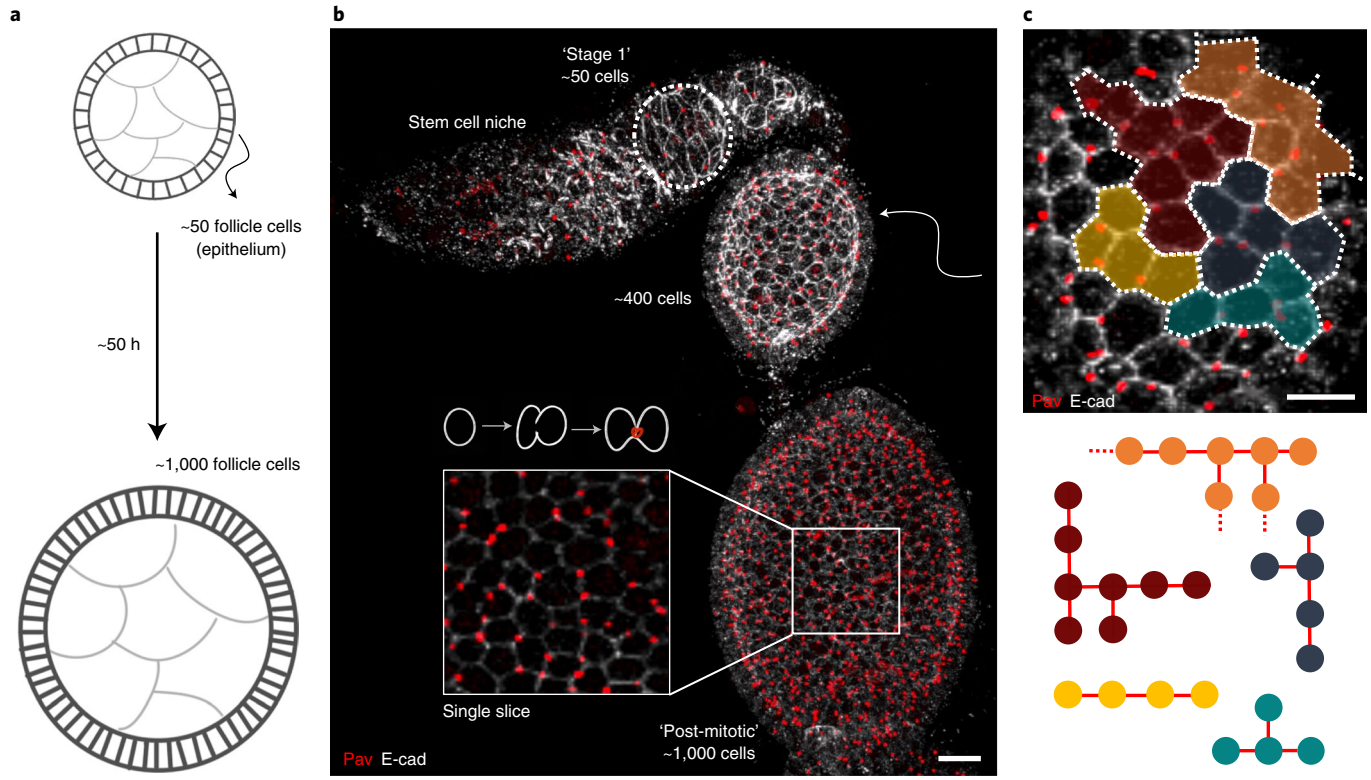


Fig. 1 | Ring canals encode the history of cell divisions, permitting reconstruction of cell lineage trees. a, Schematic of a cross-section of an egg chamber, illustrating the ~20-fold increase in follicle cell number in the epithelium through cell divisions. The epithelium remains a single-layered tissue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

To rationalize our findings, we analysed 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 and Methods), and simulations terminated when the total number of cells reached 1,000, thus matching the approximate end of mitosis[17]. All models accounted 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 and Methods). As the distribution of ring canals had been shown previously to form a network intermediate between linear and maximally branched[18], each ring canal connection between the dividing cell and its linked neighbours 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 from the multiplicative nature of growth, we implemented a simple 'Independent model' of cell divisions in which, starting from the 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], 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; in comparison, the experimental values of ~36% and ~0.81, respectively, are substantially larger (Extended Data Fig. 3 and 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 previously: 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 and Supplementary Videos 2 and 3), and staining patterns of phosphohistone H3 were found to exhibit a statistically significant difference between the experimentally observed values and the theoretically expected number of adjacent mitotic cells were entry into mitosis 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

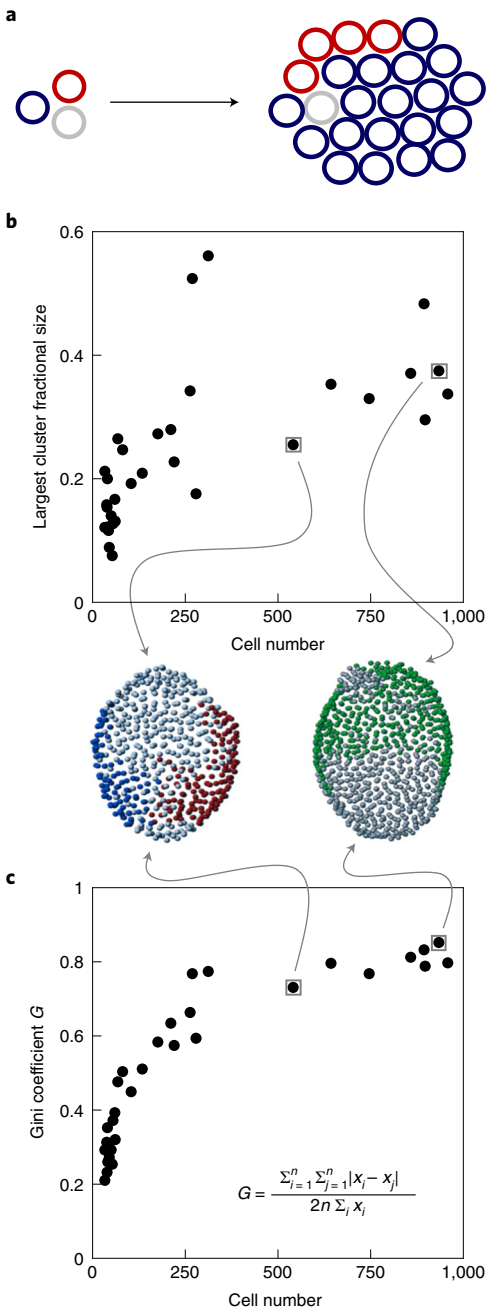


Fig. 2 | As cells proliferate, cluster sizes diverge and dominant clusters of connected cells emerge. a, Schematic illustrating the various contributions of founder cells to a proliferating tissue, leading to emergence of a dominant clone (blue). **b**,**c**, Plot of the fractional coverage of the largest cluster (**b**) and the Gini coefficient *G* (**c**) 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 (equation (1)). Cluster size statistics were obtained by analysing the complete epithelium of fluorescently labelled egg chambers across developmental stages (Fig. 1b,c and Methods). Two egg chambers are shown with their largest clusters of connected cells reconstructed, where cells are depicted as spheres (left, largest two clusters in red and blue; right, largest cluster in green).

clusters representing ~11% of the tissue with a final Gini coefficient of ~0.55, which can be compared to the experimental values of ~36% and ~0.81, respectively (Extended Data Fig. 3 and 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/Cdk1 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 (for example, 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 potentially spread and induce the neighbours' transition from a resting state 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 analysed a simple model of excitable systems that has few parameters and captures key attributes of the biological system, namely the 'Forest Fire' model[10,28,32]. 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 doing so (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_i, modelling a spontaneous entry into mitosis. The F state then transitions in the next time step to the R state; 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 neighbours, the T-to-F transition can also be induced, with probability p_t, by a neighbour 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 state R, it now acquires a new and linked neighbouring cell that is also in the refractory state, representing the second daughter cell that arises from the division, whereas 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 large portion of the parameter space (Fig. 3b,c, Extended Data Fig. 5 and Supplementary Video 4), captures the observed dynamics of cluster size divergence (Fig. 3d,e and 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. First, p_r must be notably 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 on their individual values). Second, to

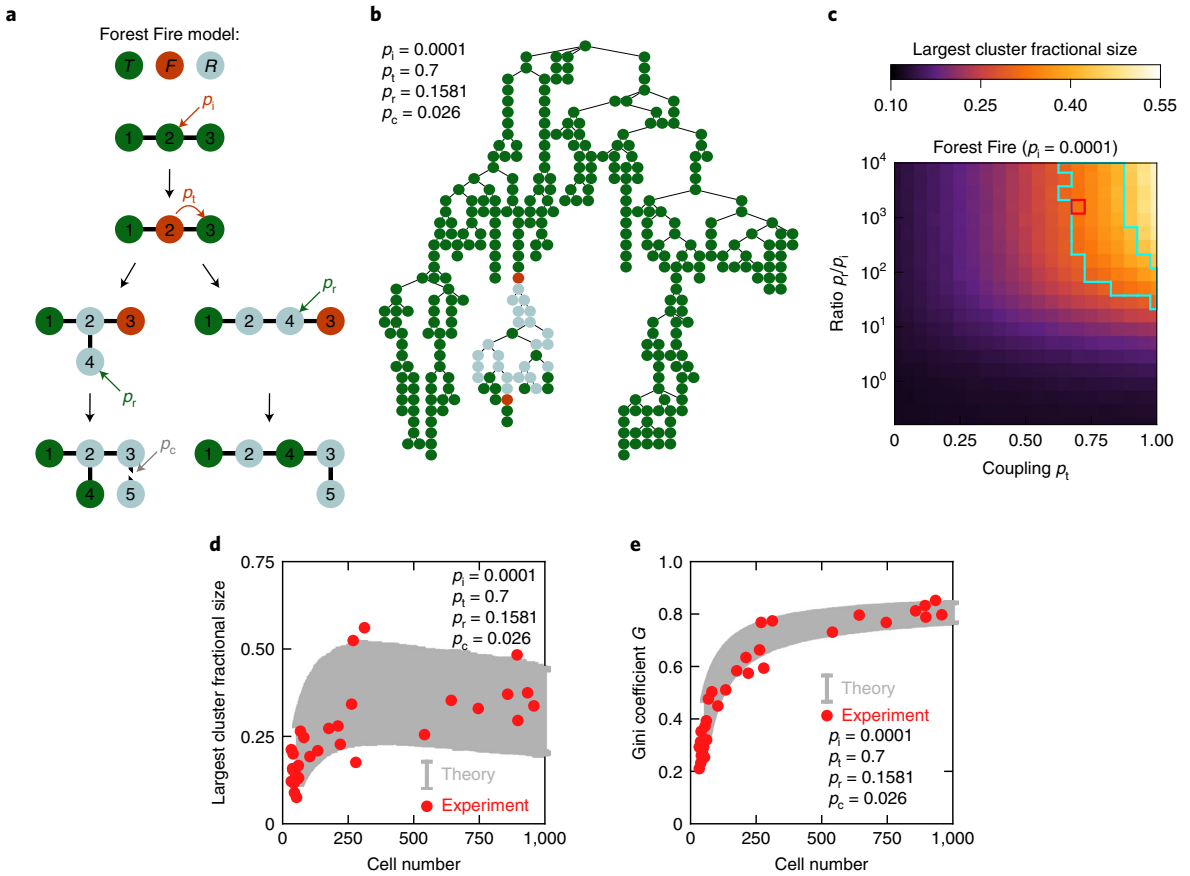


Fig. 3 | A Forest Fire model captures the observed statistics of cluster sizes and dynamics of their evolution. a, Schematic of the main events in the expanded Forest Fire model and associated probabilities. Each cluster is represented as a separate graph, individual cells as nodes on the graph and ring canals as edges. Each cell can at any point in time be in one of three states: tree (T, green; a cell capable of dividing), fire (F, red; mitosis), or refractory (R, grey; 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, recovery after division and complete division, respectively. A cell in state F that transitions to state R acquires a new neighbour, 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 (grey) 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; the 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 is 0.36, s.d. = 0.06, $n = 7$; Supplementary Table 3); Extended Data Fig. 5 shows a more extensive parameter sweep. The 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). Panels **b**, **d** and **e** use this parameter set. **d**,**e**, Predicted dynamics of the largest cluster fractional size (**d**) and the Gini coefficient (**e**) as a function of total cell number, averaged over 2,200 stochastic trajectories (grey; error bars indicate the standard deviation of the values across different simulation runs) for the same parameters as in **b**, overlaid with experimental values (red).

match the experimental findings, coupling must be of at least intermediate 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 analysed 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 substantial (for example, 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 a 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[33]. 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[34]. 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[28,35]. Second, dynamic proliferative behaviours that lead to clonal dominance can confer advantages to 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[36]; 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.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-021-01383-0.

Received: 25 June 2019; Accepted: 13 September 2021;
Published online: 1 November 2021

References

1. Gupta, V. & Poss, K. D. Clonally dominant cardiomyocytes direct heart morphogenesis. *Nature* **484**, 479–484 (2012).
2. Hall, B. A., 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).
3. Luria, S. E. & 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**, eaan0925 (2019).
5. Lopez-Garcia, C., Klein, A. M., Simons, B. D. & Winton, D. J. Intestinal stem cell replacement follows a pattern of neutral drift. *Science* **330**, 822–825 (2010).
6. Snippert, H. J. et al. Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134–144 (2010).
7. Kamimoto, K. et al. Heterogeneity and stochastic growth regulation of biliary epithelial cells dictate dynamic epithelial tissue remodeling. *eLife* **5**, e15034 (2016).
8. Sereti, K.-I. et al. Analysis of cardiomyocyte clonal expansion during mouse heart development and injury. *Nat. Commun.* **9**, 754 (2018).
9. Mikhailov A. S. in *Nonlinear Wave Processes in Excitable Media* NATO ASI Series Vol. 244 (eds Holden A. V. et al.) 127–144 (Springer, 1991).
10. Bak, P., Chen, K. & Tang, C. A forest-fire model and some thoughts on turbulence. *Phys. Lett. A* **147**, 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).
12. Clayton, E. et al. A single type of progenitor cell maintains normal epidermis. *Nature* **446**, 185–189 (2007).
13. Imran Alsous, J., Villoutreix, P., Stoop, N., Shvartsman, S. Y. & Dunkel, J. Entropic effects in cell lineage tree packings. *Nat. Phys.* **14**, 1016–1021 (2018).
14. King, R. C. *Ovarian Development in Drosophila Melanogaster* (Academic Press, 1970).
15. Margolis, J. & Spradling, A. C. Identification and behavior of epithelial stem cells in the *Drosophila* ovary. *Development* **121**, 3797–3807 (1995).
16. King, R. C. & Vanoucek, E. G. Oogenesis in adult *Drosophila melanogaster*. X. Studies on the behavior of the follicle cells. *Growth* **24**, 333–338 (1960).
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).
18. Airoldi, S. J., McLean, P. F., Shimada, Y. & Cooley, L. Intercellular protein movement in syncytial *Drosophila* follicle cells. *J. Cell Sci.* **124**, 4077–4086 (2011).
19. McLean, P. F. & Cooley, L. Protein equilibration through somatic ring canals in *Drosophila*. *Science* **340**, 1445–1447 (2013).
20. McLean, P. F. & Cooley, L. Bridging the divide. *Fly* **8**, 12–18 (2013).
21. Buckingham, M. E. & Meilhac, S. M. Tracing cells for tracking cell lineage and clonal behavior. *Dev. Cell* **21**, 394–409 (2011).
22. Nystul, T. & Spradling, A. C. Regulation of epithelial stem cell replacement and follicle formation in the *Drosophila* ovary. *Genetics* **184**, 503–515 (2010).
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).
25. Sutton, J. Gibrat's legacy. *J. Econ. Lit.* **35**, 40–59 (1997).
26. Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
27. Sawamiphak, S., Kontarakis, Z., Filosa, A., Reischauer, S. & Stainier, D. Y. R. Transient cardiomyocyte fusion regulates cardiac development in zebrafish. *Nat. Commun.* **8**, 1525 (2017).
28. Kromer, J., Khaledi-Nasab, A., Schimansky-Geier, L. & Neiman, A. B. Emergent stochastic oscillations and signal detection in tree networks of excitable elements. *Sci. Rep.* **7**, 3956 (2017).
29. Tyson, J. J. Modeling the cell division cycle: cdc2 and cyclin interactions. *Proc. Natl Acad. Sci. USA* **88**, 7328–7332 (1991).
30. Hochegger, H., Takeda, S. & Hunt, T. Cyclin-dependent kinases and cell-cycle transitions: does one fit all? *Nat. Rev. Mol. Cell Biol.* **9**, 910–916 (2008).
31. Shcherbata, H. R., Althauser, C., Findley, S. D. & 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**, 3169–3181 (2004).
32. Drossel, B. & Schwabl, F. Self-organized critical forest-fire model. *Phys. Rev. Lett.* **69**, 1629–1632 (1992).
33. Puliafito, A. et al. Collective and single cell behavior in epithelial contact inhibition. *Proc. Natl Acad. Sci. USA* **109**, 739–744 (2012).
34. Datar, S. A., Jacobs, H. W., de la Cruz, A. F. A., Lehner, C. F. & Edgar, B. A. The *Drosophila* Cyclin D–Cdk4 complex promotes cellular growth. *EMBO J.* **19**, 4543–4554 (2000).
35. Batchelor, E., Loewer, A., Mock, C. & Lahav, G. Stimulus-dependent dynamics of p53 in single cells. *Mol. Syst. Biol.* **7**, 488 (2011).
36. Roan, H.-Y., Tseng, T.-Z. & Chen, C.-H. Whole-body clonal mapping identifies giant dominant clones in zebrafish skin epidermis. *Development* **148**, dev199669 (2021).

Methods

Experiments. *Drosophila melanogaster* flies were raised under standard conditions at 25 °C and dissected using an established protocol[37]. Egg chambers were extracted by manual dissection, and images were obtained through imaging of immunostained and/or genetically fluorescently labelled samples. Experimental data for analysis of cluster size distributions were acquired by confocal microscopy of fixed egg chambers, for which entire epithelia were analysed. Cluster size distributions for each egg chamber were determined by analysing *z*-stacks of entire egg chambers (three-dimensional images) in Bitplane's Imaris[38]. 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 min, and then stained with the following primary antibodies from the Developmental Studies Hybridoma Bank: 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 488 nm, goat anti-rat 647 nm and goat anti-mouse 647 nm. DAPI (4′,6-diamidino-2-phenylindole, 1:500 dilution) was used to label nuclei. Samples were mounted in a 50–50 mixture of Aqua-Poly/Mount (Polysciences) and an optical clearing medium, RapiClear 1.47 (Sunjin Lab).

Protocol for live imaging. We used a modified version of an established protocol for live imaging of dissected egg chambers[39]. Briefly, ovaries from about three flies were dissected in Schneider's Medium (ThermoFisher no. 21720-001). Separated ovarioles were transferred to a MatTek dish (MatTek no. P35G-1.5-10-C) containing 200 µl of Schneider's Medium supplemented with fetal bovine serum (Sigma no. F4135) and insulin (Sigma no. I0516). The lid was kept on the dish to prevent drying during imaging.

Microscopy. Imaging was performed either on a Nikon A1 confocal microscope with a ×100 or ×60/1.3 NA oil objective, or on a Zeiss LSM 710 point scanning confocal microscope with a ×25/0.8 or ×40/1.2 Apochromat water objective lens. To image fixed samples, *z*-stacks (step sizes 350 nm or 500 nm) were acquired using the 405 nm diode laser, 561 nm diode-pumped solid-state laser, a 638 nm diode laser and a 488 nm 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 min and 4 min.

Reconstructing cluster sizes. We used microscopy and image processing of egg chambers with labelled 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 h, the evolution of cluster size distributions was reconstructed from fixed samples of egg chambers in which entire epithelia were analysed. We therefore acquired three-dimensional images of egg chambers at various stages of development spanning those ~50 h, during which the epithelial cell number increased ~20-fold through mitosis that occurs with incomplete cytokinesis ~97% of the time (Fig. 1a-c, Extended Data Fig. 1a,b,f and Supplementary Video 1). The youngest egg chambers analysed were 'stage 1' egg chambers with ~50 cells, in which the germline cluster is fully enveloped by the epithelium; the oldest egg chambers analysed were ~'stage 6' and 'stage 7' egg chambers, with ~600–1,000 cells, before any major changes in follicle cell morphology occurred in the latter stages. As 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). Although 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 the annotation, or filaments, or spots modules (Extended Data Fig. 1d)[38]; 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 because 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 labelled 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 ref. [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. As '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 $\mathrm{Sh} = -\sum_{i=1}^{n} \frac{x_i}{N} \log\left(\frac{x_i}{N}\right)$, where n is the total number of clusters, N is the number of cells and x_i is the number of cells in the ith clusters. The Shannon index has a maximum value of $\mathrm{Sh}_{\max} = \log n$ when cells are evenly distributed between all clusters, with less even distributions corresponding to lower values of Sh. From the Shannon index, we obtain evenness (Shannon's equitability) $J_E = \frac{\mathrm{Sh}}{\mathrm{Sh}_{\max}}$, given as the ratio between the Shannon index and the maximum possible Shannon index ($\mathrm{Sh}_{\max}$) 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, $\mathrm{Sh}_{\max} - \mathrm{Sh}$. 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 clusters. 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 were performed in Wolfram Mathematica.

Individual simulation runs for 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 are 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 neighbours 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 unlinked 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 are 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], whereas the standard deviation σ_0 is varied over several orders of magnitude, up to $\sigma_0 \approx t_0$. If the normal distribution yielded a non-positive division time during the simulation, a new time interval was selected 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 start the simulation in state T. At each time step, each cell that starts the time step in state T can transition to F spontaneously with probability p_i, and in addition with probability p_t for each linked neighbour that starts the time step in state F, representing the induction of divisions by a neighbouring 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 analysed parameter sets are all combinations of the following: p_t between 0 and 1, in increments of 0.05; p_i values of $10^{0.75-i}$, $10^{0.5-i}$, $10^{0.25-i}$ and 10^{-i} for integer

values of i between 1 and 5, inclusive, as well as $p_i = 1$; and p_r values of $5 \times 10^{0.25-i}$, 5×10^{-i}, $5 \times 10^{-0.25-i}$ and $5 \times 10^{-0.5-i}$ 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's and Hoover), we calculate the chi-squared 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

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

where the index i runs over all experimental egg chambers with more than 61 cells, N_i is the number of cells in the ith experimental egg chamber, $d^{(\alpha)}_{\text{experiment}}(N_i)$ is the value of diversity index α for the ith experimental egg chamber, whereas $d^{(\alpha)}_{\text{theory}}(N_i)$ 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. Index 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, $\langle\chi^2\rangle$, of the chi-squared coefficients for all seven diversity indices. Extended Data Fig. 6 shows $\langle\chi^2\rangle$ plots for the studied parameter space; the parameter set used in Fig. 3b,d,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_i cannot exceed 1).

The first caveat to point out in the Forest Fire simulations is that 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). Last, 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 shown only for cell numbers greater than 61. Furthermore, the theoretical values are drawn only 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 analyse 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-dividing 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 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 each individual 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 substantial, 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 extent 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]. Last, experimental studies have demonstrated the random configuration and orientation of syncytia on the egg chamber's surface[19,20], suggesting that positional biases are similarly unlikely.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Experimental data for cluster sizes in egg chambers are available at https://github.com/rozmanj/CD_ExcitableNetworks.

Code availability

All custom codes are available at https://github.com/rozmanj/CD_ExcitableNetworks.

References

37. Wong, L. C. & Schedl, P. Dissection of *Drosophila* ovaries. *J. Vis. Exp.* **1**, 52 (2006).
38. Imaris v.8.3.1 (Bitplane Inc., 2016).
39. Prasad, M., Jang, A. C.-C., Starz-Gaiano, M., Melani, M. & Montell, D. J. A protocol for culturing *Drosophila melanogaster* stage 9 egg chambers for live imaging. *Nat. Protoc.* **2**, 2467–2473 (2007).
40. Jia, D., Xu, Q., Xie, Q., Mio, W. & Deng, W.-M. Automatic stage identification of *Drosophila* egg chamber based on DAPI images. *Sci. Rep.* **6**, 18850 (2016).

Acknowledgements

We thank G. Laevsky for expert help with imaging and E. Coen for advice on how to improve the manuscript and broaden its scope. We thank E. Wieschaus, T. Schüpbach, A. Berezhkovskii, M. Krajnc, S. Dutta and C. Tarnita for helpful discussions, D. Hellwig for assistance with movie making, J. Jackson for manuscript editing and the Martin Lab (Biology, Massachusetts Institute of Technology) for lab stocks and microscopy. This work was supported by the United States National Institutes of Health (research project grant no. R01GM134204-02) (S.Y.S.) and by the Slovenian Research Agency (research core funding no. P1-0055) (J.R.).

Author contributions

J.I.A., S.Y.S. and A.K. designed the research. J.I.A. performed the experiments and analysed the data. S.Y.S. and A.K. developed the theory. J.R. implemented and analysed 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

The authors declare no competing interests.

Additional information

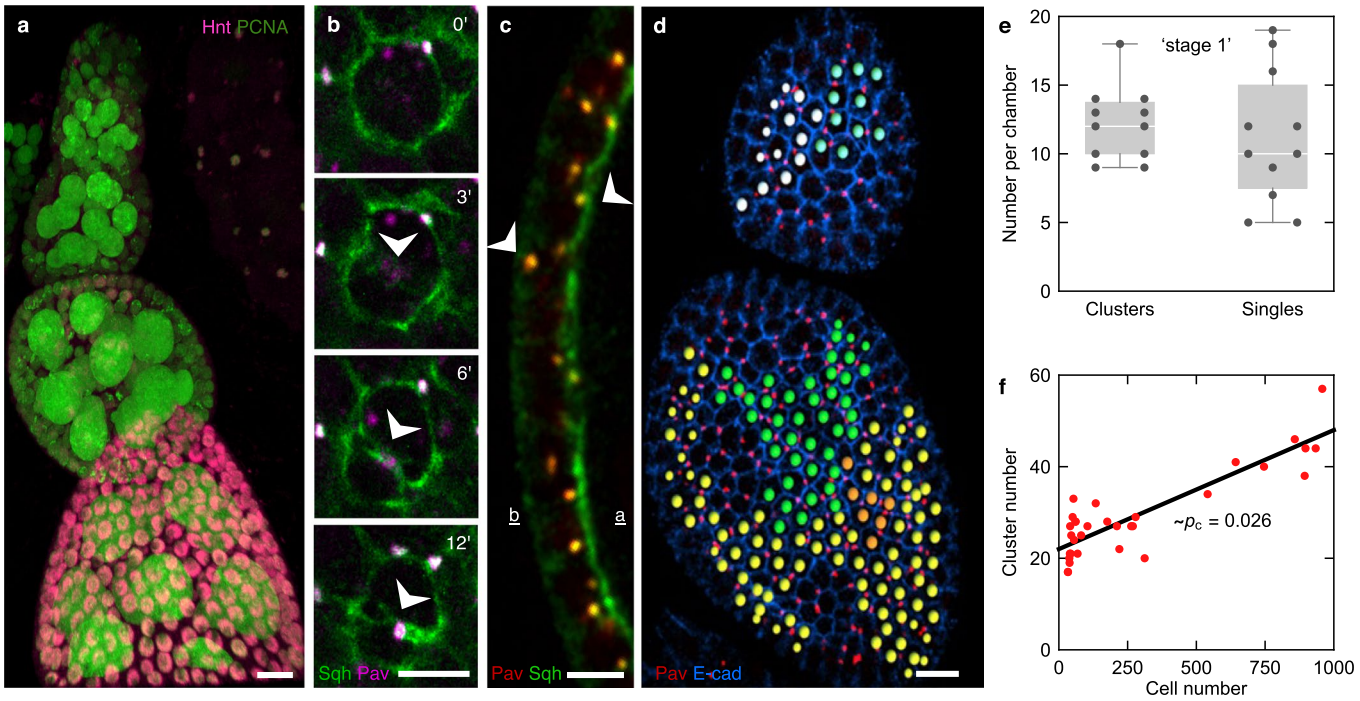
Extended data is available for this paper at https://doi.org/10.1038/s41567-021-01383-0.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-021-01383-0.

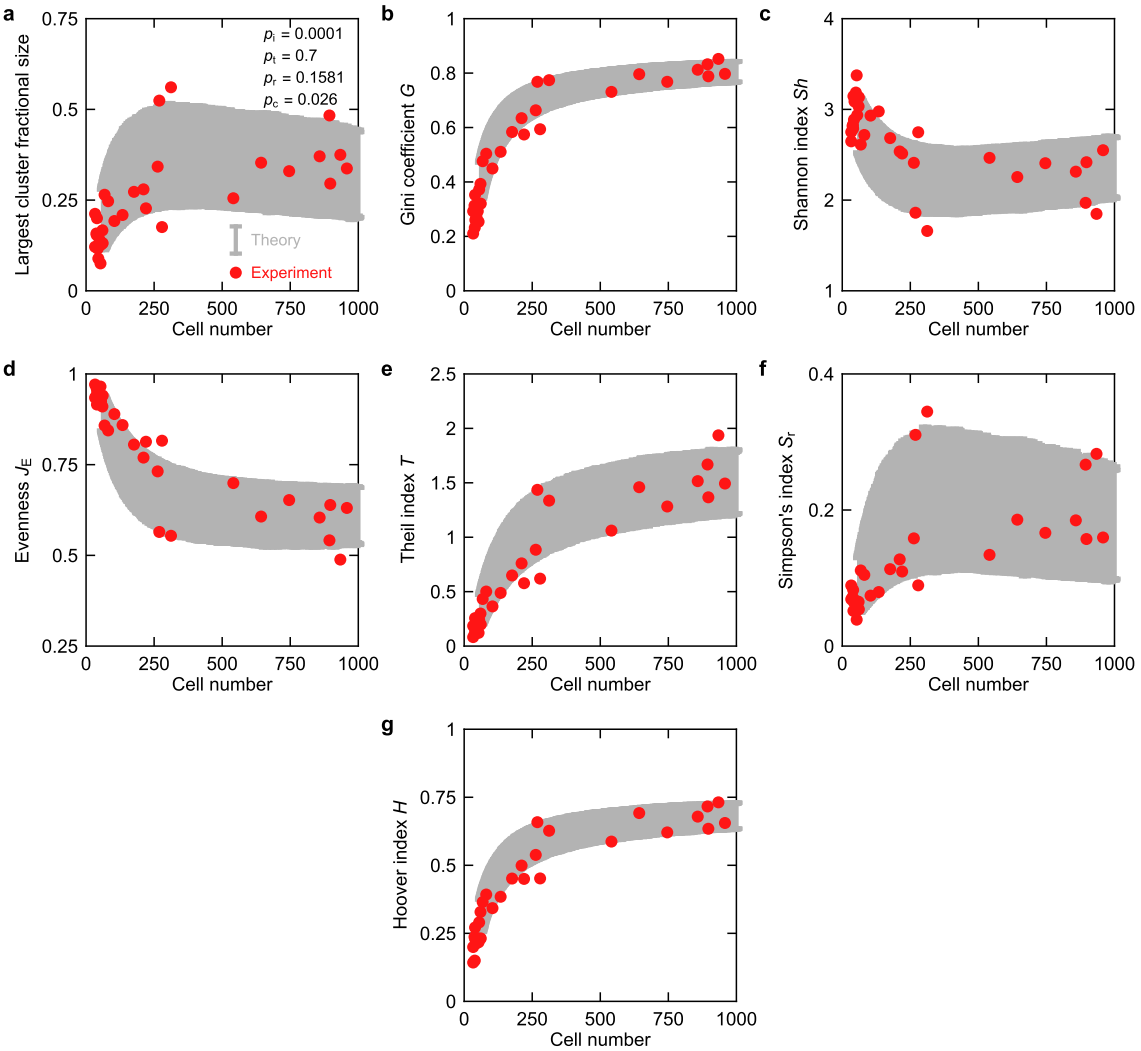
Correspondence and requests for materials should be addressed to Stanislav Y. Shvartsman.

Peer review information *Nature Physics* thanks Jordi Garcia-Ojalvo, Denise Montell and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

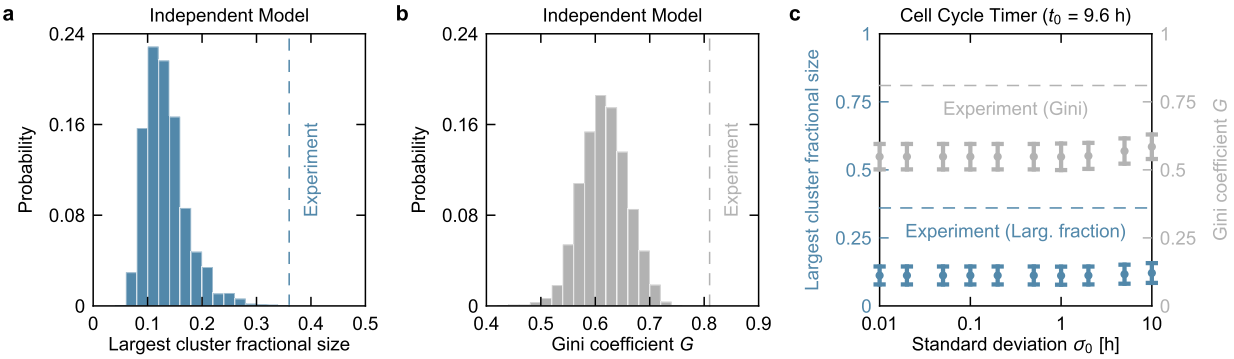
Reprints and permissions information is available at www.nature.com/reprints.



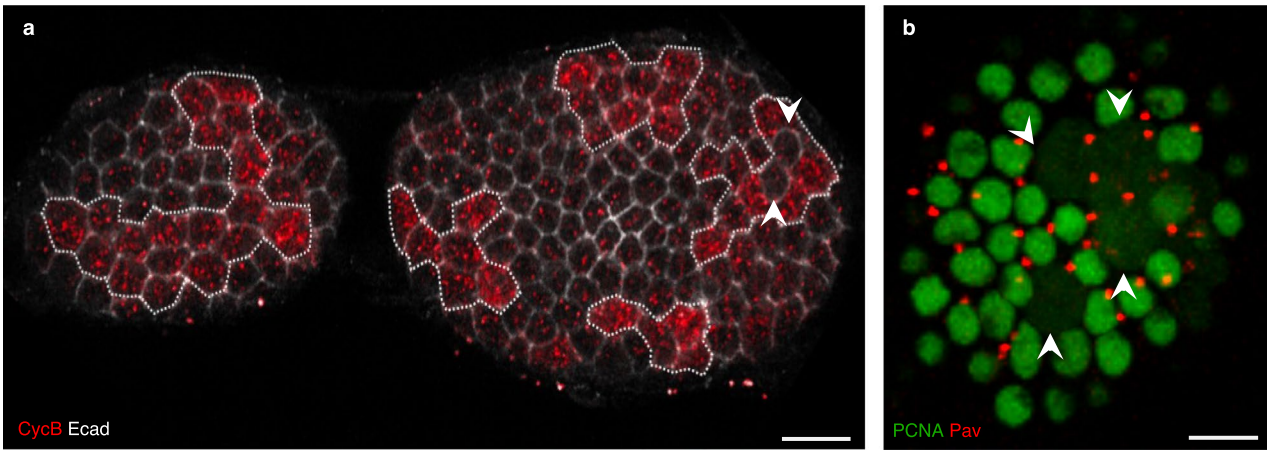
Extended Data Fig. 1 | Incomplete cytokinesis leaves daughter cells connected through stable ring canals. a, Egg chambers expressing the nuclear Proliferating Cell Nuclear Antigen (PCNA, green) and stained for Hindsight (Hnt) (magenta), which labels endocycling follicle cells that have exited mitosis[17]. *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[17,40]. 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 colour denote cells belonging to a cluster: two (white, blue) and three (green, yellow, and orange) clusters of connected 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 $p_c = 0.026 \pm 0.003$. Line shows a linear fit.



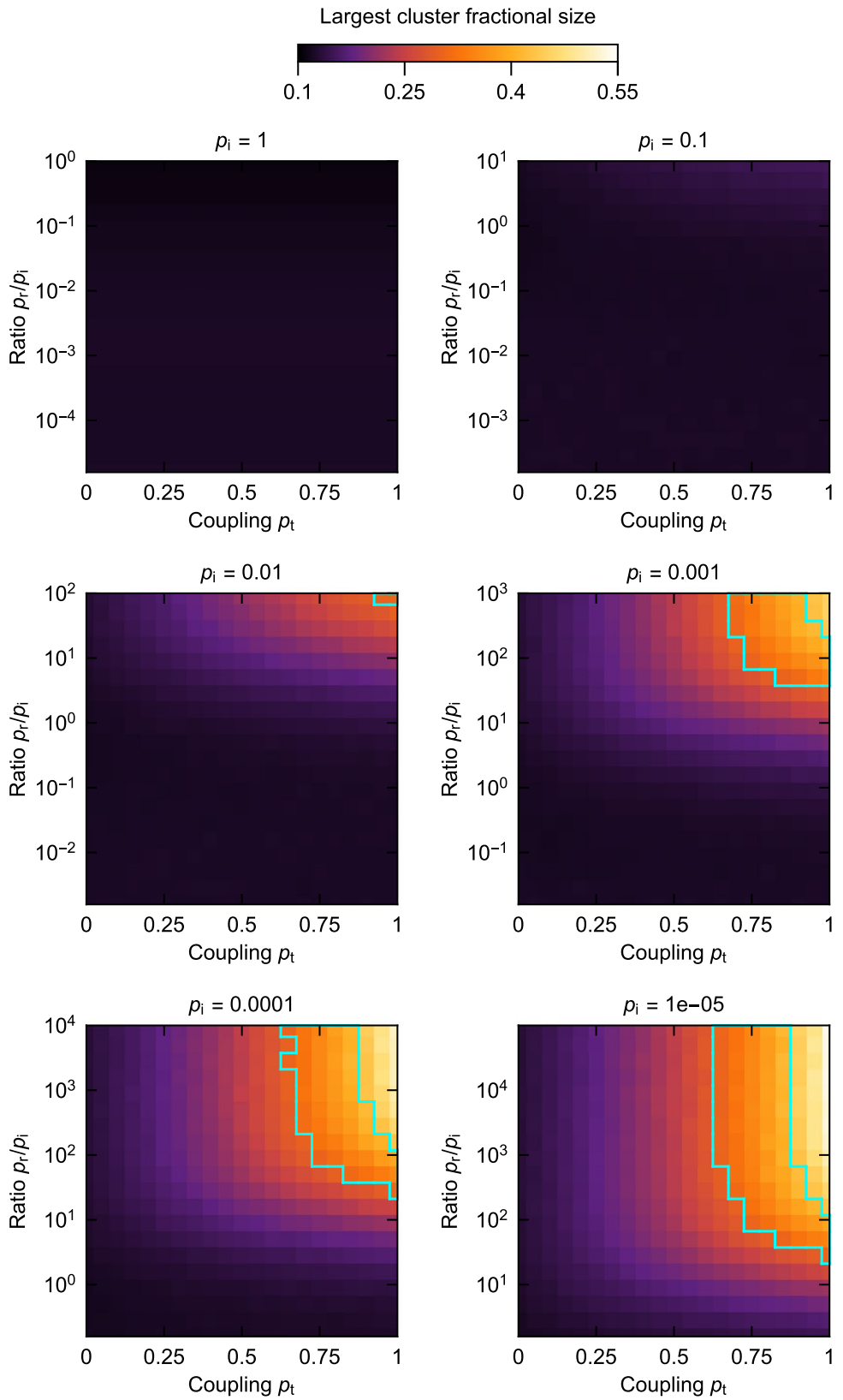
Extended Data Fig. 2 | Comparison of diversity measures used to quantify cluster size divergence. a, Largest cluster fraction (that is Berger-Parker index), **b**, Gini coefficient, **c**, Shannon index, **d**, evenness, **e**, Theil index, **f**, Simpson's index (with replacement), and **g**, Hoover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



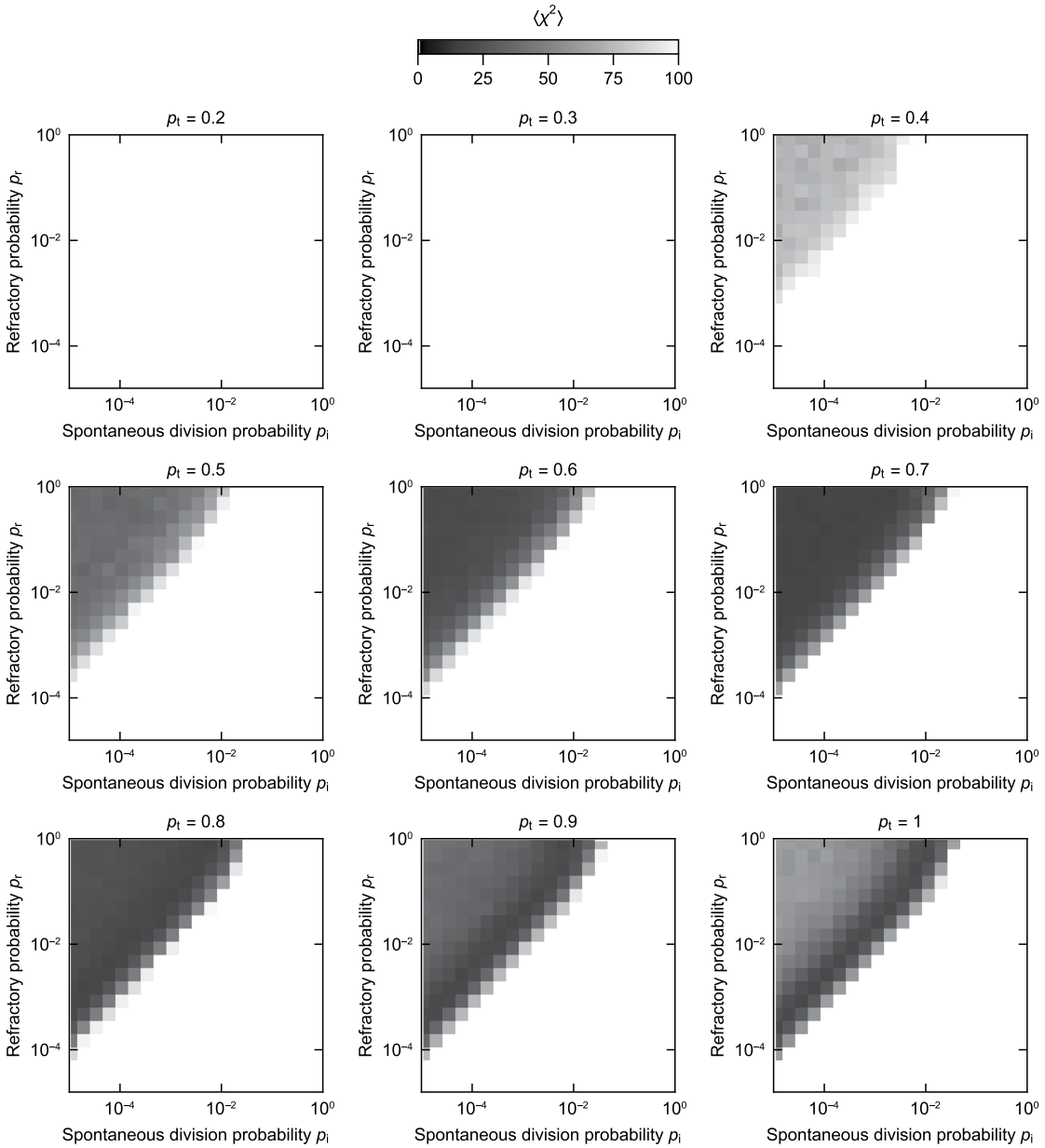
Extended Data Fig. 3 | Largest cluster fractions and Gini coefficients as obtained from the Independent and Cell Cycle Timer models of uncoupled cell divisions. a, Distribution of final largest cluster sizes as fractions of all cells in the Independent model, based on 2,200 simulation runs. Dashed lines 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 (grey) 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**.



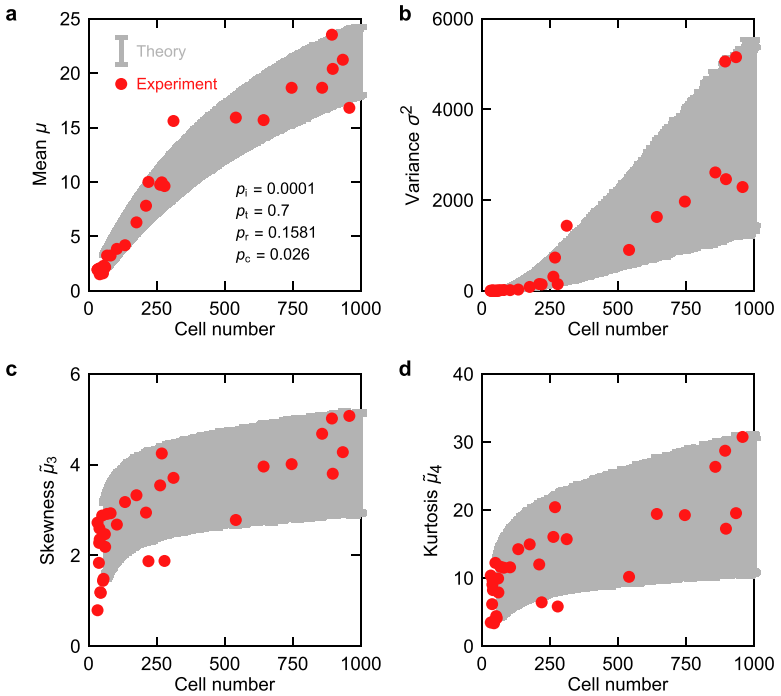
Extended Data Fig. 4 | The correlated nature of epithelial cell divisions. a, Epithelial surface of two egg chambers stained for the mitotic marker Cyclin B (CycB), showing domains (~5-15 cells) of coordinated expression. Arrowheads pointing to two adjacent dividing cells in one of those domains. **b**, 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.



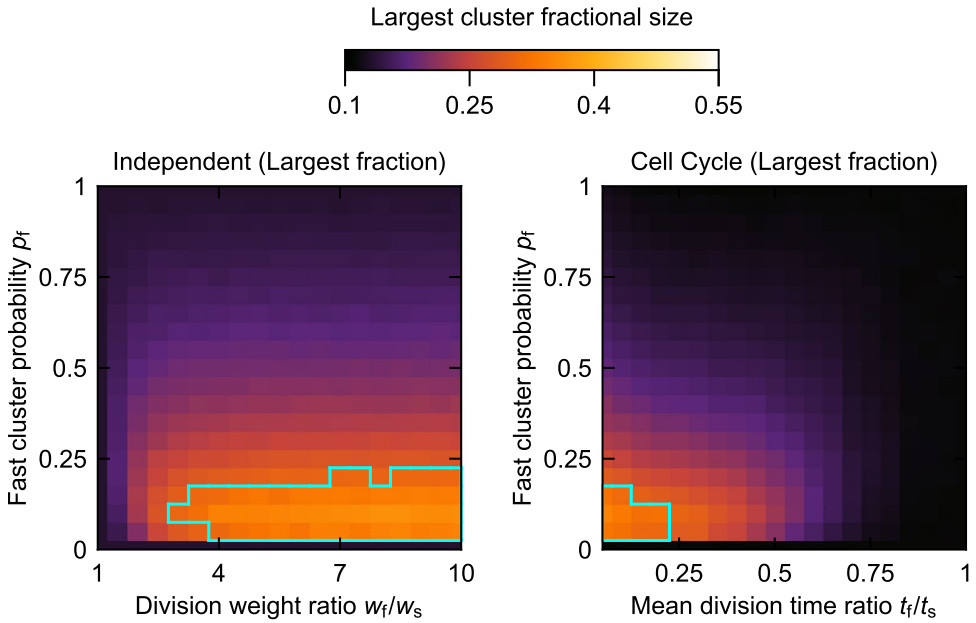
Extended Data Fig. 5 | Largest cluster sizes in the Forest Fire model. Final size fractions of the largest cluster as a function of p_t and p_r/p_i for different orders of magnitude of p_i and averaged over 2,200 simulations at each parameter set. Regardless of the value of p_i, similar average sizes 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).



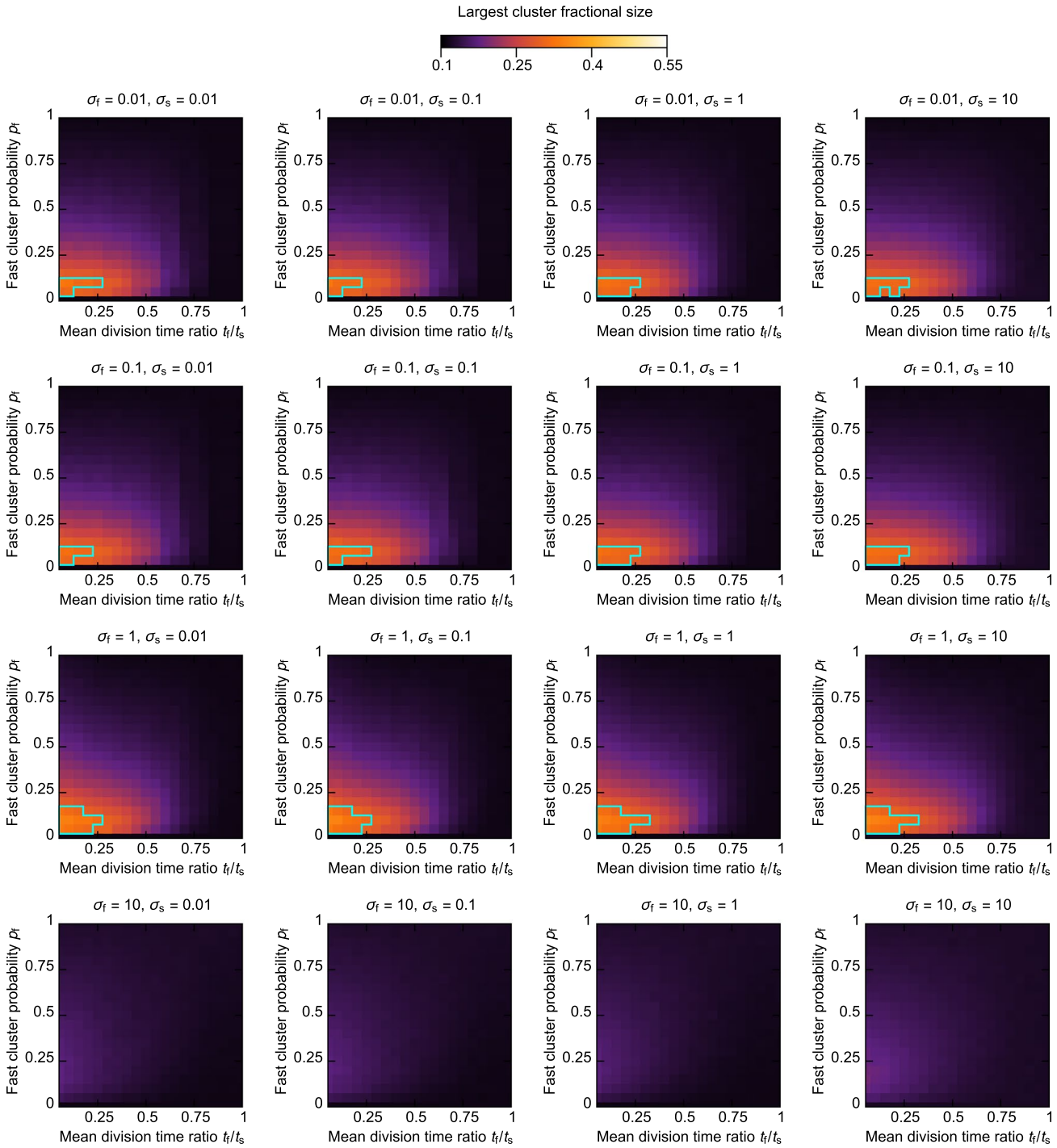
Extended Data Fig. 6 | Analysis of the Forest Fire model parameter space. Plots of the average chi-square coefficient ($\langle \chi^2 \rangle$; see Methods) of the diversity indices between the Forest Fire model simulations and the experimental data for all egg chambers with >61 cells, for different 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 $\left[\sigma^2 = \left\langle (x_i - \mu)^2 \right\rangle\right]$, **c**, skewness $\left[\tilde{\mu}_3 = \left\langle (x_i - \mu)^3 / \sigma^3 \right\rangle\right]$, and **d**, kurtosis $\left[\tilde{\mu}_4 = \left\langle (x_i - \mu)^4 / \sigma^4 \right\rangle\right]$, where x_i is the number of cells in the i^{th} cluster and $< >$ 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 (grey; parameters $p_i = 0.0001$, $p_t = 0.7$, $p_r = 0.158114$, $p_c = 0.026$), averaged over 2,200 simulation runs. Grey 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. Left panel: 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 in both panels 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). Right panel: 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 at 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).

nature research

Corresponding author(s): Prof. Stanislav Y. Shvartsman

Last updated by author(s): 09/02/2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	The raw data used in this study consisted of zstacks (3D) images of egg chambers and 3D movies of live samples; these were collected either on a Nikon or a Zeiss confocal microscope, and did not require any custom code or add-on parts for collection.
Data analysis	We used Bitplane's Imaris and its built-in modules to visualize the confocal microscopy data, to identify each cell and all cells to which it was connected through ring canals, thus extracting the sizes of the clusters from the corresponding Statistics tab for egg chambers at various stages of development. Bitplane's Imaris was also used to create the 3D visualizations and reconstructions shown in the main and supplementary figures. Fiji and Adobe Premiere Pro were used for creating and annotating the supplementary videos. No custom code was written for any of these tasks, and this information is stated explicitly in the SI. The simulations of the three models (Forest Fire, Independent, and Cell Cycle Timer) are done in C++ programming language (c++14). Some post-processing is done in Wolfram Mathematica (Version 10.3).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All simulation codes will be available for download from a public repository. The experimental data will also be publicly accessible from the same platform (we will

use github and will provide the DOI before publications). There are no restrictions on data availability.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. We used a sample size of n=32 egg chambers, which was sufficient to establish a clear trend of cluster sizes diverging and one or two dominant clusters emerging as egg chambers grow (best seen in Fig. 2b, c), against which different theoretical models could be tested. The raw data for the smallest egg chambers used as initial conditions (n = 11) and the largest chambers used to determine the average final size of dominant clusters (n = 7) are shown in Supplementary Tables 2 and 3, respectively.
Data exclusions	Five data points were excluded from our analysis because the Imaris files that contained the reconstructions of cluster sizes in the epithelia of these egg chambers had been corrupted; the raw data is present, but not the objects from the SPOTS module. One data point was excluded because we were unable to locate the corresponding file. These data points were not excluded for being outliers, and are not included in the same size (n=32).
Replication	Attempts at replication of the findings reported in this paper have been successful, and the presented results include these replicas and the uncertainties arising therefrom. The experimental data was collected over multiple sessions using the same technique for egg chamber dissection and immunostaining with antibodies. Data analysis occurred over multiple sessions as well, using the same software (Bitplane's Imaris) without issues.
Randomization	This is not relevant to our study. We collected data from n=32 egg chamber, and for each, we analyzed the entire epithelium and constituent number of clones, and clone sizes. The theoretical analysis we present in the paper relates to those data points.
Blinding	Not applicable.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	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.
Validation	Antibodies used are commercial and were either verified by others in prior (co)-staining/(co)-localization experiments or verified in separate independent studies.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	We used Drosophila melanogaster (fruit fly) as a model organism. The strains used were the following:

Laboratory animals	- a fluorescently-tagged version of Pavarotti (ubi>Pavarotti-mCherry) - fluorescently-tagged version of the Proliferating cell nuclear antigen (PCNA-GFP) - Resille-GFP (also known as P{PTT-un1}jCG8668 117-2). - Resille::GFP/Cyo; pav::mCh/Tm3 sb Standard Drosophila genetic techniques were utilized to genetically move together these three tagged proteins. The final fly is Resille::GFP/+; PCNA::GFP/pav::mCh. - fluorescently-tagged version of the E-cadherin (w; Ecad:GFP). - fluorescently-tagged version of the myosin regulatory light chain (w; sqh::GFP). - luorescently-tagged version of GAP43, a membrane marker (w; Gap43:mCh; Jupiter). - A somatic cell driver traffic jam (TJ-Gal4). - Fly-FUCCI (Fluorescent Ubiquitination-based Cell Cycle Indicator). These flies, their genotypes, and the locations from where they were obtained are all listed in Supplementary Table 1.
Wild animals	The study did not involve any wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	No ethical approval or guidance was required.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

nature portfolio

https://doi.org/10.1038/s41567-021-01383-0

Supplementary information

Clonal dominance in excitable cell networks

In the format provided by the authors and unedited

Supplementary Information

Dynamics of clonal dominance in growing 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, Slovenia.*
[3] *Faculty of Mathematics and Physics, University of Ljubljana, Ljubljana, 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
[*] Corresponding author (stas@*princeton.edu)*

Contents:

Supplementary Table 1. *Fly stocks used, their source, and figures and videos in which they appeared.*

Transgene name	Source	Relevant to
ubi>Pav::mCh[*]	from Dr. Emmanuel Derivery [1]	Figs. 1 and 2, E.D. Figs. 1 and 4, Supp. Videos 1 and 6
PCNA::GFP[**]	from Dr. Eric Wieschaus (Shvartsman lab stock at Princeton)	E.D. Figs. 1 and 4, Supp. Video 6
Resille::GFP	Shvartsman lab stock	Supp. Video 6
Resille::GFP/Cyo; pav::mCh/Tm3 sb	generated for this study	Supp. Video 6
Resille::GFP/Cyo; PCNA::GFP/pav::mCh	generated for this study[***]	Supp. Video 6
w; Ecad:GFP	Shvartsman and Martin lab stocks[*****]	E.D. Figs. 1 and 4
w; sqh::GFP	Martin lab stock	E.D. Fig. 1
w; Gap43:mCh; Jupiter:GFP	Martin lab stock	Supp. Video 2
TJ-Gal4	Martin lab stock	Supp. Video 3
Fly-FUCCI[****]	BDSC[******] #55100	Supp. Video 3

[*] *Pav* is a kinesin-like protein and is a component of stalled cytokinetic furrows [1, 2].

[**] *PCNA* is nuclear and exhibits characteristic cell cycle-dependent changes in intensity and pattern [3].

[***] Standard genetic techniques were utilized to genetically move together these tagged proteins; since *Resille::GFP* and *PCNA::GFP* localize to different cellular compartments, they can be differentiated.

[****] Fly-FUCCI expressing *E2f* and *CycB* degron-tagged GFP and mRFP, respectively, under UASp control.

[*****] Adam Martin Lab at the MIT Department of Biology

[******] BDSC: Bloomington *Drosophila* Stock Center.

Supplementary Table 2. *Experimental initial conditions for number of clusters and corresponding sizes*.

Egg chamber (EC)	Cluster size distribution
EC1	4, 4, 4, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC2	5, 5, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC3	8, 4, 4, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC4	5, 4, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC5	6, 4, 3, 3, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC6	7, 3, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1
EC7	6, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1
EC8	4, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1
EC9	8, 5, 5, 4, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC10	4, 4, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC11	7, 4, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1

Supplementary Table 3. *Experimental cluster sizes for egg chambers with >600 cells.*

Egg chamber (EC)	Cluster size distribution
EC1L	323, 153, 90, 60, 59, 26, 25, 22, 21, 17, 15, 11, 11, 9, 9, 9, 7, 6, 6, 6, 5, 5, 5, 4, 4, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC2L	350, 346, 36, 36, 18, 17, 17, 13, 10, 10, 10, 9, 8, 8, 6, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC3L	265, 198, 79, 75, 31, 30, 28, 26, 20, 20, 20, 16, 15, 9, 7, 7, 5, 4, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC4L	432, 97, 77, 63, 60, 48, 28, 13, 11, 7, 7, 5, 5, 4, 4, 3, 3, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC5L	318, 137, 88, 53, 33, 32, 32, 32, 24, 20, 11, 10, 8, 7, 6, 4, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC6L	246, 151, 49, 45, 38, 30, 25, 24, 23, 13, 11, 10, 9, 9, 7, 7, 6, 6, 5, 4, 3, 2, 2, 2, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1
EC7L	227, 115, 74, 61, 46, 16, 10, 8, 7, 6, 6, 5, 5, 4, 4, 4, 4, 4, 4, 4, 3, 3, 3, 3, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1

[1] Derivery, E. et al. Polarized endosome dynamics by spindle asymmetry during asymmetric cell division. *Nature* **528**, 280-285 (2015).

[2] Airoldi, S. J., McLean, P. F., Shimada, Y. & Cooley, L. Intercellular protein movement in syncytial *Drosophila* follicle cells. *J. Cell Sci.* **124**, 4077-4086 (2011).

[3] Zerjatke, T. et al. Quantitative cell cycle analysis based on an endogenous all-in-one reporter for cell tracking and classification. *Cell Rep.* **19**, 1953-1966 (2017).