Entropic effects in cell lineage tree packings

Jasmin Imran Alsous1,2,5, Paul Villotreix2,3,5, Norbert Stoop4,5, Stanislav Y. Shvartsman1,2 and Jörn Dunkel6,4*

Optimal packings of unconnected objects have been studied for centuries, but the packing principles of linked objects, such as topologically complex polymers or cell lineages, are yet to be fully explored. Here, we identify and investigate a generic class of geometrically frustrated tree packing problems, arising during the initial stages of animal development when interconnected cells assemble within a convex enclosure. Using a combination of 3D imaging, computational image analysis and mathematical modelling, we study the tree packing problem in Drosophila egg chambers, where 16 germ-line cells are linked by cytoplasmic bridges to form a branched tree. Our imaging data reveal non-uniformly distributed tree packings, in agreement with predictions from energy-based computations. This departure from uniformity is entropic and affects cell organization during the first stages of the animal’s development. Considering mathematical models of increasing complexity, we investigate spherically confined tree packing problems on convex polyhedra that generalize Platonic and Archimedean solids. Our experimental and theoretical results provide a basis for understanding the principles that govern positional ordering in linked multicellular structures, with implications for tissue organization and dynamics.

Packing problems are part and parcel of life. From DNA packing in the nucleus and cell assembly during embryogenesis to crystal growth, densification of granular materials and masonry, the spatial packings of fundamental building blocks play an essential role in determining the structure and function of physical and biological matter. Since Kepler’s seventeenth century conjecture on dense sphere packings, a rich body of experimental and theoretical work has helped clarify the physical and mathematical principles underlying optimal embeddings of unconnected objects. By contrast, much less is known about the packings of topologically linked structures in confined spaces, a setting relevant to the understanding of chromosome condensation, protein folding, multi-scaffold DNA-origami structures, synthetic multicellular self-assembly and the positioning of cell lineage-trees (CLTs) in embryonic cysts.

Here, we combine 3D imaging of fruit fly egg chambers with mathematical modelling to investigate a generic yet previously unexplored class of geometrically frustrated tree packing problems (TPPs). This class of problems arises during oogenesis in mammals, birds, insects and other multicellular organisms, when interconnected germline cells assemble within an enclosed domain. Our analysis shows that topological and symmetry-based entropic constraints inherent to TPPs severely restrict and bias the observed packing distributions. This finding is in stark contrast to the conventional packings of unconnected objects, implying that topological constraints can provide an efficient biophysical mechanism for controlling cell positioning during early oogenesis and embryogenesis.

Higher life forms develop through successive symmetry breaking transitions that transform a cell cluster with little obvious structure into a highly compartmentalized functional organism. While gene expression and chemically induced morphogenesis have long been studied, the role that topological and geometric constraints play in determining the positioning of cells during the initial stages of development is not yet well understood. Furthermore, physical linkages between germline cells through tubular cytoplasmic bridges have been confirmed in germline tissues of various mammals, insects, amphibians (Table 1 in ref. 1) and molluscs, and similar topological linkages play an important role in the colony formation of choanoflagellates, the closest living relatives of animals. Although the structure and function of cytoplasmic bridges have long been studied in the context of intercellular biochemical transport, their role for physical cell positioning, and tissue organization and dynamics has yet to be explored.

Cytoplasmic bridges often originate from incomplete cytokinesis, forming a hierarchical CLT network that encodes the history of cell divisions. In several species, including the fruit fly Drosophila melanogaster, these lineage trees are encapsulated into approximately spherical cysts, giving rise to a geometrically constrained TPP. A closely related packing problem with directly visible dynamical implications is encountered in the embryos of colonial Volvox algae, which are among the simplest multicellular organisms. In Volvox, the spatial embedding of the cytoplasmic bridge network along the surface of a fluid-filled vesicle determines the position of the phialopore, an opening at the anterior pole that defines the initiation point for the inversion of the embryos. This example demonstrates that tree embeddings can provide a robust physical symmetry breaking mechanism. It is therefore desirable to develop a better empirical and theoretical understanding of TPPs in convex enclosures. As we shall show below, the associated mathematical challenge involves solving the subgraph isomorphism problem of identifying all topology-preserving embeddings of a given planar tree in a convex polyhedron.

To investigate TPPs systematically in a biologically relevant model organism, we acquired 3D confocal images of D. melanogaster egg chambers, the precursors of mature oocytes (Fig. 1a). An egg chamber is a cluster of 16 germine cells (1 oocyte and 15 supporting nurse cells), generated by four synchronous divisions of a founder cell. The divisions occur with incomplete cytokinesis, so that the cells remain connected through membranous bridges called ring canals and form a hierarchical lineage tree (Fig. 1a–c). In the rounded 3D egg chamber, each of the 16...
germline cells is additionally attached to an outer epithelium, a single-layered sheet of smaller somatic cells that provides a convex hull (Fig. 1a). Using 3D confocal images with fluorescently labelled nuclei, membranes and ring canals, we identified the individual cells and ring canals in n = 121 rounded cysts (Fig. 1d). The cells and ring canals define the nodes and edges of the CLT, respectively (Fig. 1b,c). The nodes of the lineage trees can be unambiguously labelled by a cell's position in the division history (Fig. 1c; Supplementary Information). This convention is adopted throughout, with the oocyte cell v = 1 denoting the root of the tree and nurse cells v = 2, 3, ..., 16 (Fig. 1b,c). Armed with this unique tree topology, we next focus on the question of how the lineage tree is spatially embedded in 3D.

To this end, we first generated 3D membrane-based reconstructions of individual egg chambers from our imaging data (Fig. 1d). The 3D spatial organization of the cell-tree packing resembles that of an orange, in that each germline cell contacts the outer epithelium; the germline cells correspond to the segments, and the epithelium to the outer peel. Our 3D reconstructions reveal that the topological constraints impose a highly conserved layered organization of the 15 nurse cells relative to the oocyte: nurse cells separated from the oocyte by the same number of ring canals, regardless of generation, typically reside within the same layers (Fig. 1c), but a cell's relative position within a layer is found to vary between different egg chambers. By considering the 2D planar unfolding of the 3D cell-tree packing (Supplementary Information), tree states can be distinguished and uniquely labelled by their leaf sequences (Fig. 1c). Generally, for an unfolded CLT with v = 1, ..., V vertices of degree d, basic combinatorial considerations (Fig. 1b) show that there exist

\[ P = \prod_{v=1}^{V} (d_v - 1)! \]  

planar embeddings up to 2D rotational symmetry, where \( n! = n(n-1) \ldots 1 \) and \( 0! = 1 \). For the 16-vertex lineage tree of the egg chamber, we have \( d_1 \in \{4, 4, 3, 3, 2, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1\} \),
yielding \( P = 144 \) distinct embeddings. Each of these embeddings is uniquely determined by the order in which the eight leaves (terminal nodes) of the tree, given by the nodes \( v = 9, \ldots, 16 \), are arranged. This allows us to assign a unique tree-state vector \( e \) to each embedding by starting with \( v = 16 \) and then listing the leaf indices in counter-clockwise order; for example, the sequence \( e = [16, 14, 9, 13, 15, 11, 10, 12] \) uniquely identifies the embedding realized in Fig. 1c. For each tree state \( e = [16, a, b, c, d, e, f, g] \) there exists a corresponding mirror-symmetric embedding \( e' = [16, g, f, e, d, c, b, a] \). When embedded in 3D, \( e \) and \( e' \) can give rise to packings of opposite handedness, reflecting the fact that pure rotations and reflections belong to different branches of the special orthogonal group. In general, it is not a priori that \( e \) and \( e' \) are biologically equivalent. However, to reduce the complexity of the TPP, we will treat mirror-symmetric packings as mathematically equivalent in the remainder, distinguishing only the \( P' = 144/2 = 72 \) embeddings \( e_{\text{un}} \) that are not related by rotations or reflections.

We first explored how frequently specific tree states \( e \) are realized in the egg chambers. Our \( n = 121 \) experimental samples suggest that certain tree configurations appear more frequently than others (Fig. 1f). To test the likelihood of observing our experimentally measured statistics under a uniform null-hypothesis, we numerically computed 100,000 realizations of 121 samples from a uniform distribution over the 72 distinct tree states. For each instance, we recorded the number of configurations observed more than \( f_{\text{c}} = 4 \) times (Fig. 1g) and the largest highest count (Fig. 1h), with the frequency cut-off chosen to be more than twice the expected frequency \( F = 121/72 \) under a uniform distribution. The resulting histograms for these two large-deviations statistics indicate that the experimentally observed data are highly atypical under the uniformity hypothesis (red bars in Fig. 1g,h).

As we show below, mathematical analysis of graph automorphisms and a detailed statistical mechanics model both strongly favour the non-uniformity hypothesis.

To quantify metric properties of the 3D tree embeddings, we estimated the cell–cell adjacency probabilities from the 121 rounded cysts (Supplementary Information). Adjacent cells can be connected by ring canals (red edges in Fig. 1c) or not (grey edges in Fig. 1e).

The adjacency probabilities extracted from the data confirm quantitatively the highly conserved layered structure of the 3D embedding and also reveal cell permutations within a layer across different spatial embeddings (Fig. 1i). The measured adjacency probabilities and the underlying contact graph topologies provide benchmarks for testing theoretical models.

To rationalize the experimental observations and provide a mathematical framework for convex TPPs, we analysed and compared conceptually related models of increasing complexity: first, we consider tree embeddings on convex equilateral polyhedra (Fig. 2). This purely geometric framework, which neglects energetic and cell-size effects, is well suited for clarifying the role of entropic effects. Subsequently, we generalize to an energy-based model by considering CLT embeddings on graphs arising from generalized Thomson packing problems for equal and unequal spheres on spherical surfaces (Fig. 3a,b). As we shall see, the energy-based approach reproduces many qualitative characteristics of the polyhedral model but yields better agreement for the adjacency statistics. Throughout, we focus on 16-vertex models as relevant to the experiments, although the underlying ideas generalize to arbitrary vertex numbers.

A minimal geometric model accounting for basic aspects of the CLT embeddings in rounded cysts identifies the cell positions with the vertices of a convex equilateral polyhedron. Convexity is required as the cells adhere to the epithelium, which forms a convex enclosure (Fig. 1a). Equilaterality assumes approximately equal cell diameters, a simplification that will be dropped later. The most famous polyhedra of this type are the 5 Platonic and 13 Archimedean solids, which satisfy, however, additional symmetry requirements that prohibit 16-vertex realizations. We therefore consider here the more general class of convex equilateral polyhedra with non-uniform faces, known as Johnson solids. There exist 92 different Johnson solids in total, with vertex numbers ranging from 5 to 75, but only four of these have exactly 16 vertices (Fig. 2a).

By comparing with topological features of the experimentally measured adjacency networks, we find that, among those four, the packing structure of the disphenocingulum \( J_{90} \) is approximately realized in some of the cysts, whereas the other three (\( J_{29} \) and \( J_{60} \)) show a
larger or smaller number of degree-4-vertices than observed in our experiments (Fig. 3c). Analysis of the $J_{90}$ solid helps explain why certain tree packings occur more frequently than others. To illustrate the underlying entropic argument, we computed all 5,184 possible CLT embeddings on $J_{90}$ and determined the frequencies of the 72 embedding states $v$. (Supplementary Information). The tree-state histogram shows that certain tree configurations can be realized by a larger number of microstates and hence have higher entropy; such highly degenerate macrostates are more frequently observed than low-entropy states. Analogous arguments apply to the refined energy-based models discussed next.

To obtain a more accurate description of the experimentally observed adjacencies, we introduce an energy-based model that generalizes the classic Thomson problem of arranging particles with electrostatic Coulomb repulsion on a sphere. The model represents the cell $v$ as a soft sphere of radius $r_v$ at position $x_v$, and the epithelium as a spherical container of radius $R$. Ring canals connecting cells $v$ and $w$ are modelled as harmonic springs of strength $k_c$. Similarly, adhesion to the epithelium and steric repulsion between adjacent spheres without ring canal connections are described by quadratic potentials, yielding the energy

**Fig. 3** Energy-based models confirm non-uniform CLT distributions and capture experimentally measured cell–cell adjacencies. **a**. The two lowest-energy solutions $T_0$ and $T_1$ of the quadratic Thomson problem, equation (2), for equally sized spheres exhibit the same discrete symmetries as those of the electrostatic Thomson problem. For equal-size spheres, our MC simulation runs ($n=10^4$) converged to either $T_0$ or $T_1$ with probability $p_0=0.86$ and $p_1=0.14$, in good agreement with the theoretical predictions $24/(24+4) \approx 0.857$ and $4/(24+4) \approx 0.143$ based on the automorphisms of the two polyhedra. Numerically obtained tree-state histograms for equal-size spheres (grey) approximate well the exact distributions representing the $p_i$-weighted average over all 62,256 CLT embeddings on $T_0$ and 94,344 embeddings on $T_1$ (green). These results also support an entropically driven departure from a uniform distribution over cell-tree configurations. **b**. A typical realization of the $n=4,000$ simulated CLT packings with sphere volumes matched to the experimental average values for the corresponding cells, showing layered cell arrangements consistent with experiments (see Fig. 1d). The simulated tree-state histogram indicates entropically favoured configurations. **c**. Experimental adjacency-graph characteristics are best reproduced with volume-matched spheres (empty circles), whereas the equal-sphere model (blue) underestimates the vertex degree. Among the four Johnson solids (red), the disphenocingulum $J_{90}$ is closest to the experimental data. **d,e**. Adjacency probability matrices for the volume-matched energy model (d) agree with the experimental data (see Fig. 1) within a maximal relative entry-wise spectral error (Supplementary Information) of less than 10%; corresponding maximal errors for the $J_{90}$ and equal-sphere Thomson models are larger by ~5% (Supplementary Information).
The \( k_t \) term represents steric repulsion between unconnected adjacent spheres with \( |x_u - x_{uv}| < r_u + r_v \), and the \( k_b \) term epithelial adhesion. Results presented below use \( k_b = 1 \), \( k_r = 0.2 \) and \( k_t = 0.3 \), reflecting the relative strength of the interaction forces, and were found to be robust against parameter variations. To identify tree embeddings and their statistics, we minimized equation (2) numerically, employing an annealed Metropolis–Hastings Monte Carlo (MC) algorithm\(^33,34\) with 3D random initial conditions (Supplementary Information).

The basic model assumes equal cell radii, \( r_u = r_v \), in equation (2). In this case, our 10,000 MC runs always converged to one of the two polyhedra \( T_0 \) and \( T_1 \) shown in Fig. 3a. \( T_0 \) and \( T_1 \) are energetically equivalent within 0.01% and have the same symmetries as the two lowest-energy solution of the classical electrostatic Thomson problem\(^23,25,26\). Their different occurrence probabilities \( p_0 = 0.86 \) and \( p_1 = 0.14 \) in the MC simulations reflect the cardinalities 24 and 4 of their graph automorphism sets; that is, the number of ways in which vertices can be relabelled without changing adjacencies. Intuitively, the automorphism sets determine the effective degeneracies of the two low-energy solutions, after factoring out trivial global rotations and translations. The tree-state histogram obtained from the MC simulations (grey in Fig. 3a) agrees well with the exact histogram obtained by determining all 62,256 and 94,344 CLT embeddings on \( T_0 \) and \( T_1 \), respectively (green in Fig. 3a; Supplementary Information), providing further support for the hypothesis that certain CLT embeddings are entropically favoured.

The Johnson solid \( J_0 \) and the Thomson graphs \( T_0 \) and \( T_1 \) do not yet have the correct degree distribution to reproduce the experimentally measured adjacency probabilities (Fig. 3c) and, in particular, the highly conserved layered structure observed in the egg chambers (Fig. 1d). The discrepancy is resolved by adapting the sphere radii \( r_u \) in equation (2) to match the experimentally measured average volumes of the corresponding cells in each layer (Supplementary Information). As for equal radii, the tree-state distribution remains non-uniform (Fig. 3b); this non-uniformity now arises due to a combination of entropic and energetic effects as differences in the cell radii \( r_u \) lift the degeneracies of the minima of equation (2). The associated adjacency probability matrix (Fig. 3d) agrees with the experimental data (Fig. 1i) within 8% relative error (Fig. 3c; Supplementary Information), corroborating that energy models of the type (2) define a useful theoretical framework for studying biologically relevant TPPs under convexity constraints.

To conclude, although our investigation focused on the insect model organism \textit{Drosophila}, encapsulated CLTs have also been reported in the germelines of amphibians, molluscs\(^27\), birds, humans and other mammals\(^28\), suggesting that tree-packing problems play a fundamental role at the onset of oogenesis and embryogenesis\(^29\) in a wide range of organisms. Entropic constraints can favour particular cell-packing configurations, similar to conformational entropy barriers in protein folding\(^30\). Since oocytes receive essential biochemical signals from adjacent nurse cells during development\(^31\), topologically supported relative cell localization may be important for reproducible oogenesis and morphogenesis. In particular, the presence of topological links arising from incomplete cytokinesis suggests that cell ordering and rearrangement processes can deviate strongly from the soap bubble paradigm used to describe patterns of organization of simple cell aggregates\(^32\), with potentially profound implications for tissue-scale organization and dynamics. More broadly, the insights from this study can provide useful guidance for the controlled self-assembly of topologically linked microstructures in confined geometries, as realizable with modern DNA origami\(^33\) and recently developed molecular linker toolboxes for multicellular self-assembly\(^34\).

Methods
Methods, including statements of data availability and any associated accession codes and references, are available at https://doi.org/10.1038/s41567-018-0202-0.

Received: 6 September 2017; Accepted: 8 June 2018; Published online: 16 July 2018

References
1. Conway, J. H. & Sloane, N. J. A. Sphere Packings, Lattices, and Groups 3rd edn (Springer, New York, NY, 1999).
2. Hales, T. C. et al. A revision of the proof of the Kepler conjecture. Discret. Comput. Geom. 44, 1–34 (2010).
3. Weaire, D. A short history of packing problems. Forma 14, 279–285 (1999).
4. Donev, A. et al. Improving the density of jammed disordered packings using ellipsoids. Science 303, 990–993 (2004).
5. Irvine, W. T. M., Vitelli, V. & Chaikin, P. M. Pleats in crystals on curved surfaces. Nature 468, 947–951 (2010).
6. Meng, G., Paulose, J., Nelson, D. R. & Manoharan, V. N. Elastic instability of a crystal growing on a curved surface. Science 343, 654–657 (2014).
7. Majikes, J. M., Nash, J. A. & LaBeau, T. H. Competitive annealing of multiple DNA origami: formation of chimeric origami. New J. Phys. 18, 115001 (2016).
8. Uhlmann, F. SMC complexes: from DNA to chromosomes. Nat. Rev. Mol. Cell Biol. 17, 399–412 (2016).
9. Dayel, M. J. et al. Cell differentiation and morphogenesis in the colony-forming choanoflagellate Salpingoeca rosetta. Dev. Biol. 357, 73–82 (2011).
10. Haglund, K., Neizis, I. P. & Stenmark, H. Structure and functions of stable intercellular bridges formed by incomplete cytokinesis during development. Commun. Integr. Biol. 4, 1–9 (2011).
11. Johnson, N. W. Convex solids with regular faces. Can. J. Math. 18, 169–200 (1966).
12. Hayashi, T. & Carthew, R. W. Surface mechanics mediate pattern formation in the developing retina. Science 341, 647–652 (2004).
13. Höhn, S., Honerkamp-Smith, A. R., Haas, P. A., Khuc Trong, P. & Goldstein, R. E. Dynamics of a volvox embryo turning itself inside out. Phys. Rev. Lett. 114, 178101 (2015).
14. Lieberman-Aiden, E., et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science 326, 289–293 (2009).
15. Tomer, R., Khairy, K., Amat, F. & Keller, P. J. Quantitative high-speed imaging of entire developing embryos with simultaneous multiview light-sheet microscopy. Nat. Methods 9, 755–763 (2012).
16. Ohnesorge, F. & Binning, G. True atomic resolution by atomic force microscopy through repulsive and attractive forces. Science 260, 1451–1456 (1993).
17. Manoharan, V. N., Elsesser, M. T. & Pine, D. J. Dense packing and symmetry of small clusters of microparticles. Science 301, 483–487 (2003).
18. Fletcher, A. J. J. Dry Stone Walling: History and Heritage (Amberley, Stroud, 2016).
19. Manoharan, V. N. Colloidal matter: packing, geometry, and entropy. Science 349, 1253751 (2015).
20. Lauga, E. & Brenner, M. P. Evaporation-driven assembly of colloidal particles. Phys. Rev. Lett. 93, 238301 (2004).
21. Teich, E. G., van Anders, G., Klotza, D., Dhemuchadse, J. & Glotzer, S. C. Clusters of polyhedra in spherical confinement. Proc. Natl Acad. Sci. USA 113, E669–E678 (2016).
22. Doig, A. J. & Sternberg, M. J. Side-chain conformational entropy in protein folding. Protein Sci. 4, 2247–2251 (1995).
23. Glass, D. S. & Riedel-Kruse, I. H. A genetically encoded adhesin toolbox for programming multicellular morphologies and patterns. Preprint at https://doi.org/10.1011/240721 (2017).
24. Inman Alsonu, J., Villoutreix, P., Berezkhkovski, A. M. & Svartsman, S. Y. Collective growth in a small cell network. Curr. Biol. 27, 2670–2676 (2017).
25. Gilbert, S. F. Developmental Biology 8th edn (Sinauer Associates, Sunderland, MA, 2006).
26. Turing, A. M. The chemical basis of morphogenesis. Phil. Trans. R. Soc. B 237, 37–72 (1952).
27. Cartwright, J. & Arnold, J. M. Intercellular bridges in the embryo of the Atlantic squid, Loligo pealei. Development 57, 3–24 (1980).
28. Green, K. J., Viamontes, G. L. & Kirk, D. L. Mechanism of formation, ultrastructure, and function of the cytoplasmic bridge system during morphogenesis in Volvox. J. Cell. Biol. 91, 756–769 (1981).
29. Becalska, A. N. & Gavis, E. R. Lighting up mRNA localization in Drosophila oogenesis. Development 136, 2493–2503 (2009).
30. Green, K. J. & Kirk, D. L. Cleavage patterns, cell lineages, and development of a cytoplasmic bridge system in Volvox embryos. J. Cell. Biol. 91, 743–755 (1981).
31. Cromwell, P. R. Polyhedra (Cambridge Univ. Press, Cambridge, 1997).
32. Thomson, J. J. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. Philos. Mag. 7, 237–265 (1904).
33. Allen, M. P. & Tildesley, D. J. Computer Simulation of Liquids (Oxford Univ. Press, Oxford, 1989).
34. Gould, H. & Tobochnik, J. Statistical and Thermal Physics: With Computer Applications (Princeton Univ. Press, Princeton, NJ, 2010).
35. Rakhmanov, E. A., Saff, E. B. & Zhou, Y. M. In Computational Methods and Function Theory 1994 (Penang) Vol. 5 (eds Ali, R. M., Ruscheweyh, S. & Saff, E. B.) 293–309 (World Scientific, River Edge, NJ, 1995).
36. Atiyah, M. & Sutcliffe, P. Polyhedra in physics, chemistry and geometry. Milan J. Math. 71, 33–58 (2003).
37. Erber, T. & Hockney, G. M. In Advances in Chemical Physics Vol. XCVIII (eds Prigogine, I. & Rice, S. A.) 495–594 (Wiley, New York, NY, 1997).

Acknowledgements
We thank G. Laevsky for expert help with imaging, and E. Gavis, A. Spradling, T. Orr-Weaver, F. Nijhout, L. Manning, H. Mattingly, Y. H. Song, T. Stern, S. Okuda and C. Doherty for helpful discussions. This work was supported by the National Science Foundation Science and Technology Center for Emergent Behaviors of Integrated Cellular Systems (J.I.A. and S.Y.S.), the NIH R01GM107103 (S.Y.S.), the WIN programme between Princeton University and the Weizmann Institute (P.V.), a James S. McDonnell Foundation Complex Systems Scholar Award (J.D.) and an Edmund F. Kelly Research Award from the MIT Department of Mathematics (J.D.). This research was partially supported by the Allen Discovery Center programme through The Paul G. Allen Frontiers Group.

Author contributions
All authors designed the research. J.I.A. performed the experiments. J.I.A., P.V. and N.S. analysed the data. N.S. and J.D. developed the theory. N.S. performed the simulations. All authors wrote the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41567-018-0202-0.
Reprints and permissions information is available at www.nature.com/reprints.
Correspondence and requests for materials should be addressed to J.D.

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Methods

Tree state combinatorics. To derive equation (1), we start from cell \( v = 1 \) in Fig. 1b that has vertex degree \( d_1 = 4 \) and fix the edge \((1, 2)\) as a reference axis. The remaining three edges \{(1, 3), (1, 5), (1, 9)\} emanating from cell 1 can then be arranged in \((d_1 - 1)! = 3!\) different ways relative to edge \((1, 2)\). Next consider the four cells \( v = 2, 3, 5, 9 \) that are topologically connected to cell 1 and coloured in blue in Fig. 1b,c. Cell 2 has vertex degree \( d_2 = 4 \) and the three emanating edges \{(2, 10), (2, 6), (2, 4)\} can be permuted in \((d_2 - 1)! = 3!\) possible ways relative to the incoming edge \((1, 2)\). Similarly, the two edges emanating from cell 5 can be permuted in \((d_5 - 1)! = 2!\) possible ways relative to the incoming edge \((1, 5)\), and so on. Repeating this procedure for each tree layer and multiplying all the permutations yields equation (1).

Code availability. All used codes can be downloaded from https://github.com/stoopn/CLTPackings.

Data availability. The experimental data are publicly accessible at https://github.com/stoopn/CLTPackings.
Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see Reporting Life Sciences Research. For further information on Nature Research policies, including our data availability policy, see Authors & Referees and the Editorial Policy Checklist.

1. Sample size
   Describe how sample size was determined.

   The manually counted sample size n=81 was large enough to determine a 1 sigma deviation from uniformity as shown in Fig 1f, g

2. Data exclusions
   Describe any data exclusions.

   Data (three-dimensional confocal images) was only excluded in cases were it was not possible to identify and/or label the individual cells within any given germline cyst. This would have made it impossible to deduce the germline cysts’ spatial arrangement. This had the effect of reducing our potential sample size.

3. Replication
   Describe whether the experimental findings were reliably reproduced.

   No attempts at replication failed. The data was collected over multiple sessions using the same antibodies, microscope and analyzed using the same software (Bitplane’s Imaris) without any issues. Each step of the process (dissection of germline cysts, immunostaining, microscopy, labeling and reconstructions) was developed and optimized independently.

4. Randomization
   Describe how samples/organisms/participants were allocated into experimental groups.

   This isn’t relevant to our study. We collected data from 81 germline cysts, identified and labeled the cells in each, and deduced the arrangement of the cells. Our theoretical analysis in the paper relates to those data points.

5. Blinding
   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

   Not applicable

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.
6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

- The exact sample size \((n)\) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. \(P\) values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on statistics for biologists for further resources and guidance.

7. Software

Describe the software used to analyze the data in this study.

We used Bitplane’s Imaris to visualize our confocal microscopy data, to identify and to annotate the cells in the germline cysts, and to create the 3D surface reconstructions. No custom code was written for these tasks.

Matlab codes used to perform simulations and analyze the simulated as well as experimental data can be found on a public repository at: https://github.com/stoopn/CLTPackings

This includes the two excel sheets with data for adjacencies and 1d descriptors for the germline cysts

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). Nature Methods guidance for providing algorithms and software for publication provides further information on this topic.

8. Materials and reagents

Policy information about availability of materials

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

Primary antibodies:
Sheep anti-GFP (1:1000, obtained from E. Wieschaus Lab, Princeton University), mouse anti-Hts (1:500, Developmental Studies Hybridoma Bank (DSHB)), rabbit anti-PTyr (1:500, Santa Cruz Biotechnology).

Secondary antibodies:
Alexa-Fluor goat anti-rabbit 546nm, goat anti-sheep 488nm, and goat anti-mouse 647nm (1:300).

All primary antibodies were either verified in prior co-staining/co-localization experiments or verified in separate independent studies.
10. Eukaryotic cell lines
a. State the source of each eukaryotic cell line used.
   - No cell lines were used
b. Describe the method of cell line authentication used.
   - No cell lines were used
c. Report whether the cell lines were tested for mycoplasma contamination.
   - No cell lines were used
d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by ICLAC, provide a scientific rationale for their use.
   - No cell lines were used

- Animals and human research participants

Policy information about studies involving animals; when reporting animal research, follow the ARRIVE guidelines

11. Description of research animals
   Provide details on animals and/or animal-derived materials used in the study.
   - We used Drosophila melanogaster (fruit fly) as our model organism. The strains used were either wild type (WT) or Riselle-GFP tagged flies.

Policy information about studies involving human research participants

12. Description of human research participants
   Describe the covariate-relevant population characteristics of the human research participants.
   - The study did not involve human research participants