**Supplement 1:
Simulation of collective cellular behavior using an extended Dynamic Vertex Model**

Vertex-based models1-3 have been successful in describing cell shape statistics in developing embryos4-7 and glassy and jamming behavior in cultured epithelial sheets8-11. These computational models are simple yet contain the salient features that offer useful means to investigate the role of cell-cell interactions and how single cell behavior is connected to emergent collective behavior at the multicellular level.

The framework of this model is based on the observation that the cells comprising epithelial sheet tend to maintain a columnar structure and the 2D cross-section of the epithelial sheet forms a polygonal tiling in the plane1, 2. The apical surface of cells is then specified by the location of vertices, as well as the geometry of edges that connect the vertices. In this model, the edges represent the conformation of the cell-cell junctions. Unlike classical implementation of the vertex model, here we allow the edges to be curved as an emergent property (as detailed below). The cells in such a tiling emerge as *curved* polygons of different shapes with well-defined area ($A$) and perimeter ($P$).

****

Figure S1. Schematic for the dynamic vertex model which shows a part of the simulated tissue. Left, the net propulsion (solid red arrow, definition Eq. 3) on a given vertex (red dot) has a magnitude $v\_{0}$ and direction derived by the polarizations of three cells sharing the vertex, given by $\vec{n\_{1}}$, $\vec{n\_{2}}$ and $\vec{n\_{1}}$ (solid blue arrows). Right, the active force on a mid-edge vertex $k$ given by $\vec{f\_{k}}$ (solid green arrow, definition Eq. 5) which depends on the active forces on the head vertex ($\vec{f\_{k,1}}$) and tail vertex ($\vec{f\_{k,2}}$) (solid red arrows), $l\_{k}$ being the end-to-end length of the edge carrying $k$.

The classical vertex model: Beginning with the network of vertices and connecting edges, vertex models then describe the mechanical energy of each cell written as function of cell area (Ai) and perimeter (Pi) of the $i$-th cell:

$E\_{i}=K\_{P}P\_{i}^{2}+λ\_{i}P\_{i}+K\_{A}(A\_{i}-A\_{0}^{i})^{2}$ (1)

The first term describes the elastic energy required to deform the cell cortex and the contractile elements embedded within the cell cortex. This term is quadratic in cell-perimeter, $K\_{P}P\_{i}^{2}$ with elastic constant $K\_{P}$. The second term describes the energetic effects arising from cell-cell adhesion and cortical contractile tension. This term is linear in cell perimeter, $λ\_{i}P\_{i}$. The coefficient $λ\_{i}$ can be reduced in two distinct ways 4, 12: (1) by increasing the homotypic cadherin bonds on cell-cell junctions or (2) by diminishing cortical contractile tension. Both scenarios would lead to a longer perimeter for the cell4, 12.

The first two terms in Eq.1 can be combined into the well-known form4, 9: $K\_{P}(P\_{i}-P\_{0}^{i})^{2}$ where $P\_{0}^{i}=-{λ\_{i}}/{\left(2K\_{P}\right)}$, an effective preferred perimeter13 of the $i$-th cell. The third term in Eq 1 is associated with the energy cost for changing the cross-sectional area of the$ i$-th cell from its rest area $A\_{0}^{i}$ with an elastic constant $K\_{A}$. This energy arises due to cytoskeletal elasticity and cell deformations that change cell area away from $A\_{0}^{i}$ 5. Together the perimeter and area energy terms give rise to the vertex model energy function4, 5, 13 for $N$ cells:

$E=\sum\_{i=1}^{N}\left[K\_{P}(P\_{i}-P\_{0})^{2}+K\_{A}(A\_{i}-A\_{0})^{2}\right]$ (1a)

For simplicity, here we treat all the cells as being mechanically identical. We can then rewrite Eq 1a in a dimensionless form as follows,

$\tilde{E}=\sum\_{i=1}^{N}\left[(p\_{i}-p\_{0})^{2}+ \tilde{K}\_{A}(a\_{i}-1)^{2}\right]$ (1b)

wherein $K\_{P}A\_{0}$ is taken as the natural energy scale and $\sqrt{A\_{0}}$ is taken as the natural length scale. The rescaled tissue energy, cell perimeter and cell area now become: $\tilde{E}=E/(K\_{P}A\_{0})$, $p\_{i}={P\_{i}}/{\sqrt{A\_{0}}}$, $a\_{i}={A\_{i}}/{A\_{0}}$, respectively, and $\tilde{K}\_{A}= K\_{A}A\_{0}/K\_{P}$. The rescaled preferred cell perimeter becomes $p\_{0}={P\_{0}}/{\sqrt{A\_{0}}}$ which is also called the target cell shape index14. When there is mismatch between the actual length of cell boundary ($p\_{i}$) and this preferred value ($p\_{0}$), tensions arise on cell edges. If the rescaled perimeters of two cells $i$ and $j$ sharing an edge are $p\_{i}$ and $p\_{j}$, respectively, then the tension on the shared edge is given by:

$T\_{ij}=\left(p\_{i}-p\_{0}\right)+\left(p\_{j}-p\_{0}\right)$. (2)

The Dynamic Vertex Model (DVM): To describe cellular migration, the rate of change of position of the $i$-th vertex, $\vec{r}\_{i}$ , at time $t$ is represented by the overdamped equation of motion,

$Γ\frac{d\vec{r}\_{i}}{dt}=\vec{F}\_{i}^{int}+v\_{0}\sum\_{j\leftrightarrow i}^{}w\_{j}^\_{j}$ (3)

where, $Γ$ is describes frictional damping. $\vec{F}\_{i}^{int}$ describes those forces that arise due to cell-cell interactions, and set by a spatial gradient in tissue mechanical energy: $-{∂E}/{∂\vec{r}\_{i}}$. The last term represents an active motility force $\vec{f\_{i}}$ on vertex $i$ which has a magnitude $v\_{0}$. Its direction is set by a weighted average of the polarization vectors of cells adjacent to vertex $i$, unlike a flat sum over the associated cell polarizations as described before8, 15, 16. The weight factor $w\_{j} $for cell *j* (polarization $^\_{j}$) is given by: $w\_{j}={l\_{j}}/{(2ZΣ\_{j\leftrightarrow i}}l\_{j})$ where $l\_{j}$ is the sum of the lengths of the two edges shared by cell $j$ and vertex $i$, $Z$ is the connectivity of vertex $i$ and the division by 2 is done to avoid double counting. This ensures that the active force on any vertex gets the largest contribution from the neighboring cell that contains the longest edges connected to the vertex. We choose the cell polarization vectors in spirit of recent models of self-propelled particles17-22: $^\_{j}=\left(cosα\_{j}, sinα\_{j}\right)$ where the angle of polarization $α\_{j}$ follows over-damped dynamics according to:

$\frac{dα\_{j}}{dt}=η\_{j}, \left〈η\_{j}\left(t\right)η\_{k}\left(t'\right)\right〉= 2D\_{R}δ\_{jk}δ\left(t-t'\right).$ (4)

This description models the front-back polarity that drives motility in migrating cells14, 23, 24. Each polarization vector is subjected to a Gaussian random noise $η\_{j}$ with zero mean and variance $2D\_{R}$. This noise introduces a persistence to the cell-polarization with a period $1/D\_{R}$ which we keep unchanged in all our simulations. The force generation on any vertex is shown schematically in Fig. S1.

Curving cell-cell boundaries: In this expanded version of the DVM, we now allow each edge to curve, as, as occurs during pEMT, by introducing a mid-edge vertex which provides an additional degree of freedom on each edge. Like the other vertices, each mid-edge vertex is subjected to active forces derived from the cell motility forces on the neighboring vertices. For the mid-edge vertex $k$ on a given edge, the active force $\vec{f}\_{k}$ is calculated using the active forces on the head and tail vertices of the edge using the following:

 $\vec{f}\_{k}=\frac{1}{2}\left[\left(1-\frac{l\_{1}}{l\_{k}}\right)\vec{f}\_{k,1}+\left(1-\frac{l\_{2}}{l\_{k}}\right)\vec{f}\_{k,2}\right]$ (5)

where $\vec{r}\_{k}$ is the position vector of the mid-edge vertex, $\vec{f}\_{k,1}$, $\vec{r}\_{k,1}$ and $\vec{f}\_{k,2}$, $\vec{r}\_{k,2}$ are the active forces and position vectors of the head and tail vertices, respectively (Fig. S1). Also, $l\_{1}=\left|\vec{r}\_{k}-\vec{r}\_{k,1}\right|$, $l\_{2}=\left|\vec{r}\_{k}-\vec{r}\_{k,2}\right|$ represent the Euclidian distances of the head and tail vertices from the mid-edge vertex respectively, and $l\_{k}=l\_{1}+l\_{2}$, essentially the end-to-end length of the edge.

We estimate the contour length of an edge from an arc that represents the edge. The parametric equation for any point on the arc representing an edge is given by:

$\vec{R}\left(s\right)= \vec{r}\_{k,1}s{(s-1)}/{2}+ \vec{r}\_{k}\left(1-s\right)\left(1+s\right)+\vec{r}\_{k,2}s{(s+1)}/{2}$ (6)

where $-1\leq s\leq 1$. This is implemented in the surface evolver program25. Because we keep the mean cell area fixed and cell-to-cell variations of area are small, pressure differences across cell edges are negligible. Therefore, when $p\_{0} $is small, the tissue is deep in the solid state, and the edges are under high tension and hence straight. However, for large $p\_{0}$ the edge tension diminishes, and edges curve to accommodate the large perimeter.

Simulation details: We initialized each simulation using independent Voronoi tessellation of randomly placed points posing as cell centers. This gives rise to a confluent polygonal tiling that is random. Then we minimized tissue energy (Eq 1b) to find a state with energy close to the ground state energy at zero motility ($v\_{0}=0$) using the conjugate gradient protocol with respect to the vertex locations. Next, we allow finite motility i.e. $v\_{0}>0$ and performed the dynamical simulations. The vertex locations are updated by the Euler method using Eq. 3 with a timestep of $Δt= 10^{-2}τ$ where $τ={Γ}/{K\_{P}}$, the unit of time in the DVM. All lengths in DVM was measured in unit of $\sqrt{A\_{0}} $ where we use $A\_{0}=\overbar{A}$, the mean area of the cells which is maintained at unity throughout any simulation. The rotational noise on the direction of cell polarizations is given by Eq. 4 where $D\_{R}={1}/{τ}$ , kept fixed throughout the study.

We determined the effects of independently varying the magnitude of cellular motility force, $v\_{0}$ and the preferred cell perimeter, $p\_{0}$. When keeping all other parameters fixed, we found that increasing only $p\_{0}$ recapitulated observations made during pEMT, while increasing only $v\_{0}$ recapitulated observations made during UJT. Note that each of our simulations depends on a single set of $v\_{0}$ and $p\_{0}$ values which do not change during the simulation. Moreover, for each set of $v\_{0}$ and $p\_{0}$ values we ran 20 different simulations from independent initial configurations. Thus, all the error bars associated with data from the simulations (Fig. 4 and Fig. S4) are SOM, generated from these 20 different independent trajectories.

From the 20 simulations corresponding to each set of $v\_{0}$ and $p\_{0}$ values, we calculated the following average quantities across the cells in the simulated tissue: edge tension (Eq. 2), edge tortuosity (Eq. 5), aspect ratio, cell shape index8, 14 (defined for cell $i$ as $q\_{i}={P\_{i}}/{\sqrt{A\_{i}}}= {p\_{i}}/{\sqrt{a\_{i}}}$), velocities and effective diffusivities. The aspect ratio of any cell in the tissue was calculated from the eigenvalues of the shape tensor10, generated using the positions of the vertices of the cell.

In a confluent epithelial sheet the basic mode of migration is T1 transitions26 in which cells swap positions with their local neigbors27, 28. In the DVM we allow T1 rearrangements by implementing an embargo timer on each cell given by parameter $τ\_{T1}$. This sets a lower limit on the time between successive T1s involving a cell. We chose appropriate values for $τ\_{T1}$ in different simulation scenarios. For simulations with $v\_{0}>0.05$ (associated with UJT) we use $τ\_{T1}=τ$, while we use a much larger value $τ\_{T1}=10^{3}τ$ for the simulations at $v\_{0}=0.05$ (associated with pEMT where T1 processes were rarely observed in the experiments). The threshold length for T1 edge swap was fixed at $l\_{c}=0.1$ in simulation length unit. We used periodic boundary conditions in both x and y directions on our simulation box containing $N=400$ cells. All our simulations were implemented using the Surface-Evolver program25.

**References**

1. Nagai, T. & Honda, H. A dynamic cell model for the formation of epithelial tissues. *Philosophical Magazine Part B* **81**, 699-719 (2001).

2. Fletcher, A.G., Osterfield, M., Baker, R.E. & Shvartsman, S.Y. Vertex models of epithelial morphogenesis. *Biophysical journal* **106**, 2291--2304 (2014).

3. Honda, H. Description of cellular patterns by Dirichlet domains: The two-dimensional case. *Journal of Theoretical Biology* **72**, 523-543 (1978).

4. Farhadifar, R., Roeper, J.-C., Aigouy, B., Eaton, S. & Julicher, F. The Influence of Cell Mechanics, Cell-Cell Interactions, and Proliferation on Epithelial Packing. *Current Biology* **17**, 2095-2104 (2007).

5. Hufnagel, L., Teleman, A.A., Rouault, H., Cohen, S.M. & Shraiman, B.I. On the mechanism of wing size determination in fly development. *Proceedings of the National Academy of Sciences* **104**, 3835-3840 (2007).

6. Brodland, G.W. *et al.* Video force microscopy reveals the mechanics of ventral furrow invagination in Drosophila. *Proceedings of the National Academy of Sciences* **107**, 22111--22116 (2010).

7. Chiou, K.K., Hufnagel, L. & Shraiman, B.I. Mechanical Stress Inference for Two Dimensional Cell Arrays. *PLOS Computational Biology* **8**, 1-9 (2012).

8. Park, J.-A. *et al.* Unjamming and cell shape in the asthmatic airway epithelium. *Nat Mater* **14**, 1040--1048 (2015).

9. Staple, D.B. *et al.* Mechanics and remodelling of cell packings in epithelia. *Eur. Phys. J. E* **33**, 117-127 (2010).

10. Atia, L. *et al.* Geometric constraints during epithelial jamming. *Nature Physics* **14**, 613-620 (2018).

11. Malinverno, C. *et al.* Endocytic reawakening of motility in jammed epithelia. *Nat Mater* **16**, 587-596 (2017).

12. Manning, M.L., Foty, R.A., Steinberg, M.S. & Schoetz, E.-M. Coaction of intercellular adhesion and cortical tension specifies tissue surface tension. *Proceedings of the National Academy of Sciences* **107**, 12517--12522 (2010).

13. Bi, D., Lopez, J.H., Schwarz, J.M. & Manning, M.L. A density-independent rigidity transition in biological tissues. *Nature Physics* **11**, 1074--1079 (2015).

14. Bi, D., Yang, X., Marchetti, M.C. & Manning, M.L. Motility-driven glass and jamming transitions in biological tissues. *Phys Rev X* **6** (2016).

15. Czajkowski, M., Sussman, D.M., Marchetti, M.C. & Manning, M.L. Glassy Dynamics in Models of Confluent Tissue with Mitosis and Apoptosis. *eprint arXiv:1905.01603*, arXiv:1905.01603 (2019).

16. Hirashima, T. & Adachi, T. Anisotropic Cellular Mechanoresponse for Radial Size Maintenance of Developing Epithelial Tubes. *bioRxiv*, 172916 (2017).

17. Fily, Y., Henkes, S. & Marchetti, M.C. Freezing and phase separation of self-propelled disks. *Soft Matter* **10**, 2132-2140 (2014).

18. Fily, Y. & Marchetti, M.C. Athermal Phase Separation of Self-Propelled Particles with No Alignment. *Phys. Rev. Lett.* **108**, 235702 (2012).

19. Cates, M.E. & Tailleur, J. Motility-Induced Phase Separation. *Annual Review of Condensed Matter Physics* **6**, 219-244 (2015).

20. Chate, H., Ginelli, F., Gregoire, G. & Raynaud, F. Collective motion of self-propelled particles interacting without cohesion. *Phys Rev E Stat Nonlin Soft Matter Phys* **77**, 046113 (2008).

21. Garcia, S. *et al.* Physics of active jamming during collective cellular motion in a monolayer. *Proc Natl Acad Sci U S A* **112**, 15314-15319 (2015).

22. Bechinger, C. *et al.* Active Particles in Complex and Crowded Environments. *Reviews of Modern Physics* **88** (2016).

23. Mayor, R. & Etienne-Manneville, S. The front and rear of collective cell migration. *Nat Rev Mol Cell Biol* **17**, 97-109 (2016).

24. Szabo, A. *et al.* Collective cell motion in endothelial monolayers. *Physical biology* **7**, 046007 (2010).

25. Brakke, K.A. The surface evolver. *Experiment. Math.* **1**, 141-165 (1992).

26. Weaire, D. & Rivier, N. Soap, Cells and Statistics - Random Patterns in 2 Dimensions. *Contemp Phys* **25**, 59-99 (1984).

27. Bertet, C., Sulak, L. & Lecuit, T. Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature* **429**, 667-671 (2004).

28. Zallen, J.A. & Wieschaus, E. Patterned gene expression directs bipolar planar polarity in Drosophila. *Dev Cell* **6**, 343-355 (2004).