Spatial model of cell-fate choice uncovers strong links between tissue morphology and tissue regeneration

Somya Mani[1,✉] **and Tsvi Tlusty**[1,✉]

[1]IBS-CSLM

Tissues in multicellular organisms are immensely diverse: animal tissues include sheet-like epithelia, bundles of syncitial muscle fibres and the branched and interconnected nerves, while plants contain sheet-like epidermis and highly organized bundles of vascular tissue. However, at the microanatomical level, tissues are notably similar in that they are organized into distinct domains: domains are characterized by their cellular compositions and hold precise adjacency relations among each other. These morphological similarities are surprising because multicellularity has evolved multiple times independently. Separately, tissues also hold a remarkable functional similarity: across all multicellular organisms, including poor regenerators such as mammals, tissues routinely heal from injuries. The cellular organization within tissues, as well as the ability regenerate result from developmental processes: cells divide, die, differentiate and migrate according to cues they receive from their neighborhoods.

We ask two interlinked questions: What diversity of tissue morphology can simple developmental processes generate? And is tissue morphology related to tissue regeneration? We address these questions using an agent based model of cell-fate decisions where cells use simple rules to respond to their cellular neighborhoods

Our model produces a rich diversity of tissue morphologies: By simply tuning the density of cellular interactions and the propensity of cellular differentiation, we produce tissues that go from disordered and sparse to tissues organized into dense and contiguous domains. Importantly, tissue morphology was strongly linked to regeneration in the model: the ability to heal was highly enriched in densely packed, contiguous tissues. Moreover, the predominant mode of tissue healing in the model recapitulates natural mechanisms: tissues healed through the replacement of injured cells through cell-division in adjacent regions. Our work generates experimentally testable predictions on the effects of manipulating cellular interactions on tissue morphology and in turn, on tissue regeneration.

Cell-fate | cell differentiation | cell migration | cellular neighborhood | tissue domains | tissue regeneration

Correspondence: ***somyamn@gmail.com***

Introduction

Tissues are systems of cells of one or more cell-types that together perform specific functions beyond what any single cell is capable of (1). Tissue level organization is widely distributed across multicellular organisms: Among animals, Cnidarians are the most basal lineage that have tissue level organization (2). Tissues are also present in other evolutionary lineages that invented multicellularity independently: land plants (3), brown algae (4) and fungi (5). These organisms vary widely in the number and complexity of their tissues: For example, while the Cnidarian Hydra contains two tissue layers that have 3-4 cell-types (6), complex animals like humans contain at least sixty tissues that are composed of tens of cell-types(7).

In plants, whose cells are non-motile, all cell-types composing a tissue are ontogenetically closely related, for example, all cell-types in the plant leaf derive from a common progenitor in the shoot apical meristem (8). In contrast, cell-types composing a tissue in animals can originate from distinct developmental lineages: for example, nematocytes and epitheliomuscular cells in the Hydra ectoderm are produced by distinct progenitors (6), in humans, many tissues contain tissue-resident lymphocytes that are produced by hematopoietic stem cells (9).

Irrespective of the cellular complexity of tissues and developmental origins of their cell-types, cells within a tissue interact closely with one-another using molecular signals and mechanical cues. These inter-cellular interactions shape cellular decisions to divide, die, differentiate or migrate (10). These cellular decisions in turn dictate tissue form and function (11): for example, in the mammalian inner ear, interactions between ear hair cells and support cells lead to very precise patterning that are essential for hearing (12), while in plant root meristems, intercellular interactions with cells in the quiescent center regulate differentiation of surrounding cells (13). In general, tissues across organisms tend to organize into distinct microanatomical domains that are characterized by their cellular compositions and form niches for inter-cellular interactions (14–16).

In this work, we explore how the control of cell-fate decisions by inter-cellular interactions might regulate cellular organization within tissues. Additionally, cellular interactions are also essential for regeneration of tissues from injuries : For example, in mammalian muscles (17) and spinal cord (18), and in plants (19). We therefore also ask how tissue organization is related to the ability of tissues to regenerate: Importantly, we ask whether the requirement for tissue regeneration might impose constraints on tissue forms.

We approach these questions using an agent-based model of cell-fate decisions: Cells in the model are housed in a two-dimensional grid, and interactions among neighboring cells

DRAFT

result in cellular decisions to differentiate, migrate, divide or die. We sample a wide range of developmental rules to produce a rich variety of tissue-level organizations: Similarly to real tissues, model-generated tissues were arranged into distinct domains that are characterized by their cellular compositions. The model's tissues varied widely in the relative arrangement of these domains: The variety of tissues spanned those with sparse cellular distribution to full and densely occupied grids, and tissues with small, disperse domains to large, contiguous domains. Interestingly, we found tissue regeneration to be strongly linked with tissue form: almost all regenerative tissues were densely occupied with cells which formed large and contiguous domains. Moreover, the model's regenerative tissues healed predominantly via a localized response: cells lost upon injury were replaced through cell-division and differentiation of neighboring cells. Importantly, our model's predictions are amenable to experimental testing: the model's parameters represent elementary developmental processes such as cellular signaling, which can be experimentally manipulated in widely studied tissue systems across different multicellular groups: For example, basal Metazoa such as Hydra (20),*in vitro* systems such as mammlian intestinal organoids (21), plant systems such as the Arabidopsis root meristem (22) and brown algae such as *Saccharina latissima*(23).

Model of cellular interactions and cell-fates

A tissue in the model is composed of N distinct cell types. These cells are distributed across the tissue's body: In the model, the tissue is housed in a 2-dimensional grid with G^2 grid-squares and periodic boundaries (Fig.1(B)). Multiple cells are allowed to reside within any grid-square: We represent a cell of cell-type $n \in \{1,2,3,...N\}$ in a grid-square $g \in \{1,2,3,...G^2\}$ as $c(n,g)$, and $C^{n,g}$ gives the number of cells of cell-type n in the grid-square g.

The geometry of the body is given by a distance matrix D: $D(g_a,g_b) = d \in \mathbb{W}$ implies the grid-squares g_a and g_b are a distance d apart. Specifically, g_a and g_b are *adjacent* if $D(g_a,g_a) \leq 1$. In a 2-dimensional grid, each grid-square is adjacent to five grid-squares including itself (Fig.1(B)).

Cells in adjacent grid-squares are called *neighbors*. The *neighborhood* of any cell c in grid-square g is given by a N-length binary vector H: $H(n) = 1$ if a cell of cell-type n is a neighbor of cell c; $H(n) = 0$ otherwise. There are 2^N distinct possible neighborhoods, which can be arranged in the standard binary ordering: $H^0 = [0,0,0...,0]$, $H^1 = [0,0,0...,1]$, ..., $H^{2^N-1} = [1,1,1...,1]$.

By allowing multiple cells to reside within the same grid-square, our model presents a coarser picture of cellular arrangements within tissues compared to standard models where the location, geometry and neighborhood of each cell (such as in cellular automaton (24) and vertex models (25)) or even subparts of cells (such as in the cellular Potts model (26)) are explicitly described. Thus our formalism portrays a more general representation of tissues than cellular automaton models or vertex models since we do not impose any constraints on the geometries or adjacencies of individual cells (27). At the same time, the simplicity of our model as compared to the cellular Potts model and lattice-free models (such as in (28, 29)) makes it computationally feasible to sample a broad space of tissue patterns.

Tissue developmental rules. Cellular interactions are prescribed by the interaction matrix I (Fig.1(A), Methods: Generating tissue developmental rules): A cell of type n_i in square g_a interacts with a cell of type n_j in square g_b if $I(n_i,n_j) \neq 0$ and g_a and g_b are adjacent. That is, cellular interactions in the model are medium-range and a cell can interact with another cell some tens of cells away. We do not impose any constraints on the form of cellular interaction and do not distinguish between molecular and mechanical signals (30). The model's cellular interactions can represent autocrine, juxtacrine and paracrine signaling, which are known to be important during tissue development (31), for example, during embryonic development across animals (32) and during the growth and development of tissues in the plant shoot (14). The model's cellular interactions can also represent mechanical signals (33): an interaction between two cells in the model can represent the local modification of the extracellular matrix by one of the cells which impacts the behavior of the second cell. For example, during skull development in mice, differentiated osteoblasts secrete collagen into the extracellular matrix induces differentiation of nearby mesenchymal cells (34).

We call a cellular interaction *stabilizing* if $I(n_i,n_j) = 1$ and *destabilizing* if $I(n_i,n_j) = -1$. Particularly, a cell-type n_i is *intrinsically stable* if $I(n_i,n_i) \geq 0$ and *intrinsically unstable* if $I(n_i,n_i) = -1$. Intrinsically stable cells are destabilized in the presence of any destabilizing interactions and intrinsically unstable cells are stabilized in the presence of any stabilizing interactions. For example, the germline stem cells (GSC) in the worm *Caenorhabditis elegans* are intrinsically unstable and tend to differentiate into germ cells. Interaction of GSCs with distal tip cells stabilizes them by repressing their differentiation (35).

A stable cell maintains its cell-type and its position within the body, whereas an unstable cell either differentiates into a new cell-type or migrates to a new grid-square. In the model, we allow a migrating cell to move a randomly chosen distance $m \leq G$ away. In order to maintain tissue connectedness, cells in the model can only migrate to grid-squares that are either already occupied, or are adjacent to occupied squares.

Additionally, unstable cells of the same cell-type behave differently in different cellular neighborhoods (Fig.1(C)): For example, let a cell of cell-type n_i be unstable in two distinct neighborhoods H^a and H^b. In neighborhood H^a, the cell differentiates into cell-type n_j; i.e. $n_i|H^a \rightarrow [n_j, m = 0]$, while in neighborhood H^b, the cell migrates; i.e. $n_i|H^b \rightarrow [n_i, m = m_1]$. In nature, the effect of such positional inputs on cell-fate are especially apparent in migratory multipotent stem cells, such as neural crest cells in zebrafish (36).

While for the most part, plasticity of cell-fate is the purview of stem-cells, under special circumstances, such as stem cell ablation, differentiated cells can revert to a stem-like state capable of proliferation and differentiation (37). Hence in the

DRAFT

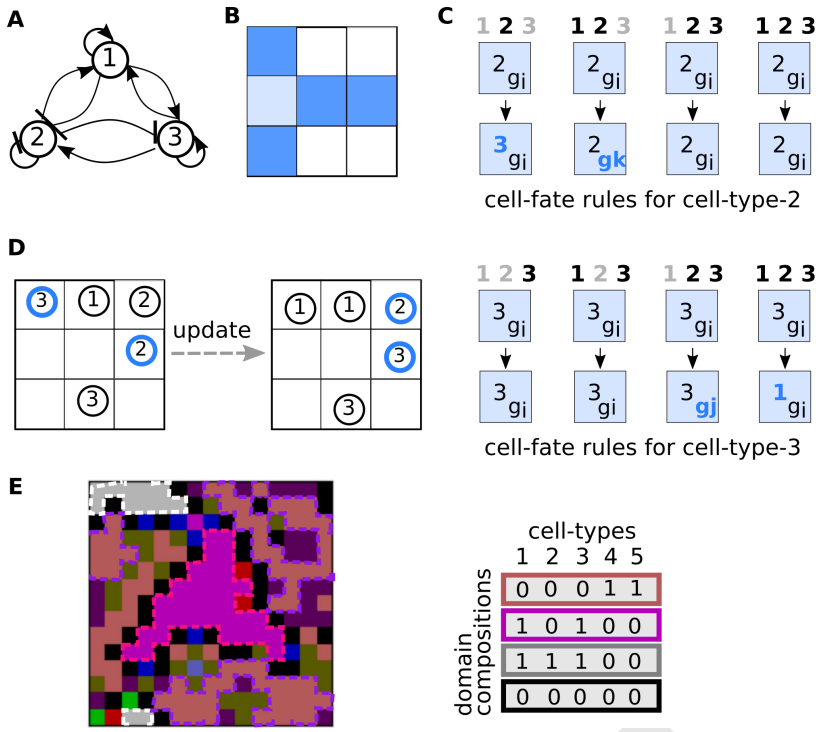

Fig. 1. An example of cellular interactions and cell-fate rules for a tissue with $N = 3$ cell-types in a $G = 3$ grid. (A) Cellular interactions: Circles represent distinct cell-types. Sharp arrows indicate stabilizing interactions and blunt arrows indicate destabilizing interactions. Cell-types *1* and *3* are intrinsically stable whereas cell-type *2* is intrinsically unstable. **(B)** Adjacent grid-squares: The grid has periodic boundaries and the dark blue squares are all adjacent to the light blue square. **(C)** An example of Cell-fate rules for cell-types that interact according to (A): Cell-type-1 is intrinsically stable and has only stabilizing interactions, therefore, it neither differentiates not migrates in any cellular neighborhood. We show here cell-fate rules for cell-type-2 (top) and cell-type-3 (bottom). Light blue squares in the first row represent the grid-square occupied by the cell of interest: The cell-type and the location of the grid-square are indicated within the square. The numbers in bold written above the square represent the cellular neighborhood: black numbers indicate the presence of the corresponding cell-type and grey numbers indicate the absence of the corresponding cell-types. Arrows point to the cell-fate of the cell of interest in the corresponding cellular neighborhood: Cellular decisions of unstable cells to differentiate or migrate are are highlighted in blue font.**(D)** A tissue represented by the grid on the left updates to the grid on the right following Cell-Fate Rules given in (C). **(E)** An example of a stable $N = 5$, $G = 15$ tissue. Colors of grid-squares indicate their cellular compositions. We call contiguous regions of the tissue that have the same composition a *domain*: Cells in the center of the domain have a cellular neighborhood identical to the domain composition, while cells at domain edges have richer cellular neighborhoods. Domains have variable sizes: we highlight with dashed borders the three large domains in this tissue. The cellular compositions of these three domains are given alongside the image: 1 indicates the presence of at least one cell of the corresponding cell-type while 0 indicates absence of the corresponding cell-type. Black grid-squares are empty.

model, we do not impose any restrictions on cell-fate potential, and instead allow hierarchical structures in cell-fate potential to emerge from the interplay of cellular interactions. Cell Fate Rules in the model comprise the set of all cell-fate transitions $n_i|H^a \rightarrow [n_j, m]$, such that $n_i, n_j \in \{1, 2, ..., N\}$, $m \leq G$ and $a \in \{0, 1, ..., 2^N - 1\}$ (see Methods: Generating tissue developmental rules).

In addition to cellular interactions, differentiation and migration, cells in the model can also divide and die (Methods:Updating tissue states). In nature, cell proliferation is often regulated by cell density (38, 39). To capture this in the model, cell division and cell death are regulated by the number of cells in grid-squares: Let $C^{all,g}$ be the total number of cells in some grid-square g. Any cell in grid-square g divides into an identical cell-type with probability $P_{div} = max(0, (10 - C^{all,g})/9)$ and a cell dies with probability $P_{death} = min(0.8, 0.02(C^{all,g} - 10))$. The fixed cell-division and death parameters we use lead to grid-squares having on average 10 cells; therefore the neighborhood of a single cell, which comprises five adjacent grid-squares (Fig.1(B)), is composed of 50 cells. These numbers can be tuned.

The daughter cell of a dividing cell in grid-square g can be born either in the same square or in some adjacent square. In nature, many cells divide asymmetrically during embryonic development, and in adults, tissue stem cells divide asymmetrically to produce daughter cells with distinct fates (40). In the model, daughter cells are identical to each other immediately after cell division, but daughter cells born in adjacent grid-squares can end up with distinct cell-fates, which allows us to capture the effects of asymmetric cell-division.

Model parameters and generating tissue development rules. We draw the interaction matrix I using parameters $0 \leq P_{den} \leq 1$ and $0 \leq P_{stable} \leq 1$: $I(n_i, n_j) \neq 0$ with probability P_{den} and $I(n_i, n_j)$ is positive with probability P_{stable}. P_{den} controls the level of inter-cellular signaling and P_{stable} controls cell-fate potential.

We draw Cell Fate Rules using parameter $0 \leq P_{diff} \leq 1$: for some cell of cell-type n_i that is unstable in neighborhoods H^a, the cell differentiates with probability P_{diff} and migrates otherwise. In neighborhood H^a, if cell-type n_i differentiates, the final cell-type n_j is chosen uniform randomly from among the N distinct cell-types. Instead if cell-type n_i migrates, the number of grid-squares it migrates is drawn uniformly from within a distance $m \leq G$. Together, the interaction matrix I and the Cell Fate Rules comprise *tissue develop-*

ment rules (Methods:Generating tissue developmental rules). Note that *tissue development rules*, once drawn, remain fixed throughout a simulation.

Finally, the parameter $0 \leq F_{adj} \leq 1$ controls the location of daughter cells: A fraction $1 - F_{adj}$ of daughter cells are born within the same grid-square as the mother cell, and a fraction F_{adj} are born in a different adjacent grid-square. F_{adj} is a complex parameter that takes into account the organization of cells within grid-squares which includes cell shape (41) and the directions of the axes of cell-division (42).

At any time-step t, the state of the tissue is completely characterized by the cellular compositions of its grid-squares $C_t^{n,g}$ for all $n \in \{1, 2, ...N\}$ and $g \in \{1, 2, ...G^2\}$. At $t = 0$, we initialize the tissue by randomly placing 20 cells of random cell-types across the connected set of grid-squares $[0, 1, 2, 3, 4]$. The state of the tissue is updated synchronously by updating the state of each cell in the tissue according to the *tissue development rules* (Fig.1(D), Methods:Updating tissue states). We simulate tissue development in the model for $T = 500$ time-steps. We explore the spatial and compositional features of tissues in the final time-step (Fig.1(E)), and analyze how tissues grow and heal from injuries.

Results

Model-generated tissues are rich in cell-types. We first analyse how Cell Fate Rules regulate the cellular compositions of model-generated tissues. As a measure of tissue complexity, we count the number of distinct cell-types composing different tissues. In our model, a majority of tissues in the final time-step contain all 5 cell-types (Fig.2(A), Supplementary Information:FigS2). As one expects, the number of tissue cell-types increases with the probability of stabilizing interactions, P_{stable} (Pearson's r = 0.15) and decreases with the propensity of cells to differentiates, P_{diff} (r = -0.41). The number of tissue cell-types also decreases with P_{den} which controls the probability that a cell-type interacts with some other cell-type (r = -0.14): Generally, increasing the number of cellular interaction partners increases the probability of destabilizing interactions.

To understand the nature of tissue cell-types, we measure their plasticity in terms of their cell-fate potentials (Fig.2(B), Supplementary Information:FigS2): from the Cell Fate Rules, we calculate cell-fate potential of any cell-type as the set of all distinct cell-types which it differentiates into across all possible cellular neighborhoods. Note that here cell-fate potential refers to the set of cell-types which any given cell-type *directly* differentiates into. Cell-fate potentials are expected to be wider if one additionally considers indirect differentiation involving multiple transitions. As expected, average cell-fate potential decreases with P_{stable} (r = -0.31), and increases with P_{diff} (r = 0.62). Average cell-fate potential increases with P_{den} (r = 0.28): Increasing cellular interaction partners in turn increases the probability that a cell responds differently in a wide variety of cellular neighborhoods.

As one can infer from the correlations with model parameters discussed above, model-generated tissues with more cell-types tend to have lower average cell-fate potential (Fig.2(C), Pearson's r between number of tissue cell-types and average cell-fate potential = -0.35). That is, most model-generated tissues tend to be rich in cell-types and the constituent cell-types tend to have low plasticity. Nevertheless, 31% (57%) of model-generated tissues contain at least one cell-type which can directly (through multiple transitions) differentiate into all 5 cell-types of the tissue (Supplementary Information:FigS3). The presence of these highly plastic cell-types in model-generated tissues is reminiscent of dedicated stem-cells that replenish cells lost upon injury from regenerative tissues (16).

Tissues vary widely in the sparsity and contiguity of cellular arrangements.

Tissues are organized into domains. In nature, the different cell-types composing a tissue are not randomly scattered, rather, they are arranged into distinct, but functionally and developmentally interconnected domains. Each tissue domain has a defined cellular composition and sits in a specific position with respect to other tissue domains. For example, there are three distinct domains in the mammalian intestinal epithelium: the crypt base, the crypt walls and the villi (15). Similarly, in plants, the shoot apical meristem is composed of three domains: the central zone, the organizing center and the peripheral zone (14).

Model-generated tissues resemble natural tissues and are organized into domains: 77% model generated tissues reach steady states where each cell is part of a stable cellular neighborhood (Supplementary Information: FigS1). Since the neighborhood of a cell extends beyond a single grid-square and encompasses also cells in adjacent grid-squares, at steady state, contiguous grid-squares tend to be arranged into domains defined by their cellular compositions (Fig.1(E)). We characterize tissues using five features that describe the spatial arrangement of cells and domains (Methods:Analysing final tissues, Supplementary Information:FigS4):

1. *Fullness*, or the number of grid-squares in the final tissue that are occupied by cells describes tissue size and sparsity (Fig.3(C)).

2. The *number of distinct tissue domains* describes how cells of different types are distributed into distinct stable cellular neighborhoods.

3. *Coverage equality* describes whether or not the tissue is dominated by a domain of one particular cellular composition, as opposed to being more uniformly covered by different domains.

4. *Dispersity* describes whether domains of identical cellular composition are present as a single connected region, or are scattered across the tissue in smaller disjoint regions (Fig.3(B)).

5. *Mixedness* describes the diversity of cellular compositions of grid-squares adjacent to any given grid-square. For example, grid-squares at the edges of

DRAFT

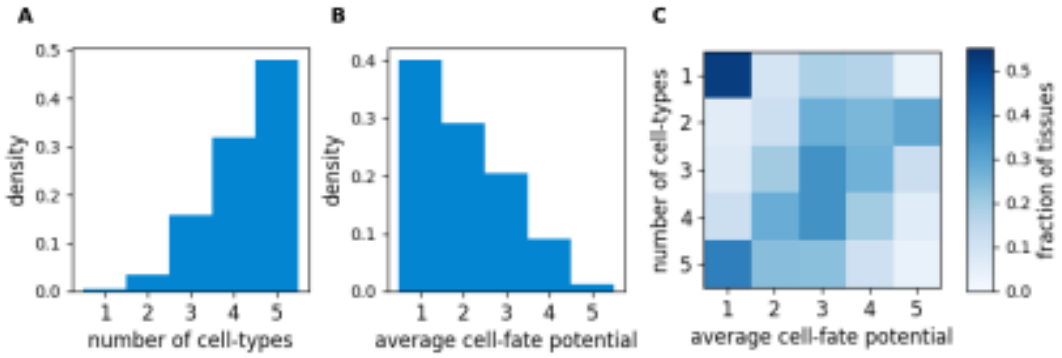

Fig. 2. Number of tissue cell-types and their cell-fate potential. (A) Histogram of the number of cell-types in model-generated tissues. **(B)** Histogram of cell-fate potential averaged across all cell-types in a tissue. **(C)** Heat map where each element represents the fraction of tissues that have a number of cell-types given by the rows with an average cell-fate potential given by the columns.

domains are more *mixed* than those in the center; in this sense, all disperse tissues are expected to have a high overall *mixedness*, but some non-disperse tissues with large but intertwined domains (resembling noodles) also have high *mixedness*.

Of these five features, *dispersity* and *mixedness* are related to clinically relevant quantities that are calculated from spatial omics studies (43).

Model-generated tissues fall into four sectors. We perform a Principal Components Analysis of all model generated tissues against these five features and find that the first two Principal Components (PC1 and PC2) capture 86% of the variance in the data (Methods:Analysing final tissues, Supplementary Information:FigS5). We also find that model parameters are significantly predictive of tissue spatial properties: we can divide the plane described by PC1 and PC2 into four sectors such that for 126 out of the 135 parameter sets tested, at least 50 out of 100 tissues generated using a single parameter set lie within a single sector (Fig.3(A)).

Note that the Principal Components themselves do not separate these tissues into distinct clusters, and tissue properties vary more or less continuously. Our separation of tissues into sectors is for instructive purposes and for ease of discussion. Tissues populating the four sectors have typical properties; appropriately, we name these four sectors *stunted* (17% of tissues), *sparse-contiguous* (16% of tissues), *full-contiguous* (48% of tissues) and *random-disperse* (19% of tissues) (Fig.3(B)). Thus, these sectors capture qualitative features of tissue morphology and spatial spread of cellular compositions.

Our model parameters are based on local cellular interactions, nonetheless they provide clues about global cellular spatial arrangements: for example, the *stunted* sector (Fig.3(E)) is populated by tissues whose cells tend to be non-interactive (low P_{den}) and intrinsically stable (high P_{stable}). Additionally, growth of these tissues is hampered because daughter cells tend to be born within the same grid-square as the mother cell rather than an adjacent square (low F_{adj}) and the levels of cell migration are low (high P_{diff}) (Fig.3(D), first column). In contrast, the *random-disperse* sector (Fig.3(F)) is populated by tissues with highly interactive and unstable cells that are highly migratory and tend to produce daughter cells in adjacent grid-squares (Fig.3(D), fourth column). An example of a natural tissue resembling model-generated *random-disperse* tissues is animal blood (44).

However, most natural tissues tend to be contiguous: For example, in animals, epithelial tissues form sheets (45) and mammalian striated muscles form bundles (46); in plants, the leaf epidermis forms sheets (47), while the leaf spongy parenchyma forms a sparse but contiguous domain (48). In this sense, most natural tissues resemble tissues in the *sparse-contiguous* and *full-contiguous* sectors (Fig.3(G-H)). Model parameters that generate tissues in the *sparse-contiguous* and *full-contiguous* sectors are more moderately distributed than in the *stunted* and *random-disperse* sectors (Fig.3(D), second and third columns).

Tissue development resembles a sequence of sector crossings. In real organisms, tissue development is expressed as a sequence of intertwined morphogenetic and cell-fate changes. For example, some of the first morphological changes during human embryo development are compaction of the ball of cells during Morula, followed by cavitation. These morphological changes are coupled with cell-fate changes: compaction is coupled with cellular differentiation into the inner cell mass and the trophectoderm, while cavitation is coupled with further differentiation of the inner cell mass into the epiblast and the hypoblast (49). These cell-fate decisions are not cell autonomous; instead, all through this process, different cell-types communicate with each other and play an important role in regulating each other's cellular decisions (50).

In contrast with animals, plants cells are immobile and morphogenetic changes take place entirely through cell growth and division. Despite this fundamental difference, plant and animal tissue development are described in a similar manner: During *Arabidopsis thaliana* embryonic development, the embryo progresses through the dermatogen, globular and heart-shaped stages. These morphological changes are coupled with cell-fate changes: the epidermal lineage is produced at the dermatogen stage, ground and vascular tissue lineages as well as shoot and root apical meristems are pro-

DRAFT

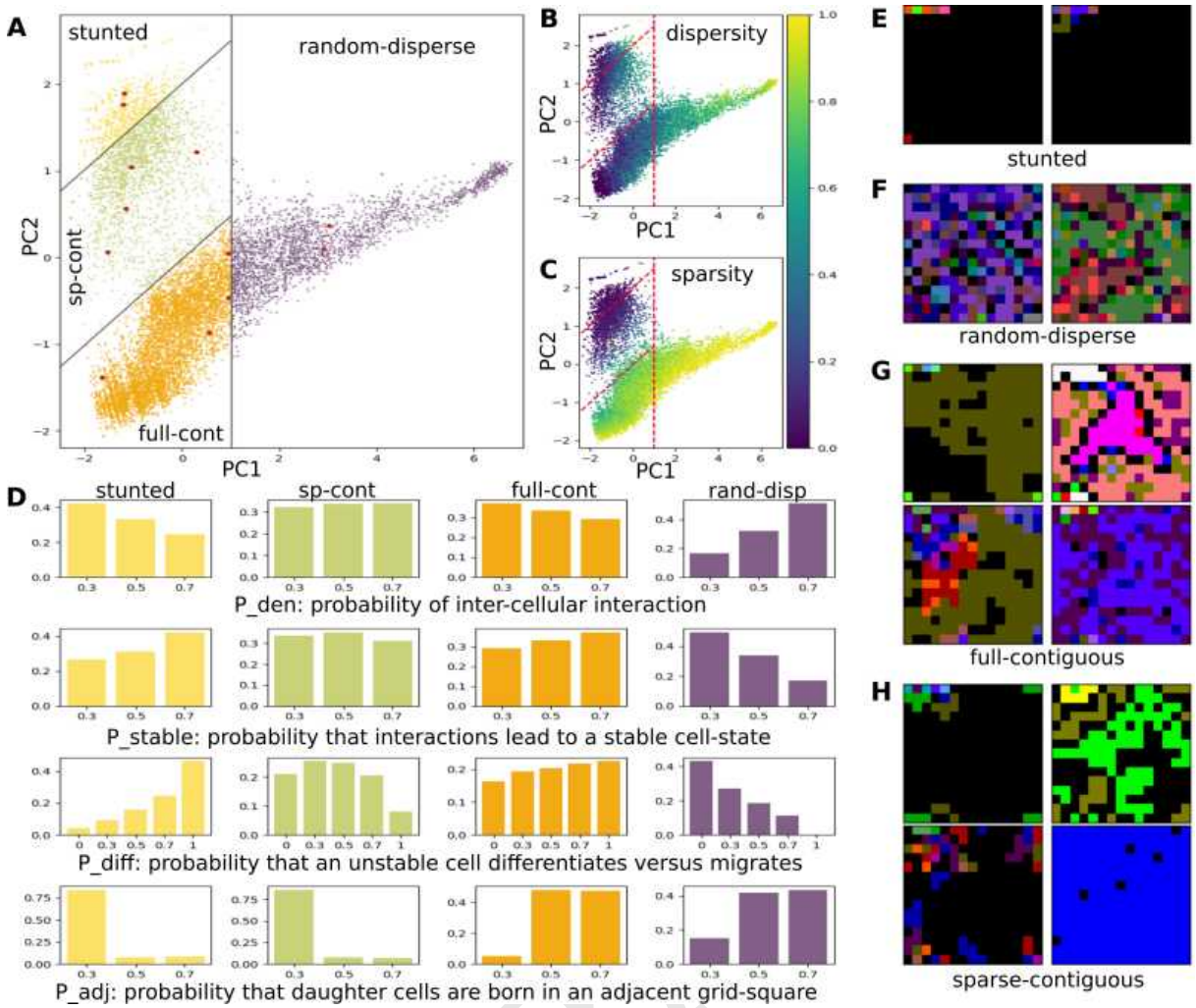

Fig. 3. Sectors define qualitatively distinct tissues. (A) Spread of model generated tissues across PC1 and PC2. Each point represents the state of a model-generated tissue at the final time-point. Tissues belonging to the 4 different sectors are indicated by different colors (yellow = stunted, green = sparse-contiguous, orange = full-contiguous, purple = random-disperse). Lines indicate borders between sectors. Red points represent examples of tissues from each sector shown in (E-H). **(B)** Values of *dispersity* across model generated tissues. **(C)** Tissue *fullness*, which we calculate as the number of occupied grid-squares, across model generated tissues. **(D)** Histograms for model parameters (rows) encoding tissues belonging to different sectors (columns). **(E-H)** Representative model-generated tissues belonging to different sectors: **(E)***Stunted* tissues show very little growth; **(F)** the *random-disperse* sector contains tissues in which the grid is densely populated by small contiguous domains. **(G)** the *full-contiguous* sector contains tissues in which the grid is densely populated by large contiguous domains; **(H)** the *sparse-contiguous* sector contains diverse forms of tissues. Depending on the distance of the tissue from the sector boundaries, it can resemble tissues in the stunted sector (top-left image), or in the full-contiguous sector (bottom-right image), or be a sparse and disperse tissue (bottom-left image). Importantly, this sector contains tissues where the grid is sparsely populated, yet the domains are contiguous.

duced at the globular stage, and the first plant organs, cotyledons are produced at the heart stage (51).

In our model, the four sectors capture tissue morphology (full versus sparse) as well as cellular compositions and spatial arrangements of tissue domains. Therefore, we express tissue development as sequences of sector crossings (Fig.4, Methods:Analysing tissue developmental trajectories, Supplementary Information:FigS6). In order to do this, we project the states of growing tissues onto the two Principal Components that describe tissues in their final states. As expected, almost all tissue developmental trajectories begin in the *stunted* sector (only 11 out of 13500 trajectories begin in the sparse-contiguous sector). Overall, 97.6% of tissues stably reach a final sector and make their last sector crossing within 450 time-steps. Most tissue developmental trajectories are simple and short with 1.7 ± 1.5 crossings and take on average 21.3 ± 27.5 time-steps to reach the final sector, while 13.5% trajectories involve oscillations (the red trajectory in Fig.4(B) oscillates between the *full-contiguous* and the *random-disperse* sectors). In particular, non-oscillatory trajectories that lead to *full-contiguous* tissues involve 2.2 ± 1.0 crossings and take 36.5 ± 20.7 time-steps to reach the final sector (Fig.4(D)).

For tissues that end up in different sectors, we summarize their average tissue developmental trajectories as transition matrices across the four sectors (Fig.4, insets): For each tissue's developmental trajectory, we count the fraction of time-steps the tissue either crosses into a different sector or stays within its current sector. For example, we see from the inset in Fig.4(D) that all *full-contiguous* tissues have a *sparse-contiguous* intermediate stage. A majority of these trajectories involve only two sector crossings: *stunted*–>

sparse-contiguous–> full-contiguous, with multiple update steps within each of the intermediate sectors (Fig.4(D), red and blue trajectories). A small number of trajectories transition via the *random-disperse* sector and some of these trajectories involve oscillations between the *random-disperse* sector and the final *full-contiguous* sector (Fig.4(D), black trajectory).

Full and contiguous tissues heal well.

Simulating tissue injury and recovery. In the model, we test the regenerative ability of tissues by simulating injuries to final time-step tissues and assessing their ability to return to their original states (Methods:Analysing tissue healing). We perform three kinds of injuries:

1. *single square-single cell-type* – we ablate all cells of a single randomly chosen cell-type from a single randomly chosen grid square (Fig.5, first row)
2. *single square-all cells* – we ablate all cells of any cell-type from a single randomly chosen grid square (Fig.5, second row)
3. *adjacent squares-single cell-type* – we ablate all cells of a single randomly chosen cell-type from all grid-squares adjacent to a single randomly chosen grid square (Fig.5, third row)

After injuring a tissue, we allow it to heal by updating the injured tissue for 100 time-steps using the same Cell Fate Rules originally used during tissue development. We assess the ability of a tissue to heal by comparing the distance between the state of the tissue before injury and immediately after injury against the distance between the state of the tissue before injury and after 100 time-steps of healing: We say that a *tissue heals* if distance from uninjured tissue decreases; the *tissue degrades* if this distance increases; the *tissue does not heal* if this distance neither increases nor decreases. We perform 25 independent injuries of each of the three types for each tissue and report the fraction of injuries for which the tissue heals (*recovery-fraction*: Fig.5, first column), does not heal (*no-recovery-fraction*: Fig.5, second column), or degrades (*degradation-fraction*: Fig.5, third column).
In order to gain some intuition for the process of tissue healing, consider how cell-fate in the model depends on the cellular neighborhood; i.e. the set of all cells present in all adjacent grid-squares. For cells present in a grid-square in the middle of a domain of defined cellular composition, ablation of one or more cells in a single adjacent grid-square does not change its neighborhood, and is thus unlikely to influence its behavior. In other words, ablating cells in a single grid-square, as happens in *single square-single cell-type* and *single square-all cells*, only perturbs the tissue if this grid-square is at the edge of a domain. Tissues that have high *mixedness* are expected to have large proportions of grid-squares on the edges of domains; in keeping with our intuition, the *degradation-fraction* for injuries where cells are ablated from single grid-squares is most strongly correlated with *mixedness* out of all tissue structural features: Pearson's *r* between *degradation-fraction* and *mixedness* for *single square- single cell-type* injuries = 0.53 (p-value=0.0) and for *single square-all cells* injuries = 0.51 (p-value=0.0). For 55.6% of all tissues, when either a single cell-type or all cells are ablated from a single grid-square, tissues tend to neither heal nor get worse, and the *no-recovery-fraction* is greater than 0.5 (green points in Fig.5 second column, rows 1 and 2).
In contrast, *adjacent squares-single cell-type* ablations are bound to change cellular neighborhoods for cells in the grid-square in the center of the ablation. We find that the ability to heal from *adjacent squares-single cell-type* injuries was enriched in the *full-contiguous* sector: only 16.8% of all tissues have a recovery-fraction ≥ 0.5 for these injuries, and 99.4% of these tissues are in the *full-contiguous* sector. This result reveals an intimate connection between tissue morphology and regeneration, and highlights evolutionary constraints on tissue architecture stemming from the requirement of tissues to heal. We expect that these constraints apply universally to all multicellular groups despite their independent origins (52).

Mechanism of tissue recovery in the model. In order to understand the general mechanism of tissue healing in the model, we tested how model parameters are correlated with recovery-fraction for *adjacent squares-single cell-type* injuries (Supplementary Information:FigS7): While P_{den}, P_{stable} and P_{diff} show modest correlations (*r* = -0.11, 0.13, 0.24, respectively), the ability to heal is strongly correlated with F_{adj} (*r* = 0.44). This suggests that the mechanism of tissue healing in the model depends on cell division in adjacent grid-squares replenishing lost cells in the injury sites. This mechanism is notably similar to tissue healing in plants, where cells adjacent to injuries switch their cell division axes in order to replace lost cells (22). Such replacement of damaged cells through division of cells in adjacent domains is also see in mammalian intestinal epithelium (15) [AND LIVER – add ref].
We wanted to also test how tissue healing in *metazoa* may be different from other multicellular groups: metazoan cells have the special ability to move, while cells in plants, multicellular fungi and algae are non-motile. In our model, cellular motility is controlled by parameter P_{diff}: tissues generated with $P_{diff} < 1$ contain motile cells, while tissues generated with $P_{diff} = 1$ lack motile cells. We find that tissues that lack motility have a greater ability to heal than those with motile cells: 29.7% of tissues with $P_{diff} = 1$ have a recovery-fraction ≥ 0.5 for *adjacent squares-single cell-type* injuries. This result is in line with observations of the remarkable degree to which plants are regenerative in comparison with animals in general (53). Our model additionally suggests that high regeneration ability should also be present in the less well-studied multicellular groups of fungi and algae, whose cells are also non-motile.

Discussion

Multicellular tissues come in an enormous variety of forms: animal tissues include epithelial sheets (41), bundles of mus-

DRAFT

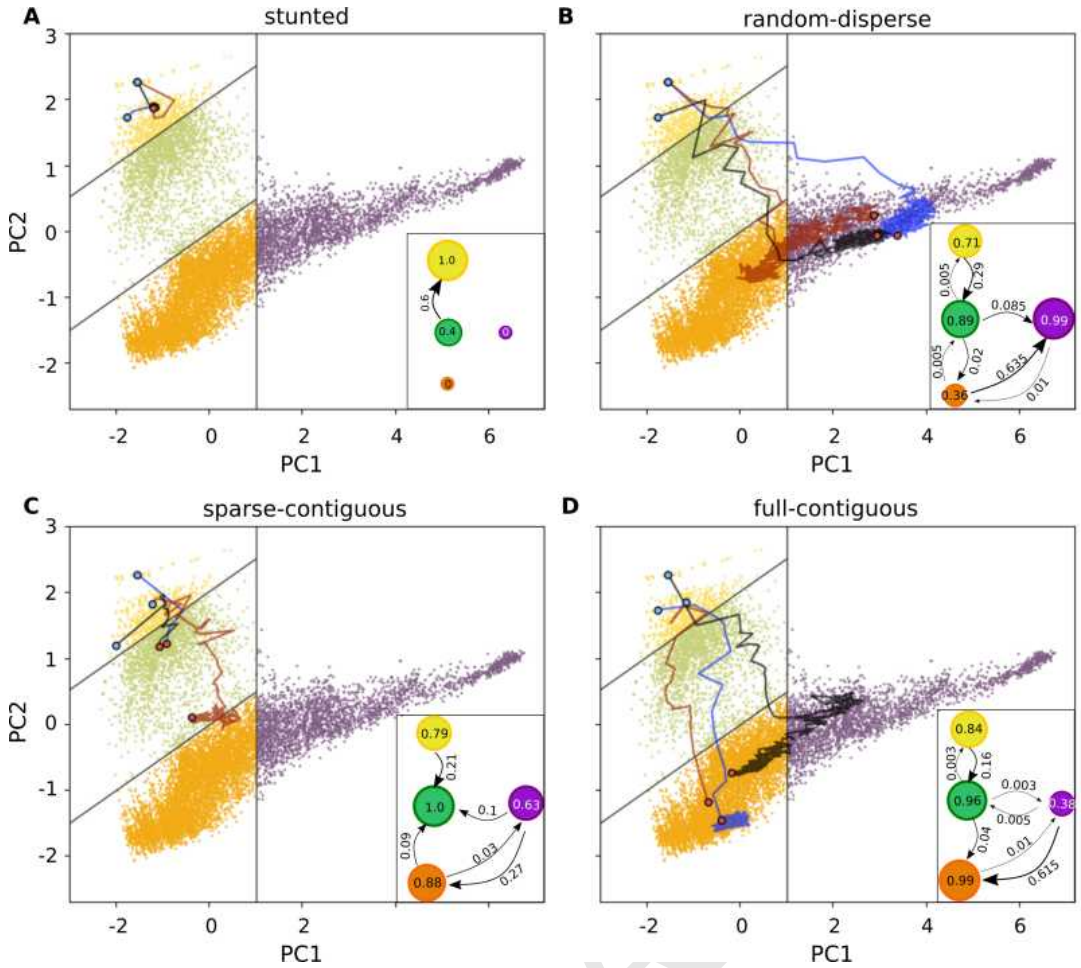

Fig. 4. Tissue growth trajectories and sector crossings. Representative tissue growth trajectories projected onto PC1 and PC2 for tissues whose final states fall into (A) *stunted*, (B) *random-disperse*, (C) *sparse-contiguous*, (D) *full-contiguous* sectors. Blue circles represent the initial state of tissues; red circles represent final states of tissues; colored lines indicate growth trajectories. Insets: Graphs indicating probabilities of sector crossings. Nodes (colored circles) represent different sectors (yellow = stunted, green = sparse-contiguous, orange = full-contiguous, purple = random-disperse). Numbers on edges (arrows) indicate the fraction of updates in a trajectory that take a tissue from the source to the target sector averaged across all trajectories for (A) *stunted*, (B) *random-disperse*, (C) *sparse-contiguous*, (D) *full-contiguous* tissues. Numbers within circles represent the average fraction of updates that stay within the corresponding sector.

cles (46) and disorganized and fluid blood (44), while tissues in plants include epidermal sheets (47), bundles of vascular tissue (54) and the sparse yet connected spongy mesophyll in leaves (48). Despite the apparent diversity, tissues are very similar to each other at the micro-anatomical level: Tissues tend to be arranged into contiguous domains or zones that are characterized by their cellular compositions. A particularly salient example of tissue domains are stem cell niches which are the cellular neighborhoods where stem cells reside, and signals from the niche regulate stem cell proliferation and differentiation (16).

In this work, we ask why tissue structures are so similar across organisms despite the multiple independent origins of multicellularity. Through agent-based modeling, we explore two distinct possibilities: First, that the diversity of tissue structures that developmental processes can produce are limited, and second, that tissue structures prevalent in nature may provide some selective benefits to organisms. We find a close association between model-generated tissue structures with their ability to regenerate, thus indicating that tissue regeneration could be one such selective benefit.

We model tissues as a collection of cells interacting locally with other cells in the neighborhood, and the model's tissues develop when cells make decisions to differentiate, migrate, divide or die based on inter-cellular interactions. We explored here tissues housed in two-dimensional grids where cells in any given grid square interact with all other cells in adjacent grid-squares. That is, our model trades the ability to explicitly encode the exact cellular neighbors of any given cell in order to overcome constraints on the geometry of single cells that are inherent in more traditional models, such as the vertex model (25). Concurrently, our framework is much simpler than models such as the cellular Pott's model (26) or voronoi tesselation models (28), thus retaining the ability to survey vast spaces of tissue developmental rules.

Our framework is also flexible, and we highlight two important ways to expand the model: First, while two and three-dimensional grids may be appropriate to study many solid tissues, tissues such as the plant spongy mesophyll (48), or the branched animal lung epithelium (55) can be better represented with more irregular geometries. Second, the model can be expanded by taking into account external signals,

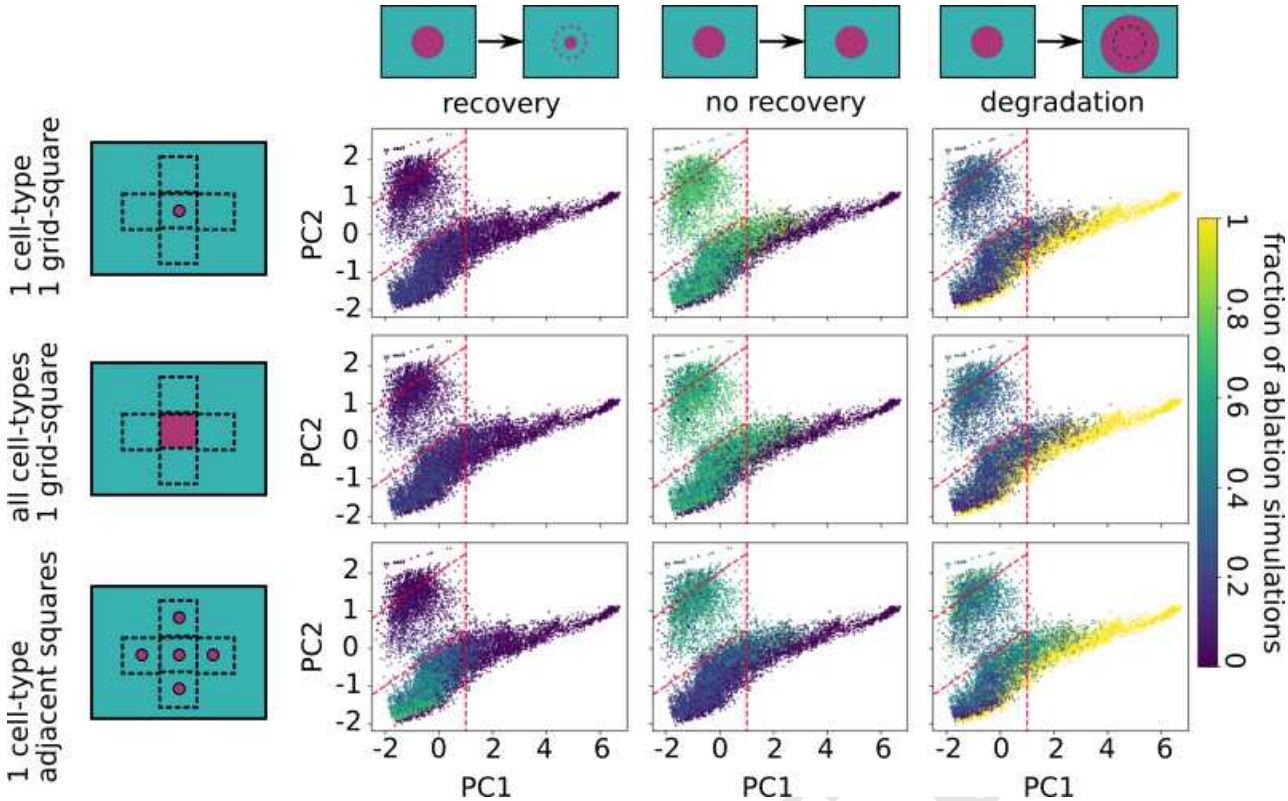

Fig. 5. Tissue regeneration. Points in the scatter plots indicate tissues. Colors of points indicate the fraction of times across the 25 injury simulations that the tissue showed recovery (left column), no recovery (middle column) or degradation (right column). Rows indicate the kind of injury: *single square-single cell-type* (first row), *single square-all cells* (second row), *adjacent squares-single cell-type* (third row).

which could represent morphogen gradients that form an integral aspect of animal (56) as well as plant development (57).

Overall, our simple spatial model of cell-fate captures essential features of tissue growth, form and function. Importantly, our model maps how varying tissue cellular interactions can lead to changes in tissue architecture, which in turn is closely related to the ability of tissues to heal. Particularly, the ability to heal is enriched in *full-contiguous* tissues. These inter-related tissue properties allow experimental tests of the model's predictions: For example, our model predicts that increasing the parameter P_{den}, which controls the density of inter-cellular interactions, should lead to more disperse tissues (Fig.3(D), first row), and disperse tissues tend to have low regeneration capacity (Fig.5). To experimentally test this prediction, one could tune the density of cellular interaction by manipulating the types of receptors expressed on tissue cell-types and observe the ensuing effects on tissue architecture and tissue regeneration. In addition, our results highlight how a need for tissue regeneration may impose selective pressures that mold tissue morphology across all multicellular organisms. Moreover, the mechanism of healing in the model through replacement of damaged cells by cell-division in adjacent regions; that is the model suggests a local mechanism of healing, which recapitulates the process of healing in plant as well as animal tissues.

Methods

A. Sampling the space of tissue developmental parameters. We considered tissues with $N = 5$ cell-types in a $G \times G$ 2-dimensional grid with periodic boundaries, where $G = 15$. The geometry of the tissue is described by the grid's adjacency matrix A, where $A(g_i, g_j) = 1$ if grid-squares g_i and g_j are adjacent to each other. In a 2-dimensional grid, each square is adjacent to 4 other squares, as shown in Fig.1(B). We uniformly survey parameter values from the following sets: $P_{den} \in \{0.3, 0.5, 0.7\}$, $P_{stable} \in \{0.3, 0.5, 0.7\}$, $P_{diff} \in \{0, 0.3, 0.5, 0.7, 1\}$ and $P_{adj} \in \{0.3, 0.5, 0.7\}$. For each combination of parameter values, we draw 100 replicate *tissue development rules*, and overall, we examine tissues resulting from 13,500 distinct rules. For 77% of these simulated tissues, cellular compositions of all grid-squares reach steady state within 400 time-steps.

B. Generating tissue developmental rules.

Interaction matrix. The interaction matrix I is a matrix of size $N \times N$ where

$$I(n_i, n_j) = 0 \text{ if } Ber(P_{den}) = 0 \text{ and implies that cell } n_j \text{ does not interact with } n_i,$$
$$I(n_i, n_j) = 1 \text{ if } (Ber(P_{den}) = 1) \bigwedge (Ber(P_{stable}) = 1) \text{ and implies that } n_i \text{ stabilizes } n_j,$$
$$I(n_i, n_j) = -1 \text{ if } (Ber(P_{den}) = 1) \bigwedge (Ber(P_{stable}) = 0) \text{ and implies that } n_i \text{ destabilizes } n_j$$

Note that the interaction matrix for each tissue is randomly

generated according to parameters P_{den} and P_{stable} and remains fixed for the duration of the simulation.

Cell Fate Rules. Let $H^a = 0,1^N$ be an N-length binary vector describing a cellular neighborhood, such that $H^a(n) = 1(0)$ indicates the presence (absence) of cell-type n in neighborhood H^a. An N-length vector $Stable^{H^a}$ encodes the stability of cell-types present in the neighborhood H^a.
$Stable^{H^a}(n_i) = 1$ and cells of cell-type n_i are stable in neighborhood H^a if:

- it is intrinsically stable and receives no destabilizing interactions: that is, $I(n_i, n_i) = 0 or 1$, and $\nexists n_j s.t. (H^a(n_j) = 1) \bigwedge (I(n_j, n_i) = -1)$

- it is intrinsically unstable and receives stabilizing interactions: that is, $I(n_i, n_i) = -1$, and $\exists n_j s.t. (H^a(n_j) = 1) \bigwedge (I(n_j, n_i) = 1)$

$Stable^{H^a}(n_i) = -1$ and cells of cell-type n_i are unstable in neighborhood H^a if:

- it is intrinsically stable and receives destabilizing interactions: that is, $I(n_i, n_i) = 0 or 1$, and $\exists n_j s.t. (H^a(n_j) = 1) \bigwedge (I(n_j, n_i) = -1)$

- it is intrinsically unstable and receives no stabilizing interactions: that is, $I(n_i, n_i) = -1$, and $\nexists n_j s.t. (H^a(n_j) = 1) \bigwedge (I(n_j, n_i) = 1)$

$Stable^{H^a}(n_i) = 0$ for cell-types not present in neighborhood H^a. Unstable cells either differentiate or migrate. We encode whether a cell of cell-type n_i differentiates in neighborhood H^a in an N-length binary vector $Diff^{H^a}$: $Diff^{H^a}(n_i) = 1$ if $(Stable^{H^a}(n_i) = -1) \bigwedge (Ber(P_{diff}) = 1)$.
Finally, for cells present in H^a, we encode Cell Fate Rules as a pair of look-up tables: A differentiation table $DiffTab^{H^a}$ and a migration table $MigTab^{H^a}$.

- $DiffTab^{H^a}(n_i) = n_j$ indicates that cell-type n_i differentiates into cell-type n_j in neighborhood H^a. If $(Stable^{H^a}(n_i) = -1) \bigwedge (Diff^{H^a}(n_i) = 1)$, n_j is randomly drawn from among the N cell-types. On the other hand, if $(Diff^{H^a}(n_i) = 0) \bigvee (Stable^{H^a}(n_i) = 1)$, then $n_j = n_i$.

- $MigTab^{H^a}(n_i) = m$ such that $m \leq G$ indicates that cell-type n_i in neighborhood H^a migrates to a grid-square a distance m away from its current grid-square. If $(Stable^{H^a}(n_i) = -1) \bigwedge (Diff^{H^a}(n_i) = 0)$, m is uniform randomly drawn from $0, 1, 2, ...G$. On the other hand, if $(Diff^{H^a}(n_i) = 1) \bigvee (Stable^{H^a}(n_i) = 1)$, then $m = 0$.

Note that although Cell Fate Rules for cellular differentiation and cell migration are randomly generated according to the parameter P_{diff}, they remain fixed for the duration of the simulation.

C. Updating tissue states. At any time-step, the state of the tissue is completely characterized by the cellular compositions of its grid-squares: $C_t^{n,g}$ gives the number of cells of cell-type n for $n \in 1, 2, 3, ...N$ in grid-square g for $g \in 1, 2, 3, ...G^2$ at time-step t. At $t = 0$, we initialize the tissue by randomly placing 20 cells of random cell-types across the connected set of grid-squares $0, 1, 2, 3, 4$. At each time-step of the simulation, the state of the tissue is updated through sequential steps of cell death, cell division, and cell differentiation and migration.

Cell death. In the tissue at time t, the total number of cells in a grid-square g is given by $C_t^{all,g} = \sum_{n=1}^{N} C_t^{n,g}$. Then the probability of cell-death in this grid-square is given by

$$P_{death}^g = min(0.8, 0.02(\frac{C_t^{all,g} - 10}{)}) \text{ if } C_t^{all,g} > 10,$$
$$P_{death}^g = 0 \text{ otherwise.}$$

That is, there is no cell-death in grid-squares with 10 or fewer cells, and the probability that any cell dies increases linearly as the number of cells in the grid-square increases beyond 10. We draw the number of cells of cell-type n in grid-square g that die in this time-step, $DeadCells_t^{n,g}$ from the binomial distribution $Bin(C_t^{n,g}, P_{death}^g)$. Then, $RemainingCells^{n,g} = C_t^{n,g} - DeadCells_t^{n,g}$ gives the cells of cell-type n that remain in grid-square g.

Cell division. After cell death, the total number of cells remaining in some grid-square g_i is given by $RC_t^{all,g_i} = \sum_{n=1}^{N} RemainingCells_t^{n,g_i}$. Then the probability of cell-division in this grid-square is given by $P_{div}^{g_i} = max(0, \frac{10 - RC_t^{all,g_i}}{9})$. That is, there is no cell-division in unoccupied grid-squares and in squares with 10 or more cells, and the probability that any cell divides decreases linearly as the number of cells in the grid-square increases up to 10.
We draw the number of cells of cell-type n in grid-square g_i that divide in this time-step, $DividingCells_t^{n,g_i}$ from the binomial distribution $Bin(RC_t^{n,g_i}, P_{div}^{g_i})$. $DividingCells_t^{n,g_i}$ is then the number of daughter cells of cell-type n produced in grid-square g_i. A fraction $(1 - P_{adj})$ of these daughter cells are born within grid-square g_i. Note that this fraction can include all the daughter cells produced. The rest of the daughter cells are assigned to randomly chosen grid-squares adjacent to g_i.
Let $D_{g_i}^{n,g_j}$ be the number of daughter cells of cell-type n produced by division of cells in grid-square g_i that are born in grid-square g_j. For any g_j not adjacent to g_i, $D_{g_i}^{n,g_j} = 0$ for any cell-type $n \in 1, 2, .., N$.
Then, $TC_t^{n,g_j} = RemainingCells_t^{n,g_i} + \sum_{g_i=1}^{G^2} D_{g_i}^{n,g_j}$ gives the total number of cells of type n in grid-square g_j after cell-division.
Note that in the model, the probability of cell-division is the same for all cells of all cell-types in a given grid-square. Additionally, two cells of the same cell-type in the same grid-square may divide differently: Depending on P_{div}^g, it is possible that only one of these cells divides. In case both these cells divide, depending on F_{adj}, it is also possible that the daughters are produced in different grid-squares.

DRAFT

Cell differentiation and migration. Let N-length binary vector H_{g_i} represents the cellular neighborhood for grid-square g_i in TC_t.

$$H_{g_i}(n) = 1 \text{ if } (TC_t^{n,g_j} > 0) \bigvee ((TC_t^{n,g_j} > 0) \bigwedge (A(g_i, g_j) = 1)).$$

Cell migration and differentiation rules for cells in grid-square g_i can then be calculated using Cell Fate Rules: We encode the identity of migrating cell-types in an N-length binary vector $Migcells_t^{g_i}$ such that

$$Migcells_t^{g_i}(n) = 1 \text{ if cell-type } n \text{ migrates and } MigTab_n^{H_{g_i}} > 0 \text{ and}$$
$$Migcells_t^{g_i}(n) = 0 \text{ if cell-type } n \text{ is non-migratory and } MigTab_n^{H_{g_i}} = 0$$

In order to find the new locations of cells migrating from grid-square g_i, we use matrix powers of the adjacency matrix, A^m: The set of all grid-squares m steps away from g_i is $J_{g_i}^m = [g_j$ such that $A^m(g_i, g_j) > 0]$. For each cell in g_i that migrates a distance m away, we randomly choose its new grid-square from $J_{g_i}^m$. We encode updated locations of migrating cells in $CellMig$: number of cells of cell-type n originally in grid-square g_i which have migrated to grid-square g_j is given by $CellMig_{n,g_j}^{g_i}$. Note that $CellMig$ only contains the updated locations of cells that migrate, i.e., for all cells of any cell-type n such that $Migcells_t^{g_i}(n) = 1$.

Non-migrating cells are either stable in the current cellular neighborhood H_{g_i}, or they differentiate. Rules governing the fate of non-migrating cell-types are given by $DiffTab^{H_{g_i}}$. We encode updated cell-types for non-migrating cells in $CellDiff$: $CellDiff^{n,g_i} = \sum_{n_j=1}^{N} (TC_t^{n_j,g_i} > 0)(\delta(DiffTab_{n_j}^{H_{g_i}}, n_i))(\delta(MigTab_{n_j}^{H_{g_i}}, 0))$. That is, $CellDiff^{n,g_i}$ is the number of non-migrating cells in g_i which differentiate into cell-type n.

The state of the tissue in the next time-step is then given by: $C_{t+1}^{n,g_i} = DiffCells^{n,g_i} + \sum_{g_j=1}^{G^2} CellMig_{n,g_i}^{g_j}, \ \forall n \in [1, 2, ..N]$.

In this work, we synchronously update each grid-square at each model time-step, and perform 500 updates for each tissue.

D. Analysing final tissues. We characterize tissue morphology based on how tissue domains (regions of uniform cellular composition within tissues) are arranged relative to each other. In order to do this, we measure five quantities:

Fullness: Let $C_{500}^{any,g}$ describe whether grid-square g contains any cells, or is empty: $C_{500}^{any,g} = 1 - \delta(\sum_{n=1}^{N} C_{500}^{n,g} = 0)$. We calculate *tissue fullness* as the number of grid-squares that are occupied by cells is then given as $fullness = \sum_{g=1}^{G^2} C_{500}^{any,g}$.

Number of unique domains: We express the cellular composition of any occupied grid-square as a N-length binary vector $CellComp^g$, where $CellComp^g(n) = 1$ if $C_{500}^{n,g} > 0$. $CellComp^g$ can take on one of 2^N values depending on the presence or absence of the N cell-types. We arrange these 2^N distinct cellular compositions according to the standard binary ordering. Tissues in the model tend to be composed of contiguous domains of uniform cellular composition, such that the cellular compositions of grid-squares within a single domain are identical to each other. Therefore, the number of unique cellular compositions across grid-squares gives us the number unique tissue domains (nud). In this work, the number of cell-types $N = 5$, therefore the maximum value of $nud = 2^5$, and $nud = 1$ for completely uniform tissues with a single domain.

Coverage equality: Let $UniqDomains$ be the set of cellular compositions of unique domains. And let $IdenSquares$ be an nud length vector where $IdenSquares(i)$ is the number of grid-squares with cellular composition identical to the ith unique domain. We say that a tissue has high *coverage equality* if each unique domain cellular composition is represented by roughly equal number of grid-squares; i.e., $EqualCover(i) = \frac{G^2}{nud} \forall i \in [1, 2, .., nud]$. At the other end, for a tissue with extremely unequal distribution of cellular compositions across the grid, $UnequalCover(1) = G^2$ and $UnequalCover(i) = 0 \forall i in [2, 3, .., nud]$. Let $dmax$ be the Euclidean distance between the vectors $UnequalCover$ and $EqualCover$. For any given tissue, let d_{tissue} be the Euclidean distance between $IdenSquares$ and $EqualCover$. We calculate its *coverage equality* as:

$$CoverageEquality = 1 - d_{tissue}/dmax.$$

Dispersity: Any set of grid-squares of identical composition, say $UniqDomains(i)$, is either completely contiguous, or splits into smaller subsets of contiguous regions. Contiguity of grid-squares can be assessed using the adjacency matrix A: A subset of grid-squares is contiguous if it forms a connected component, i.e., any grid-square in the subset can be reached from any other grid-square in the subset by traversing the grid. We encode the number of contiguous components with cellular compositions identical to $UniqDomain(i)$ as $IndenComponents(i)$. We call a tissue *disperse* if grid-squares with identical cellular compositions are split into multiple distinct connected components. We measure the dispersity of domains with a cellular composition $UniqDomains(i)$ as $DomainDispersity(i) = \frac{IndenComponents(i)-1}{IdenSquares(i)-1}$. We measure $dispersity$ of the tissue as the average $DomainDispersity$; i.e., $dispersity = \frac{\sum_{i=1}^{nud} DomainDispersity(i)}{nud}$.

Mixness: Finally, we examine the local neighborhood of each grid-square by counting $GridMixness_i$: the number of unique cellular compositions across the set of grid-squares adjacent to g_i, including itself. For a 2-dimensional grid, $GridMixness_i$ is at most 5. We describe a G^2 length vector $MixedSquares$, where $MixedSquares(i) = \frac{GridMixness-1}{5}$, such that for a grid-square in the middle of a contiguous domain, $MixedSquares(i) = 0$. We also assign $MixedSquares(i) = 0$ if grid-square g_i is unoccupied. We define tissue $mixedness$ by averaging $MixedSquares$ over all grid-squares.

We characterize the morphology of any model-generated tissue in the final time-point by the set of five num-

DRAFT

bers: $fullness$, nud, $CoverageEquality$, $dispersity$ and $mixness$.

Principal Components and sectors. These 5 properties are not independent of each other; particularly, $mixness$ and $dispersity$ are closely related quantities (Pearson's correlation coefficient = 0.74). To find relevant dimensions across which properties of model-generated tissues in the final time-point ($t = 500$) vary, we perform Principal Components Analysis. We find that the first two Principal Components (PC1 and PC2) explain 85.6% of the variance, and we now characterize tissues based on PC1 and PC2.

In order to assess how tissue properties (PC1 and PC2) change with model parameters, and for ease of discussion, we further split the space of tissues into *sectors*, such that tissues generated using the same set of model parameters ($P_{den}, P_{stable}, P_{diff}, P_{adj}$) are likely to all fall within the same *sector*. Based on our raw assessment of model-generated tissues, we choose to divide the space of tissues into *four* sectors. Sector boundaries are straight lines: boundaries between the *stunted* and *sparse-contiguous*, and the *sparse-contiguous* and *full-contiguous* sectors are characterized by their slopes and y-intercepts. The boundary separating the *random-disperse* sector is vertical, and is characterized by its x-intercept. We find optimal values for the slopes and intercepts of sector boundaries which maximize the number of parameter sets for which at least 50% of generated tissues lie within the same sector. For the sector boundaries we use, 50% of generated tissues lie within the same sector for 93.3% of the 135 parameter sets sampled in this study.

E. Analysing tissue developmental trajectories.

For each tissue, for each intermediate time-point t during its development, we calculate the five quantities that describe its morphology: $fullness_t$, nud_t, $CoverageEquality_t$, $dispersity_t$ and $mixness_t$. We scale each of these properties: For example, $scaled - fullness_t = \frac{fullness_t - MeanFinalFullness}{SDFinalFullness}$; where $MeanFinalFullness$ is the average $fullness$ across all tissues at $t = 500$, and $SDFinalFullness$ is the standard deviation of $fullness$ across all tissues at $t = 500$. We then project scaled values of properties of tissues at intermediate time-points onto the Principal Components (PC1 and PC2) that describe model-generated tissues in the final time-point. This allows us to place developing tissues at intermediate time-points into the four *sectors*: *stunted, sparse-contiguous, full-contiguous and random-disperse*.

For each tissue $tissue_i$ such that $i \in [1, 2, ..., 13500]$, let the the sector assignment at time t be $update_t^{tissue_i} = x$, for $x \in$ [*stunted, sparse-contiguous, full-contiguous, random-disperse*]. We describe tissue developmental trajectories in terms of sequences of sector crossings: $update_{x,y}^{tissue_i}$ is the number of update steps during its development in which the tissue transitions from a sector x to a sector y, where $x, y \in$ [*stunted, sparse-contiguous, full-contiguous, random-disperse*]. For example, $update_{stunted,sp-contig}^{tissue_i}$ is the number of times $tissue_i$ transitions from the *stunted* to the *sparse-contiguous* sector, and $update_{sp-contig,sp-contig}^{tissue_i}$ is the number of update steps in which $tissue_i$ stays within the *sparse-contiguous* sector.

We then estimate *sector transition probabilities* for each sector: For all tissues which end up in sector $z \in$ [*stunted, sparse-contiguous, full-contiguous, random-disperse*] at $t = 500$,

$$update_x^z = \sum((\sum_{t=1}^{500} update_t^{tissue_i})\forall tissue_i \text{ such that } update_{t=500}^{tissue_i} = z)).$$

$$\text{And, } update_{x,y}^z = \sum(update_{x,y}^{tissue_i}\forall tissue_i \text{ such that } update_{t=500}^{tissue_i} = z).$$

$$\text{Then, } TransitionProb_{x,y}^z = \frac{update_{x,y}^z}{update_x^z}.$$

F. Analysing tissue healing.

Producing injuries. For each tissue, we test its ability to heal by simulating three kinds of injuries:

InjuryA: For each tissue, we perform 25 independent injuries of this type. For some tissue among the 13500 model-generated tissues, let $InjusryA_{t=0}^{n,g}(i) = C_t^{n,g} = 500$, where $i \in [1, 2, ...25]$ represents the injury-id, and $n \in [1, 2, ...N]$, $g \in [1, 2, ...G^2]$.

In this type of injury, we ablate all cells of one particular cell-type from a single grid-square. To do this, we randomly choose a non-empty grid-square g_r and a cell-type n_r such that $InjuryA_0^{n_r,g_r}(i) > 0$, and set $InjuryA_1^{n_r,g_r}(i) = 0$.

InjuryB: For each tissue, we perform 25 independent injuries of this type. For some tissue among the 13500 model-generated tissues, let $InjusryB_{t=0}^{n,g}(i) = C_t^{n,g} = 500$, where $i \in [1, 2, ...25]$ represents the injury-id, and $n \in [1, 2, ...N]$, $g \in [1, 2, ...G^2]$.

In this type of injury, we ablate all cells of all cell-type from a single grid-square. To do this, we randomly choose a non-empty grid-square g_r such that $InjuryA_0^{all,g_r}(i) > 0$, and set $InjuryA_1^{n,g_r}(i) = 0 \forall n \in [1, 2, ...N]$.

InjuryC: For each tissue, we perform 25 independent injuries of this type. For some tissue among the 13500 model-generated tissues, let $InjusryC_{t=0}^{n,g}(i) = C_t^{n,g} = 500$, where $i \in [1, 2, ...25]$ represents the injury-id, and $n \in [1, 2, ...N]$, $g \in [1, 2, ...G^2]$.

In this type of injury, we ablate all cells of one particular cell-type from all grid-squares adjacent to a given grid-square. To do this, we randomly choose a non-empty grid-square g_r and a cell-type n_r such that $InjuryA_0^{n_r,g_r}(i) > 0$, and set $InjuryA_1^{n_r,g_{adj}}(i) = 0 \forall g_{adj}$ such that $A(g_r, g_{adj}) <= 1$.

We calculate the magnitude of an injury as the Euclidean distance between the uninjured tissue ($Injury_{t=0}^{n,g}(i)$) and the tissue immediately after injury ($Injury_{t=1}^{n,g}(i)$): $InjuryMagnitude_{t=1}(i) = \sum_{n=1}^{N}\sum_{g=1}^{G^2}(Injury_{t=0}^{n,g}(i) - Injury_{t=1}^{n,g}(i))^2$.

Tissue recovery. In order to simulate tissue recovery, we update the state of injured tissues for 100 time-steps using the same Cell Fate Rules that were used to simulate tissue development as described above in Updating tissue states. For each tissue, we test how the magnitude of the injury changes during recovery for the 25 replicate injuries of each of the three

DRAFT

kinds of injuries (InjuryA, InjuryB, InjuryC): For the ith replicate injury, the magnitude of the injury after 100 time-steps of recovery is given by $InjuryMagnitude_{t=100}(i) = \sum_{n=1}^{N}\sum_{g=1}^{G^2}(Injury_{t=0}^{n,g}(i) - Injury_{t=100}^{n,g}(i))^2$.
We say that a tissue is healing if $InjuryMagnitude_{t=100}(i) < InjuryMagnitude_{t=1}(i)$, the tissue shows no recovery if $InjuryMagnitude_{t=100}(i) = InjuryMagnitude_{t=1}(i)$, and the tissue is degrading if $InjuryMagnitude_{t=100}(i) > InjuryMagnitude_{t=1}(i)$.
For each tissue, for each injury type, we measure the fraction of replicate injuries out of 25 that result in healing ($recovery - fraction$), no recovery ($no - recovery - fraction$), or tissue degradation ($degradation - fraction$).

Acknowledgements

We were funded by Institute for Basic Science, grant number: IBS-R020-D1.

Supplementary Note 1: Supplementary Information

The supplementary file contains the following figures:

- FigS1: Histogram for the number of time-steps to tissue steady-state.
- FigS2: Scatter-plots showing the dependence on model parameters of number of distinct tissue cell-types, average cell-fate potential.
- FigS3: Scatter-plots showing the dependence on model parameters of maximum cell-fate potential of tissue cell-types.
- FigS4: Heatmaps displaying dependence on model parameters of the 5 spatial properties using which we characterize model-generated tissues.
- FigS5: Scatter-plots showing the dependence on model parameters of Principal Components which represent tissue spatial properties.
- FigS6: Histograms for the number of sector crossings, number of time-steps to reach the final sector.
- FigS7: Heatmaps displaying dependence of tissue healing in response to injuries of the type *neighboring-squares-single cell-type* on model parameters.

Bibliography

1. Miri Adler, Noa Moriel, Aleksandrina Goeva, Inbal Avraham-Davidi, Simon Mages, Taylor S Adams, Naftali Kaminski, Evan Z Macosko, Aviv Regev, Ruslan Medzhitov, et al. Emergence of division of labor in tissues through cell interactions and spatial cues. *Cell Reports*, 42(5), 2023.
2. Ulrich Technau and Robert E Steele. Evolutionary crossroads in developmental biology: Cnidaria. *Development*, 138(8):1447–1458, 2011.
3. Marina Oliva and Ryan Lister. Exploring the identity of individual plant cells in space and time. *New Phytologist*, 240(1):61–67, 2023.
4. Ioannis Theodorou and Bénédicte Charrier. Brown algae: Ectocarpus and saccharina as experimental models for developmental biology. In *Handbook of Marine Model Organisms in Experimental Biology*, pages 27–47. CRC Press, 2021.
5. László G Nagy, Peter Jan Vonk, Markus Künzler, Csenge Földi, Máté Virágh, Robin A Ohm, Florian Hennicke, Balázs Bálint, Á Csemetics, Botond Hegedüs, et al. Lessons on fruiting body morphogenesis from genomes and transcriptomes of agaricomycetes. *Studies in Mycology*, 104(1):1–85, 2023.
6. Thomas W Holstein. The hydra stem cell system–revisited. *Cells & Development*, 174: 203846, 2023.
7. Ian A Hatton, Eric D Galbraith, Nono SC Merleau, Teemu P Miettinen, Benjamin McDonald Smith, and Jeffery A Shander. The human cell count and size distribution. *Proceedings of the National Academy of Sciences*, 120(39):e2303077120, 2023.
8. Maya Bar and Naomi Ori. Leaf development and morphogenesis. *Development*, 141(22): 4219–4230, 2014.
9. Elaheh Hashemi and Subramaniam Malarkannan. Tissue-resident nk cells: development, maturation, and clinical relevance. *Cancers*, 12(6):1553, 2020.
10. Clare M O'Connor, Jill U Adams, and Jennifer Fairman. Essentials of cell biology. *Cambridge, MA: NPG Education*, 1:54, 2010.
11. Raul Ramos, Benjamin Swedlund, Anand K Ganesan, Leonardo Morsut, Philip K Maini, Edwin S Monuki, Arthur D Lander, Cheng-Ming Chuong, and Maksim V Plikus. Parsing patterns: Emerging roles of tissue self-organization in health and disease. *Cell*, 187(13): 3165–3186, 2024.
12. Roie Cohen, Liat Amir-Zilberstein, Micha Hersch, Shiran Woland, Olga Loza, Shahar Taiber, Fumio Matsuzaki, Sven Bergmann, Karen B Avraham, and David Sprinzak. Mechanical forces drive ordered patterning of hair cells in the mammalian inner ear. *Nature communications*, 11(1):5137, 2020.
13. Claudia Van Den Berg, Viola Willemsen, Giel Hendriks, Peter Weisbeek, and Ben Scheres. Short-range control of cell differentiation in the arabidopsis root meristem. *Nature*, 390 (6657):287–289, 1997.
14. Christophe Gaillochet, Gabor Daum, and Jan U Lohmann. O cell, where art thou? the mechanisms of shoot meristem patterning. *Current opinion in plant biology*, 23:91–97, 2015.
15. Hans Clevers. The intestinal crypt, a prototype stem cell compartment. *Cell*, 154(2):274–284, 2013.
16. Elaine Fuchs and Helen M Blau. Tissue stem cells: architects of their niches. *Cell stem cell*, 27(4):532–556, 2020.
17. Hiroyuki Koike, Ichiro Manabe, and Yumiko Oishi. Mechanisms of cooperative cell-cell interactions in skeletal muscle regeneration. *Inflammation and Regeneration*, 42(1):48, 2022.
18. Faith H Brennan, Yang Li, Cankun Wang, Anjun Ma, Qi Guo, Yi Li, Nicole Pukos, Warren A Campbell, Kristina G Witcher, Zhen Guan, et al. Microglia coordinate cellular interactions during spinal cord repair in mice. *Nature communications*, 13(1):4096, 2022.
19. Kaoru Sugimoto, Haruka Temman, Satoshi Kadokura, and Sachihiro Matsunaga. To regenerate or not to regenerate: factors that drive plant regeneration. *Current opinion in plant biology*, 47:138–150, 2019.
20. Jan U Lohmann, Ingrid Endl, and Thomas CG Bosch. Silencing of developmental genes in hydra. *Developmental biology*, 214(1):211–214, 1999.
21. Xuan Zheng, Max A Betjes, Pascal Ender, Yvonne J Goos, Guizela Huelsz-Prince, Hans Clevers, Jeroen S van Zon, and Sander J Tans. Organoid cell fate dynamics in space and time. *Science advances*, 9(33):eadd6480, 2023.
22. Petra Marhava, Lukas Hoermayer, Saiko Yoshida, Peter Marhavỳ, Eva Benková, and Jiří Friml. Re-activation of stem cell pathways for pattern restoration in plant wound healing. *Cell*, 177(4):957–969, 2019.
23. Ioannis Theodorou and Bénédicte Charrier. The shift to 3d growth during embryogenesis of kelp species, atlas of cell division and differentiation of saccharina latissima. *Development*, 150(21):dev201519, 2023.
24. Yih Lee, Stylianos Kouvroukoglou, Larry V McIntire, and Kyriacos Zygourakis. A cellular automaton model for the proliferation of migrating contact-inhibited cells. *Biophysical journal*, 69(4):1284–1298, 1995.
25. Reza Farhadifar, Jens-Christian Rõper, Benoit Aigouy, Suzanne Eaton, and Frank Jülicher. The influence of cell mechanics, cell-cell interactions, and proliferation on epithelial packing. *Current biology*, 17(24):2095–2104, 2007.
26. Athanasius FM Marée, Verônica A Grieneisen, and Paulien Hogeweg. The cellular potts model and biophysical properties of cells, tissues and morphogenesis. *Single-cell-based models in biology and medicine*, pages 107–136, 2007.
27. James M Osborne, Alexander G Fletcher, Joe M Pitt-Francis, Philip K Maini, and David J Gavaghan. Comparing individual-based approaches to modelling the self-organization of multicellular tissues. *PLoS computational biology*, 13(2):e1005387, 2017.
28. Martin Bock, Amit Kumar Tyagi, Jan-Ulrich Kreft, and Wolfgang Alt. Generalized voronoi tessellation as a model of two-dimensional cell tissue dynamics. *Bulletin of mathematical biology*, 72:1696–1731, 2010.
29. Sebastian A Sandersius and Timothy J Newman. Modeling cell rheology with the subcellular element model. *Physical biology*, 5(1):015002, 2008.
30. Mihaela Pavlicev and Günter P Wagner. Reading the palimpsest of cell interactions: what questions may we ask of the data? *iScience*, 2023.
31. Erick Armingol, Hratch M Baghdassarian, and Nathan E Lewis. The diversification of methods for studying cell–cell interactions and communication. *Nature Reviews Genetics*, pages 1–20, 2024.
32. Norbert Perrimon, Chrysoula Pitsouli, and Ben-Zion Shilo. Signaling mechanisms controlling cell fate and embryonic patterning. *Cold Spring Harbor perspectives in biology*, 4(8): a005975, 2012.
33. Chii J Chan, Carl-Philipp Heisenberg, and Takashi Hiiragi. Coordination of morphogenesis and cell-fate specification in development. *Current Biology*, 27(18):R1024–R1035, 2017.
34. Yiteng Dang, Johanna Lattner, Adrian Lahola-Chomiak, Diana Alves-Afonso, Elke Ulbricht, Anna V Taubenberger, Steffen Rulands, and Jacqueline Tabler. Self-propagating wave drives noncanonical antidurotaxis of skull bones in vivo. *bioRxiv*, pages 2023–07, 2023.
35. Dana T Byrd and Judith Kimble. Scratching the niche that controls caenorhabditis elegans germline stem cells. In *Seminars in cell & developmental biology*, volume 20, pages 1107–1113. Elsevier, 2009.
36. Tatiana Subkhankulova, Karen Camargo Sosa, Leonid A Uroshlev, Masataka Nikaido, Noah

DRAFT

Shriever, Artem S Kasianov, Xueyan Yang, Frederico SLM Rodrigues, Thomas J Carney, Gemma Bavister, et al. Zebrafish pigment cells develop directly from persistent highly multipotent progenitors. *Nature Communications*, 14(1):1258, 2023.
37. Yongchang Yao and Chunming Wang. Dedifferentiation: inspiration for devising engineering strategies for regenerative medicine. *NPJ Regenerative medicine*, 5(1):14, 2020.
38. Yilin Fan and Tobias Meyer. Molecular control of cell density-mediated exit to quiescence. *Cell reports*, 36(4), 2021.
39. Mariana Pavel, Maurizio Renna, So Jung Park, Fiona M Menzies, Thomas Ricketts, Jens Füllgrabe, Avraham Ashkenazi, Rebecca A Frake, Alejandro Carnicer Lombarte, Carla F Bento, et al. Contact inhibition controls cell survival and proliferation via yap/taz-autophagy axis. *Nature communications*, 9(1):2961, 2018.
40. Bharath Sunchu and Clemens Cabernard. Principles and mechanisms of asymmetric cell division. *Development*, 147(13):dev167650, 2020.
41. Kara L McKinley, Nico Stuurman, Loic A Royer, Christoph Schartner, David Castillo-Azofeifa, Markus Delling, Ophir D Klein, and Ronald D Vale. Cellular aspect ratio and cell division mechanics underlie the patterning of cell progeny in diverse mammalian epithelia. *Elife*, 7:e36739, 2018.
42. Taryn E Gillies and Clemens Cabernard. Cell division orientation in animals. *Current biology*, 21(15):R599–R609, 2011.
43. Yuzhou Feng, Tianpei Yang, John Zhu, Mabel Li, Maria Doyle, Volkan Ozcoban, Greg T Bass, Angela Pizzolla, Lachlan Cain, Sirui Weng, et al. Spatial analysis with spiat and spasim to characterize and simulate tissue microenvironments. *Nature Communications*, 14(1):2697, 2023.
44. Yosuke Nagahata, Kyoko Masuda, Yuji Nishimura, Tomokatsu Ikawa, Shinpei Kawaoka, Toshio Kitawaki, Yasuhito Nannya, Seishi Ogawa, Hiroshi Suga, Yutaka Satou, et al. Tracing the evolutionary history of blood cells to the unicellular ancestor of animals. *Blood, The Journal of the American Society of Hematology*, 140(24):2611–2625, 2022.
45. Clare E Buckley and Daniel St Johnston. Apical–basal polarity and the control of epithelial form and function. *Nature reviews Molecular cell biology*, 23(8):559–577, 2022.
46. Kavitha Mukund and Shankar Subramaniam. Skeletal muscle: A review of molecular structure and function, in health and disease. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 12(1):e1462, 2020.
47. Daniel T Zuch, Siamsa M Doyle, Mateusz Majda, Richard S Smith, Stéphanie Robert, and Keiko U Torii. Cell biology of the leaf epidermis: Fate specification, morphogenesis, and coordination. *The Plant Cell*, 34(1):209–227, 2022.
48. Aleca M Borsuk, Adam B Roddy, Guillaume Théroux-Rancourt, and Craig R Brodersen. Structural organization of the spongy mesophyll. *New Phytologist*, 234(3):946–960, 2022.
49. Shifaan Thowfeequ and Shankar Srinivas. Embryonic and extraembryonic tissues during mammalian development: shifting boundaries in time and space. *Philosophical Transactions of the Royal Society B*, 377(1865):20210255, 2022.
50. Takumi Okubo, Nicolas Rivron, Mio Kabata, Hideki Masaki, Keiko Kishimoto, Katsunori Semi, May Nakajima-Koyama, Haruko Kunitomi, Belinda Kaswandy, Hideyuki Sato, et al. Hypoblast from human pluripotent stem cells regulates epiblast development. *Nature*, pages 1–3, 2023.
51. Alma Armenta-Medina, C Stewart Gillmor, Peng Gao, Javier Mora-Macias, Leon V Kochian, Daoquan Xiang, and Raju Datla. Developmental and genomic architecture of plant embryogenesis: from model plant to crops. *Plant Communications*, 2(1), 2021.
52. Kenneth D Birnbaum and Alejandro Sánchez Alvarado. Slicing across kingdoms: regeneration in plants and animals. *Cell*, 132(4):697–710, 2008.
53. Momoko Ikeuchi, Yoichi Ogawa, Akira Iwase, and Keiko Sugimoto. Plant regeneration: cellular origins and molecular mechanisms. *Development*, 143(9):1442–1451, 2016.
54. Margot E Smit, Cristina I Llavata-Peris, Mark Roosjen, Henriette van Beijnum, Daria Novikova, Victor Levitsky, Iris Sevilem, Pawel Roszak, Daniel Slane, Gerd Jürgens, et al. Specification and regulation of vascular tissue identity in the arabidopsis embryo. *Development*, 147(8):dev186130, 2020.
55. Johannes C Schittny. Development of the lung. *Cell and tissue research*, 367:427–444, 2017.
56. Maria Romanova-Michaelides, Zena Hadjivasiliou, Daniel Aguilar-Hidalgo, Dimitris Basagiannis, Carole Seum, Marine Dubois, Frank Jülicher, and Marcos Gonzalez-Gaitan. Morphogen gradient scaling by recycling of intracellular dpp. *Nature*, 602(7896):287–293, 2022.
57. Yangjie Hu, Moutasem Omary, Yun Hu, Ohad Doron, Lukas Hoermayer, Qingguo Chen, Or Megides, Ori Chekli, Zhaojun Ding, Jiří Friml, et al. Cell kinetics of auxin transport and activity in arabidopsis root growth and skewing. *Nature Communications*, 12(1):1657, 2021.

DRAFT