



Edinburgh University

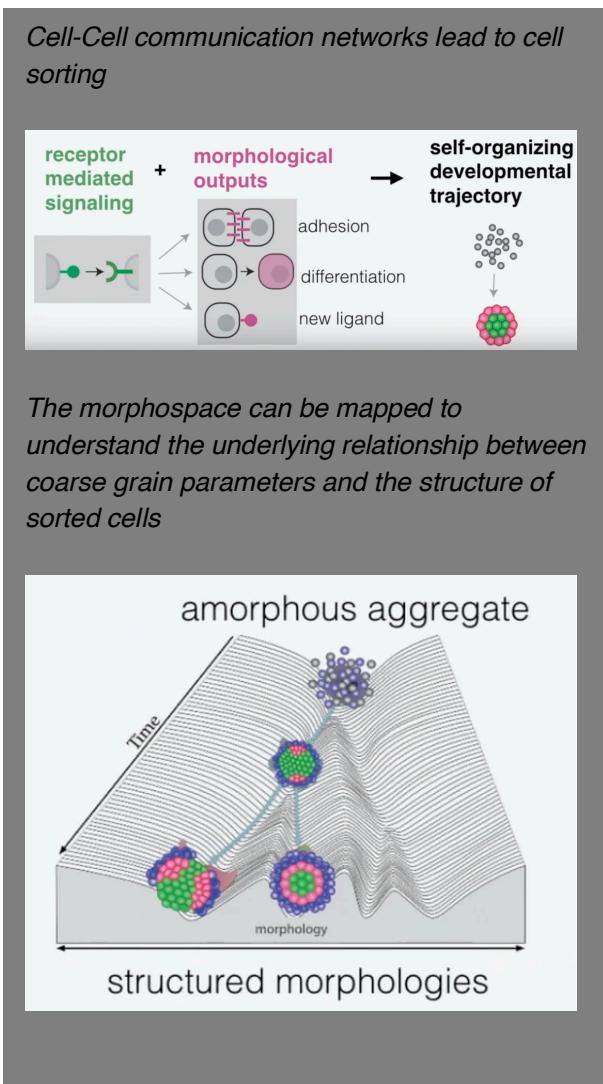
Senior Honours Project

Modelling self-assembly of tissues in silico

Author Hector Crean

Supervisor Davide Michieletto

Graphical Abstract



In Brief

Genetically encoded algorithms control individual and collective cell behaviour. We aim to reconstruct these programs in silico to understand their logic. In vitro and in vivo studies have identified many genes that control cell-cell signalling and cell morphology. Most development systems, however, use contact cell-cell signalling interactions to induce morphological responses. We consequently explored whether simple synthetic circuits, in which morphological changes are driven by contact cell-cell signalling interactions, could suffice to generate self-organizing multicellular structures.

We based our computation signalling network on the synnotch receptor system, which induces user-specified transcription programs in response to detecting molecular signals from neighbours. We chose changes in cell adhesion (caused by varying expression of cadherin) as a morphological effector. It was hoped that we could identify key hallmarks of natural development systems in our simulations: self-organisation into multi-layered structures; divergence of genetically identical cells into distinct types, and symmetry breaking. Our success in finding these features was limited.

Computational modelling allows us to track the trajectories of these multi-cellular systems through morphospace (space of forms). It reveals that cell-signalling allows a more thorough exploration of the morphospace, than is possible with mechanical induced sorting only. Plotting the energy landscape corresponding to each point in the morphospace would allow an understanding of stable and meta-stable states in the morphospace, and how tissues transition between them. As in protein folding, Evolution has chosen parameters (cell adhesion, transcription rate etc.) that lead to appropriate cell sorting,

Chapter 1: Introduction:

1.1 Background

Cells can differentiate into specific cell types that spatially self-organise into tissue architectures. This is achieved by cells communicating, and inducing the expression of proteins, in one-another. In embryology, there is specific interest in how regulatory programs encoded in the DNA of a single fertilized cell can integrate together into the gross organisation of a whole organism. Toda et Al. have noted that it is still an open question as to how compact DNA programs encode algorithms that allow individual cells to construct complex macrostructures ([Toda et al., 2019](#), [Velazquez et al., 2018](#)). To address this question, there has been an attempt to construct in silico models of cell morphogenesis. Elowitz et al. have particular advocated the principle of ‘build[ing] life to understand it’: coding multicellular collective behaviour from the bottom up may provoke insight ([Elowitz et al., 2010](#)).

Of course, it is important that in silico models reconstruct and work with observations in the biological realm. Cell-cell regulatory networks have frequently been studied in vitro, using methods such as custom light sheet microscopy, to track cell divisions, individual cell fates, and dynamic spatial distribution of cells during development ([McDole et al., 2018](#)). Our model will constantly have to make reference to this research. Of primary importance is finding a cell ‘chassis’ in the biological realm capable of executing genetically encoded functions, but simple enough to model ([Sasai et al., 2013](#)). Along with this ‘hardware’, we must choose appropriate ‘software’ (genetic programs) that can drive morphogenetic behaviour (self-assembly, differentiation, migration). A key question is the minimal componentry necessary to program multicellular assemblies. It is a matter of philosophical debate as to whether these ‘toy models’ are useful facsimiles of biological systems: biologists often suggest them to be too coarse grain to have any real use. In view of this, a theme of this report will be whether a small range of parameters can accurately be used to model biological systems. While this report is primarily focused on the biological realm, it is equally applicable in studying animal collective behaviours, such as flock formation and sociality of ants, which also arise in the absence of central control, and are generated by local interaction among individuals.

1.2 Simulating tissue via Cellular Potts Model

In the Cellular Potts Model (CPM), a cell monolayer is represented on a two-dimensional lattice of size L by L, where in our case L=60. Lattice points are each assigned with a ‘spin’ (or ‘cell index’), $0 \leq \sigma \leq N_{\text{cells}}$, where N_{cells} is a variable that controls the number of cells in the simulation. Lattice points on the periphery of one another, with the same spin, are seen collectively as a cell ([Voss-Bohme et al., 2012](#)). Spin $\sigma=0$ generically denotes fluid sites, but we will be simulating a confluent monolayer (packing fraction unity), and so we will need no recourse to this. The cells evolve dynamically via a Monte Carlo algorithm based on an energy function, H

$$H = \sum_{\langle x, x' \rangle} J(\sigma(x), \sigma(x')) + \lambda \sum_{i=1}^N (a_i - A_i)^2$$

where x and x' denote lattice sites, a_i and A_i are the current and target area of the i -th cell (which we set to 16). The first term in the expression represents ‘interfacial effect’, and the second ‘approximate

area conservation'. A third term, the cell motility, is often added, but will not be considered here. J is chosen, in the most simple cases, such that:

$$J(\sigma, \sigma') = \begin{cases} 0, & \text{for } \sigma = \sigma' \\ \alpha, & \text{for } \sigma \neq \sigma' \end{cases}$$

where $\alpha > 0$ determines the interfacial energy between two cells. The magnitude α affects cell shape and the flexibility of the cell boundary ([Szabo et al., 2010](#)). The CPM, each iteration, randomly picks L^2 lattice points, and attempts to change the lattice points' spin to the spin of one of its neighbours. This change of spin will be accepted based on the Metropolis test, related to the associated change of energy, ΔH .

The lattice structure of these models offer ease in determining which cells are neighbouring, which are just any pairs in adjacent lattice sites. On the other hand, the model does not permit great spatial flexibility of cells: they will always be pixelated ([Merks, 2015](#)). The vertex model can be used as an alternative to the CPM, which natively allows for more complex cell shapes ([Kabla, 2012](#)).

A number of other features must be incorporated into our CPM for later functionality. We define an array, referred to as 'Type', which maps cell spins to specific cell 'types'. Each 'type' will have specific properties associated with it. During each sweep of the algorithm, furthermore, we count the number and type of cell boundaries, and store the information in an array, 'HeterotropicBoundary'. This allows us to later analyse cell sorting. Modules allowing us to calculate the Mean Squared Distance of cells from one another, the radius of gyration/asphericity of cells, as well as number of cells of each type were added for a similar purpose. A code was written allowing a visual representation of cells spins, cell adhesiveness and cell types in gnuplot. This code will all be available in the public domain.

Chapter 2: Mechanical effectors of cell sorting

2.1 Phase Space within which to do our modelling.

The phase (glassy/fluid) of biological tissues is contingent on their position in parameter space. Of particular interest is the phase space of dense tissues, whose cells are at confluence (packing fraction equal to zero), which is relevant to the study of wound healing, embryonic development, and cancer metastasis. In the glassy phase, dense tissues exhibit caging, dynamical heterogeneities and viscoelastic behaviour. The cells are positionally ordered and their motion is heavily constrained. In the fluid phase, on the other hand, cells are motile and the position order drops. Parameters which control the solid-liquid transition have been explored by pursuing systematic investigation of the tissue phase diagrams, via *in silico* models.

Several metrics correlate with glassy dynamics, such as sub-diffusive behaviour. M. Chiang and D. Marenduzzo constructed an *in silico* model of dense tissues using the CPM, analysing mean square displacement of cell trajectories and the effective diffusivity as a function of various control parameters ([Chiang et al., 2016](#)). They inferred that *adhesion/stiffness*, α , and *cell motility*, P , are the key parameters in determining the rigidity of the tissue, with motility particularly triggering a transition between the solid to fluid regime. This project, building on previous work of Chiang and Marenduzzo, was able to confirm these results. For the sake of simplicity, we assumed no cell motility ($P = 0$), so that the only control parameter was α , linked to interfacial tension, γ , of the cell. This characterises

the competition between cell-cell adhesion and cortical tension. In the vertex model, a control parameter linked to shape is a key sorting parameter, which is extremely useful; it enables biologists to determine experimentally the mechanical properties of tissues through imaging cell membranes. This is not available within the CPM, and is a potential drawback of this modelling framework.

In order to analyse the relationship between α and the phase of the tissue, we can calculate the MSD of the cells:

Mean square displacement of cells vs. time on a log-log plot for $\alpha_{stiff}=1-2$

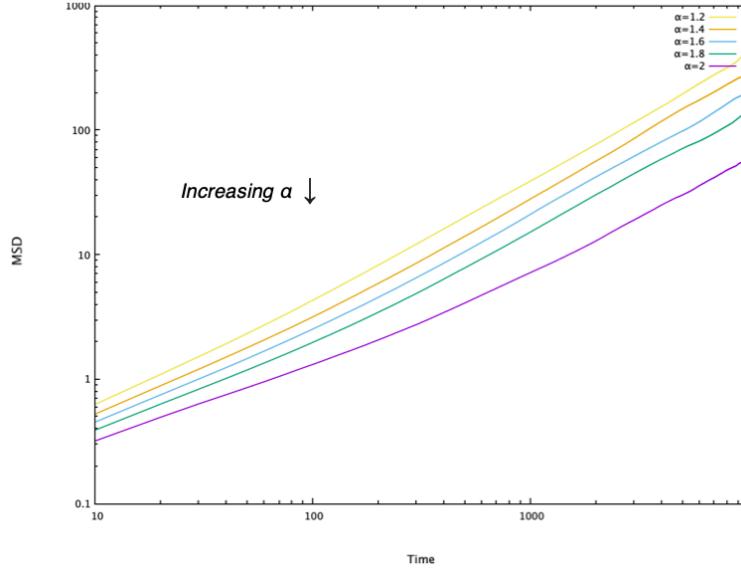


Figure 1:

It has been computationally prohibitive to plot this relationship till $t=100,000$, as has been done by Chiang and Marenduzzo. Their research reveals that the difference between the final MSD value for $\alpha = 1.2$ (yellow curve) and for $\alpha = 1.4$ (orange curve) is more than a decade and is significantly larger than the difference between the final MSD values for higher α . Their data shows that for $\alpha < 1.4$, the cells are diffusive (as the MSD increases linearly with time). For $\alpha > 1.4$, the system remains sub-diffusive, indicated by a gradient that is less than one.

In sum, the system displays fluid-like behaviour below $\alpha = 1.4$, and glassy-like behaviour above it.

This gives us an indication of how our choice of α will affect the phase space of the tissue in subsequent modelling. Clearly, more thorough analysis is required, particularly with regard to how mixtures of cells with different values of α will impact on the collective tissue phase. We must also explore the coupling between adhesion and the timescales of rearrangement of cells. More importantly, however, these results vindicate the use of the CPM for modelling cells: it correctly predicts glassy behaviour in dense model tissues, in agreement with experimental observations. We do, however, need to perform experiments *in vitro* to gain a more precise understanding of the α values that replicate the structure of typical tissues.

2.2 Mixtures of cells with different α values

Having explored how the phase of tissues can be controlled by a parameter related to cell adhesion/stiffness, α , we focused on analysing glass transitions in mixtures of cells with different properties (variable interfacial tension within the population) to refine our model such that it would be applicable to real cell monolayers. At a later date, it would be of interest to model situations where cells can grow/shrink, divide/die, and intercalate. These are cell behaviours vital in the biological realm. Equally, as has been noted, the generality of the study could be increased by considering suspensions of cells of different packing fractions, where the “P’eclet number” (relative importance of self-propulsion and cell diffusion) is likely to be a governing parameter. For the moment, however, these complications are unnecessary.



Figure 2: Common cell behaviours

We modelled two populations of cells, half with α_{stiff} , and the other half having α_{soft} . The cell sorting was visualised.

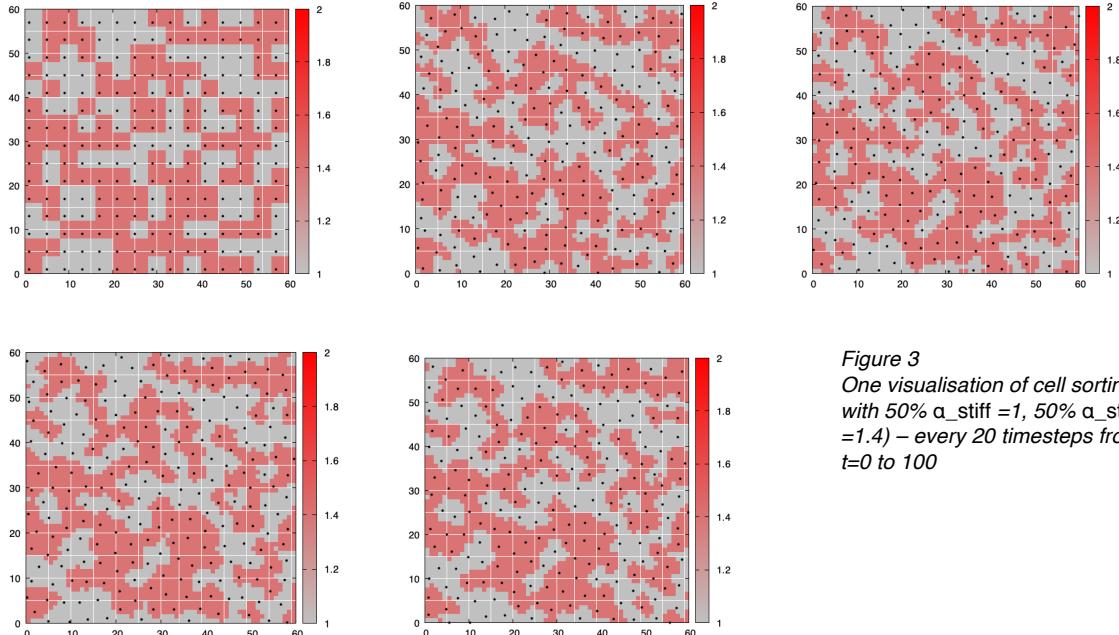


Figure 3
One visualisation of cell sorting,
with 50% $\alpha_{\text{stiff}} = 1$, 50% $\alpha_{\text{stiff}} = 1.4$ – every 20 timesteps from $t=0$ to 100

We can also visualise the system at in a longer time limit (time $t=10,000$) for other α_{stiff} values, where $\alpha_{\text{soft}} = 1$. This gives an impression of the ‘steady-state’ cluster formations. The timescale of getting to an ‘steady-state’ is influenced by the adhesiveness of the cells.

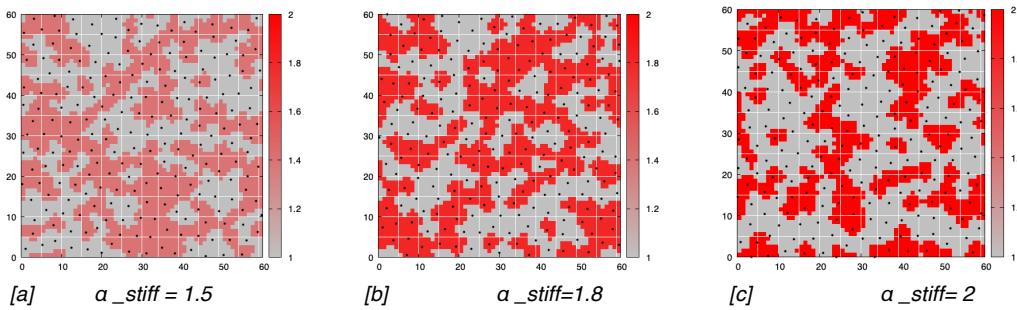


Figure 4

There is clearly phase separation (or “sorting”) into clusters. We can investigate the phase-space of these simulations by, again, calculating the MSD. We shall maintain α_{soft} at 1, and vary f (fraction of ‘soft’ to ‘stiff’ cells) and α_{stiff} .

Mean square displacement of cells vs. time on a log-log plot for $\alpha_{\text{stiff}}=1-2$, with $f=0.25$

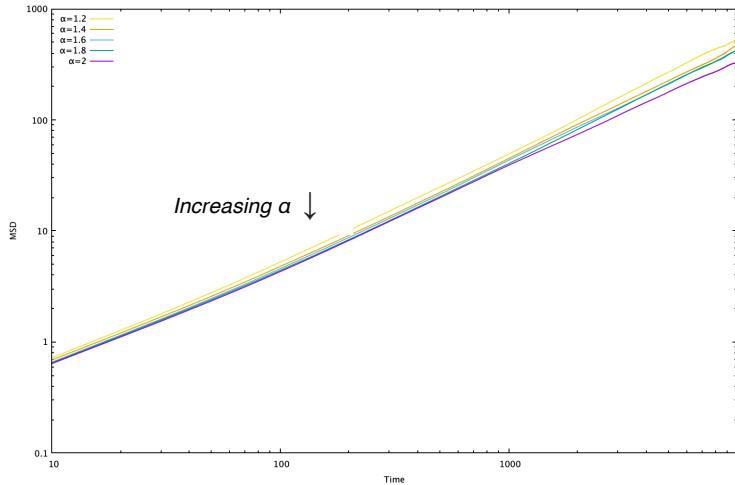


Figure 5
Unsurprisingly, when the fraction of stiff cells is relatively low, the impact on the phase of the tissue is subtle. MSD for exclusively α_{soft} cells is more than double that for a mixture of α_{soft} and α_{stiff} ($=2$) cells in a ratio of 3:1 over 10,000 time steps. This suggests that a small amount of adhesive cells can 'seed' the formation of clusters, which impact on the collective tissue phase.

Mean square displacement of cells vs. time on a log-log plot for $\alpha_{\text{stiff}}=1-2$, with $f=0.5$

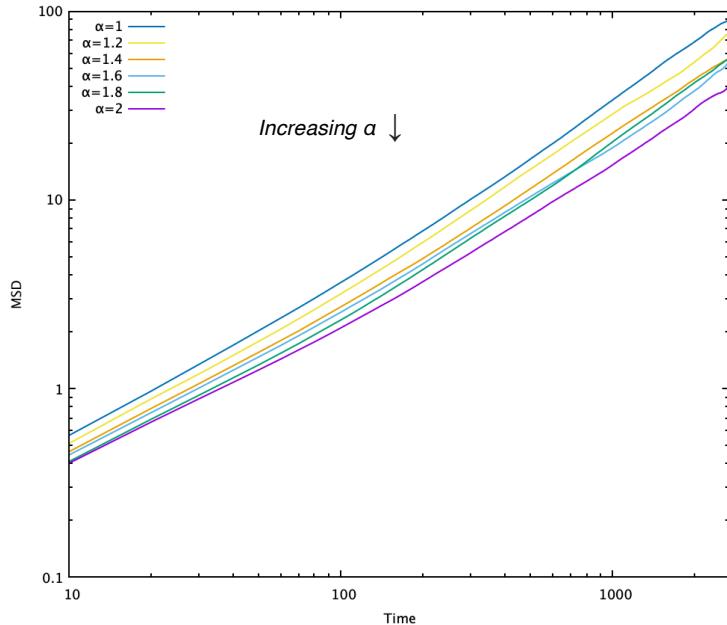


Figure 6
This suggests that in mixtures of cells, $\alpha_{\text{stiff}}=1.4$ is still a transition point from fluid to glassy. The MSD for values of $\alpha_{\text{stiff}}=1.4-1.8$ are almost indistinguishable, while the region between $\alpha_{\text{stiff}}=1.8-2.0$ appears to constitute another change in phase.

This seems to be due to structural difference between how clusters of cells pack. As can be seen in the above diagrams, with $\alpha_{\text{stiff}}=1.8$, the cells of similar adhesion form large connected networks. On the other hand, when $\alpha_{\text{stiff}}=2$, the more adhesive cells form long, isolated clusters. These appear to represent 2 distinct phases.

Mean square displacement of cells vs. time on a log-log plot for $\alpha_{\text{stiff}}=1-2$, with $f=0.75$

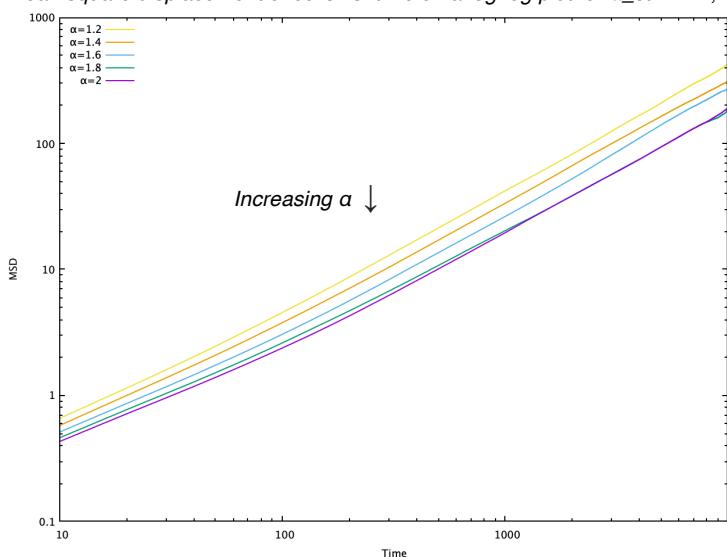


Figure 7

The MSD of cells with $\alpha_{\text{stiff}} > 1.6$ appear to be in a rigid phase, where there movement is fairly similar.

Later, we will be modelling tissues with up to 4 cell types, each with different values of α . The number of permutations of cell adhesiveness, and ratios of cell types, is large.

As can be seen, different regions of parameter space are correlated with different tissue types. In particular, our results seem to reproduce the two basic forms of embryonic: epithelial and mesenchymal cells. Epithelial cells are compact tissues with strong intercellular contact, thereby acting as physical barriers to other cells and molecules. They often organise into sheets with stable neighbourhoods. This is reminiscent of figure 4 [c], where we see segments of more adhesive cells form sheets that make the collective tissue more solid, and suggests that even simple models can reproduce biological phenomena. An epithelial-mesenchymal transition occurs when adhesion between cells decreases, enabling mesenchyme to migrate along the extracellular matrix. We postulate that this occurs in the region of $\alpha = 1.4-16$, depending on the proportion of cell types with different adhesiveness.

In sum, mechanical properties of cells alone can induce cell sorting. Before continuing, it should be recognised that mechanical input signals have been shown to also influence cellular processes. For instance, fluid shear stress inhibits endothelial cell proliferation ([Steward et al., 2015](#)). Mechanical forces thus represent an ersatz cell signalling channel. Next, we shall consider how generic ‘cell-cell’ circuits work alongside differential cell adhesion to control cell-sorting.

Chapter 3: Cell Circuits

3.1 Signalling mechanisms between cells

Over long range, cells can communicate by the diffusion of morphogens through the intercellular space (Autocrine/Paracrine signalling). At short range, cells often communicate in direct contact by the ligand of one cell connecting with a receptor of another (Juxtracrine signalling). Transduction of the signalling leads to the induction or repression of target genes through activation or repression of transcription factors ([Johnson et al., 2017](#)). This can result in morphological outputs such as adhesion, differentiation, pattern formation, proliferation, and fate decision (terminal or plasticity).

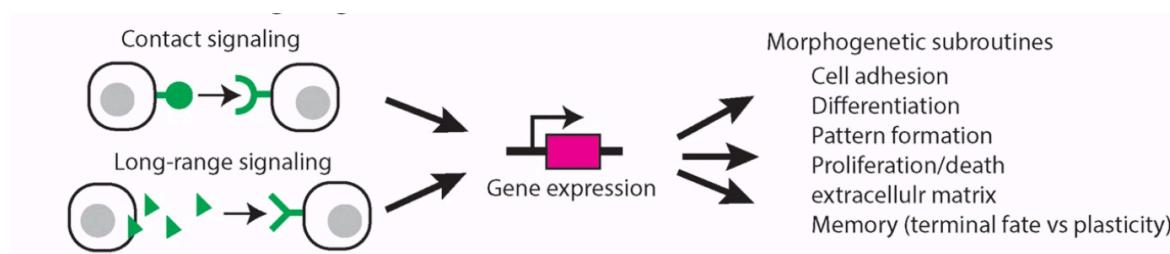


Figure 8 Cell signalling channels

Finding an appropriate chassis for our cell circuit

Continuing on from our modelling of how stiffness/adhesion of cells leads to cell-sorting, we looked to create *in silico* toy models in which signalling led to two distinct morphological outputs: expression of cadherin, which leads to greater cell-cell adhesion, and the differentiation of the cell into a different

'type'. For later purposes, we also wanted to be able to model the expression of a new ligand. Our simulation in silico required a biological analogue, which could be tested in vitro. Morsut et Al. have engineered mouse fibroblast cells with a series of cell-cell communication circuits. These 'SynNotch' circuits induce the expression of cadherin (adhesion molecule) as well as fluorescent proteins (Morsut et al., 2016). These cell circuits are contact dependent, with orthogonal signalling pathway, making them particularly simple to model. When in contact, sender cells trigger SynNotch receptors on corresponding receiver cells, which induces the production of E-Cadherin and GFP. Receiver cells, consequently, cluster due to their higher affinity for each other than for sender cells.

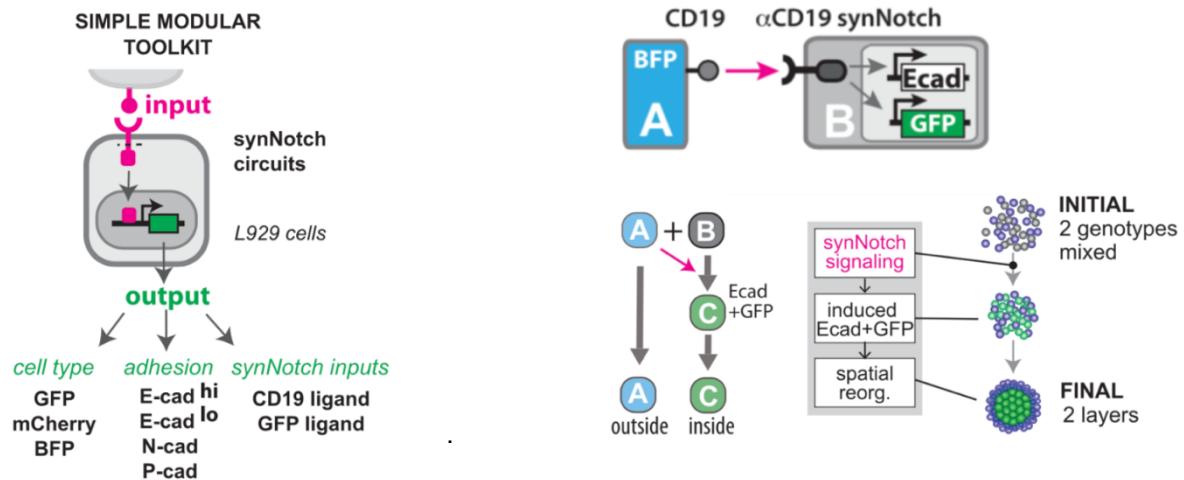


Figure 9

[a]

Molecular components used for assembly of simple morphological circuits. Two synNotch ligand-receptor pairs (surface ligands CD19 and GFP) were used for cell signaling, three fluorescent proteins as markers of "differentiation," and several cadherin molecules (expressed at different levels) as morphological outputs.

[b]

*In an initial experiment, one group of cells acted as signal **sender cells** and expressed a ligand (CD19 protein) on their surfaces. These cells also expressed a blue fluorescent protein (BFP) for tracking purposes. The second group of cells acted as signal **receiver cells** and expressed a SynNotch receptor capable of binding to the sender cells' CD19 ligand. In response to binding, these receiver cells activated both E-cadherin and a green fluorescent protein (GFP).*

Fluorescent proteins are used to tag cells, and are a metric for cell type. For instance, Morsut et Al. used 'mCherry', which is a member of the mFruits family of monomeric red fluorescent proteins (mRFPs). Equally, they use green fluorescent proteins (GFP) that exhibit bright green fluorescence. Both these fluorescent proteins are used as reporters of gene expression.

More complex, multi-layered, cell circuits can be obtained. Morsut et Al. modified sender and receiver cells such that receiver cells produced a ligand on activation, which could in turn interact with a receptor on the sender cells. The sender cells, upon binding with this ligand, produce a low level of E-cadherin. Playing with different adhesion molecules, a wide array of symmetry/asymmetric structures can be generated. The structures produced exhibit self-organisation, divergence of genotypically identical cells into distinct 'types', and symmetry breaking – all key hallmarks of natural development systems.

3.2 Implementation in silico of adhesion matrix

The term J in our Hamiltonian, accounting for interfacial tension between cells in the bulk, was previously a square matrix of σ 's, with zeros as in the leading diagonal,

$$\text{adhesion matrix} \quad \begin{matrix} & \sigma_1 \dots \sigma_n \\ J = & \begin{pmatrix} J_{1,1} & J_{1,n} \\ \vdots & \\ \sigma_n & J_{n,n} \end{pmatrix} \end{matrix} \quad \text{cell types}$$

Figure 10 Cell adhesion matrix, J , where sigmas refer to the cell index

Our adhesion matrix, J , becomes more complicated with the introduction of cell circuits. There are different adhesion strengths between different cell types, depending on the strain and density of cadherin expressed on the cell surface. It is non-trivial to work out this adhesion matrix, and Zhang et. Al have discussed various models for this interaction (Zhang et al. 2011, Steinberg et al., 1994, Foty et al., 2005).

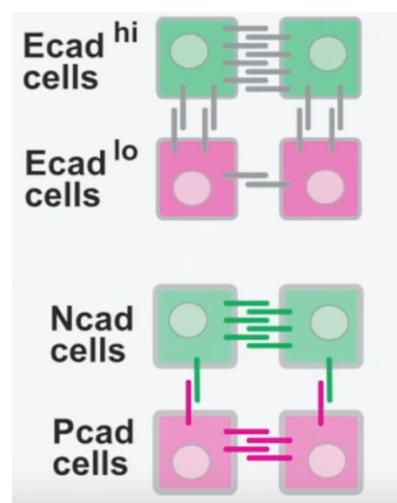


Figure 11
Different cells have different adhesions to other cells, due to both varying expression and type of Cadherin on their surface. In this case Ecad, Ncad and Pcad, all from the superfamily of cadherin molecules, are being expressed.

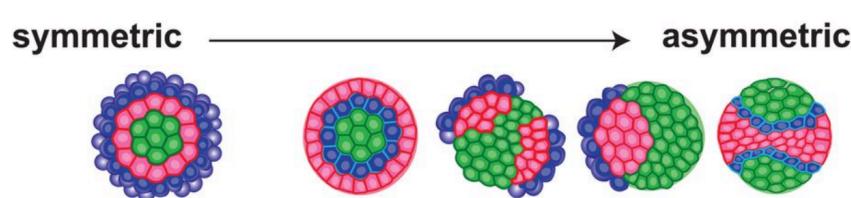


Figure 12
It has been observed in vitro that a different adhesion matrix between cells affects the (a)symmetry of the resulting tissue structure.

3.3 Implementation in silico of cell-cell contact signalling

Cell-to-cell communication networks comprise both intra- and intercellular processes. In order to simplify modelling, intracellular signal transduction networks are generally treated as “black boxes” with specified input-to-output response relationships. This abstracts molecular detail yet still captures essential dynamic properties of the system (Thurley et al., 2018, Qian et al., 2018).

Cells have a cell-specific density of ligands/receptors on their surface. For two adjacent cells, the signalling strength between them is a function of the number of ligands/receptors in contact. Varying degrees of contact are correlated with differing signalling strengths. In the biological realm, furthermore, the response of a cell to this signalling is an integral of the signal strength over time, and so is not just dependent on the instantaneous ligand-receptor contact, but the history of this contact in the recent past. In addition, rather than certain ligand-receptor interactions leading precisely to a morphological outputs, the same ligand-receptor bonding can lead to different outputs, only differentiated by the magnitude of the signal strength as well as the time over which the signalling occurs. For the sake of simplicity, our *in silico* model will assume a one-to-one correspondence between signalling and a morphological output.

Morsut et Al. have derived a phenomenological equation relating the cell response to the amount of ligands in contact with receptors in order to perform their simulation; it has been observed *in situ*, and it is then constructed via various shaping parameters, such that the integration of the function gives the appropriate relationship between ‘response’ and gene activation.

$$\frac{dR}{dt} = \frac{\alpha}{1 + e^{-(S-\beta)/\varepsilon}} - \frac{1}{K} R \quad \text{where} \quad S = \sum_{SN} \frac{L * \text{contact surface area with } \sigma}{\text{sur}(\sigma)}$$

- S is the amount of ligands in touch that the sender cell has, and the amount of surface of the receiving cells that see ligands
- α, β and ε are shaping parameters.

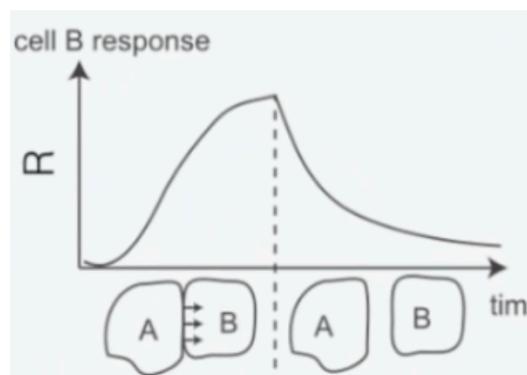


Figure 13 Distributions are used to phenomenologically model cell input-response relationships. Cells gain more ‘points of response’ when greater surface area is in contact.

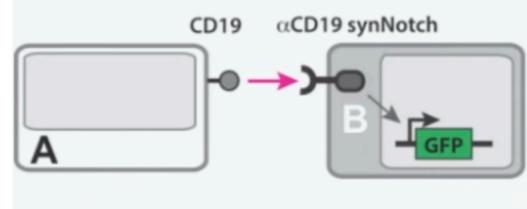


Figure 14 Simple cell-contact signalling

Cell-cell contact dependent signalling can be implemented via a points-based system, where cells gain ‘points’ for having one of its lattice points being in contact with the lattice point of another cell. For this purpose, we can construct an array tracking the ‘heterotrophic boundary length’, for each cell. This comprises a 6 by 225 array, with details of the number of contact points each cell has with other cell types at a given moment. This information can then be used to calculate the point at which a cell has had sufficient stimulation to express cadherin.

An additional concern arises when considering the two methods by which morphological changes are activated: gene inhibition and gene activation. Gene activation, for instance, has 6 main forms (in terms of the biomechanics), relating to how the RNA/DNA polymerase acts. This means that gene

activation is often not permanent. If cell A can cause a change in cell B to state B', B' can transition back to B after a ‘decay time’, if it does not receive sufficient signalling through binding with cell A’s receptors. The duration of gene activation can be altered by playing with stability of transcription factors. We must therefore also incorporate in our simulation a mechanism whereby cells can transition back to their original state.

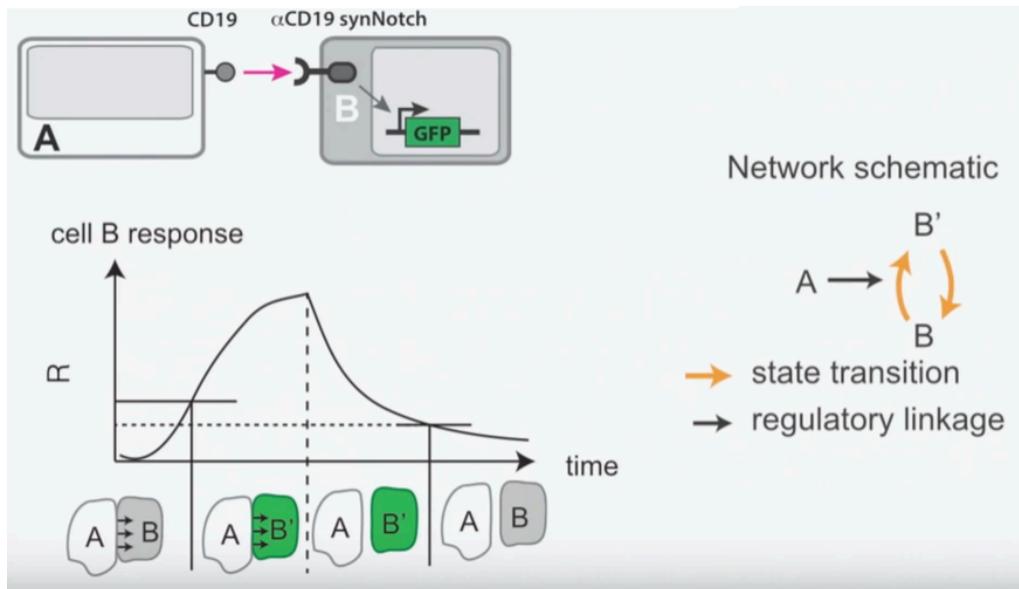


Figure 15 Gene expression is not permanent. Cell B' can transition back to B after a ‘decay time’, if it does not receive sufficient signalling from A.

While we did incorporate the functionality to control cell switching via heterotopic boundary lengths ,it was decided that the relevant features could be elegantly modelled more simply. During the metropolis step of the CPM, we introduced a ‘gene switch’ step. If the spins of two adjoining lattice points are different (indicating different cell types), a random number between 0-1 is generated. If this is less than the ‘k_on’ number, then cell that the lattice point belongs to will change type. In this way, it simulates the fact that more cell contact with sender cells leads to larger ligand exposure, and so greater probability of a receiver cell changing type. Although it only incorporates information from one timestep, the speed of the simulation is fast.

Chapter 4: Measuring cell sorting in cells with one-layer cell-circuits

4.1 Equilibrium number of cell types

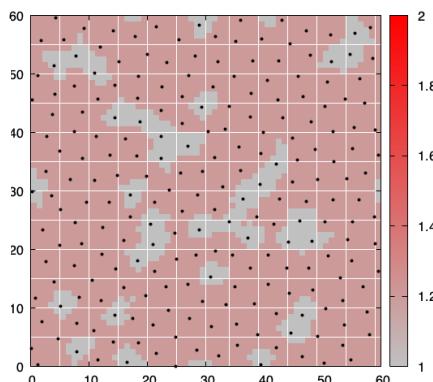
To get an impression of how cell adhesion effects cell sorting, we plotted the graphs that visually represented how cells with specific α values, and cell types, clustered. Our cell circuit was the simplest possible, with B cells (red) turning to C cells (green) in the presence of A cells (grey). A and C cells have an equal adhesiveness, referred to as α_{stiff} , while B has the baseline adhesiveness, α_{soft} , where $\alpha_{soft} < \alpha_{stiff}$.

We modelled across a long time period, such that cell sorting would reach a steady-state.

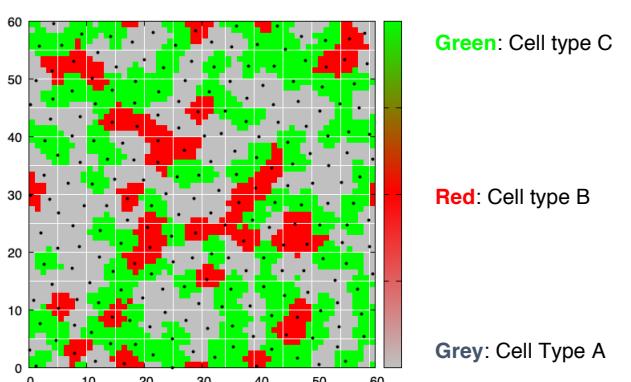
a_stiff VALUES OF CELLS

a_stiff 1.2

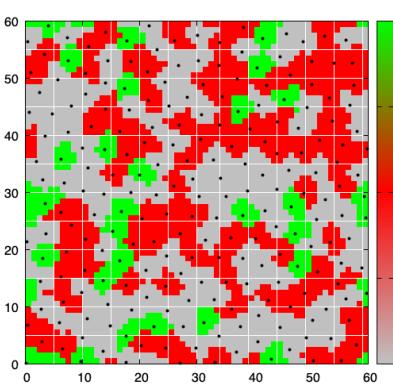
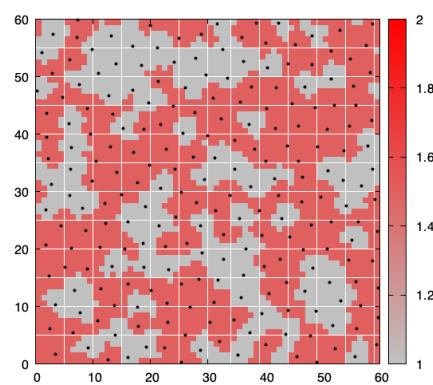
Alpha values are of cells are visually depicted by a grey-red colour palette. Lighter grey represents alpha=1, and bright red is alpha=2.



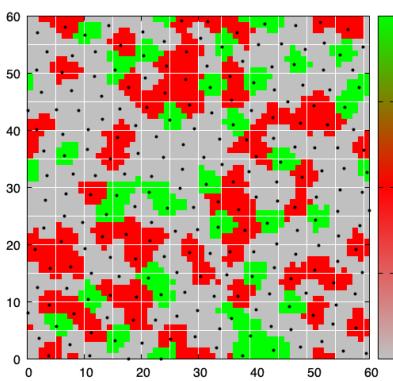
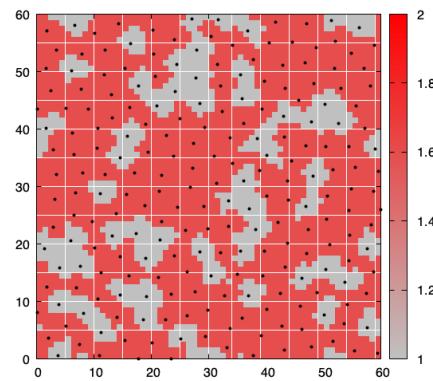
CELL TYPES



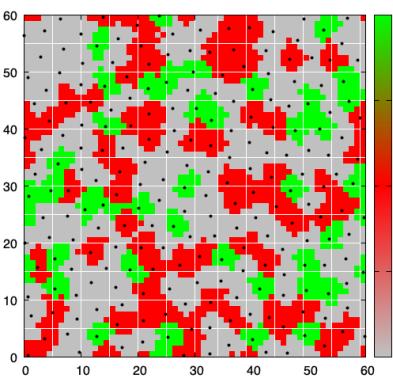
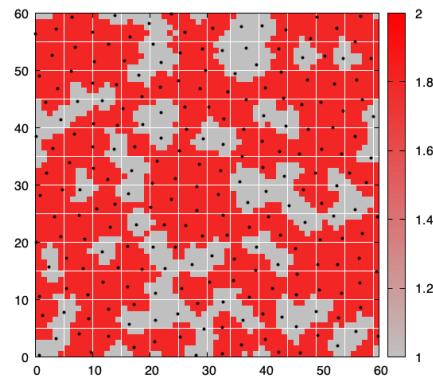
a_stiff 1.4



a_stiff 1.6



a_stiff 1.8



a _stiff 2.0

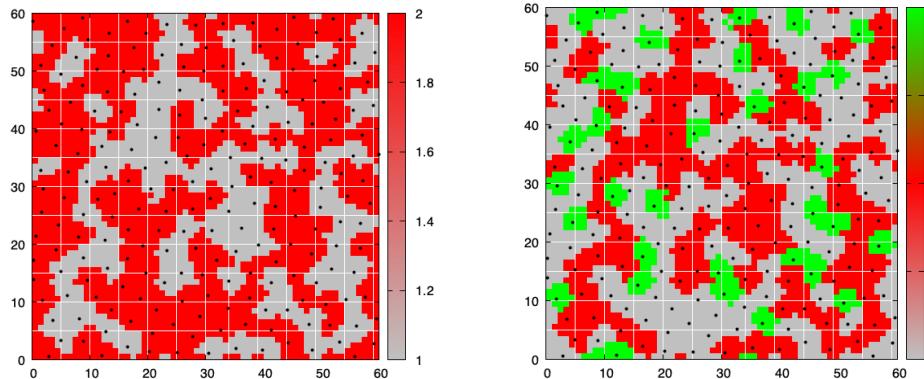


Figure 16: Initial ratio of type A to type B cells, $f=0.5$; probability of A-B contact leading to cell switch, $k_{on} = 0.05$; probability of cell type C switching back to B type, $k_{off} = 0.05$

This extremely simple set-up produced some cell-sorting. Cell types A and C, which both have the same value of alpha, appear to cluster together. Additionally, we might note that there are less C cells in the higher alpha regime. Presumably, this is caused by type A cells clustering more at higher α_{stiff} , and thus having less surface exposed to interact with type B cells. In order to substantiate this finding, we compute fraction of cell types in the long time limit, for different values of α_{stiff} , as well as a parameter (k_{off}) that controls the likelihood of a cell type C to switch back to a cell type B.

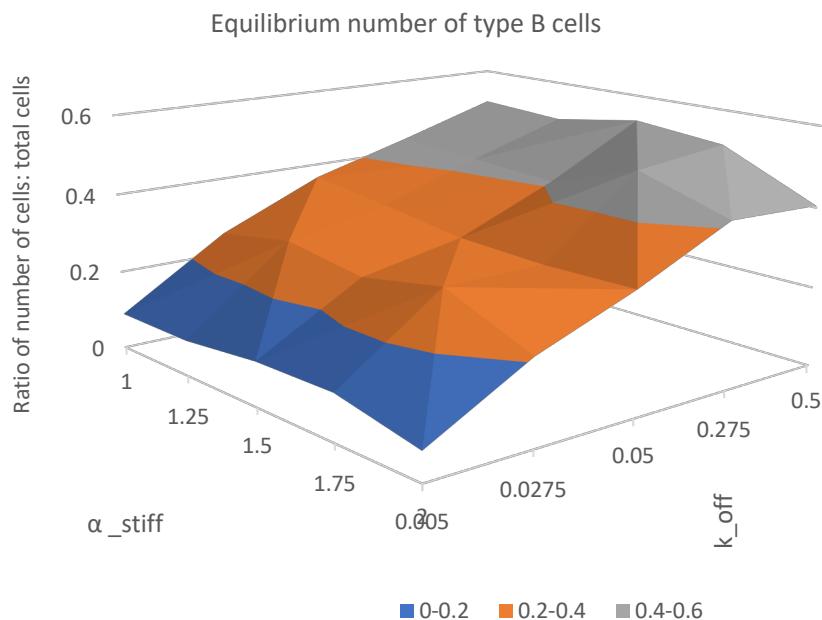


Figure 17

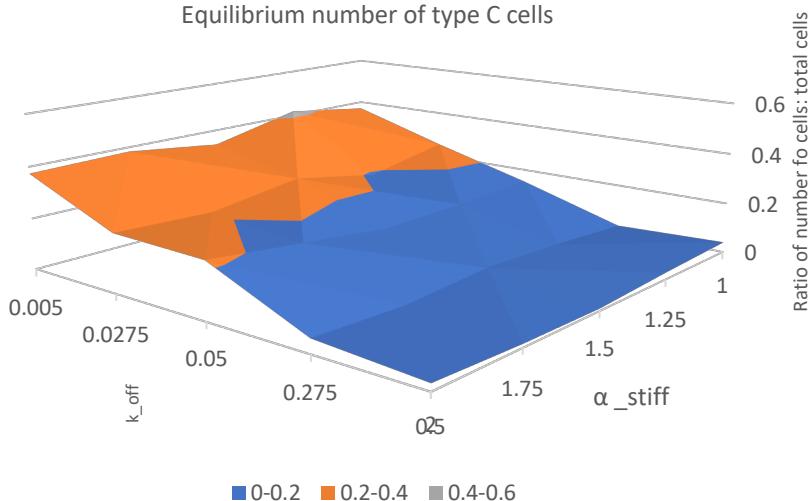


Figure 18

It was thought that higher α_{stiff} would cause A type cells to cluster together, and thus reducing their surface contact with B type cells. Less contact between A and B cells would reduce the number of B cells converted to C cells. While this is somewhat evident from the graphs, the parameter k_{off} clearly has a much greater impact on the equilibrium number of type B and C cells. In other words, the mechanical properties of cells clearly have an influence on cell sorting, but the cell circuit themselves are key drivers to changing the structure of cell formations. This shall be explored more fully later, as we consider more complex cell circuits.

4.2 Cell shape

An important observable to track is the cell shape. What shapes are the most common? Can this be quantified? Can we quantify how long and thin as against short and thick the cells are?

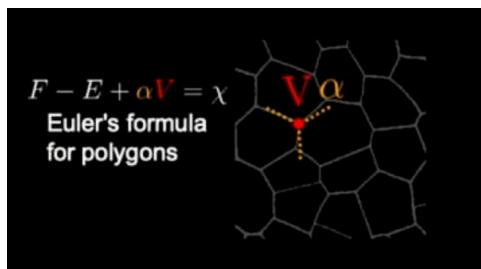


Figure 19 The benefits of implementing a vertex model include the ability to describe the shape of a tissue based on changes in cell shape and rearrangements
In vertex models, it has been noted that Euler's formula for polygons, $F - E + \alpha V = \chi$, implies that epithelial cells will have an average of 6 neighbours.

It has already been noted that the vertex model is more suitable for measuring cell-shape. Within the Cellular Potts Model, there are more generic methods of exploring cell shape (Velazquez et al., 2018). The gyration tensor, S_{mn} , can describe the second moments of position of lattice points that make up a cell.

$$S_{mn} \stackrel{\text{def}}{=} \frac{1}{2N^2} \sum_{i=1}^N \sum_{j=1}^N (r_m^{(i)} - r_m^{(j)}) (r_n^{(i)} - r_n^{(j)})$$

The gyration tensor is a symmetric, and so a coordinate transform can be used to diagonalize the matrix.

$$\mathbf{S} = \begin{bmatrix} \lambda_x^2 & 0 & 0 \\ 0 & \lambda_y^2 & 0 \\ 0 & 0 & \lambda_z^2 \end{bmatrix}$$

These diagonal elements are the principal moments of the gyration tensor, and can be combined to give several parameters that describe the cell shape. The squared radius of gyration is the sum of the principal moments:

$$R_g^2 = \lambda_x^2 + \lambda_y^2$$

The asphericity is the ratio of the principle moments:

$$b = \frac{\lambda_y}{\lambda_x}$$

We can plot the radius of gyration and aspericity of cell types in different regions of parameter space.

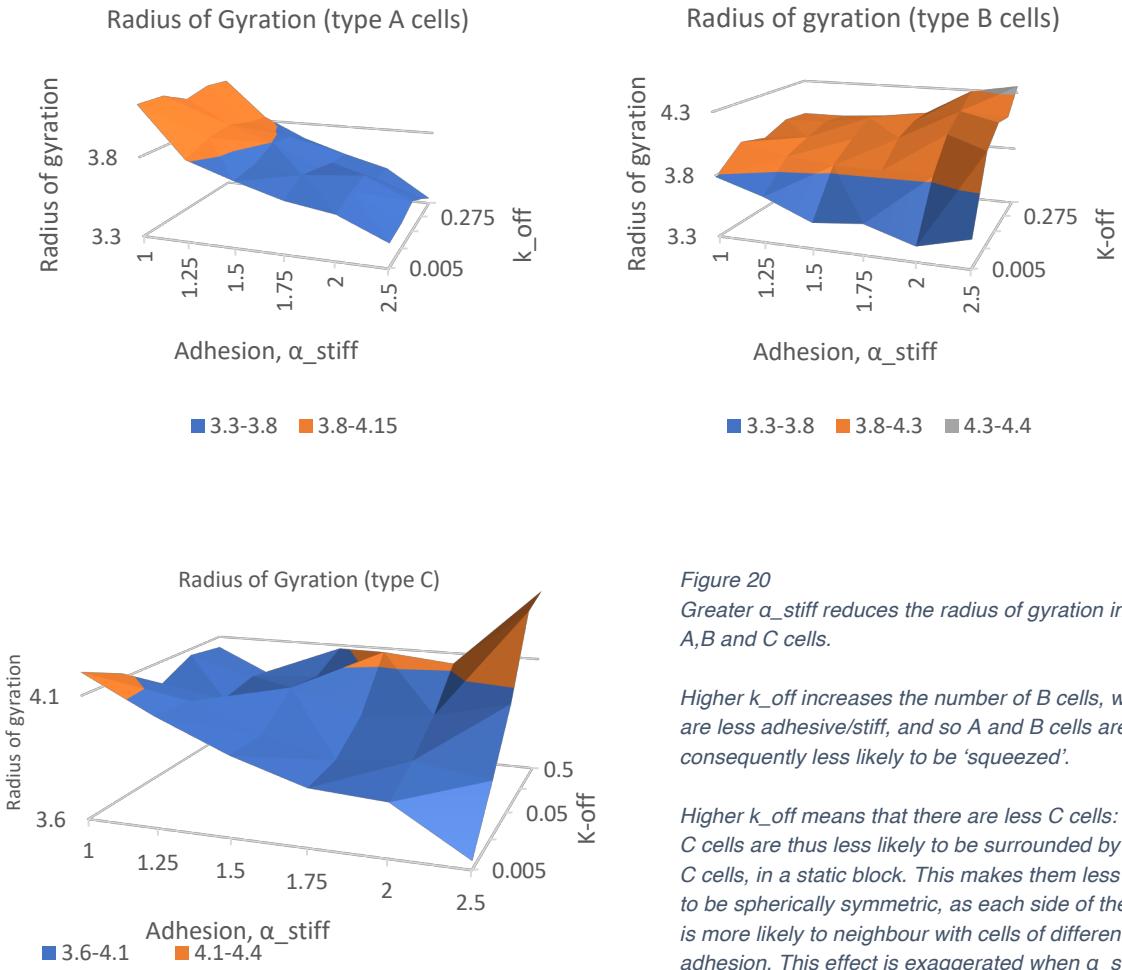


Figure 20

Greater α_{stiff} reduces the radius of gyration in type A, B and C cells.

Higher k_{off} increases the number of B cells, which are less adhesive/stiff, and so A and B cells are consequently less likely to be 'squeezed'.

Higher k_{off} means that there are less C cells: these C cells are thus less likely to be surrounded by other C cells, in a static block. This makes them less likely to be spherically symmetric, as each side of the cell is more likely to neighbour with cells of differential adhesion. This effect is exaggerated when α_{stiff} is large.

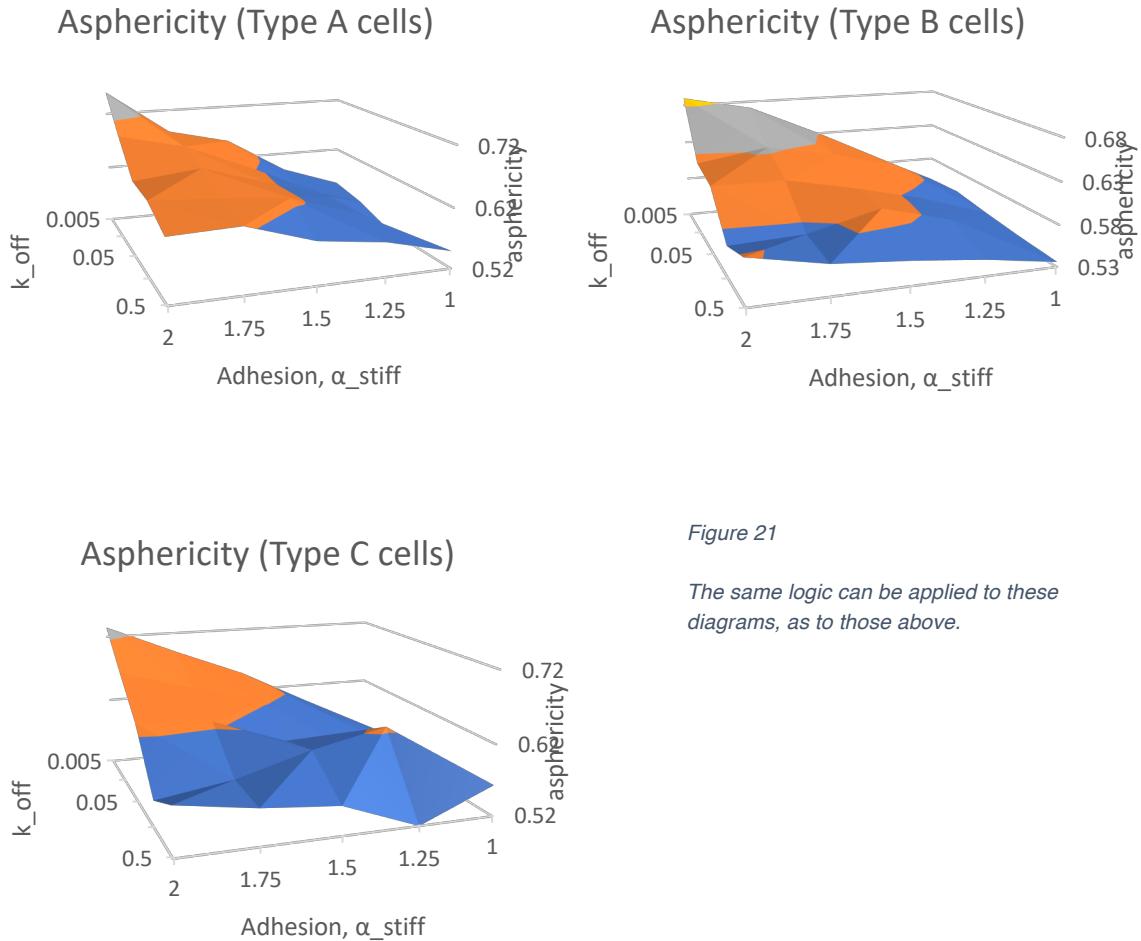


Figure 21

The same logic can be applied to these diagrams, as to those above.

4.3 Heterotypic boundary length

We need to find a way to quantify the type of sorting, and its thoroughness. The most thorough sorting is arguably one in which the boundary length between two different cell ‘types’ is at a minimum. Between this state, and a completely disordered, random state, there are also many configurations. As a consequence, we will computationally detect clustering of a population of cells by investigating the ratio of boundary types, quantifying this as a function of $\alpha_{stiff}/\alpha_{soft}$. In the literature, this is often referred to as, the ‘heterotypic boundary length’ (HBL): the total contact length between cells with different adhesion levels, measured in pixels. Other useful descriptors of these systems are the symmetry/asymmetry, elongation and cavitation of the clusters.

The HBL is defined as:

$$L_h = \sum_{\vec{i}, \vec{i}' \text{ neighbors}} \left(1 - \delta(N_{\sigma(\vec{i})}, N'_{\sigma'(\vec{i}')} \right).$$

In an ideal, fully sorted configuration, cells expressing the higher levels of cadherins will cluster together and round up into a solid sphere, surrounded by successive spherical shells of cells expressing successively lower levels of cadherins.

We plotted the heterotrophic boundary lengths for a range of α_{stiff} values from 1-4. As can be seen (figure 22), a change in cell-adhesiveness only marginally affected the resulting structure formed. It became clear that, with simple cell circuit, the dominant parameter that changed sorting was the $k_{\text{off}}/k_{\text{on}}$ parameters, which controlled the rate of morphological change of a B cell to a C cell. For instance, if we make a cell 10 times more likely to transition from B to C, the fully sorted structure is radically different.

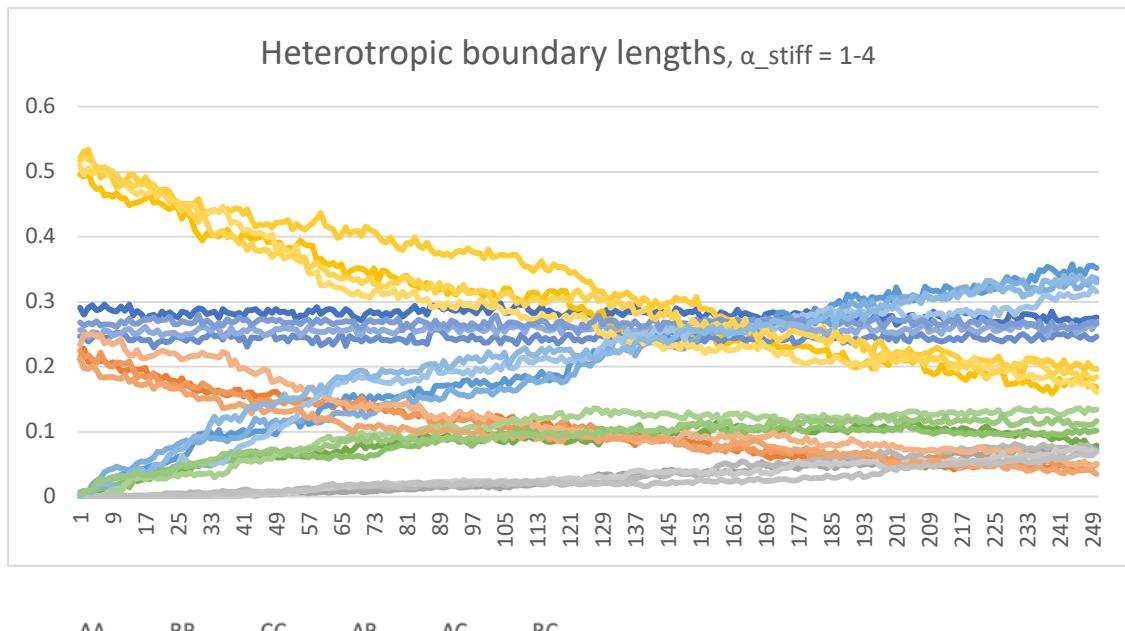


Figure 22

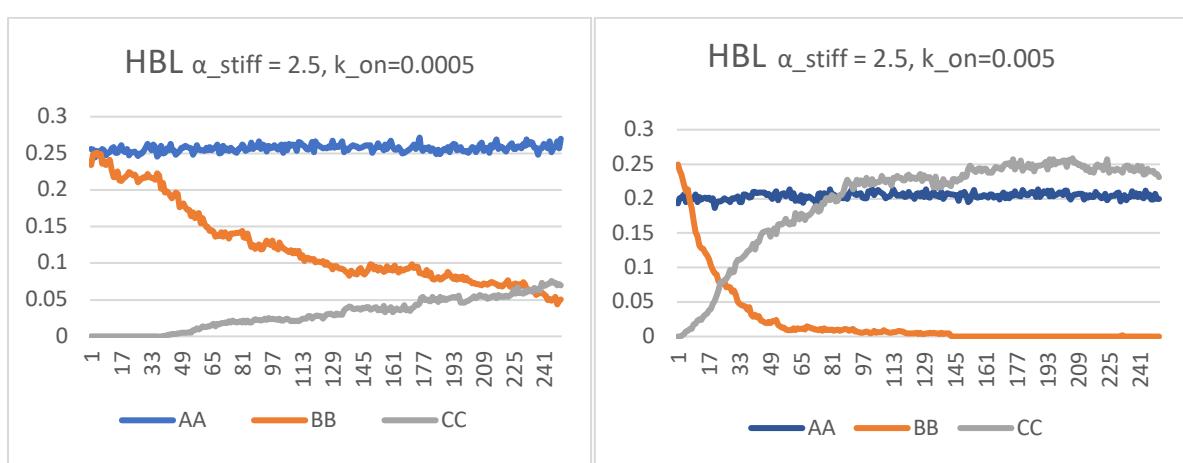


Figure 23 In this one level cell-circuit, the switch rate on/off dominates the way in which cells cluster. Changing α_{stiff} does not radically change the pattern into which the cell sorts into, just the speed of this process.

Chapter 5: Higher order cell circuits

5.1 Two Layer circuits

We can build complexity by layering additional cell circuits onto the previous one. Previously we had cell types A, B, and C, where a cell B would transition to a C cell, if it were in contact with an A cell. We added a fourth type, D, where cells change to D cells if they are in contact with a C cell.

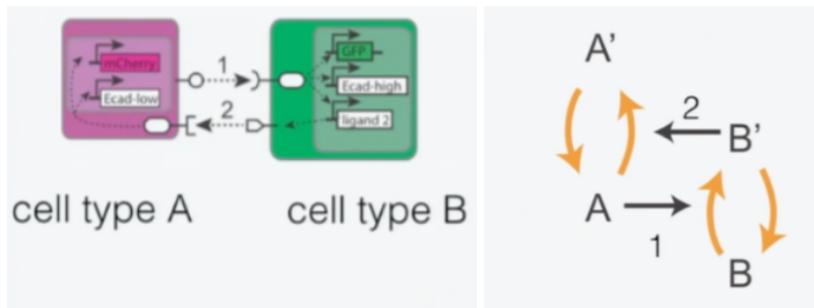


Figure 24
B cells transform to B' (C cells),
change cadherin expression,
and then present a ligand that
can interact with A cells

In addition, the simulation was modified such that each cell type has a different adhesiveness, which has an effect on the form of the adhesion matrix. In the accompanying code, these are referred to as α_{soft} (B cell), α_{stiff} (A cell), $\alpha_{stiffer}$ (D cell), $\alpha_{stiffest}$ (C cell), where the adhesiveness/stiffness of the cells is given by: $\alpha_{soft} < \alpha_{stiff} < \alpha_{stiffer} < \alpha_{stiffest}$. Two additional control parameters (k_{on2} and k_{off2}) are added to control for the switch rate between type A and type D cells.

	A	B	C	D
A	low	low	med	med
B		low+	med	med
C			med	med+
D				high

Figure 25
Changing the form of the
Adhesion Matrix has large
ramifications on the cell sorting.

We modelled this 3 -layer signalling network. At $t=500$ (when the system has reached steady state), the sorting looks as follow for one example:

Cell Types:

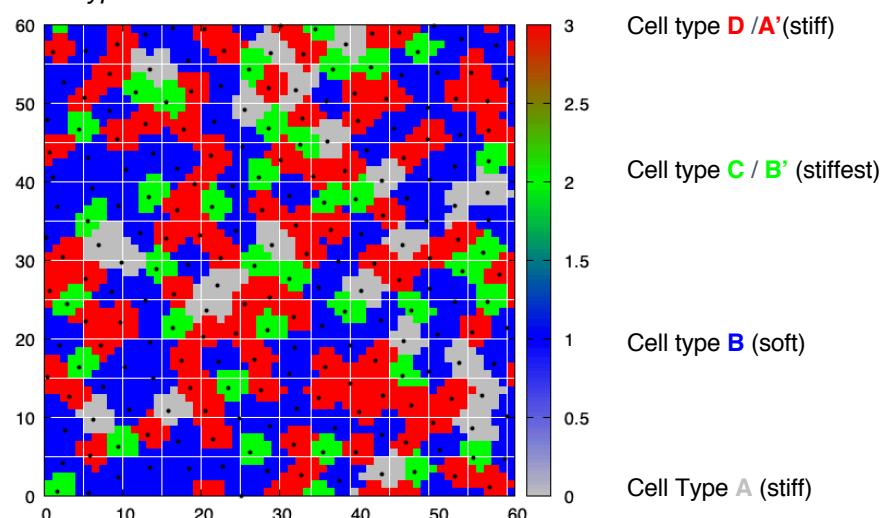


Figure 26

Cell adhesiveness:

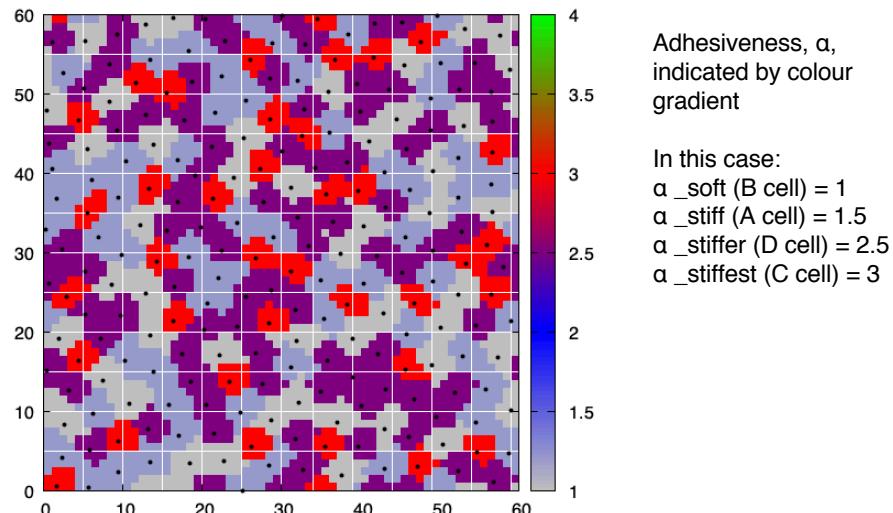
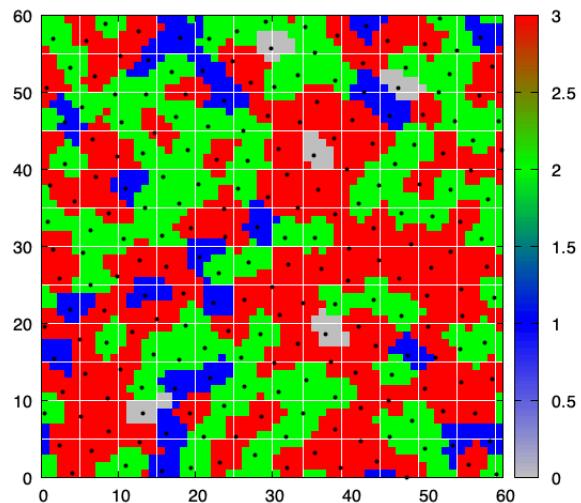
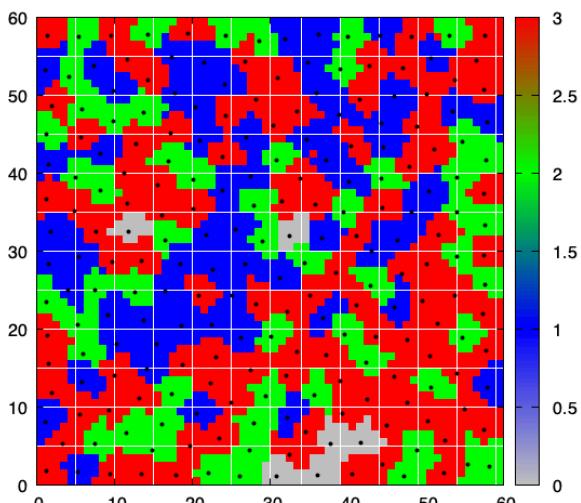


Figure 27

We can vary the parameters, and gain different sorting patterns. Large changes can be gained by just making minor tweaks to k_{on} , k_{off} , $k_{\text{on}2}$, $k_{\text{off}2}$:



'Green' cells are the most adhesive, and are often surrounded by the second most adhesive type: 'red' cells.



'Green' and 'red' cells, the most adhesive, clump together, and are surrounded by blue cells, which are the least adhesive.

Chapter 6: Conclusion and future direction of research

In effect, a much wider array of sorting structures can be gained from the combination of cell circuits and cell adhesion, as opposed to cell adhesion alone. The parameter space in which we do our modelling has been shown to replicate dense tissues in the biological realm, although more work has to be done to correlate tissue types to different values of the parameter α .

Going forward, there are a number of paths we can take:

1) Mapping the trajectory through morphospace

We can systematically make plots for different steady states in the long-time limit, and compare the heterotypic boundary lengths for various parameters, mapping these morphologies into a series of 2d plots. Dynamically, we can also track the trajectory through morphospace of, for instance, C cell-C cell and D cell-D cell clustering, observing the temporal evolution of system. A systematic study of this reveals that small changes in adhesion matrices make for large change in structures and trajectories. As an example, we have plotted the trajectory through morphospace of a system with parameters: $\alpha_{\text{soft}} = 1$, $\alpha_{\text{stiff}} = 1.3$, $\alpha_{\text{stiffer}} = 2$, $\alpha_{\text{stiffest}} = 2.5$, $f = 0.5$, $k_{\text{on}} = 0.005$, $k_{\text{off}} = 0.001$, $k_{\text{on2}} = 0.005$, $k_{\text{off2}} = 0.001$

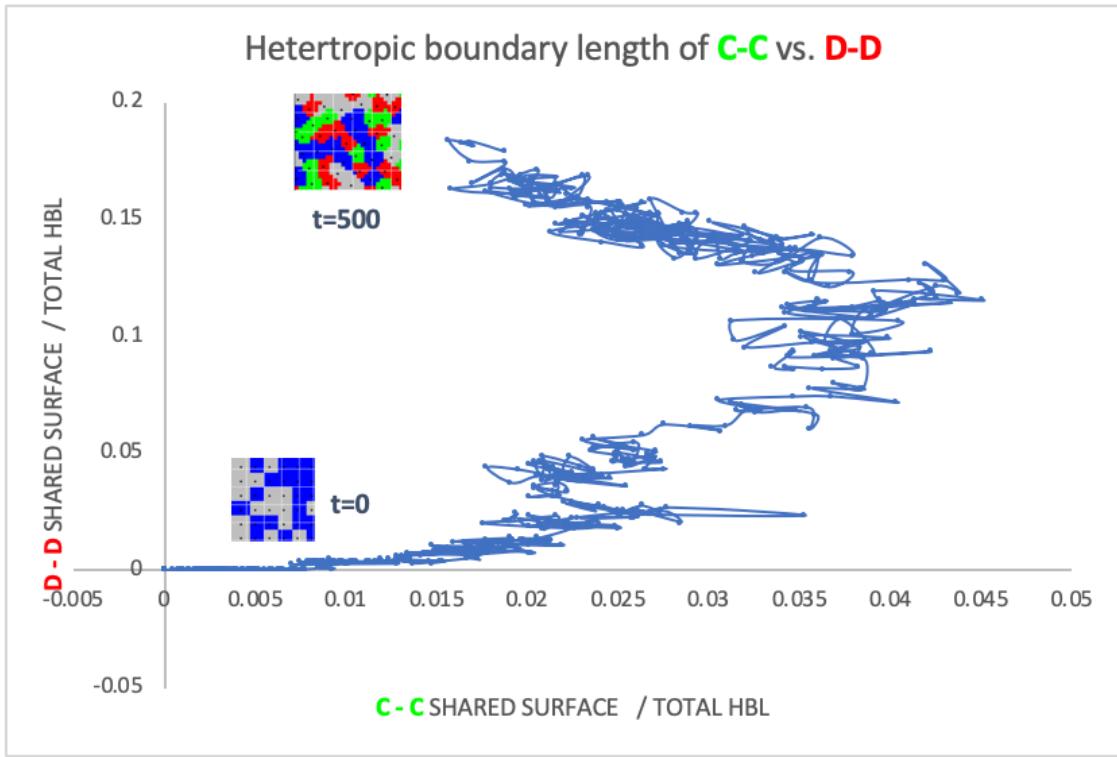


Figure 30

Repeating this process for different values of the parameters, it is possible to build a picture of the morphospace that is determined by 2 layer cell-circuits, with cells of different adhesiveness.

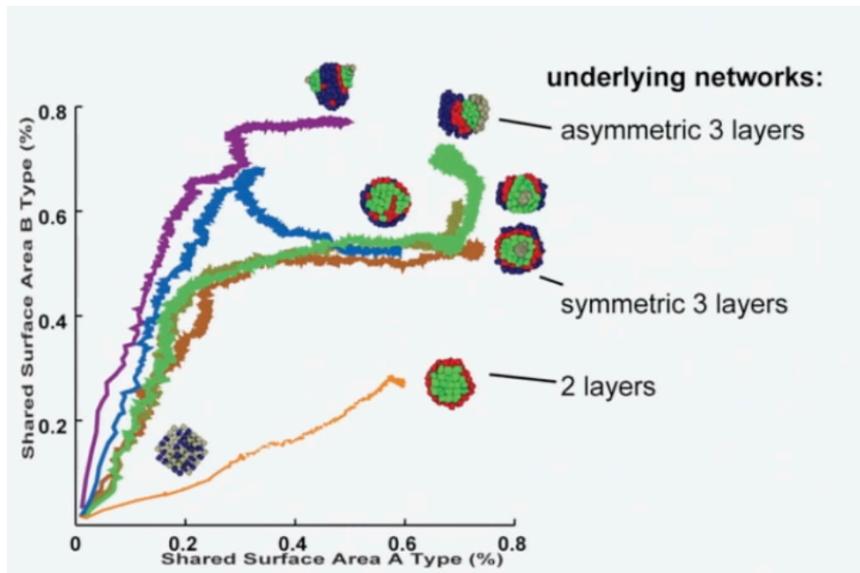


Figure 31

Mapping the trajectories of heterotypic boundary lengths, proposed by Leonardo Morsut

[Source: Programming self-organization of multicellular structures with synthetic cell-cell signalling coupled to morphogenetic effectors, Leonardo Morsut, Harvard CMSA, Published on Jan 16, 2019]

Unfortunately, our simulation does not clearly replicate cell clustering of the *in vitro* experiments by Leonard Morsut et al. We suspect that this is due to one (or a combination) of the following:

- Lack of sophistication in our adhesion matrix, stemming from a primitive understanding of the biology underpinning cadherin adhesion
- Lack of inclusion of a motility term in our modelling, P, we shall discuss in the appendix.
- Misuse of boundary conditions. We used periodic boundary conditions, but perhaps reflecting boundary conditions would have been a better choice.
- Perhaps the size of the lattice is too large. Cells cannot move quickly in the cellular potts model, so we potentially need to either increase the fluidity of the simulation (so cells can more easily move), increase the self-motility (i.e. diffusion) of the cells, or increase the timeframe over which the simulation is done, in order to see any true cell sorting.

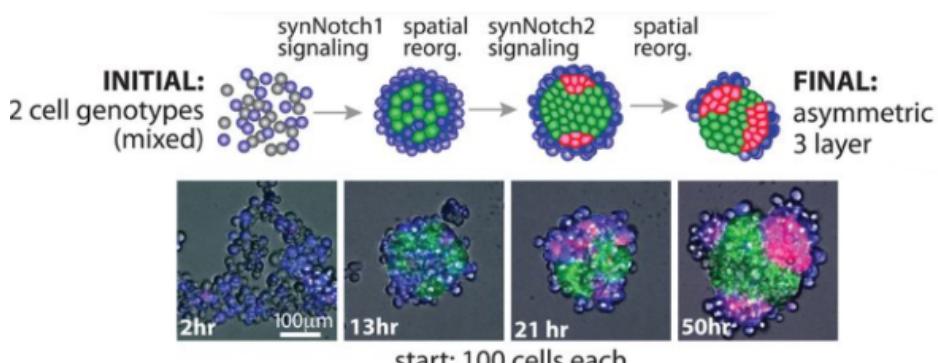


Figure 32

Images from experiment of cell sorting. Two major features that our *in silico* model fails to reproduce is cell layering and evidence of asymmetric or symmetric cell distributions.



As our modelling is still rough-hewn, making a systematic mapping of the trajectories of the Heterotypic boundary lengths in morphospace is not worthwhile at the current time.

2) Defining the energetics of the system

Different configurations of cells are correlated with different energy states of the system. Each boundary type between cells has a different energy. Signalling and adhesion both influence the shape of this potential landscape. For a future project, it would be useful to create a Pseudo-potential landscape describing the dynamics of system. Differing cell-cell signalling changes the shape of potential landscape. We can explore how the existence of multiple stable configurations is modulated via changes of signalling parameters. Through computational modelling, we can show how these trajectories allow the multicellular system a more exhaustive exploration of the space of forms (morphospace), when compared to systems with signalling only, or with sorting only.

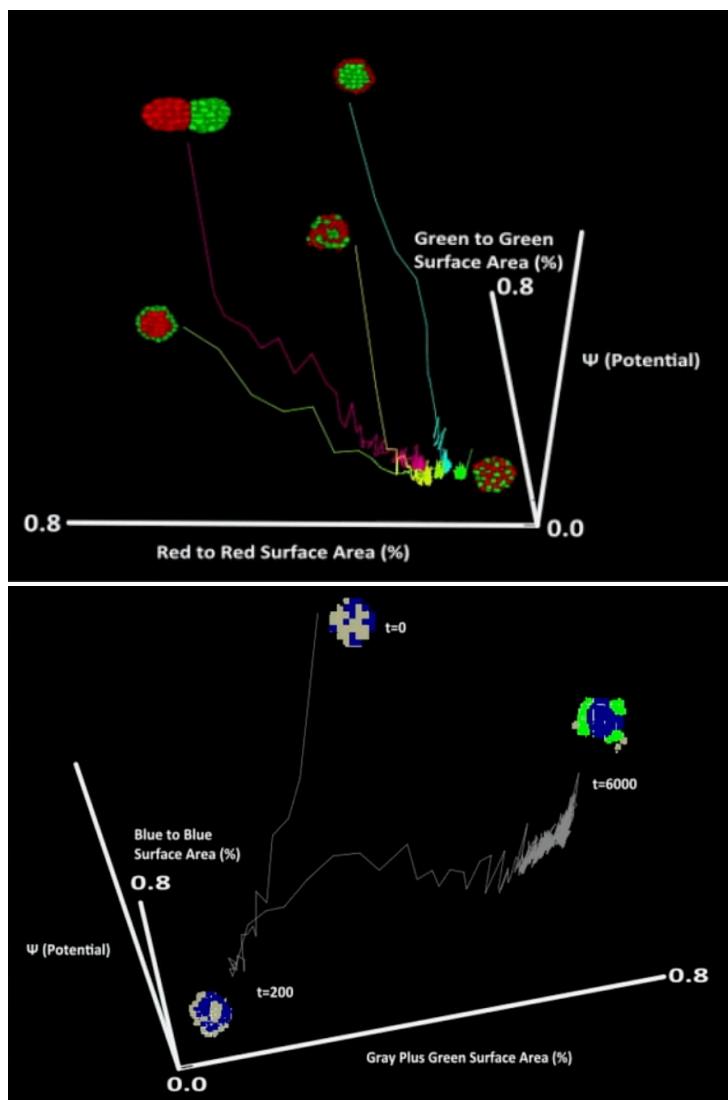


Figure 33 a, b
By constructing potential landscapes, we can explore how the existence of multiple stable configurations is modulated via changes of signalling parameters. For instance, if signalling were very fast, the system would not be able to explore intermediary states; if signalling is slow in comparable to cell movement, intermediate states would be explored by the stem. These intermediary states may be crucial to structure development, if these intermediaries nucleate growth. These potential landscapes are reminiscent of Waddington landscapes, and give an indication of how tissues can 'evolve.'

3) Cell fate and structure formation

As a next step, we could integrate cell fate into our simulation to nucleate the development of functional organoids. For instance, if cells could die, we could construct model for how lumens form. Indeed, recent studies have suggested that cell sorting using cadherin cell circuits was a key driver of evolutionary change in the earliest organisms. This research has primarily focused on Volvocine green algae, which represents the “evolutionary time machine” model for studying multicellularity, because they encompass the whole range of evolutionary transition of multicellularity from unicellular *Chlamydomonas* to >500-celled *Volvox* ([Kawai et al., 2013](#)).

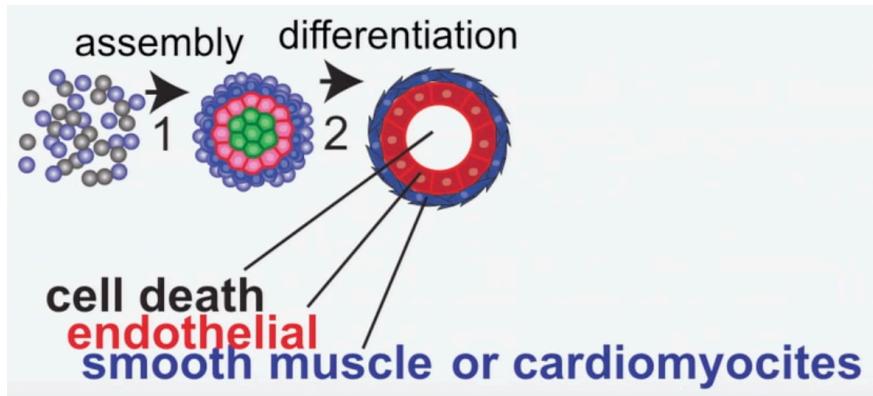


Figure 32 Differential adhesion leads to the most adhesive cells collecting in the centre, with successive layers of less adhesive cells. The centre cells then die, leading to a central cavity. This shows a mechanism by which a common structural motif (a lumen) can form through simple cell sorting.

Bibliography

- [Anderson et al., 2006] Alexander Anderson, Chaplain Mark A. J. and Rejniak Katarzyna A, Single-Cell Based Models in Biology and Medicine, Springer, (2007)
- [Chiang et al., 2016] M. Chiang and D. Marenduzzo, Glass transitions in the cellular Potts model, *EPL* 116 28009, (2016)
- [Foty et al., 2005] R.A Foty, Steinberg, The differential adhesion hypothesis: a direct evaluation., *MS Dev Bio* 278: 255–263, (2005)
- [Johnson et al., 2017] Marion B. Johnson, Alexander R.March, Leonardo Morsut, Engineering multicellular systems: Using synthetic biology to control tissue self-organization, *Curr Opin Biomed Eng.* 163–173., (2017)
- [Kabla et al., 2012] Kabla A. J., J., Collective cell migration: leadership, invasion and segregation, *R. Soc. Interface*, 9 , 3268, (2012)
- [Toyooka et al., 2013] Hiroko Kawai-Toyooka, Yuki Hamamura, Tetsuya Higashiyama, Akira Noga, Masafumi Hirano, Bradley J. S. C. Olson, Hisayoshi Nozaki , The Simplest Integrated Multicellular Organism Unveiled, *PLOS ONE* 8(12): e81641, (2013)
- [Merks, 2015] Merks, R, Cell-based modeling. In Encyclopedia of Applied and Computational Mathematics, pp. 195– 201, Springer, (2015)
- [McDole et al., 2018] K. McDole, L. Guignard, F. Amat, A. Berger, G. Malandain, L.A. Royer, S.C. Turaga, K.Branson, P.J. Keller, Toto imaging and reconstruction of post-implantation mouse development at the single-cell level, *Cell*, vol. 175, pp. 859-876, (2018)
- [Sasai, 2013] Y. Sasai, Cytosystems dynamics in self-organization of tissue architecture, *Nature*, 493 pp. 318-326 (2013)
- [Morsut et al., 2016] Leonardo Morsut, Kole T. Roybal, Xin Xiong, Russell M. Gordley, Scott M. Coyle, Matthew Thomson, Wendell A. Lim, Engineering Customized Cell Sensing and Response Behaviors Using Synthetic Notch Receptors, 2016, *Cell* 164, 780–791, (2016)
- [Qian et. Al, 2018] Yili Qian, Cameron McBride, and Domitilla Del Vecchio, Programming cells to work for us, <https://doi.org/10.1146>, (2018)
- [Steinberg et al., 1994] M.S Steinberg, Takeichi M, Experimental specification of cell sorting, tissue spreading, and specific spatial patterning by quantitative differences in cadherin expression., *Proc Nat Ac Sc USA* 91: 206–209 (1994)
- [Steward et al., 2015] Robert L. Steward Jr, Cheemeng Tan, Chao-Min Cheng Philip R.Le Duc, Cellular force signal integration through vector logic gates, *Journal of Biomechanics*, Volume 48, Issue 4, Pages 613-620, (2015)
- [Thurley et al., 2018] Kevin Thurley, Lani F. Wu, Steven J. Altschuler, Modelling Cell-to-Cell Communication Networks Using Response-Time Distributions, *Cell Systems* 6, 355–367, (2018);

[Szabo et al., 2010] Szabó A., Unnep R., M̄ehes E., Twal W. O., Argraves W. S., Cao Y. and Czirók A., Collective cell motion in endothelial monolayers, *Phys. Biol.*, 7 046007, (2010)

[Teague et al., 2016] B.P. Teague, P. Guye, R. Weiss, Synthetic morphogenesis, *Cold Spring Harb Perspect Biol*, 8, pp. 1-16, (2016)

[Toda et al., 2019] Satoshi Toda, Jonathan M.Brunger, Wendell A.Lim, Synthetic development: learning to program multicellular self-organization, *Current opinions in System biology*, (2019)

[Velazquez et al., 2018] Jeremy J. Velazquez, Emily Su, Patrick Cahan, Programming Morphogenesis through Systems and Synthetic Biology, *Cell Press: Tissue Engineering* volume 36, issue 4, (2018)

[Voss-Bohme et al., 2012] Anja Voss-Bohme, Multi-Scale Modeling in Morphogenesis: A Critical Analysis of the Cellular Potts Model, *PLOS ONE* 7(9): e42852. (2012)

[Zhang et al., 2011] Ying Zhang , Gilberto L. Thomas, Maciej Swat, Abbas Shirinifard, James A Glazier, Computer Simulations of Cell Sorting Due to Differential Adhesion, *PLOS ONE* 6(10): e24999, (2011)

Thanks

I would like to express my huge thanks to my supervisor, Davide Michieletto, for his enthusiasm and patience. He made the whole process incredibly enjoyable.

