Discrete explorations of multicellular growth and morphogenesis

Roeland M.H. Merks*†‡

January 10, 2012

1 Introduction

It was a new, and somewhat scary adventure for me to start working in a mathematical environment: after my undergraduate studies in biology, I had worked on problems in theoretical biology in computer science, biology, and physics institutes, but never had I been so close to "real" mathematicians. I joined CWI via the Netherlands Consortium for Systems Biology in September 2008, and became part of MAC-4, the Life Sciences group. Would the mathematicians not frown upon our simulation methods? Our methods help develop new hypotheses for biologists, show how complex biological phenomena can be driven by simple underlying rules, and also produce beautiful pictures, but they do not produce any real proofs or theorems.

My worries were quickly taken away after having received a warm welcome at CWI, and after the introduction to Jan Karel Lenstra. The original Escher print in Jan Karel's office made me feel more at ease among mathematicians. The mathematics of life and biological morphogenesis is prominent in the work of M.C. Escher: apart from the metamorphosis from fish to bird in Jan Karel's office, Escher made many lizard tessellations, and of course the *Pedalternorotandomovens centroculatus articulosus* - the cute rolling creature similar to the extant pill millipede (Figure 1). Later on I learned that Jan Karel had been one of the main proponents of building a Life Sciences group at CWI, and also helped to establish the Netherlands Institute for Systems Biology that ensures that the models, simulation, and methods are developed in collaboration with experimental biologists, so they are applied to "real life" applications.

1.1 Biological morphogenesis

What kind of "real life" applications does my group work on? Here I would like to focus on morphogenesis, the mechanisms driving biological growth and form. A classic

*Centrum Wiskunde & Informatica, MAC-4 (Life Sciences), Science Park 123, 1098 XG Amsterdam
†Netherlands Consortium for Systems Biology / Netherlands Institute for Systems Biology, Science Park 123, 1098 XG Amsterdam
‡merks@cwi.nl

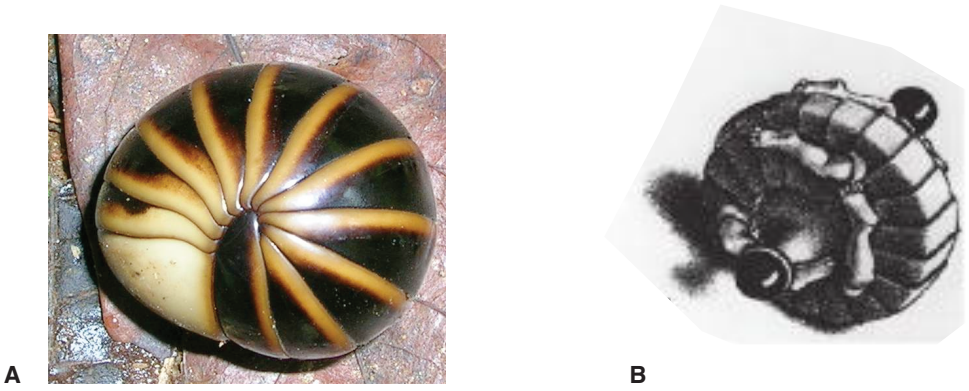

Figure 1: A "real life" and "mathematical" creature side by side. **A.** Pill millipede, possibly from order *Glomerida*. **B.** *Pedalternorotandomovens centroculatus articulosus*, or *wentelteefje* by M.C. Escher.

mathematical study of morphogenesis is Alan Turing's seminal paper "The chemical basis of morphogenesis"[16]. Turing showed how periodic chemical patterns, *e.g.*, the signals eliciting the growth of fingers or the stripes of a zebra, could arise from a homogeneous initial condition of two chemical substances, the *morphogens* A and B,

$$\begin{aligned} \frac{\partial A}{\partial t} &= a(A-h)+b(B-k)+D_A\frac{\partial^2 A}{\partial x^2}, \\ \frac{\partial B}{\partial t} &= c(A-h)+d(B-k)+D_B\frac{\partial^2 B}{\partial x^2}, \end{aligned} \tag{1}$$

with a, b, c, and d production or degradation rates for morphogen A and B. h and k are the homogeneous steady states for A and B, and D_A and D_B are the diffusion coefficients for morphogen A and B. Turing showed that uniform "stationary waves with finite wave-length", that is stripes or spots, are a steady state of this system if one of the morphogens diffuses much faster than the other.

Turing beautifully showed how patterns can arise out of "nothing" in a strip of tissue consisting of identical cells exchanging chemicals. My group builds upon Turing's pioneering work, and attempts to explain morphogenesis as a problem of self-organization of moving, growing and dividing cells. The cells in a multicellular organism follow simple rules, encoded by the information in their DNA. These rules guide cell behavior: how does a cell move, how strongly does it adhere to surrounding cells and to extracellular materials, what materials does it secrete itself, etc. Perhaps more importantly, the cellular rules determine how a cell will respond to signals from its neighbors, and to signals the cells have deposited in the extracellular space. In a typical response, they would "switch on" a different set of cellular rules: they could, for example, change their shape, start or stop moving, or tighten the adhesion to adjacent cells. Such shifts in cell behavior occur, *e.g.*, during "cell differentiation".

Using cell-based computer simulations, a form of agent-based simulations in which the agents are biophysically detailed models of cells [9, 11], we predict the emergent

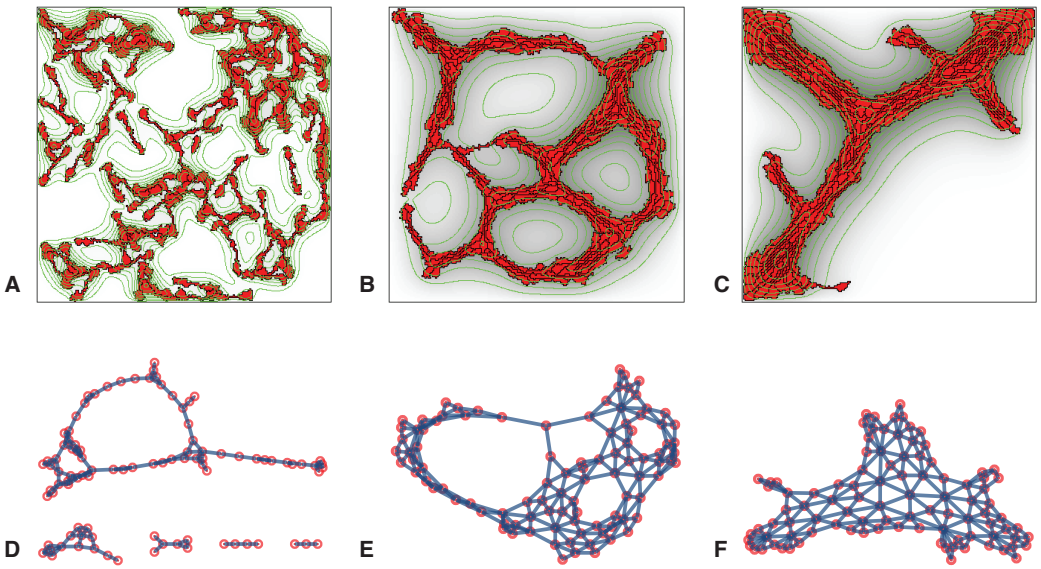

Figure 2: Snapshots from a Cellular Potts simulation of *de novo* formation of a vascular network from dispersed endothelial cells (panels **A-C**) and corresponding graph representations (panels **D-F**). Snapshots taken at 50 (**A**, **D**), 2000 (**B**, **E**), and 20000 (**C**, **F**) Monte Carlo steps.

collective behavior following from the stochastic behaviors of many, discrete cells, that follow a discrete set of cellular rules. In this way, we have studied how cells form blood vessels [7, 12], and how other cells with different rules form plants [8, 10].

1.2 Biological tissues are graphs

Did I just write "discrete"? Discrete cells, discrete rules? Yes, cells are discrete objects. Yet, many of the simulation methods in morphogenesis rely on continuum methods, including PDEs. This is a very powerful approach if we are interested in phenomena at large scales in tissues consisting of more-or-less identical cells, so we can safely "smoothen" away the individual cells. It allows us to make use of a wealth of tools in dynamic systems theory: find bifurcations, do linear stability analyses... But at smaller scales, and in more heterogeneous tissues, we cannot ignore biology's innate discreteness. For this reason, cell-based simulation methods model cells as individual entities. A challenge in cell-based models, however, is that the analytical tools currently lack the level of sophistication of tools for continuum models. To what extent could such methodology be developed using the tools available to discrete mathematicians? Cells are discrete objects, so could discrete mathematics methods be of use in analyzing multicellular morphogenesis?

In this contribution, I will explore the use of graph theory in analyzing the cellular

configurations coming from our simulations of blood vessel growth. As a first step, we transform the simulated biological tissue into a graph that represents the adjacency relations of cells. That is simple: cells become nodes. If two cells are adjacent to each other, an edge connects them. Here I only focus on topology; that is, I throw away the locations of cells, and only keep the adjacency relations. Figure 2 shows an example of such a transformation. In the remainder, I will first briefly outline the cell-based algorithms we use to simulate blood vessel growth, and then outline some preliminary explorations in the use of graph theoretic measures for characterizing blood vessel sprouting. I will end by discussing some future developments in the discrete mathematics of biological morphogenesis.

2 Methods

2.1 Hybrid cellular Potts model

We modeled cell motility at a mesoscopic level using the Cellular Potts model (CPM) [4, 5]. The CPM is a lattice-based Monte-Carlo approach that describes biological cells as spatially-extended patches of identical lattice indices $\sigma(\vec{x})$ on a square or triangular lattice, where each index uniquely identifies, or labels a single biological cell. Connections (links) between adjacent lattice sites $(\vec{x_1}, \vec{x_2})$ of unlike index, $\sigma(\vec{x_1}) \neq \sigma(\vec{x_2})$, represent bonds between apposing cell membranes, where the bond energy is $J(\vec{x_1}, \vec{x_2})$, assuming that the types and numbers of adhesive cell-surface proteins determine J. A penalty increasing with the cells deviation from a designated target volume $A_{\text{target}}(\sigma)$ imposes a volume constraint on the simulated cells. We define the pattern's effective energy:

$$H = \sum_{(\vec{x_1},\vec{x_2})} J(\sigma(\vec{x_1},\vec{x_2}))(1-\delta(\sigma(\vec{x_1}),\sigma(\vec{x_2})))+\lambda_A \sum_{\sigma}(a(\sigma)-A_{\mathrm{T}}(\sigma))^2+\lambda_L \sum_{\sigma}(l(\sigma)-L_{\mathrm{T}}(\sigma))^2, \tag{2}$$

where $(\vec{x_1}, \vec{x_2})$ is a pair of adjancent lattice sites, $a(\sigma)$ is the current area of cell σ, and $A_{\mathrm{T}}(\sigma)$ is its target area. The parameter λ_A represents the cell's resistance to compression, and δ is the Kronecker delta. The last term represents a cellular length constraint, with $l(\sigma)$ representing the current length of cell σ, L_{T} a target length, and λ_L a parameter setting the cell's rigidity along its long axis [7].

To mimic the cytoskeletally-driven cell extensions and retractions that drive cell motility, we randomly choose a source lattice site $\vec{x_1}$, and attempt to copy its index $\sigma(\vec{x_1})$ into a randomly-chosen adjacent lattice site $\vec{x_2}$. During a Monte Carlo Step (MCS) we carry out N copy attempts, with N the number of sites in the lattice. We calculate how much the effective energy change, ΔH, that would occur if we performed the copy, and accept the attempt with a Boltzmann-like probability,

$$P(\Delta H) = \{e^{\frac{-\Delta H}{M(\sigma)}}, \Delta H \geq 0; 1, \Delta H < 0\}, \tag{3}$$

where $M(\sigma)$ defines an intrinsic random cell motility M of cell σ, that drives energetically unfavorable movements.

We assume that endothelial cells secrete a morphogen $c(\vec{x},t)$ that diffuses and degrades in the extracellular matrix,

$$\frac{\partial c}{\partial t} = \alpha(1 - \delta(\sigma(\vec{x}),0) - \epsilon\delta(\sigma(\vec{x}),0)c + D\frac{\partial^2 c}{\partial x^2}, \tag{4}$$

with $\delta(\sigma(\vec{x}),0) = 0$ inside cells and $\delta(\sigma(\vec{x}),0) = 1$ otherwise. Here α is the rate at which cells secrete the morphogen c, ϵ denotes the morphogen's degradation rate, and D is the morphogen diffusion. We further assume that cells chemotact towards the morphogen: they preferentially move towards higher concentrations of c. To do so, we include an additional chemical effective-energy change, $\Delta H_{\text{chemotaxis}}$, at each copy attempt [15] from site $\vec{x_1}$ into adjacent site $\vec{x_2}$,

$$\Delta H_{\text{chemotaxis}} = -\mu(c(\vec{x_2} - \vec{x_1})). \tag{5}$$

In this way the algorithm favors cellular extensions into the direction of higher concentrations of the morphogen. We solve the partial differential equations for chemoattractant diffusion and degradation (Eq. 4) numerically using a finite-difference scheme on a lattice matching the CPM lattice, with $\Delta x = 2\ \mu m$. We use 15 diffusion steps per Monte Carlo step, with $\Delta t = 2\ s$, thus equating a Monte Carlo step to 30 s real time.

2.2 Cellular networks as graphs

To study a cellular pattern as a graph, we first construct a graph. We represent patches of identical indices $\sigma = i$ (*i.e.*, a biological "cell") by node i. Nodes i and j share an edge $e(i,j)$ if,

$$\exists_{(\vec{x_1},\vec{x_2})} : \sigma(\vec{x_1}) = i \wedge \sigma(\vec{x_2}) = j \wedge \sigma(\vec{x_1}) \neq \sigma(\vec{x_2}). \tag{6}$$

So cell i and cell j share an edge $e(i,j)$ if there exists an adjacent pair of lattice sites $(\vec{x_1},\vec{x_2})$ such that $\sigma(\vec{x_1}) = i$ and $\sigma(\vec{x_2}) = j$. We exclude self-connected nodes.

The resulting graphs can be studied using standard tools from graph theory. To do so, we made use of the graph tools implemented in *Mathematica 8* and the algorithms in the *Mathematica Graph Utilities Package*. In this pilot study, we looked only at the above defined adjacency relations of cells. In more extensive studies the nodes can be given attributes, including cell type, surface area, or cellular locations. No additional notation is required to represent these node attributes, because the node indices uniquely identify the cells.

3 Results

3.1 Simulation of *de novo* formation of a blood vessel network

Figures 2A-C show an example of blood vessel network formation in our model, with an appropriate set of parameters (Table 1 lists the parameters used). After the initial formation of the chemical gradients, the cells move towards one another and form a

$\lambda_A = 50$	$A_{\mathrm{T}} = 400\ \mu m^2$
$\lambda_L = 5$	$L_{\mathrm{T}} = 120\ \mu m$
$J_{\mathrm{cell,medium}} = 20$	$J_{\mathrm{cell,cell}} = 40$
$M(\sigma) = 50$	$D = 1.0 \times 10^{-13}\ m^2 s^{-1}$
$\alpha = 1.8 \times 10^{-4}\ s^{-1}$	$\epsilon = 1.8 \times 10^{-4}\ s^{-1}$
$\mu = 1000$	
Grid size: 200×200 lattice sites	Number of cells: 100

Table 1: Parameters used for the *de novo* vasculogenesis simulation studied in this paper

network structure closely resembling real-life blood vessel networks in animals and in cell cultures [6]. Also note that the initial fine-grained network slowly coarsens over time, a phenomenon also seen in cell cultures [6]. A movie of this simulation is available at `http://www.youtube.com/watch?v=VbeRFcl6RD4`. For more details on the mechanisms driving such blood vessel sprouting, the relation to actual blood vessel growth, and extensive parameter studies of this model, see our previous and ongoing work [12, 6, 7]. Figures 2D-F show the graph representations corresponding with the blood vessel networks.

3.2 Variability of graph measures marks morphological events

What graph measures could we use to analyze the morphological transformation from a set of disconnected cells, to a fine-grained network, and into a coarse-grained network? To detect the first occurence of a vascular network, the first obvious measure to look at is the number of connected components in the graph: the number of connected components drops rapidly during the first steps of the simulation, as the cells move towards one another (Figure 3A). After around 300 Monte Carlo steps all cells have connect to one another.

After the first connections between cells have been made, the blood vessel network slowly coarsens over time; a good measure of graph complexity should capture this coarsening process. Also, we would like to have a good definition of the number of branches in the simulated networks. We first identified the *eccentricity* of the vertices in the graph, where the eccentricity of a node is the longest of all shortest paths to all other nodes in the graph. We expect nodes in the branches to have high eccentricity. Another useful measure is the *graph periphery*, defined as the set of vertices with maximum eccentricity. We noticed that many branches typically contained a set of connected peripheral nodes; so simply counting the peripheral nodes overestimates the number of branches. We therefore defined the number of *connected peripheral subgraphs* as the number of connected components in the graph periphery. We plot this measure in Figure 3B as a function of time. The number of connected peripheral subgraphs jitters over time, suggesting that it is very sensitive to the small variations in morphology due to our Monte-Carlo simulation algorithm.

Interestingly, between time steps 20000 and 40000 and after time step 80000 we observed a temporary drop of the values and variability of the number of connected pe-

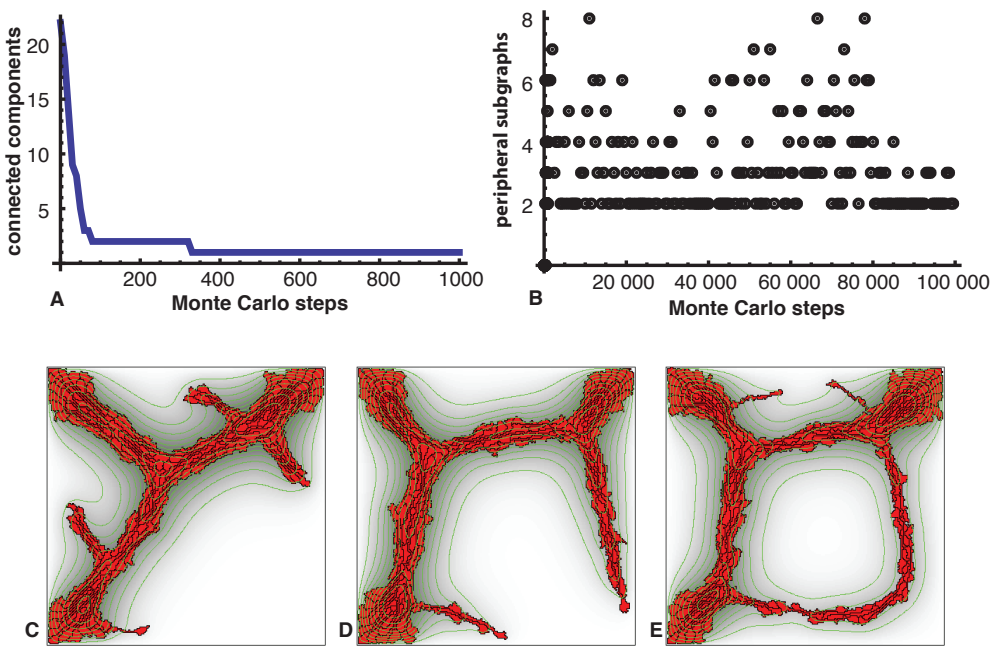

Figure 3: **A** Number of connected components of the vascular network over time (in Monte Carlo steps). **B** Number of connected peripheral subgraphs in the vascular network over time. **C** Snapshot of simulation at 20000 Monte Carlo steps, **D** at 30000 Monte Carlo steps, and **E** at 40000 Monte Carlo steps.

ripheral subgraphs. What happened in our simulation during this period? Figures 3C-E show three snapshots of our simulation at 20000, 30000, 40000 timesteps. Two large sprouts form; at around step 40000 they touch and form a large "hole" in the morphology. It steadily shrinks and disappears around step 60000. Once the two sprouts have reconnected the number of connected peripheral components increases again. A similar event occurs after timestep 80000, whereas we have not observed other, similar events in the simulation. Thus these observations suggest that the variation of the number of connected peripheral subgraphs may help to identify morphological transformations in our simulations. We are now exploring a) if other measures more reliably identify the number of branches in these networks, and b) if indeed the *time-variation* or *structural stability* of graph measures may be a useful measure for other dynamical morphological features as well.

3.3 Network complexity: community partition structure

Changes in the variability of the eccentricity-based graph measures may help to identify morphological events, as we have seen in the previous section. However, exactly because of their high variability we cannot use them for the morphological analyses of static vascular networks. We next explored the use of community structure measures [14]. A

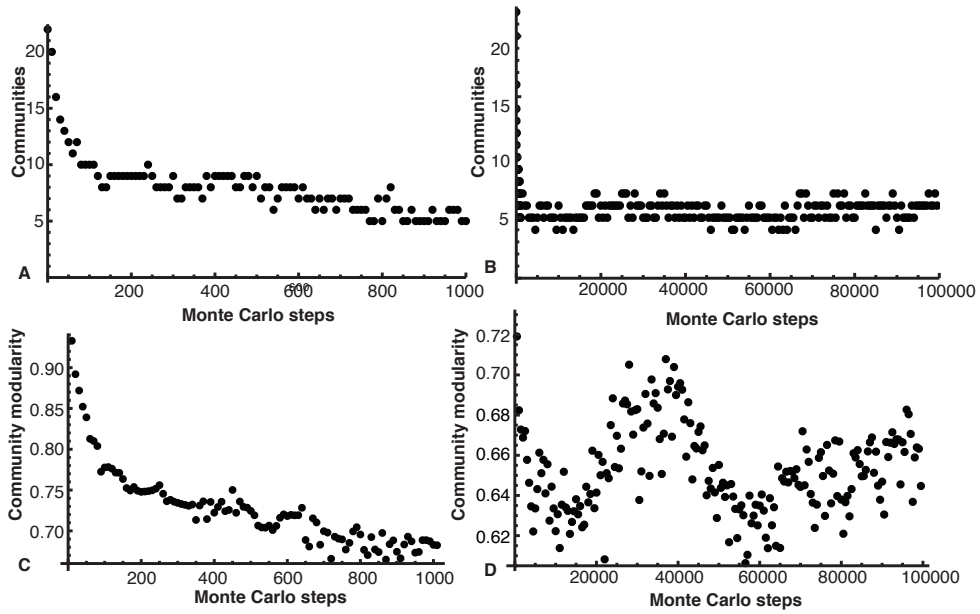

Figure 4: Community structure of vascular networks **A-B** Optimal number of community partitions detected in the vascular network over time. **A** Monte Carlo steps 0 to 1000; **B** Monte Carlo steps 0 to 100000. **C-D** Community modularity (Q) of optimal community partition of vascular networks. **C** Monte Carlo steps 0 to 1000; **D** Monte Carlo steps 0 to 100000.

community in a graph is a subset of the graph that is more strongly connected internally than with the other nodes in the graph. The *community modularity* Q of a partition of the nodes V in communities V_i, is defined as

$$Q = \sum_{i=1}^{k} (e_{ii} - a_i^2), \tag{7}$$

with e_{ii} denoting the fraction of all edges in the graph internal to community i, and a_i denoting the fraction of edges running between community i and the other nodes. The community modularity is a number smaller than 1, with larger values of Q corresponding to stronger community structure. To calculate community partitions and community modularities we use the algorithm by Clauset [2] implemented in Mathematica 8.

Figure 4A shows the number of detected communities corresponding with the highest value of Q for the first 1000 Monte Carlo steps of the vasculogenesis simulation. Panel B shows the development of the number of communities for MCS 0 to 100000. During the first 1000 Monte Carlo steps the number of communities decreases from about twenty to about five communities, corresponding with the coarsening of the network pattern. Over a longer timescale, the number of communities in the simulated vascular network fluctuates around six. Thus, the number of communities nicely characterizes the slow

coarsening of the vascular network, but once the network has settled on a characteristic size, the community structure does not capture the dynamic, topological transformations of the network.

Figures 4C-D show the community modularity Q corresponding with the optimal community partition as a function of time. During the first 1000 Monte Carlo steps of our simulation, Q drops gradually from a value of above 0.9 to a value of about 0.7, corresponding with the expected reduction in community structure during network formation and coarsening. Interestingly, over a longer time scale the community modularity continues to fluctuate. The modularity drops to a value of around 0.6 around Monte Carlo step 20000 and peaks at 0.7 near Monte Carlo step 40000, after which it drops steadily until time step 60000. Possibly this variation of the community modularity corresponds with the formation of two large sprouts between Monte Carlo steps 20000 and 40000, the reconnection of the sprouts and the subsequent shrinking and closure of the resulting hole at Monte Carlo step 60000. Thus formation and closure of holes are clearly indentified. Our initial explorations suggest that changes in optimal values of the community modularity Q identify both the initial network coarsening, and more long-term, topological transformations of the vascular network.

4 Discussion

In this contribution I presented some initial explorations on the use of discrete mathematical methodology for the analysis of biological morphogenesis. I first identified the number of connected peripheral nodes for each frame of a vascular network simulation. Although this number itself varied too much to provide useful information about the structure of the vascular networks, I found that close to major morphological events its temporal variation dropped, for example during the formation of large sprouts and the fusion of these sprouts. This suggests that the robustness of a graph measure against small modifications of the graph structure can provide useful information [1, 3], even if the graph measure itself does not. Thus, a differential graph theory might be useful for the study of multicellular morphogenesis and other dynamic networks [3].

I next looked at the evolution of community modularity in a simulated vascular network. I found that the number of communities in an optimal community partition characterizes the initial coarsening of the simulated vascular network well, whereas the subsequent formation and fusion of sprouts are not well captured. However, the community modularity associated with the optimal community partition correlates well with both the initial network coarsening and the subsequent formation and fusion of sprouts.

The presented methodology indeed is a useful extension of our existing toolbox for analyzing angiogenic sprouting, based on computational geometry and mathematical morphology [13, 7, 12]. However, the real power of using graph representations of morphogenetic topological transformations may only be unleashed once we will use it to state and solve generalized problems of biological morphogenesis in formal mathematical terms. Assuming that evolution of multicellular organisms strives for optimality, we can pose various interesting questions about optimal mechanisms of biological morphogenesis. An example: assume a two-dimensional (or n-dimensional) piece of tissue, in

which cells with diameter 1 communicate via contact-dependent ("juxtacrine") signals. That is, cells must be in contact in order to pass a signal. What is the fastest way that a signal can spread through the tissue?

If the cells do not move relative to one another (epithelial tissues, plants) the answer is trivial: if a signal spread from a cell to its neighbors in time Δt, the signal spreads over a volume of radius r in time $r\Delta t$. But what if the cells shift positions and mix? How much faster or slower will the signal spread? What is the optimal way and speed for the cells to mix in order to attain optimal spread of the signal? Suppose all cells must have been in contact with one another in order for a biological mechanism to proceed. For a finite piece of biological tissue, what is the optimal mixing strategy, under the condition that cells can only exchange positions with their neighbors? How does this depend on the topology of the tissue; *e.g.*, does it differ between branched tissues (*e.g.*, lungs, kidney, gland tissue) and flat layers of tissue? If one cell type can invade a tissue of another type, how will the mixing strategy determine the spreading speed? What is the optimal way for a flat tissue to change into a branched topology? How should they locally change their adjacency relations? Do we expect a different strategy if cells can mix, as in most animal tissues, and if they cannot, as in plants and some animal tissues? And how do these insights compare with biological observation: is biology optimal?

Using cell-based computer simulations and biological experiments it is possible to get useful insights into these questions. But what is still missing from cell-based modeling approaches in biology, is a formal differential graph theory of dynamic, discrete representations of growing and developing tissues. So, Jan Karel, if you are still wondering what to work on now you can spend more time on discrete mathematics, I hope to have convinced you that the theoretical biology of multicellular morphogenesis is full of fun and challenging problems. I look forward to discuss these with you!

5 Acknowledgments

Erik van Dijk and Margriet Palm are thanked for their suggestion to use graph representations for analyzing tissue patterns; Stefan Canzar is thanked for discussion, and András Szabó is thanked for discussion and critical reading of the manuscript. This work was cofinanced by the Netherlands Consortium for Systems Biology (NCSB), which is part of the Netherlands Genomics Initiative / Netherlands Organisation for Scientific Research, and by the Division for Earth and Life Sciences (ALW) with financial aid from the Netherlands Organization for Scientific Research.

References

[1] J. Bryden, S. Funk, N. Geard, S. Bullock, and V. A. A. Jansen. Stability in flux: community structure in dynamic networks. *J. Roy. Soc. Interface*, 8(60):1031–1040, 2011.

[2] A. Clauset. Finding local community structure in networks. *Phys. Rev. E*, 72:026132, 2005.

[3] G. Ghoshal and A.-L. Barabasi. Ranking stability and super-stable nodes in complex networks. *Nat. Commun.*, 2:394, 2011.

[4] J. Glazier and F. Graner. Simulation of the differential adhesion driven rearrangement of biological cells. *Phys. Rev. E*, 47(3):2128–2154, 1993.

[5] F. Graner and J. A. Glazier. Simulation of biological cell sorting using a two-dimensional extended Potts model. *Phys. Rev. Lett.*, 69(13):2013–2016, 1992.

[6] R. M. H. Merks and J. A. Glazier. Dynamic mechanisms of blood vessel growth. *Nonlinearity*, 19(1):C1–C10, 2006.

[7] R. M. H. Merks, S. V. Brodsky, M. S. Goligorksy, S. A. Newman, and J. A. Glazier. Cell elongation is key to in silico replication of in vitro vasculogenesis and subsequent remodeling. *Dev. Biol.*, 289(1):44–54, 2006.

[8] R. M. H. Merks, Y. Van de Peer, D. Inzé, and G. T. S. Beemster. Canalization without flux sensors: a traveling-wave hypothesis. *Trends Plant Sci.*, 12(9):384–90, 2007.

[9] R. M. H. Merks and J. A. Glazier. A cell-centered approach to developmental biology. *Physica A*, 352(1):113–130, 2005.

[10] R. M. H. Merks, M. Guravage, D. Inzé, and G. T. S. Beemster. Virtualleaf: An open-source framework for cell-based modeling of plant tissue growth and development. *Plant Phys.*, 155(2):656–66, 2011.

[11] R. M. H. Merks and P. Koolwijk. Modeling morphogenesis in silico and in vitro: Towards quantitative, predictive, cell-based modeling. *Math. Mod. Nat. Phenom.*, 4(5):149–171, 2009.

[12] R. M. H. Merks, E. D. Perryn, A. Shirinifard, and J. A. Glazier. Contact-inhibited chemotaxis in de novo and sprouting blood-vessel growth. *PLoS Comp. Biol.*, 4(9):e1000163, 2008.

[13] A. Mezentsev, R. M. H. Merks, E. O'Riordan, J. Chen, N. Mendelev, M. S. Goligorsky, and S. V. Brodsky. Endothelial microparticles affect angiogenesis in vitro: role of oxidative stress. *Am. J. Physiol. Heart Circ. Physiol.*, 289(3):H1106–14, 2005.

[14] M. Newman and M. Girvan. Finding and evaluating community structure in networks. *Phys. Rev. E*, 69(2):026113, 2004.

[15] N. Savill and P. Hogeweg. Modelling morphogenesis: From single cells to crawling slugs. *J. Theor. Biol.*, 184(3):229–235, 1997.

[16] A. Turing. The chemical basis of morphogenesis. *Phil. Trans. Roy. Soc. B*, 237:37–72, 1952.