Evolution and Morphogenesis of Differentiated Multicellular Organisms: Autonomously Generated Diffusion Gradients for Positional Information

Johannes F. Knabe[1], Maria J. Schilstra[2] and Chrystopher Nehaniv[1,2]

[1]Adaptive Systems and [2]BioComputation Research Groups
Centre for Computer Science & Informatics Research
University of Hertfordshire, AL10 9AB Hatfield, United Kingdom
j.f.knabe@herts.ac.uk

Abstract

Development is the powerful process involving a genome in the transformation from one egg cell to a multicellular organism with many cell types. The dividing cells manage to organize and assign themselves special, differentiated roles in a reliable manner, creating a spatio-temporal pattern and division of labor. This despite the fact that little positional information may be available to them initially to guide this patterning. Inspired by a model of developmental biologist L. Wolpert, we simulate this situation in an evolutionary setting where individuals have to grow into "French flag" patterns. The cells in our model exist in a 2-layer Potts model physical environment. Controlled by continuous genetic regulatory networks, identical for all cells of one individual, the cells can individually differ in parameters including target volume, shape, orientation, and diffusion. Intercellular communication is possible via secretion and sensing of diffusing morphogens. Evolved individuals growing from a single cell can develop the French flag pattern by setting up and maintaining asymmetric morphogen gradients – a behavior predicted by several theoretical models.

Introduction

The development of multicellular organisms from a single fertilized egg cell has fascinated humans at least since Aristotle's speculations more than 2000 years ago (34). In the more recent past our understanding of how interacting genes direct developmental processes has greatly increased (31; 12; 34). Cell differentiation, the inducing effects of intercellular signaling, changes in cell form like contraction, the self-organizing properties of adhesion and cell sorting in animal morphogenesis (13) are among the important principles better understood now. And although every cell is controlled by a Genetic Regulatory Network (GRN), the resulting multicellular dynamics are also strongly influenced by physical constraints.

Development has also caught the attention of computer scientists. Traditionally their evolutionary algorithms (EAs) would neglect development for a relatively direct mapping from genotype to phenotype. To overcome problems with these EAs, many models have been proposed that incorporate development in some way – reviews of these are given in Stanley and Miikkulainen (28), Kumar and Bentley (21).

One of the earliest researchers looking for a theoretical explanation of how cells in a developing embryo could establish their different roles was Turing (30). He proposed a general symmetry breaking mechanism via the setting up of chemical gradients with reaction diffusion systems. Somewhat later, Wolpert (32, 33) came up with the very illustrative French flag model as an attempt to explain how morphogen gradients could give cells positional information as a general biological process. "Stem cells" placed along a given morphogen gradient would only have to read the morphogen contraction at their position and react to threshold values to decide whether they are in the blue, white or red part of the flag. Note that this assumes the existence of a gradient but does not explain how such a gradient could be set up by the cells. Jaeger and Reinitz (16) proposed a revised French flag model which to some degree takes the dynamic, feedback-driven nature of pattern formation into account.

These theoretical models have inspired the present work. From a single cell, placed in the middle of a 60×40 pixel grid, a multicellular organism has to grow. The organism not only has to span the grid but its cells also have to express different colors in different areas (blue, white, red – from left to right).

Several other "French flag" inspired implementations exist which are related to our work to varying degrees. Miller (25) evolved cells to match a 12×9 French flag pattern, with a particular focus on self repair in the evolved individuals. There was a concept of cell division, however the possibility of overwriting of neighboring cells during division might have been crucial for the outcomes. This work was extended by Federici (9), who evolved individuals to match less regular 9×6 patterns. An interesting analysis of why evolved developmental individuals are often also fault tolerant is given in follow-up work (10).

More recently, Devert et al. (8) evolved neural network controlled cells to match 32×32 pixel patterns, without division (cells would be everywhere from the beginning).They focused on robustness and an adaptable stop criterion – an individual was considered final when a stable state was

reached.
Crucially, in all these models cells were homogenous in size, namely 1 pixel on a grid each. Also, cell-cell communication was predefined as directed between nearest neighbors (with fixed neighbors).
Bongard and Pfeifer (3) and Bongard (2) use a spatial, developmental system ("Artificial Ontogeny") to create impressive critters, however their model is not very biologically faithful as the units their creatures are composed of have a pre-specfied cylindrical shape (different length possible) and complex internal elements like motoric joints. Symmetry breaking activation is induced by application of morphogens to the opposing ends of all new units.
Hogeweg (14) has used a model were cells also have spatial extent and can move relative to each other. Cell types were under control of a boolean GRN, where the GRN's expression pattern was interpreted as one of the predefined cell types. Differential gene expression was initiated by two prescheduled asymmetric cell divisions (i.e. one gene would be on in one daughter cell and off in the other). Communication was direct between nearest neighbor cells; a cell had the states of two nodes from neighboring cells as input to two of its own nodes. Unlike in the other mentioned works, selection was not for particular cell arrangements, but the number of exhibited cell types was used as fitness criterion – an evolutionary algorithm would select for individuals that expressed many different types. The cell type would determine some cell properties like adhesion (the possible adhesion values were pre-specified), so that certain multicellular forms could be observed as an evolutionary byproduct.
On the contrary, in our model there are no predefined communication channels; cells can only sense morphogen concentrations. Morphogens are excreted by cells and diffuse on the grid without preference for direction. Cells can actively aim to adopt heterogenous sizes and shapes – the importance of which for morphogenesis has been shown by other researchers, e.g. Merks et al. (23), Zajac et al. (35). A cell's GRN also individually and independently controls other cell properties like morphogen secretion and cadherin expression. So there is no *a priori* notion of cell types.

Genetic Regulatory Network (GRN) model

In (20), where our proposed GRN model was first described in its basic form, we used it to evolve single-celled biological clocks with the circadian rhythm abstracted to a sinusoidal wave or other periodic function. GRNs producing such cyclic behavior in response to various periodic environmental stimuli could easily be evolved. Reproducing the phase of their input as well as the production of the inverse or shifted phase was demonstrated[1], however in that investigation every evolutionary run had only one of these objectives. In later experiments we showed that it is possible to integrate two functionalities in one GRN instantiated in different contexts within a multicellular entity (18). However, *differentiation* was still induced by a given signal instead of evolving through the interaction of GRNs. Here GRNs of this type are used for the first time as "control units" of *spatially extended cells.* [2]
Every GRN consists of proteins and a genome made up of genes. Gene activation is controlled by regulatory sites (cis-sites or cis-modules), each composed of – possibly – several protein binding sites. Depending on the attachment of proteins to the binding sites the corresponding cis-modules positively or negatively influence the production of (not necessarily different) proteins. In molecular biology, proteins acting in such a way are called Transcription Factors (TFs). In our model all proteins are potentially regulatory and there are no restrictions on recurrence. A main difference from the `Biosys` GRN model by Quick et al. (27) is that there can be any number of cis-modules per gene and every cis-module can have any number of protein binding sites. This is to model a second, non-linear, level of regulation: Molecular biologists have found that TFs not only show additive behavior in influencing a gene's transcription. Some TFs interact with each other or even form protein-protein compounds, resulting in synergistic changes to their influence, see e.g. (Schilstra and Nehaniv; 7).In logical terms (but note that values are actually continuous) one can think of this grouping of inputs as an OR of ANDs. The AND level certainly constitutes a canalizing function in the sense of Kauffman (17) as a single zero value there causes the whole term to be zero no matter what the other stimuli are. In summary the model, as compared to previous models, is designed to facilitate the evolution of complex dynamics, coming a little closer to nature than previous models in terms of regulatory logic, where "5-10 regulatory sites are the rule that might even be occupied by complexes of proteins" (1) and non-linear synergetic effects are possible (7).
The following subsections describe our GRN model. For a more formal description and analysis of the model and its representation please see (20; 19).

Genetic Representation

The genome is represented as a string of base four digits, encoding several genes and some global parameters of the network. Digits 0 and 1 are *coding* digits that may be involved in regulation or protein coding. To differentiate between a sequence of coding bits, a cis-module boundary and a gene boundary the genetic alphabet was increased to four values, with digit 2 delimiting the end of a cis-module and digit 3 delimiting the end of a gene. In the version of the model used here there is a predefined number $2^4 = 16$ of different protein types, so that always four bits encode a protein type.

[1] For results from those experiments see also http://panmental.de/GRNclocks/.

[2] Please see the associated web page at http://panmental.de/ALifeXIflag for more results, videos, and the full source code.

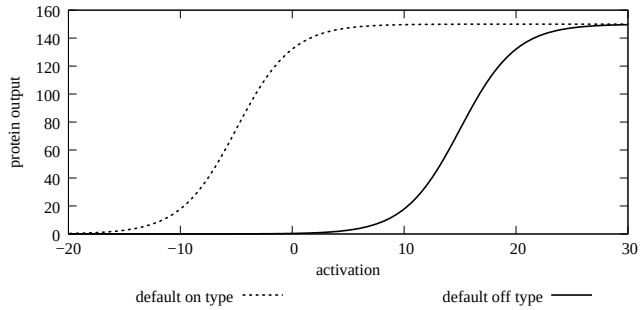

Figure 1: **Activation Types.** Every gene produces proteins according to the cumulative activation level of its cis-modules and its activation type: either even when no activation is present ("default on" - left) or only with positive activation ("default off" - right).

In the experiments described here we used a fixed number of genes, namely twenty, to facilitate analysis. After parsing the genome into genes, the last four coding digits of every gene determine its output behavior, a number of bits for the protein type produced and the last bit for the gene's activation type, which can be "default on" – active unless repressed or "default off" – silent until activated by regulatory sites,see fig. 1. The genome also encodes several evolvable variables global to the GRN. These are the *protein-specific decay rates* (four bits for every protein, indexing into a fixed look-up table of values), the global *binding proportion* (also four bits indexing into a look-up table, but identical for all proteins), and finally the global *saturation value* (three bits indexing to a look-up table, again identical for all proteins). These latter variables especially facilitate changes in the strength and timing of gene expression important in the evolution of multicellular individuals, cf. Buss (4).

Regulatory Logic

The model is run over a series of discrete time steps, its lifetime. In each time step initially a fraction of the free proteins, determined by the global binding proportion parameter, are bound to matching sites. The fraction of proteins available for binding is assigned to the binding site that has the same binary code as the protein. If there is more than one binding site competing for the same protein the fraction is equally distributed between all matching sites. In this process all protein binding sites are treated the same, regardless of the cis-module they belong to. Calculation of every gene's activation level is done by adding (activatory) or subtracting (inhibitory) the values per cis-module but only the lowest value of bound protein per cis-module is used to allow for non-linear effects. The cumulative activation level of all cis-modules then serves as input to one of two activation functions, depending on the gene's type, "default on" or "default off" as shown in fig. 1. The output of the gene's activation function is added to the unbound concentration of that gene's output protein type. After this calculation the concentrations of all unbound proteins are, if necessary, reduced to the global saturation value and all proteins, free or bound, are decayed by the protein-specific rate. Finally environmental input occurs by increasing the unbound concentration of certain proteins by some value and output by reading protein concentration values. Simple scaling is used to map stimulus input levels from the signal range to a protein concentration, and *vice versa* for output protein levels.

Spatial model

The Cellular Potts Model (CPM) has been introduced by Glazier and Graner (13), who developed it to simulate differential adhesion driven cell arrangement. Since then CPM has been used for a variety of cell-level modeling tasks, recently reviewed by Merks and Glazier (24). Although quite complex models have been realized with CPM, like the development of a cellular slime mold by Marée and Hogeweg (22), the cell level CPM simulation has not been combined with a GRN controlling cell parameters individually, without predefined cell types, before.

CPM is a two level system: On the lower level there is a grid of pixels, while the higher level cells consist of any number of lower level pixels. Cells have properties like target volume, shape, etc. and deviations from these targets incur energy penalties. Every pixel on the other hand has an integer value assigned, designating it as belonging to the cell with that identifier or zero if it is part of the "medium" (empty space). The system changes by trying to copy over one pixel's value to a randomly chosen neighbor pixel (morphogen concentration values are not copied, instead they diffuse in a separate process described below). In every time step ("Monte Carlo Step") one copy attempt is undertaken for every grid pixel on average. Copying is limited by energy constraints: Changing a pixel will change the energy properties of one (if either pixel is part of the medium) or two cells. The overall energy E is the weighted sum of all constraining properties. Copying is accepted with probability:

$$P = \begin{cases} e^{-(\Delta E - \delta)/kT} & \Delta E \geq -\delta \\ 1 & \Delta E < -\delta \end{cases}$$

In the work presented here we set offset $\delta = 0.0$, Boltzmann constant $k = 1.0$ and temperature $T = 2.0$. A two-dimensional, non-toroidal 60×40 pixel grid was used.

We use a flexible open source CPM implementation called CompuCell3D, see (6; 5) for implementation and formalism details. Through a modular plugin structure (where usually one plugin adds one energy constraint) it is easy to write extensions for the existing software (standard plugins include volume, surface, mitosis and connectivity constraints).

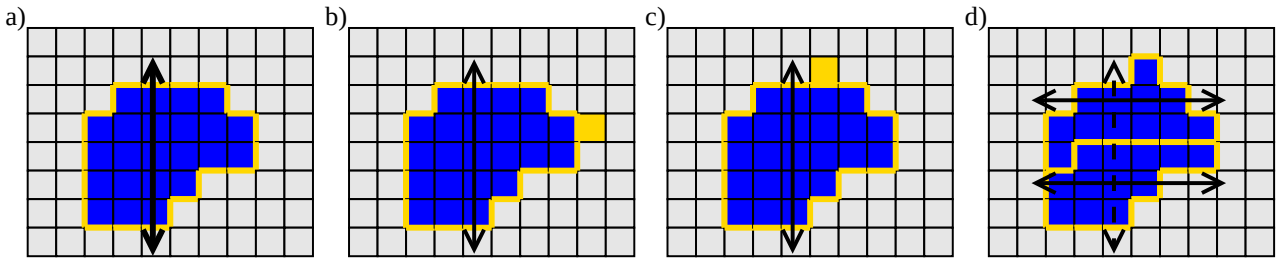

Figure 2: **Cell change example.** The blue cell in a) has volume 23 and the orientation shown by the arrow. Assuming for this cell a target ratio of 1 between the length in arrow direction and the length in the direction orthogonal to the arrow. In b) the yellow pixel is tested for becoming part of the cell. However, as this would change the ratio from 6:5 to 7:5, the transaction is energetically unfavorable. The yellow pixel in c) brings us closer to the target length ratio so it is more likely to be accepted (target volume and other constraints permitting). In d) the cell reached the mitosis volume of 24 so it split evenly along an axis orthogonal to its orientation. The daughter cells inherit an orientation rotated by 90 degrees (which can be continuously modified by GRN dynamics).

GRN controller

A GRN is able to control the cell's volume via a protein level mapping to the ratio of the size required for mitosis. So a maximal protein level meant the cell would try to grow until mitosis took place while a zero protein level would initiate shrinking and usually lead to apoptosis within a few steps. The actual volume of the cell could however differ from the target size due to neighboring cells for example. To take this and the fact that externally enforced size differences can affect the behavior of biological cells (11; 15) into account, the actual volume ratio served as input to the GRN, i.e. determined another protein's level.
Two protein levels were used to determine a cell's color. The continuous protein levels were interpreted as boolean values for this (above 0.5 threshold: true, below threshold: false). The color however does not correspond to cell type as in other models, where the type would determine cell parameters. All other parameters are controlled by the GRN independent of the color chosen. Also the color was not chosen once and for all but protein levels would be interpreted in every time step anew.
The GRN could also control the expression of "cadherins", factors that influence adhesion to other cells expressing them, i.e. it could be to some degree energetically favorable for cells with adhesion to "stick together". The adhesion strengths of the three expressible cadherin proteins were prespecified (see web page for table).

Diffusion Morphogens diffuse and decay on the underlying grid (substrate) – so it does not immediately matter for diffusion whether a cell is present or not. A cell can however increase the concentration of a morphogen on the pixels it consists of. Unlike earlier models with directed cell-cell communication mechanisms, where a cell receives as input some output (often state) of its direct neighbors, this allows for long distance communication. On the other hand meaningful communication might be harder to evolve this way as it is less directed and morphogens do not correspond to other cell variables.
The diffusion plugin "FlexibleDiffusionSolver" we used comes already with CompuCell3D and is explained (especially the numerical approximation) in its manual (6) so we only give a brief description here. The concentration c_i of morphogen i with diffusion constant d_i and decay constant k_i changes as:

$$\frac{\partial c_i}{\partial t} = d_i \nabla^2 c_i + k_i c_i + secretion_{ixy}$$

The term $secretion_{ixy}$ is the increase of morphogen i in pixel (not cell) xy. Here we used two morphogens, with the constants $d_1 = 0.2$, $k_1 = 0.009$ and $d_2 = 0.2$, $k_2 = 0.003$.
A cell receives the average morphogen concentration of the pixels it consists of as input (determining one protein level per morphogen). Two other protein levels determined the secretion of the corresponding morphogen, realized as increasing the average concentration of the cell's pixels.

Cell Shape Control Apart from the plugin which manages GRN control over parameters determining cell dynamics we developed another major new module: CellShapeControl. Earlier work successfully used a plugin to model cell elongation along the cell's longest axis (23; 35). However, in these earlier works elongation followed the cell's longest axis, so orientation was due to initial random effects. In our model the orientation is partially inherited and partially under the GRN's control (see mitosis section below for details). Furthermore, better results were found during evolution when not the length directly was under GRN control but the ratio of the length along the orientation axis to the length along the orthogonal axis, see fig. 2 a)-c).

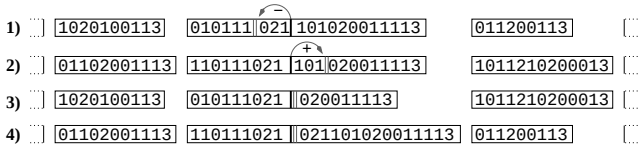

Figure 3: **Gaussian offset crossover.** Genomes of (1) parent 1, (2) parent 2, (3) offspring 1, (4) offspring 2. Only the compartment chosen for crossover and two neighboring genes are shown. Both children get digits up to the crossover point (solid bar) from their respective parent, but then continue in the other parent's genome with opposite gaussian-distributed offsets (-3 and $+3$, respectively, here).

Mitosis

Mitosis was initiated automatically once the cell reached a volume of 24 pixels, implemented as equal split along the axis orthogonal to the cell's orientation, see panel d) in fig. 2. The heritable part of the cell's orientation was shifted by 90 degrees and target volume set to half the parent's value in both daughter cells. Also, protein levels were divided equally.

Initial State In the experiments reported here we neglected "maternal factors", i.e. the first cell started without spatial extension at a size of 1 pixel, in the middle of the grid. All protein levels were set to zero for the GRN and no morphogen was present.

Evolutionary setup

A standard Genetic Algorithm with elitism, tournament selection and replacement was used. An evolutionary run lasted 250 generations containing 250–300 individuals[3]. The initial population started with one cis-module per gene and one protein binding site per cis-module, all coding bit values being randomly assigned; in network terms the nodes were randomly connected, with at most one incoming arc.

Selection

Later generations are formed by carrying over the best-performing individual of the last generation automatically and the other individuals are replaced by offspring. To generate each pair of offspring, 15 (not necessarily different) individuals of the prior generation are chosen randomly and of these the best two selected to be "parents".

[3]The variable number of individuals is due to our particular setup where clustering software Condor operated on student lab machines, so that sometimes PC usage would interrupt computations. In order not to wait for the last results to arrive we always started 500 individuals in batches of 20 but would progress to the next generation as soon as more than half of them were back. Due to batch size and step timing it could however happen that more than 260 jobs arrived before the next check.

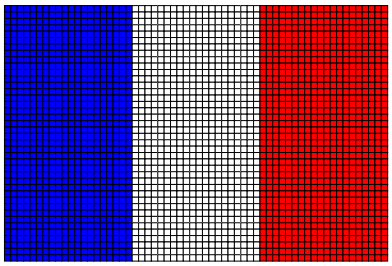

Figure 4: **French flag.** Evolutionary target was a 60×40 pixel three striped pattern.

Variability

A single-point crossover between the parent genomes occurred 90 percent of the times and every coding bit was flipped with a mutation probability of one percent. As there could be a variable number of cis- and of protein binding sites per gene their lengths will vary, so a standard bit-string crossover could change their numbers drastically. To conserve all but (at most) one of the genes as basic building units, the genomes of the parents were divided into compartments: one compartment for every gene and one compartment for the global variables. Then (with a probability of 0.9) a single compartment was chosen for crossover and in this compartment a point allocated for crossover.[4] Every pair of compartments was aligned regardless of the lengths of other compartments, as indicated in fig. 3. This process is inspired by the biological mechanism known as synapsis, the pairing of homologous chromosomes where mostly "similar" sectors pool together. To achieve variable length genes, the unequal crossing-over observed in biology is mimicked: When crossing over from parent 1's genome to the second parent's genome copying does not necessarily continue at the same position of parent 2's genome but is shifted by an offset (see fig. 3).

This offset is randomly drawn from a gaussian distributed random variable with mean 0 and standard deviation 4. The relatively large number four was chosen to increase the chance of duplicating genetic information, the importance of which was already pointed out by Ohno (26) for the evolution of biological complexity. Ohno put emphasis on whole-genome duplications while it is now, with better techniques, becoming ever clearer that "both small- and large-scale duplication events have played major roles" (29, page 320ff). Note that the offset point is limited to stay within the boundaries of the compartment, hence if crossover point + offset is smaller/larger than the left/right boundary it is set to the corresponding boundary value. So the number of 2s (cis-modules) might increase by crossover – mutation was only applied to coding digits (0s and 1s) – but not the number of 3s as these are the compartment boundaries. When

[4]This is why 'at most' one gene is changed: The crossover point could be zero or equal to the gene's coding length.

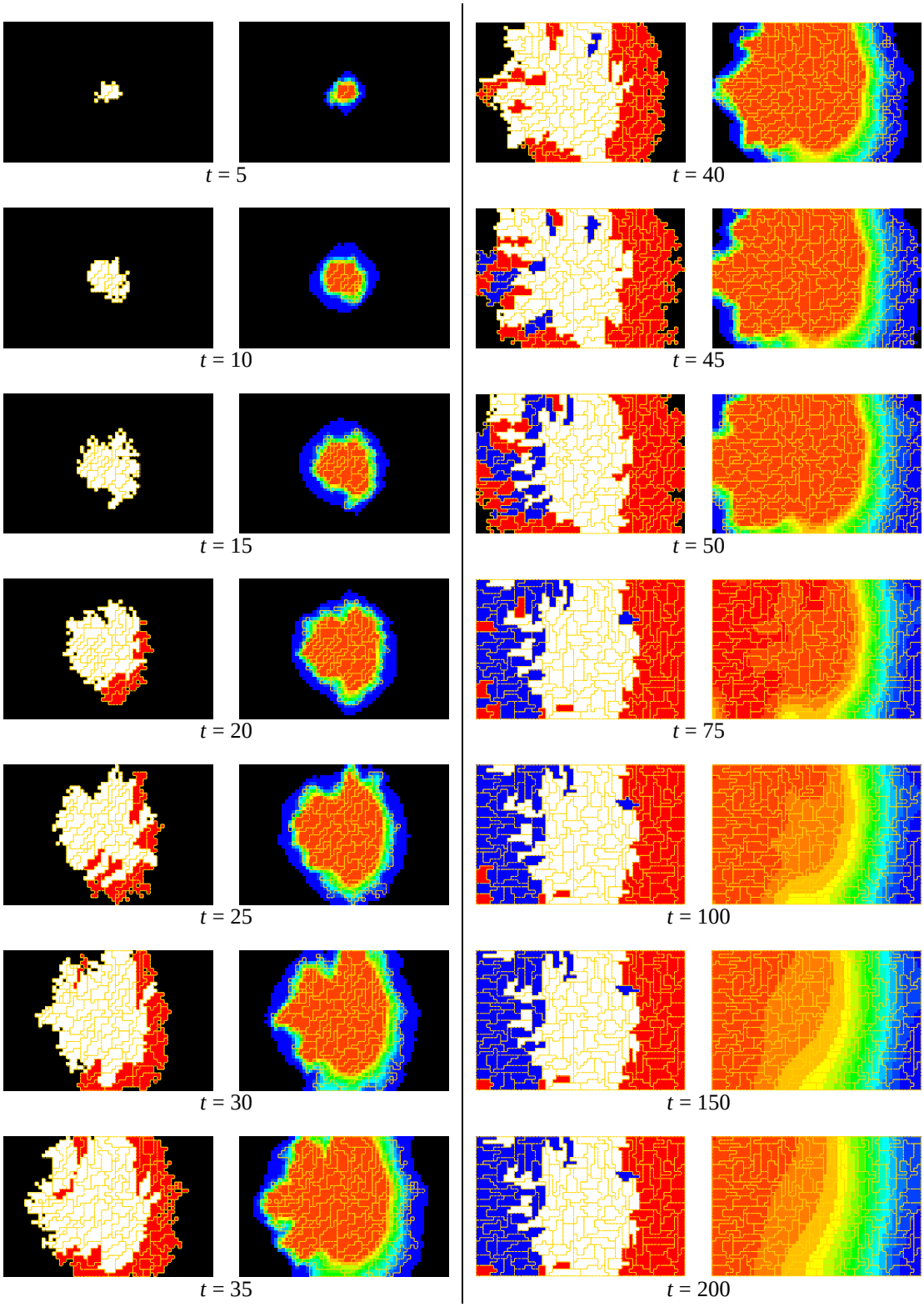

Figure 5: **Morphogenesis example.** The left picture shows the colors of the cells while the right one shows adjusted morphogen level (red=high to blue=low), respectively. Cell outlines are given in yellow, t is the time step the snapshot was taken at. See online version for color. In the left column it becomes quite clear that morphogen diffuses on the underlying grid as you can see low levels of it in areas without cells. Similarity to the target French flag pattern at $t = 200$ is 87 percent.

crossover occurs in the part encoding for global parameters the offset is always set to 0 as offsets would be meaningless here.
These processes allow both neutral crossover and mutational changes, as degenerate cis-modules (i.e. too short to bind a protein) have no regulatory effect. Additionally this means that genes could become dysfunctional, in a similar manner to the so called pseudo-genes found in nature, e.g. if there were not a single cis-module and the gene had an activation type of "off by default".

Fitness Evaluation

The lifetime of each individual was 200 time steps. Its fitness was defined simply as the match of the cell formation at the last time step compared with the 60×40 pixel target pattern comprising the entire spatial field, see fig. 4. Formally, the difference d between individual i's result pattern R^i and the target T, both of size $w \times h$, is:

$$d(R^i, T) = \frac{1}{wh} \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} |sgn(R^i_{xy} - T_{xy})| \in [0, 1]$$

Note that this measure does not specify the number, size or position of cells, only the lower level pixels are taken into consideration. Accordingly, the correctness of the match or similarity s is simply $s(R^i, T) = 1 - d(R^i, T)$.
One refinement we introduced was that the difference between the patterns would be multiplied by four minus the number of colors present in the final individual. An individual using all three possible colors would get a factor of one, while individuals using fewer colors would be penalized with a higher factor. The aim of this was to "encourage" individuals early on in evolution to use all colors – without this measure runs were prone to get stuck in local maxima with huge single-color individuals.
To improve robustness, every individual was run 10 times with different random seeds and the overall fitness calculated as the average of these repetitions.

Results and Discussion

The proposed model combines for the first time spatially extended cells with GRN control over individual cell parameters like size, shape, adhesion, morphogen secretion and orientation without predefined cell types. Evolved multicellular organisms were able to autonomously set up an asymmetric morphogen gradient and organize into a close match of the French flag pattern, so the model will certainly be of good use for future research.
A majority of evolutionary runs achieved final pattern matches of over 75 percent[5]. In particular, some individuals used morphogens to set up, and maintain, a gradient across the grid, as did the example in fig. 5. The cells would usually quickly divide and spread over the grid early on during an individual's lifetime while becoming more stable later – pattern similarities at time step 210 were very close to the originally measured fitness at time step 200. A detailed assessment of intra- and extracellular dynamics is unfortunately beyond the scope of this paper.
The fact that the initial cell in the first time step had an absolute orientation of zero degrees is clearly biologically unrealistic. The reason was simply that this made assessing fitness easier. We do not think this is a major drawback as when we started the initial cell with an orientation of 90 degrees and rotated the grid and target pattern by 90 degrees as well, the similarity achieved was just as good.
Due to the inherent randomness in the spatial process we can probably not expect a perfect pattern match on such a low level – nobody expects identical twins to be identical down to the cell level. However statistical analysis of many runs is needed to quantify these effects. Also it would be desirable to investigate robustness as well as evolvability of the system. Self-repair properties after perturbations to development would be expected as for example Miller (25) observed in his work.
In the future we plan to use more natural fitness measures and to extend the evolution of morphogenesis to more complex, three dimensional forms instead of 2D patterns (as the name CompuCell3D suggests this is relatively straightforward with the software used). Also, biological cells are even more flexible and irregular in the forms they can take; for example it is known that contraction, leading to a wedge like shape, is important for gastrulation (34). It would be interesting to see if evolvability is increased further when the GRN gets more control over cell shape.

[5] Seven out of the ten runs conducted with the particular configuration of plugins and parameters described here ended with similarities above 0.75. Recall, this is the average match over 10 repetitions with different random seeds.

Acknowledgments

We would like to thank the CompuCell3D.org team for making their software available as open source project. The free distributed computing software Condor (http://cs.wisc.edu/condor/) made it possible to compute the fitness of many individuals in parallel. Also much recommended are the well documented algorithms provided by GeometryAlgorithms.com, which were very useful.

References

Banzhaf, W. (2003). On the Dynamics of an Artificial Regulatory Network. In *Advances in Artificial Life, 7th European Conference, ECAL'03*, volume 2801 of *Lecture Notes in Artificial Intelligence*, pages 217–227. Springer Verlag.

Bongard, J. (2002). Evolving modular genetic regulatory networks. In *Proc. of the Congress on Evolutionary Computation 2002*, pages 1872–1877, Washington, DC, USA. IEEE Computer Society.

Bongard, J. C. and Pfeifer, R. (2001). Repeated structure and dissociation of genotypic and phenotypic complexity in artificial ontogeny. In L. Spector et al., editor, *Proc. of the Ge-*

netic and Evolutionary Computation Conference (GECCO-2001), pages 829–836, San Francisco, CA, USA. Morgan Kaufmann.

Buss, L. W. (1987). *The Evolution of Individuality*. Columbia University Press, New York.

Cickovski, T., Aras, K., Swat, M., Merks, R. M. H., Glimm, T., Hentschel, H. G. E., Alber, M. S., Glazier, J. A., Newman, S. A., and Izaguirre, J. A. (2007). From genes to organisms via the cell: A problem-solving environment for multicellular development. *Computing in Science and Eng.*, 9(4):50–60.

CompuCell3D (2007). Software package and manual, http://compucell3d.org/.

Cornish-Bowden, A. (2001). *Fundamentals of Enzyme Kinetics*. Portland Press, London.

Devert, A., Bredeche, N., and Schoenauer, M. (2007). Robust multi-cellular developmental design. In *GECCO '07: Proc. of the 9th Annual Conference on Genetic and Evolutionary Computation*, pages 982–989, New York. ACM.

Federici, D. (2004). Increasing evolvability for developmental programs. In Miller, J., editor, *Proc. of the Workshop on Regeneration and Learning in Developmental Systems, WORLDS 2004*.

Federici, D. and Ziemke, T. (2006). Why are evolved developing organisms also fault-tolerant? In S. Nolfi et al., editor, *SAB*, volume 4095 of *Lecture Notes in Computer Science*, pages 449–460. Springer Verlag.

Folkman, J. and Moscona, A. (1978). Role of cell shape in growth control. *Nature*, 273(5661):345–349.

Gerhart, J. C. and Kirschner, M. (1997). *Cells, Embryos And Evolution*. Blackwell Publishing.

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

Hogeweg, P. (2000). Shapes in the shadow: Evolutionary dynamics of morphogenesis. *Artificial Life*, 6(1):85–101.

Huang, S. and Ingber, D. E. (2000). Shape-dependent control of cell growth, differentiation, and apoptosis: Switching between attractors in cell regulatory networks. *Experimental Cell Research*, 261(1):91–103.

Jaeger, J. and Reinitz, J. (2006). On the dynamic nature of positional information. *BioEssays*, 28(11):1102–1111.

Kauffman, S. A. (1993). *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press.

Knabe, J. F., Nehaniv, C. L., and Schilstra, M. J. (2006a). Evolutionary robustness of differentiation in genetic regulatory networks. In Artman, S. and Dittrich, P., editors, *Proc. of the 7th German Workshop on Artificial Life 2006 (GWAL-7)*, pages 75–84. Akademische Verlagsgesellschaft Aka, Berlin.

Knabe, J. F., Nehaniv, C. L., and Schilstra, M. J. (2008). Genetic regulatory network models of biological clocks: Evolutionary history matters. *Artificial Life*, 14(1):135–148.

Knabe, J. F., Nehaniv, C. L., Schilstra, M. J., and Quick, T. (2006b). Evolving biological clocks using genetic regulatory networks. In L. M. Rocha et al., editor, *Artificial Life X: Proc. of the Tenth International Conference on the Simulation and Synthesis of Living Systems*, pages 15–21. MIT Press/Bradford Books.

Kumar, S. and Bentley, P., editors (2003). *On Growth, Form and Computers*. Elsevier Academic Press, London.

Marée, A. F. M. and Hogeweg, P. (2001). How amoeboids self-organize into a fruiting body: Multicellular coordination in *Dictyostelium discoideum*. *Proc. Natl. Acad. Sci. USA*, 98(7):3879–3883.

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

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

Miller, J. F. (2004). Evolving a Self-Repairing, Self-Regulating, French Flag Organism. In K. Deb et al., editor, *Genetic and Evolutionary Computation - GECCO 2004*, Lecture Notes in Computer Science, pages 129–139, Berlin, Heidelberg. Springer Verlag.

Ohno, S. (1970). *Evolution by Gene Duplication*. Springer Verlag.

Quick, T., Nehaniv, C. L., Dautenhahn, K., and Roberts, G. (2003). Evolving Embodied Genetic Regulatory Network-Driven Control Systems. In *Advances in Artificial Life, 7th European Conference, ECAL'03*, volume 2801 of *Lecture Notes in Artificial Intelligence*, pages 266–277. Springer Verlag.

Schilstra, M. J. and Nehaniv, C. L. Bio-logic: Gene expression and the laws of combinatorial logic. *Artificial Life*, pages 121–134.

Stanley, K. O. and Miikkulainen, R. (2003). A taxonomy for artificial embryogeny. *Artificial Life*, 9(2):93–130.

Taylor, J. S. and Raes, J. (2005). Small-scale gene duplications. In Gregory, T. R., editor, *The Evolution of the Genome*. Elsevier Academic Press.

Turing, A. M. (1952). The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72.

West-Eberhard, M. J. (2003). *Developmental Plasticity and Evolution*. Oxford University Press.

Wolpert, L. (1969). Positional information and the spatial pattern of cellular differentiation. *Journal of Theoretical Biology*, 25:1–47.

Wolpert, L. (1996). One hundred years of positional information. *Trends in Genetics*, pages 359–364.

Wolpert, L., Jessell, T., Lawrence, P., Meyerowitz, E., Robertson, E., and Smith, J. (2007). *Principles of Development*. Oxford University Press, Oxford, 3rd edition.

Zajac, M., Jones, G. L., and Glazier, J. A. (2003). Simulating convergent extension by way of anisotropic differential adhesion. *Journal of Theoretical Biology*, 222(2):247–259.