EVOLUTION OF 3D DEVELOPMENT CONTROLLED BY A GENE REGULATORY NETWORK: THE COMPLEXITY OF THE SEARCH SPACE AND EVOLVABILITY

MICHAŁ JOACHIMCZAK[1]

BORYS WRÓBEL[2]

Computational Biology Group, Department of Genetics and Marine Biotechnology
Institute of Oceanology, Polish Academy of Sciences
Powstańców Warszawy 55, 81-820 Sopot, Poland
[1]*mjoach@gmail.com*
[2]*bwrobel@iopan.gda.pl*

The question of what properties of biological systems allow for efficient evolutionary search in complex fitness landscapes (evolvability) is one of the central interests both for the research in the field of evolutionary biology and artificial life. We hope to be able to address this issue by using a model of 3-dimensional multicellular development in which cell fate is determined by differential gene expression in each cell. The development relies on an indirect mapping between the genotype and the morphology (the phenotype). Cell differentiation is allowed by positional information provided by diffusible factors, and the state of the gene regulatory network coded by the genome determines the cell fate (such us division, death, growth). The connectivity in this network is determined by the proximity of sequences of genetic elements in N-dimensional space. One can imagine these sequences as points in space which approach or move away from each other as the genomes evolve. Changing the number of dimensions of the sequence space allows to ask directly the questions about the effect of the complexity of the search space on the efficiency of the evolutionary search. Our results show that when a genetic algorithm is used, this efficiency is not significantly affected even when the space has high dimensionality.

Keywords: artificial embryogeny; gene regulatory network; evo-devo; complexity of the fitness landscape; neutrality; evolvability.

1. Introduction

The field of artificial life (AL) allows to ask experimental questions about perhaps the most important issue in the field of evolutionary and developmental biology (evo-devo): how the increase in complexity of gene regulatory networks is related to morphological innovations. To be able to investigate this matter, an AL model should have certain properties of biological realism, for example, one should not limit the size of the genome or the number of connections in gene regulatory network (GRN).

We propose a model of embryogenesis in which multicellular development in 3D space is controlled by a GRN, in some aspects similar to models introduced by Eggenberger ([4–7]). Each cell in our model has the same genome and thus the

topology of the network is the same in each cell, but the expression of genes in different cells may vary. Development starts from a single cell and in each discrete time step of the simulation cells can divide or die. The fate of the cell depends on its history and its coordinates in 3D space (which are real values, no grid is used). Cell location affects gene expression indirectly, by determining the repertoire of perceived diffusible products (produced by other cells) and diffusible factors present in the environment.

The main types of elements in the genome are "promoters" and "genes". "Genes" are the elements that act as if coding functional biological molecules. This is consistent with the biological nomenclature [9]. In biological genomes, segments of nucleic acids that code functional products (genes) are preceded by promoters. These are the places at which transcription starts: the construction of an RNA molecule based on the nucleotide sequence of the genome (usually, DNA). This RNA molecule can have a function by itself or its sequence may be (in part) translated to a functional molecule built of amino acids (a protein). The function of an RNA or a protein molecule may be structural/mechanical, catalytic or regulatory. In particular, many regulatory proteins bind to DNA in the vicinity of promoters to regulate transcription.

In biological systems, the activity of promoters is a prime determiner of the level of transcripts and thus of the functional molecules in the cell, but not the only one. Gene expression is regulated also during the subsequent steps (post-transcriptional and post-translational). This is determined by the affinity of the regulating factors (RNAs, proteins) to the sequences of regulated transcripts and polypeptides, which ultimately correspond to the sequences in the genome. The "promoters" in our model are a metaphor for all these regulatory regions (cis-regulators) and the way the affinity is determined is a conceptual simplification of the biological reality, in which, for example, the recognition of a nucleotide sequences by proteins that bind nucleotide acids is determined by the chemical properties of the two molecules.

Most of existing models of artificial embryogeny follow a many-to-one scheme of multiple regulatory regions (promoters) and one gene/product (e.g. [1, 4–7, 13]). However, in biological genetic systems, both prokaryotic and eukaryotic [9], many-to-many relationship is more common. Clustering of several genes in so-called operons is not, as originally thought, restricted to bacteria (for a recent review, see [2]). Operons allow for co-transcription from one promoter of functionally related genes (for example, involved in the same biochemical process) and thus co-regulation. Such co-transcription occurs in eukaryotes [2,9] and similar logic applies to multiple transcripts sharing common regulatory regions (enhancers, silencers [9]). Furthermore, many proteins are in fact multifunctional and consist of several parts (protein domains) that are functionally and structurally independent.

One of the features of biological systems is that regulation occurs at different levels. In our model three types of interactions are used: internal product-promoter, morphogen-promoter and morphogen-receptor. Differential expression of receptors with affinity to different morphogens allows for directing cellular growth toward

or against morphogen sources. Morphogens can be produced by the cells of the developing embryo, but some are not: gradients of these factors (external factors) are instead a feature of the environment. This is the mechanism that is crucial for cell differentiation after the first division (a symmetry breaking mechanism), but the gradients are present throughout the development. A similar mechanism is known to direct the initial stages of insect embryogenesis (for a popular introduction, see [3]).

In this preliminary work, we use a genetic algorithm to find gene networks that can regulate multicellular development of non-trivial 3D shapes (trivial shapes such as ellipsoids evolve easily in our model). In particular, we are interested in finding out how the complexity of the search space influences the efficiency of the search by the evolutionary process (the ability to find high-scoring morphologies). To measure the similarity between an evolved shape and the target we use a voxel-based approach: cubical voxels are marked as either inside or outside the shape and voxels occupied by cells are counted [10]. Thus densely packing possibly overlapping cells (which can happen when no grid is used) provides no fitness advantage.

2. Detailed model description

2.1. *The genome and genetic elements*

The genome in our model is a list of genetic elements. Its size is not fixed: element deletions and insertions/duplications are allowed during evolution. Each element in the genome is characterized by a sequence of $N+2$ real values. One value determines the element type and one allows to modify the weight of the connections in the GRN (to modify the affinity of a given sequence to others) and to determine its sign (excitatory or inhibitory). The connectivity of a given element in the GRN depends on the remaining N values, or rather on the Euclidean distance in the N-dimensional sequence space to the sequences of other elements. Connections in the GRN (directed edges) are created only if this distance is below a certain threshold. The affinity between two elements is determined by the distance and the value of the modifiers and increases if the distance gets shorter (0 distance is a perfect match). In terms of evolutionary landscapes, one can imagine (or actually visualise if $N \leq 3$) that as genomes evolve (and the element sequences change), points in N-dimensional sequence space that correspond to the elements approach one another or move away.

2.2. *A many-to-many relationship between promoters and genes: regulatory units*

In our model one or several promoters followed by one or several genes define a regulatory unit: the basic building block of the gene regulatory network. To calculate the activation level of a regulatory unit, we first compute the activity of all its

promoters:

$$p_i = \sum_{k=1}^{K} L_k w_{k,i} \,. \tag{1}$$

where p_i is the activity of a given promoter, K is the total number of regulatory factors in the genome (e.g. internal or external products, see below), L_k denotes the perceived level of the factor k, and $w_{k,i}$ is the promoter-factor affinity:

$$w_{k,i} = \begin{cases} d_{k,i} \leq 5 : \mathrm{sgn}(m_k m_i) \frac{2|m_k m_i|(5-d_{k,i})}{10 d_{k,i} + |m_k m_i|} \,; \\ d_{k,i} > 5 : 0 \,. \end{cases} \tag{2}$$

where $d_{k,i}$ is the Euclidean distance between the sequences of the promoter i and the gene of factor k, while m_k and m_i are the values of their modifier fields. Constants are chosen so the affinity is 0 (no interaction) when the distance is larger than 5 and at a maximum (10) when the distance is 0. The maximum distance for interaction is an element of biological realism: without it, elements far away from each other could interact in principle. Its introduction allows also to simplify the calculations (otherwise all the possible connections in the GRN would exist). The actual value can be viewed as a factor of scale in our system. A maximum value for affinity prevents biologically unrealistic effects: without it, for a very high affinity connection, expression of a regulator at even an extremely small level could still have a huge effect on a regulated unit.

The main purpose of the modifier fields is to allow to determine whether an interaction is inhibitory or excitatory. This depends on the signs of modifiers (inhibitory when different). As a consequence, a product of a given element may have a positive or a negative effect on different regulatory units, which captures the function on many biological factors that may act as either activators and inhibitors depending of the context. Additionally, $|m_k m_i|$ affects the shape of the (hyperbolical) affinity curve: the function becomes more convex as $|m_k m_i|$ decreases and is approximately linear for large $|m_k m_i|$.

Two types of promoters allowing for two regulatory mechanisms are introduced: additive and multiplicative. The activation level for the set Ω of genes in a given regulatory unit is:

$$L_\Omega = f_A(\prod_{i=0}^{I} p_{m,i} \sum_{j=1}^{J} p_{a,j}) \,. \tag{3}$$

where I and J denote the number of multiplicative and additive promoters (respectively) and $p_{m,1..i}$ and $p_{a,1..j}$ describe their activations. The actual indexing of the promoters starts from 1, $p_{m,0}$ is reserved for the identity element of multiplication ($p_{m,0} = 1$). Finally, f_A is a sigmoidal function with an inflection point at 0.5, bounded by 0 and 1:

$$f_A(x) = \frac{1}{1 + e^{-\omega(x-0.5)}} \,. \tag{4}$$

where ω describes the steepness of the sigmoid and does not change during the evolution or development ($\omega = 10$ was used). A sigmoid shape of the function relating gene expression to the expression of regulatory proteins is a common feature in biological genetic switches, reflecting e.g. cooperativity of binding of proteins to DNA.

The presence of a multiplicative promoter in a regulatory unit results in a strict requirement for the presence of a product with an affinity to it, otherwise the unit is not expressed. Such "all-or-nothing" regulation is quite common in biological systems and is difficult to otherwise incorporate when only additive units are used (such as in classical perceptron neural networks).

2.3. *Gene products*

We introduce three types of products that can be coded by genes in a regulatory unit: internal products, external products and receptors. The internal and external products affect the promoters in exactly the same way, but only external products (morphogens) can bind to promoters (and also receptors) in other cells than the one that produces them. The interactions between the receptors and morphogens influence the vector of cell division by shifting it toward or away from the morphogen source. Since the cells may differ in the pattern of gene expression, also their set of active receptors may be different, allowing each cell to orient its division in relation to the pattern of morphogens that are available at a given moment. The affinity between morphogens and promoters or receptors is defined by Eq. (2). In addition, we simulate simple diffusion: the perceived concentration of the morphogen in a given cell depends on the distance to the source and the level of the morphogen in the source (level of production). The level of morphogen m perceived by the cell c is:

$$L_{c,m} = \sum_{i=1, i\neq c}^{I} l_{i,m} \frac{1}{1 + D_{i,c}} \,. \tag{5}$$

where I denotes the number of cells of a developing embryo (potential producers), $D_{i,c}$ is the distance from the cell c to a source cell i in the 3D space of the developing organism, and $l_{i,m}$ is the level of morphogen m this source cell produces. The actual value of $l_{i,m}$ is delayed in time, depending on the distance from the source. This allows to simulate some of spatiotemporal effects of diffusion.

2.4. *Other types of genetic elements*

Only promoters and genes form regulatory units in our model, but we define three additional classes of genetic elements. These additional elements include, first of all, what can be considered the inputs and the outputs of the GRN: external factors and effectors, respectively.

External factors behave exactly the same as gene products: all can interact with promoters. Like products, they can also be divided into factors that diffuse from

some source and into non-diffusible factors. The difference between the factors and the products is that the perceived levels of the factors do not ultimately depend on the production by some cell and are provided as a part of the environment of the developing embryo. Diffusible factors are emitted from four specific points in 3-dimensional space in which the embryo develops. Four points is the minimal number that makes it possible for each cell to locate itself via 3D trilateration, provided each of these morphogens actually has affinity to some promoter(s) or receptor(s). The perceived level of the diffusible factors in a particular cell depends on the Euclidean distance to the source, in the same way as the level of external products (Eq. 5). This is the mechanism that allows for the initial break in the symmetry of cell divisions, a metaphor for the way such break occurs in insect embryogenesis [3].

Non-diffusible factors include three factors which provide historic/temporal information: a time signal (increases linearly from 0 and reaches 1 when the maximum time for the development is exhausted), a generation counter (incremented in each daughter cell after division), and the energy depletion level (increases from 0 to 1 as the energy of the cells goes down from its initial value to zero, decreasing at each cell division). Additional two signals are: a factor whose level is the same in every cell throughout the whole development (a "1" signal) and a sparseness signal. The "1" signal can be used by the GRN as a simple threshold for any of its regulatory units. The sparseness signal depends linearly on the number of cells in a given cell's direct proximity, up to maximum (1) for eight cells. Thus the signal reaches maximum for a relatively small number of neighbours (arbitrarily chosen), providing a simple way for a cell to detect when it is still part of a sparsely packed region.

If external factors can be viewed as the inputs, effectors are the outputs of the GRNs. Effectors are actually equivalent to single promoters attached to a specific developmental subprogram. Products or factors that have affinity to the effectors (Eq. 1) will regulate them.

The effectors either correspond to specific actions a cell can take as the embryo develops or allow to modify cell-specific parameters of the developmental process by a value corresponding to the activation level (Eq. 1). Such parameters are: the cell radius, the distance at which the daughter is put from the mother cell after division, internal division vector length, and internal division vector angles. In contrast, the effectors corresponding to cell actions act in a 0-1 fashion, not incrementally: when the activation level of such effector goes over a preset threshold, a cell divides, dies or freezes (in a frozen cell the GRN state is no longer updated).

The last type of genetic elements are pseudogenes. These are genes that are simply ignored when calculating the GRN, which means that a mutation in the sequence of a pseudogene does not affect the phenotype. The existence of this class of elements allows for a particular sequence to be temporarily shielded from the selective pressures. A genetic operator that changes the type of an element may convert an element of any type to a pseudogene and vice versa: a pseudogene to an element of any other type.

2.5. *Development*

As the cells divide, a new cell is placed by default next to its mother, but can be also placed further, at the distance controlled by the activation of a designated effector. This is a mechanism that in principle allows to mitigate dense cell packing resulting from divisions, since the cells in this model do not push each other away (we explore a model with such cell interactions elsewhere [10]). The radius of the daughter cell can also be set at division using a dedicated effector (no activity means default value, maximum activity translates into a twofold increase in the cell diameter).

The direction in which new cell is placed is determined by the division vector, which is controlled by two mutually supporting mechanisms. The evolutionary process can take advantage of any or both mechanisms to shape the development. Each mechanism uses one of two auxiliary vectors that are maintained for each cell. One is the internal division vector, inherited from the mother and possibly rotated at division. The other is the external division vector, named so because it direction is determined by the repertoire of perceived morphogens (external products and diffusible external factors). The sum of these two vectors gives the division vector. The division vector determines in which direction from the mother the daughter cell will be put, while the distance effector determines the distance from the mother.

The first mechanism draws from the mechanism used in 3D L-systems [11]. The direction of this vector is set after division and depends on the expression of angular effectors. The length of the internal vector depends on the activation of a separate designated effector.

The second mechanism relies on the external division vector, and thus on the morphogen-receptor interactions. The vector is oriented towards or away from morphogen sources. The overall effect is a sum of interactions of all receptors in the given cell with all morphogens produced by every source:

$$\vec{V_c} = \sum_{r=1}^{R} \sum_{m=1}^{M} \sum_{s=1, s\neq c}^{S} l_r w_{r,m} L_{c,m} \vec{\delta}_{s,c} \,. \tag{6}$$

where R denotes the total number of receptors in the genome, M the total number of external products and external factors defined in the genome, S is the number of sources (cells and four sources for positional external factors), l_r is the expression level of the receptor r in the cell (Eq. 3), $w_{r,m}$ is the morphogen-receptor affinity (Eq. 2), $L_{c,m}$ is the perceived level of the morphogen, and $\vec{\delta}_{s,c}$ is the normalised vector from the given cell to the source.

The non-diffusible external factors that give historic/temporal information (time, generation counter, energy level) provide a way to control the cell divisions. In particular, each cell division has some energetical cost, so the cell energy can be exhausted by rapid divisions. There is also a limit on the total energy that can be used during the development of the whole individual. Uncontrolled cell divisions cause the energy to be exhausted at early developmental stages, and thus result in low fitness. If not for this mechanism, uncontrolled cell divisions in some individuals

in the population would be a serious computational problem.

2.6. *Genetic algorithm*

We use a genetic algorithm in which a new generation is formed by copying 5 genomes without mutation (elitism), 145 with mutation, and 150 with mutations and multi-point crossover between genomes of different sizes. Tournament selection is used to choose candidate genomes for the next generation. Elite genomes can wander through the neutral regions in the sequence space (which may allow for a more efficient evolutionary search [8,12]): after a neutral mutation an elite genome always overwrites the previous version.

Initial generation is formed from a population of randomly created individuals. Each is created by inserting elements representing all possible effectors and external factors at the beginning of the genome and six regulatory units with random size are formed by sampling the number of regulatory elements from a normal distribution $N(3,3)$, and the number of products from $N(2,2)$. The probability of choosing an additive promoter was 2.5 times more than for choosing a multiplicative promoter, the ratio for products was 3:3:1 (internal:diffusible:receptor). The values of modifiers were drawn from $N(0,20)$. Random sequences for each element were obtained by drawing a vector with random direction in the N-dimensional space with the length taken from a uniform distribution with boundaries at 0 and 20. Thus the average distance between these random sequences does not depend on the dimensionality of the sequence space.

The genetic operators are designed to work on the level of the genetic elements rather than single bits or real values, because the genetic element is the basic unit of heredity in our model. Each operator has a predefined probability of occurrence per element in a genome and per generation. The first group of operators consists of those that act within individual genetic elements. First of all, the modifier field of the element can be disturbed with a value drawn from $N(0,1)$. Coordinates in the N-dimensional sequence space are mutated by adding a vector with random direction in the N-dimensional space and random length (taken from $N(0,1)$). This ensures that an average disturbance of a sequence does not dependent on the sequence dimensionality. It also allows for any point in sequence space to be reached with single mutation, albeit only small coordinate modifications have high probability. These operators are a metaphor of point mutations or short deletions/insertions in the coding or regulatory sequences in biological genomes.

The remaining operators cause inversion of the sign of the modifier (this causes a set of connections in the network to change from inhibitory to excitatory and vice versa) and changes of element types. All type changes are allowed, this includes a direct change a receptor into a morphogen or a promoter to a product (and vice versa), while conserving the sequence.

The second group of operators act on the level of entire elements (element deletion, insertion of a randomly created element or duplication of an existing one) and

on the level of the entire genome: deletion of a segment of the genome with random start and end point, and a duplication of such a segment to a random position in the genome.

The growth of nucleic acid-based genomes is to some extent kept in check by a higher probability of deletions than insertions. The ratio used in this work (about 2:1) still allows for the presence of the elements in the genome that can be mutated or deleted without affecting fitness (neutral elements). The presence of such elements, like the presence of pseudogenes, allows temporal shielding of some sequences from the selective pressures, and thus bolder movements in the sequence space, perhaps allowing for more frequent appearance of innovations beneficial from the point of view of natural selection [8, 12].

3. Results and Discussion

Our model builds on the ideas presented in several seminal papers by Eggenberger and other authors (see [4–7] and the references therein). However, in this previous work the cells grow on the grid and the structure is further reshaped by controlling the forces between neighbouring cells. In contrast, in the model presented here the cells divide freely in 3D space and gene expression can influence size of the cells and their spread. In the works of Eggenberger which introduce GRNs with the affinity based on the similarity of real numbers, only one dimension is used.

In this preliminary work we decided to find out how the complexity of the search space affects the evolvability. Since our model allows us to use any number of dimensions to represent sequences of genetic elements, we wanted to know how increasing the number of dimensions would affect ability to evolve target shapes and whether, perhaps, there is an optimal number of dimensions for our model.

In principle, one could expect that higher dimensionality will negatively affect the efficiency of the evolutionary search as it increases the size of the search space. However, there are reasons to expect beneficial effects from higher dimensionality. More dimensions allow for more neutrality in the mutational process. Neutrality can be beneficial as populations stuck in a local sub-optimum may still wander around in neutral directions in the fitness landscape searching for paths that allow for increase in fitness [8, 12]. The exact impact of the neutrality on evolvability is subject of intensive research and multiple studies confirm both positive and negative effects (see [8] for a recent review).

We investigated the effect of increasing the dimensionality by challenging the genetic algorithm with the task of evolving two different asymmetric target morphologies shown on Fig. 1a and Fig. 1c. For each shape and for each dimension tested we repeated the evolutionary run 20 times. Fitnesses of best individuals obtained in each run are plotted against the number of dimensions on Fig. 2a and Fig. 2b. Best individuals found across all dimensions are shown on Fig. 1b, 1d. Overall, the dumbbell shaped target was considerably more difficult to evolve.

Fitnesses obtained after each run for a given number of dimensions show high

variation (Fig. 2). This suggests that the genetic algorithm often gets stuck in local sub-optima. We do not observe any apparent effect of high dimensionality. If any, it is slightly detrimental, but even increasing the number of dimensions to 16 did not have a significant effect on evolvability.

Recombination between genomes is commonly used to allow populations to escape from local maxima. In other words, the expected effects of allowing recombination should go in the same direction as the increased neutrality. We have thus run a series of experiments using the stem-cap shape as a target (Fig. 1c) with recombination disabled. The results (Fig. 2c) are very similar to those observed for the genetic algorithm with recombination (Fig. 2b). To sum up, there was no effect of disabling crossing over, and here also no benefit of high dimensionality could be observed.

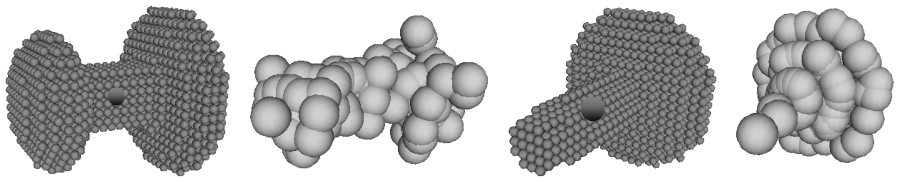

Fig. 1: Target shapes (a,c) and corresponding best evolved individuals from all experiments (b,d) Position of initial cell during growth can be seen on target shapes.

Galván-López and Poli [8] note that the controversy on whether neutrality helps or hinders evolutionary search may stem from the considerable variability in the problems and the lack of a common view on how neutrality can be increased. This makes the comparison of results obtained with different models rather tricky, especially since at the present stage our results are only preliminary. Little if any effect of the dimensionality of the sequence space on evolvability in our experiments suggest that the possible benefits from increased neutrality may be overshadowed by the drawbacks of the increased search space [8]. However, the search for efficient GRNs in our model does not depend only on the sequence mutations. The fact that we do not see significant difference between the results for 1D and 16D sequence space may also indicate that the actual search spaces do not differ all that much in complexity, which remains very large in all cases.

It remains to be seen if a more efficient search in the genome space could be promoted by introducing a more complex population structure, allowing for temporal divisions of the population with subsequent exchange of the genetic information. We plan to address these questions in further research on the interplay between evolution and development using the model presented here. However, we believe

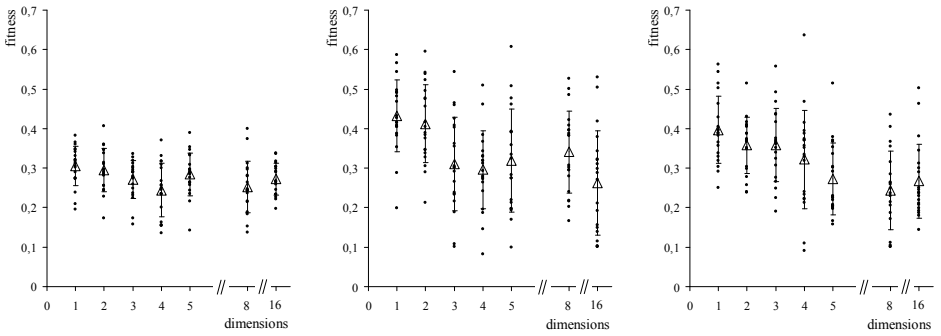

Fig. 2: The efficiency of the evolutionary search in highly-dimensional fitness landscapes. Figures show the fitnesses of the highest-scoring individuals in each of the 20 independent runs of the genetic algorithm after 1000 generations in the fitness landscape of different complexity (number of dimensions of the sequence space). Triangles are the mean fitnesses of the best individuals from all the runs; bars indicate standard deviation.

that these issues should be investigated in a setting more biologically realistic than a genetic algorithm, a setting in which the evolution is fuelled by competition for limited resources in a simulated world with a spatial structure. We hope that such spatial structure, allowing for temporal separation of subpopulations and simulating events similar to biological speciation, would permit for the investigation of morphological innovations and to ask questions about robustness of the GRNs, the issue of epistasis, and the statistical properties of the evolved GRNs. The results presented here provide us with important cues for further improvements for our model, as well as form a case against expecting intuitive interpretation of benefits of neutrality to automatically hold in an artificial life model.

4. Acknowledgements

The computational resources used in this work were obtained through the support of the Polish Ministry of Science and Education (project N303 291234), the Tri-city Academic Computer Centre (TASK) and the Interdisciplinary Center for Molecular and Mathematical Modeling (ICM, University of Warsaw; project G33-8).

References

[1] Beurier, G., Michel, F., and Ferber, J., A morphogenesis model for multiagent embryogeny, in *Proceedings of the ALife X: The Tenth International Conference on the Simulation and Synthesis of Living Systems* (MIT Press, Cambridge, MA, 2006), pp. 84–90.

[2] Blumenthal, T., Operons in eukaryotes, *Briefings in Functional Genomics and Proteomics* **3** (2004) 199–211.

[3] Carroll, S., *Endless Forms Most Beautiful: The New Science Of Evo Devo And The Making Of The Animal Kingdom* (WW Norton & Company, 2005).
[4] Eggenberger-Hotz, P., Evolving morphologies of simulated 3D organisms based on differential gene expression, in *Proceedings of the Fourth European Conference on Artificial Life* (MIT Press, Cambridge, MA, 1997), pp. 205–213.
[5] Eggenberger-Hotz, P., Exploring regenerative mechanisms found in flatworms by artificial evolutionary techniques using genetic regulatory networks, in *The Congress on Evolutionary Computation, CEC '03*, Vol. 3 (2003), pp. 2026–2033.
[6] Eggenberger-Hotz, P., Genome-physics interaction as a new concept to reduce the number of genetic parameters in artificial evolution, in *The Congress on Evolutionary Computation, CEC '03*, Vol. 1 (2003), pp. 191–198.
[7] Eggenberger-Hotz, P., Asymmetric cell division and its integration with other developmental processes for artificial evolutionary systems, in *Artificial Life IX: Proceedings of the Ninth International Conference on the Simulation and Synthesis of Living Systems* (MIT Press, Cambridge, MA, 2004), pp. 387–393.
[8] Galván-López, E. and Poli, R., An empirical investigation of how and why neutrality affects evolutionary search, in *GECCO '06: Proceedings of the 8th annual conference on Genetic and Evolutionary Computation* (ACM, New York, NY, 2006), pp. 1149–1156.
[9] Gerstein, M. B., Bruce, C., Rozowsky, J. S., Zheng, D., Du, J., Korbel, J. O., Emanuelsson, O., Zhang, Z. D., Weissman, S., and Snyder, M., What is a gene, post-ENCODE? History and updated definition., *Genome Research* **17** (2007) 669–681.
[10] Joachimczak, M. and Wrobel, B., Evo-devo *in silico*: a model of a gene network regulating multicellular development in 3D space with artificial physics, in *Proceedings of the Eleventh International Conference on Artificial Life* (MIT Press, Cambridge, MA, 2008), in press.
[11] Prusinkiewicz, P. and Lindenmayer, A., *The algorithmic beauty of plants* (Springer-Verlag, New York, NY, 1996).
[12] Shipman, R., Shackleton, M., and Harvey, I., The Use of Neutral Genotype-Phenotype Mappings for Improved Evolutionary Search, *BT Technology Journal* **18** (2000) 103–111.
[13] Steiner, T., Olhofer, M., and Sendhoff, B., Towards shape and structure optimization with evolutionary development, in *Proceedings of the ALife X: The Tenth International Conference on the Simulation and Synthesis of Living Systems* (MIT Press, Cambridge, MA, 2006), pp. 70–76.