Predicting the Timing of Critical Periods in Development

Oliver E.G Winks and Luc Berthouze
Evolutionary and Adaptive Systems Group
University of Sussex
Brighton BN1 9QH, UK
{O.Winks,L.Berthouze}@sussex.ac.uk

Abstract

Critical periods are specific periods in the development of a living organism during which there is an increased sensitivity to external perturbations. Such perturbations result in a developmental trajectory significantly different from what is considered the norm. This paper is concerned with the question of whether the presence and timing of a critical period can be predicted from the developmental profile without perturbation. Taking rate of change as a measure of this profile, we put forward the hypothesis that critical periods will occur when the rate of change is greatest. Using a simplified model of differential gene expression, cellular mechanics and evolution, we evolve organisms with different developmental profiles, test the presence of critical periods by systematically exposing the developing system to an external perturbation at different times between runs, and correlate the timing of these critical periods with the proposed developmental measure. We discuss the implications of our findings.

1. Introduction

In many respects, development can be defined as a series of transitions between structures of increasing complexity. This is especially true in the early stages of development when, in the space of just a few weeks, the fertilised egg goes from single cell to highly differentiated embryo through mitotic cell division. In this paper, our focus is on *critical periods*, that is, periods during development in which the developing system is unusually sensitive to some perturbation. The existence of critical periods is well documented in developmental biology and is one of Wilson's six principles of teratology (Wilson, 1973). Critical periods are reported in different organisms and in response to different types of perturbation. For example, exposure to ionizing radiation in early human organ development has been shown to interfere with cell signalling and disrupt normal cell differentiation (Anderson et al., 2000). In Zebrafish, exposure to a micro-gravity environment between 24 and 72 hours post fertilisation causes deficiencies in the vestibular system (Moorman et al., 2002).

Development, especially cellular development and morphogenesis, can be viewed as the product of complex interactions among many components. A useful perspective from which to consider the problem of critical periods is that of Waddington's epigenetic landscapes. Waddington was perhaps the first scientist to adopt a dynamic systems perspective of development, understanding that genes do not encode phenotypes, but rather, through interaction with each other and the environment encode a process from which phenotypes emerge. Waddington (1943) used the analogy of a landscape to visually depict development. Evolution, he argued, sculpts the landscape to stabilise some developmental trajectories and remove others. Using his analogy, genes are pictured as pegs that are connected to an overhead, flexible 'canvas' by guy-ropes. The guy-ropes represent gene interaction, and pull on the canvas to create a landscape featuring valleys and hills through which a ball, that represents the developmental process, can roll (Figure 1). The ball will naturally follow the valleys in the landscape, at the end of which are particular phenotypes, for example red or white eyes in *Drosophila*. However, as the ball rolls through the landscape it is susceptible to environmental perturbations which cause the decending ball to change trajectory (bifurcate) by rolling into different valleys. Taking critical period as such a bifurcation, we hypothesise that critical periods correlate with greatest rate of developmental change.

Models of Natural Development

The model described in this paper draws inspiration from the many examples of developmental models in the Artificial Life and Adaptive Systems literature. Bongard and Pfeifer implemented genetic reg-

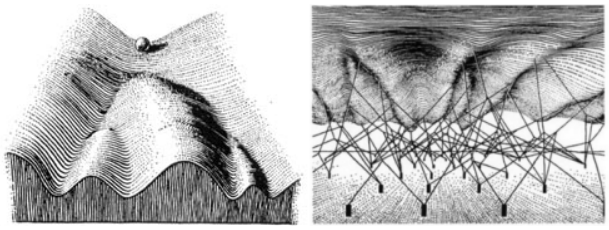

Figure 1: Waddington's epigenetic landscape. The right panel shows a view of the landscape from below with pegs representing genes and guyropes representing gene interactions.

ulatory networks (GRNs) to grow complete agents – morphology as well as neural control system (Bongard and Pfeifer, 2001, 2003). They evolved agents to perform some pre-defined task such as reaching and moving a block, or moving across a flat plane. The body was made of spherical units that could grow, divide and attach to six other body units via simulated joints. They were controled by a neural network that grew with the body. Eggenberger (1997) studied morphogenesis by utilising a GRN to control the functions of simulated cells. Stewart et al. (2005) also used a GRN to study morphogenesis. Stewart and colleagues controlled cellular functions with protein concentrations produced by the GRN. The cells could divide, die, change the direction they were facing and communicate with each other using protein diffusion gradients. Stewart and colleagues hand-designed the GRN and found several strategies for developing different kinds of cell structures.

Whilst all of these studies investigated development, and specifically morphogenesis, none of them mentioned or discussed the presence of critical periods. However, Bongard and Pfeifer (2003) did report that mutations expressed earlier in development tended to have a larger effect on the final agent, which could be related to a critical period in the agents development.

Aims

The aims of this study are two-fold. The first aim is to show that critical periods do occur in the type of complex dynamical systems typically used in simulations of natural development. This, as far as the authors are aware, is a novel piece of research and is important because of the potential for critical periods to confound research results and conclusions. The second aim is an extension on the positive outcome of the first. If critical periods do occur, then how does one predict the timing of their presence within the developmental process? In recent years artificial development and morphogenesis have been employed to create both virtual and real-life robots, this trend is likely to continue. Critical periods represent specific windows of time in which a developing system is highly sensitive to perturbation. For those who use artificial development and morphogenesis to create complex systems - such as robots - predicting these periods is not only useful for the succesful prevention of developmental failure but also in the analysis of experimental results.

2. Method

2.1 The Model

The simulation presented in this paper models the early stages of embryonic development. Specifically, it models the development of cellular structures (morphogenesis) and cell differentiation. The model is formed of two components, the genetic component and the cellular component. The genetic component simulates gene expression and genetic regulation by use of a Genetic Regulatory Network. Artificial transcription factors and proteins are synthesised that excite and inhibit genes in the network. The cellular component simulates several cell functions that make it possible to grow cellular structures composed of cells of different types. These functions are controlled by specific proteins created by the genetic regulatory network. The model exists in 3D space allowing for complex 3D structures to emerge over a fixed time period. There is a constant amount of energy in the system that cells consume so that energy consumption is a limiting factor that stops the physically impossible scenario of infinite growth.

Artificial evolution is used to create genomes that develop an organism of specific structure. Since different developmental profiles are needed to establish the generality of any finding, genome size is varied to allow organisms to develop differently into their final structure. Hand-designing the genomes would provide more control over the developmental profile of each organism, however, in practice it is difficult and time consuming.

2.2 Genetic Component

In nature, genes are responsible for creating proteins through genetic transcription. Special molecules (ribosome) translate genetic code into strings of amino acids which fold into proteins. Proteins are required in all cellular functions, but some proteins – called transcription factors – affect the transcription of genes in the genome. Transcription factors bind to the promoter region of genes that are sensitive to them. Once bound the transcription factor either increases or decreases the likelihood that the gene gets transcribed. This is the basic mechanism by which genes interact with one another; the synthesis of a transcription factor by one gene can affect the expression of all other genes in the genome. It is this interaction that the genetic component aims to model.

The genome in this model is represented by nodes that are interconnected with directed, weighted connections as shown by Figure 2.

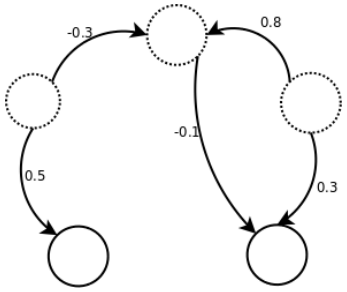

Figure 2: Representation of a genome: The nodes represent genes (dashed lines represent regulatory genes, solid lines represent structural genes) and the directed weighted connections represent the promoter regions of the nodes they are connecting to.

Each gene in the network is characterised by several variables: an activity y, a time-constant τ, a bias θ, and a type. Genes can be one of two types, structural genes or regulatory genes. The activity of the gene represents how much protein it is producing. If it is a structural gene then the protein controls some cellular function as well as influencing other genes in the network. Several structural genes can control the same function but a single structural gene can only control one function. If it is a regulatory gene then it can only influence other genes. The simulation was run for a fixed number of time steps, and at each time step every node in the network was simultaneously updated according to the following equation:

$$\frac{dy_i}{dt} = \frac{1}{\tau}\left(-y_i + E\sum w_{j,i}\sigma\left(y_j + \theta_j\right) + I\right) \qquad (1)$$

with

$$\sigma\left(y+\theta\right) = \frac{1}{1+e^{-(y+\theta)}} \qquad (2)$$

where E is the energy in the system, $w_{j,i}$ is the weight of the connection between node j and node i, and I is the external input into the node. The function σ is a sigmoid function of the j^{th} node's activity. If the gene is structural, and its activation rises above a predefined threshold value (2.5), then the cellular function that the genes product controls is activated. All cellular functions are performed over a predefined number of time steps (between 2 and 10 depending on the function) during which all structural genes that control the function receive a strong negative input which represents protein consumption. These time spans have no biological meaning and were arbitrarily selected.

Equation 1 shows the genome's activity to be dependent on the energy in the system. Energy was kept constant and evenly distributed between every cell in the organism. Therefore, the energy received by each cell decreased with the introduction of more cells to the organism.

2.3 Cellular Component

The cells in the model are spheres situated in an infinite 3D universe. All cells have the same radius (0.5 units) and can perform the same functions. In nature cells perform many different functions which are dependant on the type of cell they are. In this model only basic functions that are related to morphogenesis are simulated. The functions are as follows:

Cell Division: When a cell divides it makes a copy of itself and places the copy one and a half radius lengths away from its centre position in the direction of its mitotic spindle (described below). If that location is already occupied then no daughter cell is produced. The daughter cell's genome is initialised with the values of its mother cell. It also inherits the mitotic spindle orientation from its mother. This function takes 10 time steps to complete.

Cell Death: The cell is removed from the universe freeing up a location for another cell to divide into. This function takes 5 time steps to complete.

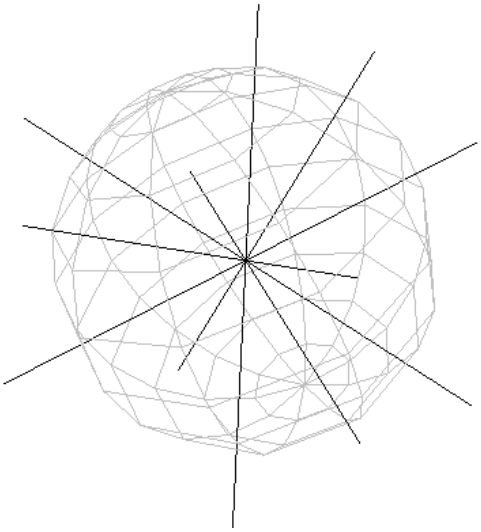

Figure 3: The 12 directions in which the mitotic spindle can point.

Cell Spindle and Cell Orbit: The cell's mitotic spindle points to one of twelve positions on the cell's surface (see Figure 3). The twelve positions are the corners of three mutually orthogonal squares centred at the cell's centre. The spindle location can be changed in one of two ways. Firstly the spindle location can be moved forward or backward one position on the same orbit, in which case the spindle

is moved to the next corner of the square it is currently on. Secondly the orbit can change, in which case the spindle moves to the same corner of the next square. Both of these functions require 2 time steps to complete.

Cell Signalling: Cell signalling is the principal mechanism by which cells differentiate and organise themselves into sub populations. Cells use concentration gradients of morphogens – proteins that can diffuse through cell membranes and induce signal responses in other cells – to provide spatial information to cell populations. Morphogens induce or suppress the expression of genes at different concentration thresholds, changing the dynamics of the genome within the receiving cell. This causes cells to perform different functions, or the same functions at different frequencies, to each other depending on their spatial location. Wolpert et al. (2007) showed how a population of hypothetical cells can differentiate using concentration gradients of morphogens to resemble a French flag. Since a cell requires three concentration gradients to pinpoint itself in 3D space, in this model cells can produce three morphogens. Morphogen diffusion is simulated by using a 3D Gaussian function centred at the position of the cell from which it originates:

$$a = \frac{1}{\sqrt{(2v\pi)^3}} e^{-\frac{1}{2}\frac{\|\vec{c}-\vec{n}\|^2}{v}} \quad (3)$$

where c is the current cell's position, n is the neighbouring cell's position and v is the gaussian function's variance.

At each time step the variance of the Gaussian function is increased until either the concentration of the morphogen falls to zero or the signal function is performed, in which case the variance is set back to its initial value. When a signal function is performed the corresponding morphogen concentration is increased by a small amount every time step for the duration of the function. As well as diffusing, all morphogens decay until their concentration falls to zero.

As well as being either a regulatory or structural gene, genes can also be 'input' genes. At every time step, each cell calculates the amount of morphogen it receives from its neighbouring cells:

$$s_m = c_m a \quad (4)$$

where s_m is the signal strength of morphogen m from the neighbouring cell and c_m is the concentration of morphogen m.

This value s_m is then fed into the input genes. Signal functions require 10 time steps to complete.

Cells that perform different functions, or the same functions at different frequencies, due to different dynamics in genetic regulation are considered to be different cell types. In this model, cell types are represented by the concentration of morphogens that a cell is producing. The type is depicted visually by colour. Each morphogen concentration maps directly to a channel in the RGB colour model with red controlled by morphogen 1, green by morphogen 2, and blue by morphogen 3.

2.4 Evolution

To create genomes with different developmental profiles, we used an adaptation of the microbial genetic algorithm (Harvey, 1996) to evolve genotypes satisfying a general enough fitness function. The algorithm was as follows:

1. Create an initial random population of genomes.
2. Run each individual for a fixed number of time steps.
3. Calculate each individuals fitness based on a prespecified fitness function.
4. Create a small sub-population (the elite population) of the fittest individuals.
5. For each non-elite individual, select a random elite and infect the non-elite with it.
6. Mutate the infected individual.
7. Repeat from step 2 for a fixed number of generations.

The genome for the evolutionary algorithm was simply an encoded version of the Genetic Regulatory Network described earlier in this paper. The network was encoded by a set of floating-point values in the range (-1, 1). This set was divided in to equally sized subsets, one subset for each gene in the network. An encoding for a single gene took the form $[function, morphogen, \tau, \theta, w_0 ... w_n]$ where τ and θ represent the time-constant and bias of the gene. The encoded values were scaled between (1.1, 10) and (-4, 4) respectively. Function and morphogen encode the cellular function that the gene governs, if any, and which morphogen that gene is sensitive to, if any. These values map on to the function or morphogen by dividing the full range (-1, 1) into equally sized segments representing the cellular function or morphogen respectively. An extra segment was included to represent no cell function (regulatory gene) and 'no morphogen sensitivity' (see Figure 4). The last set of values (w_0 ... w_n) encode the weights of the gene to every other gene in the network including itself. Therefore, if there are 9 genes in the Genetic Regulatory Network there are 9 values in the set. The weights were scaled between (-2, 2).

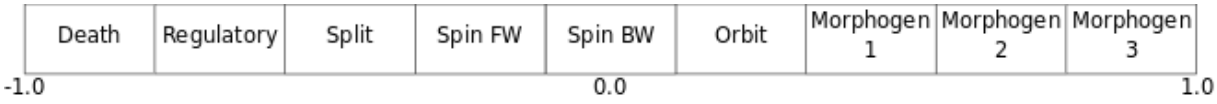

Figure 4: Mapping the gene encoding for cell function.

Several populations containing 40 individuals were evolved using the same fitness function. In an attempt to evolve organisms with different developmental profiles, each population had a different number of genes. The fittest individuals from these populations were removed after 500 generations and used in the experiments described in the results section. Each individual in the population was run for 300 time-steps before its fitness was calculated. The fitness function measured the affinity of the final organism to a target structure. A simple method was developed to specify arbitrary target shapes. By placing simple 3D shapes (spheres and cuboids) in the virtual universe and assigning target cell types to them, most structures could be specified. The particular target structure used for the fitness function was a set of 5 concentric spheres with a gradated target type, from strongly blue (type 3) in the centre, to strong green (type 2) on the outer most layer. Each sphere had a radius that was 2 cell unit larger than the sphere it encompassed, with the inner sphere beginning with a radius of 2 cell units (see Figure 5).

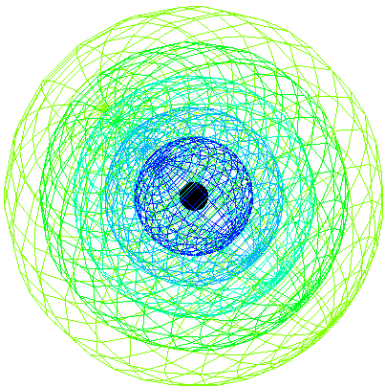

Figure 5: Target structure of the fitness function. The black sphere represents the zygote. The spheres are shaded with the colour of their target cell type.

A collision test between a cell and every sphere was performed to determine the sphere in which the cell resided. The affinity a_c of each cell to the encompassing sphere's target type was computed by:

$$a_c = \frac{3b_c(-1.5 + 2(|h_c - h_s|))}{\sqrt{c}} \quad (5)$$

where b_c is the brightness of the cell's colour, h_c is the hue of the cell's colour, h_s is the hue of the sphere's target colour and c is the number of cells in the organism. The normalisation by $\sqrt{c}$ balances the fitness attained by cell affinity to a target type and the fitness gained by cell division. The total fitness f of the organism was then simply defined as the sum of each cells affinity plus a symmetry modifier:

$$f = \sum a_c + (1 - \frac{\|\vec{m}\|}{d}) \quad (6)$$

where m is the 3D centre of mass vector and d is the maximum distance the centre of mass of an organism could possibly travel within a developmental run. In other words, the term in bracket encourages organisms to grow in place rather than shift towards the boundaries of the space.

Results and Discussion

Five genomes were evolved,with a genome size ranging from 13 to 17 genes (hereafter referred to as genomes 1 to 5 respectively). All produced organisms with differentiated layers of cells roughly matching those specified in the fitness function (i.e. blue cells in the centre of the organism getting progressively more green to the outer cell layer). Figure 6 shows snapshots of one of these organisms at different stages of development.

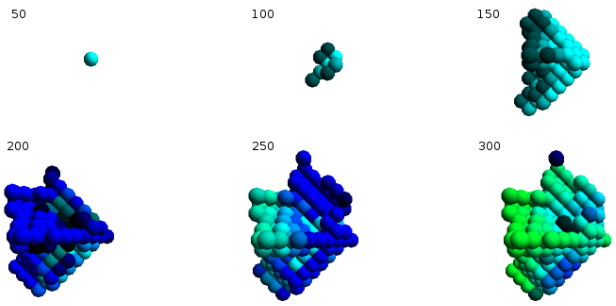

Figure 6: Snap-shots at 50 time-step intervals of the development of evolved genome with 15 genes.

2.5 Testing of critical periods

To locate critical periods in the development of these organisms, 135 developmental runs of each organism were performed. In each developmental run, the organism was subjected to a perturbation. This perturbation consisted of reducing the strength of cell signalling by introducing a parameter p in Equation

4, thus:

$$s_m = pc_m a \quad (7)$$

This perturbation lasted for 30 time-steps, i.e., 10% of the overall development. The fitness of the organism after development was then calculated. At each new run, the window was moved along the developmental run in 2 time-step increments. Ten series of experiments were run for each organism with p varying between 0.05 and 0.5, i.e., a reduction of signal strength between 50% and 95%.

For each organism, a developmental profile of the organism without perturbation was obtained by computing the absolute value of the rate of change of the average activity of the GRN across all cells. A critical period was defined as any drop in fitness of more than 2 times its standard deviation. Likewise, our hypothesis that critical periods correlate with greater rate of change was tested by identifying periods in the developmental profile of the organism when the rate of change was greater than 2 times its standard deviation after discarding the period during which the organism only had one cell – since the perturbation affects the signalling protein p, it cannot have any bearing on the developmental profile of the organism until after the first mitotic cycle. The correlation between critical period and rate of change was assessed using a test of cross-correlation, with confidence intervals set to 5% [1]. The test of cross-correlation provides two elements of information: (a) whether there is correlation between critical periods and periods of greater rate of change, (b) an estimate of the lag between critical period and rate of change.

2.6 Results

The middle panels (colour plots) in Figures 7 - 11 show raster plots resulting from the 10 runs in each experiment. The y axis denotes the perturbation strength p whilst the x axis (shared by the four top panels) represents the **onset** time of the perturbation window (the perturbation itself lasting 30 time-steps). The value of the fitness function was colour mapped on a red-to-white head-spectrum so that light colours denote high fitness and red colours denote low fitness. In each figure, the first (from the top) and fourth panels show two related measures of the organism's development without perturbation. The fourth graph shows the time-series of the activity of each of the genes that control the 8 cellular functions for each cell ('split' for cellular division, 'death' for cell death, 'sp FW' for spin forward, 'sp BW' for spin backward, 'orbit' for cell orbit, 'mrph i' for the i-th morphogen). The top graph shows the absolute value of the rate of change of the average activity of these genes. The dashed line (dotted line) denotes two (one) standard deviations from the mean. The second (from the top) panel of each graph represents the number of cells in the organism through development. The bottom graph displays the result of the cross-correlation. Its x axis denotes the lag so that peaks to the right of zero denote a lag of the critical period with respect to the rate of change. The dotted blue lines denote the confidence interval.

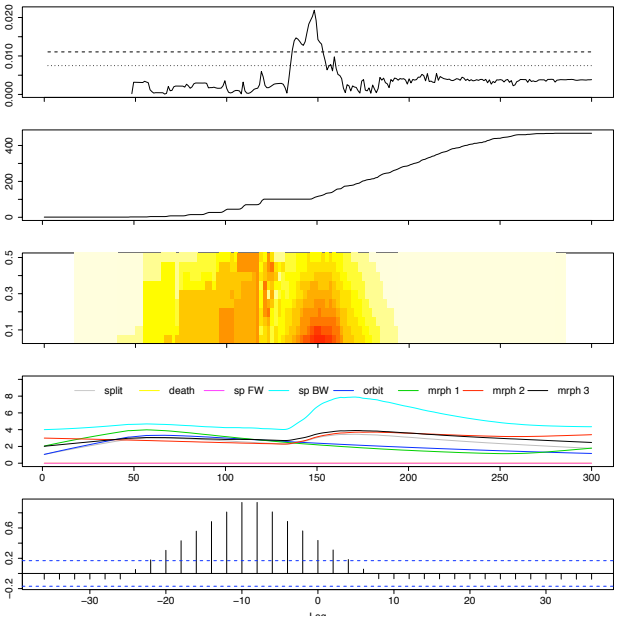

Figure 7: Results for the 13-gene organism. See text for details.

As shown by Figures 7-11, 3 of 5 organisms display critical periods around the 150th time step (and a second one for organism 2 around the 180th time step). The loss of fitness correlates with the severity of the perturbation (the lower the p, the higher intensity). Each critical period features a bar of intense colour flanked by a gradient to the background colour, showing that there is a focused period of time (approximately 40 time-steps) during which the perturbation has a noticeable effect on the outcome of development, and that this effect is reduced as the onset time of the perturbation is moved away from this period. The presence of these critical periods is consistent with the working hypothesis of this paper. A statistically significant correlation exists between the period of greater change of rate (as identified by the regions in which the rate of change is greater than the threshold) and the critical periods. The cross-correlation shows significant peaks, with a lag within the window of the perturbation. As shown by the fourth panel in each graph, the period of greater rate of change is clearly associated with marked changes of directions of some of the genes in the regulatory network (e.g., gene 'sp BW' in the 13-

[1] These confidence intervals assumed uncorrelated series. This assumption will be checked in the final manuscript, and a more accurate estimate will be obtained using a Monte Carlo significance test.

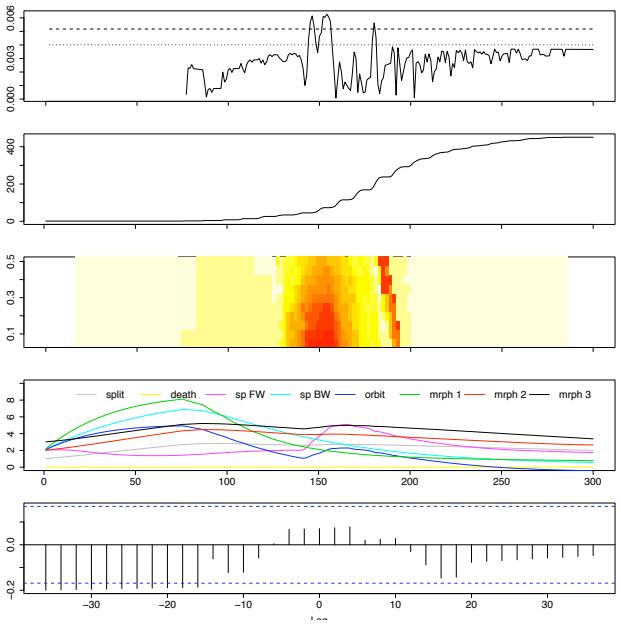

Figure 8: Results for the 14-gene organism. See text for details.

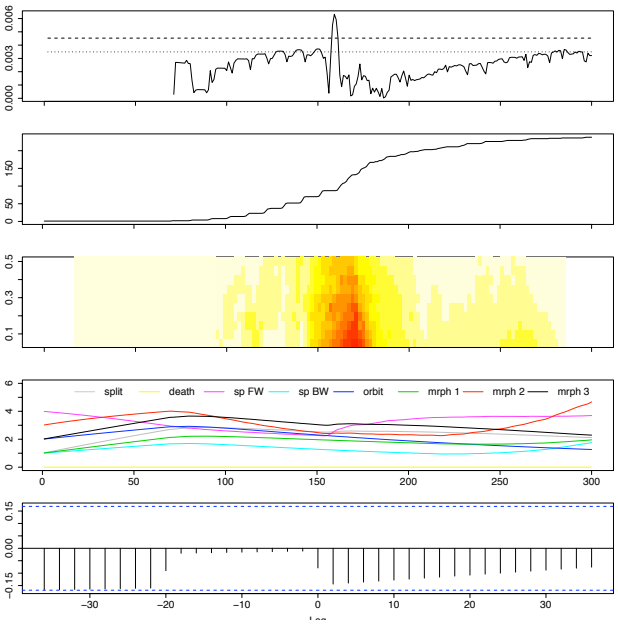

Figure 9: Results for the 15-gene organism. See text for details.

gene organism, 'sp FW' in the 14-gene and 15-gene organisms). These findings are consistent with recent findings in developmental biology that morphological variables are non monotonous in time (Cherdantsev and Tsvetkova, 2005).

Organisms 4 and 5 present a different picture. Neither of these genomes appear to have a critical period similar to that observed in the first three organisms. The absence of a critical period in organism 4 (Figure 10) is consistent with our working hypothesis. As shown by the top graph, there are no statistically meaningful peaks in neither rate of change nor cross correlation. Organism 5 exhibits a very irregular pattern of sensitivity to the perturbations early in development (between 30 and 150 time-steps). Interestingly, the cross-correlation reveals a statistically significant correlation between these periods of sensitivity and the developmental profile although none of these peaks in rate of change are significant unless a less stringent threshold of one standard deviation is considered.

Why did organisms 4 and 5 not show any critical period. A possible answer could be that of the genome size. Since evolution has more parameters to tune for these genomes, more generations might be needed to shape the developmental process. If this explanation holds, it may suggest that critical periods are not an inherent product of complex developmental systems, but rather the product of the evolutionary process on these systems.

3. Conclusion

As an initial step in the study of critical periods, the model described in this paper was deliberately simple and therefore suffers from many limitations. Firstly, the model is completely deterministic. This limitation will be addressed by adding noise to gene activity in the genetic regulatory network and in doing so, increased statistical power will be obtained. Secondly, the size of the perturbation window may have been too large in this study. By reducing the window size it may be possible to locate critical periods more precisely than we have in this paper (again, increasing statistical power). There are several questions that needs answers: (a) will different types of perturbations trigger different critical periods, with a different timing? Will critical periods become more defined and severe with more evolution? A simple experiment for this would be to evolve one population in increments of say, 100 generations, taking the fittest individual from the population and investigating the critical periods in its development. Finally, the study must be broadened to other developmental processes.

Nevertheless, this paper has achieved both of its aims by firstly showing that developmental models, similar to those in the artificial life and adaptive systems literature, do indeed exhibit critical periods. Secondly it has provided evidence of a correlation between critical periods and developmental rate of change. As such, it is a first step toward the goal of a general method for predicting critical periods in developmental systems. The ability to predict the timing of critical periods would have broad impli-

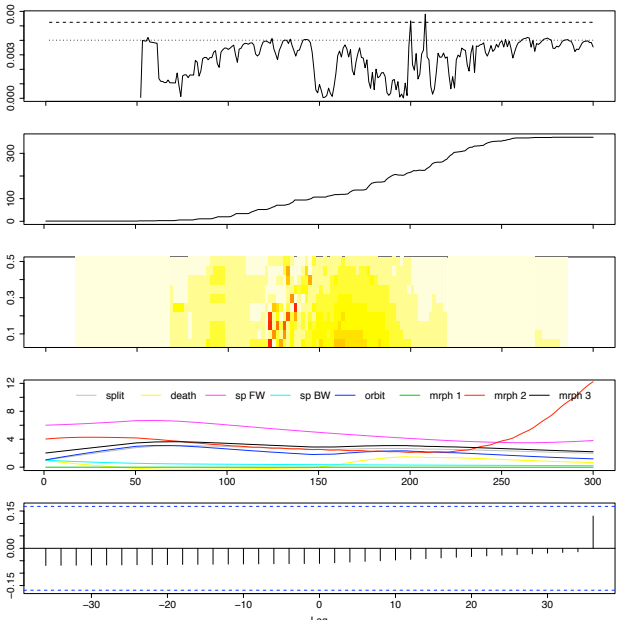

Figure 10: Results for the 16-gene organism. See text for details.

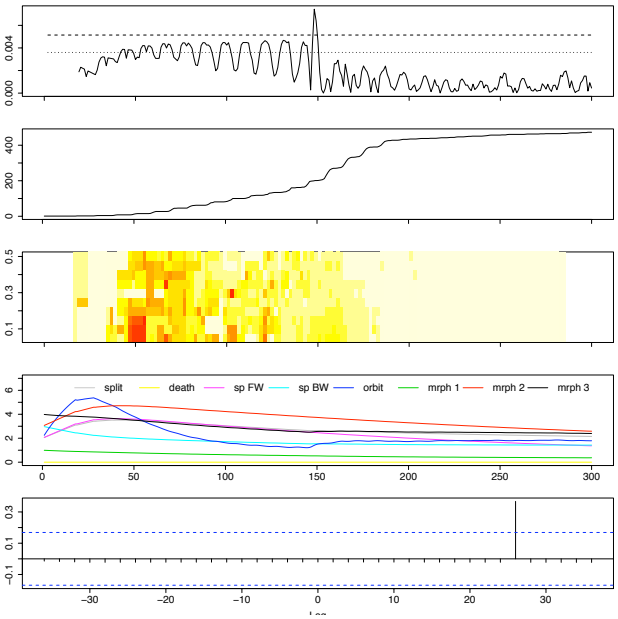

Figure 11: Results for the 17-gene organism. See text for details.

cations not only in the clinical domain (in particular, the study of teratogens) but also in the study of artificial developmental and adaptive systems. For example, an intriguing question is whether the controversial idea of critical periods in the acquisition of language could be framed in terms of periods of greatest rates of change in a model of language development.

Acknowledgements

The authors are grateful to Inman Harvey for useful discussions.

References

Anderson, L. M., Diwan, B. A., Fear, N. T., and Roman, E. (2000). Critical windows of exposure for children's health: Cancer in human epidemiological studies and neoplasms in experimental animal models. *Environmental Health Perspectives Supplements*, 108.

Bongard, J. C. and Pfeifer, R. (2001). Repeated structure and dissociation of genotypic and phenotypic complexity in artificial ontogeny. In *The Genetic and Evolutionary Computation Conference*, pages 829–836.

Bongard, J. C. and Pfeifer, R. (2003). Evolving complete agents using artificial ontogeny. *Morpho-functional Machines: The New Species (Designing Embodied Intelligence)*, pages 237–258.

Cherdantsev, V. G. and Tsvetkova, N. V. (2005). Dynamics and variability of early morphogenesis in the loach according to observations of individual developmental trajectories. *Russian Journal of Developmental Biology*, 36(3):171–180.

Eggenberger, P. (1997). Evolving morphologies of simulated 3d organisms based on differential gene expression. In Husbands, P. and Harvey, I., (Eds.), *Fourth European Conference on Artificial Life*, pages 205–213.

Harvey, I. (1996). The microbial genetic algorithm. http://www.cogs.susx.ac.uk/users/inmanh/.

Moorman, S. J., Cordova, R., and Davies, S. A. (2002). A critical period for functional vestibular development in zebrafish. *Developmental Dynamics*, 223:285–291.

Stewart, F., Taylor, T., and Konidaris, G. (2005). Metamorph: Experimenting with genetic regulatory networks for artificial development. In *Eighth European Conference on Artificial Life*.

Waddington, C. H. (1943). Organisers and genes. *American Midland Naturalist*, 30:811–812.

Wilson, J. G. (1973). *Environment and Birth Defects (Environmental Science Series)*. London: Academic Press.

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