Int. J. Appl. Math. Comput. Sci., 2004, Vol. 14, No. 3, 351–361

EVOLVING MORPHOGENETIC FIELDS IN THE ZEBRA SKIN PATTERN BASED ON TURING'S MORPHOGEN HYPOTHESIS

CARLOS PERALES GRAVÁN, RAFAEL LAHOZ-BELTRA*

* Department of Applied Mathematics, Faculty of Biological Sciences
Complutense University of Madrid, Madrid 28040, Spain
e-mail: lahozraf@bio.ucm.es

One of the classical problems of morphogenesis is to explain how patterns of different animals evolved resulting in a consolidated and stable pattern generation after generation. In this paper we simulated the evolution of two hypothetical morphogens, or proteins, that diffuse across a grid modeling the zebra skin pattern in an embryonic state, composed of pigmented and nonpigmented cells. The simulation experiments were carried out applying a genetic algorithm to the Young cellular automaton: a discrete version of the reaction-diffusion equations proposed by Turing in 1952. In the simulation experiments we searched for proper parameter values of two hypothetical proteins playing the role of activator and inhibitor morphogens. Our results show that on molecular and cellular levels recombination is the genetic mechanism that plays the key role in morphogen evolution, obtaining similar results in the presence or absence of mutation. However, spot patterns appear more often than stripe patterns on the simulated skin of zebras. Even when simulation results are consistent with the general picture of pattern modeling and simulation based on the Turing reaction-diffusion, we conclude that the stripe pattern of zebras may be a result of other biological features (i.e., genetic interactions, the Kipling hypothesis) not included in the present model.

Keywords: mammalian coat pattern, morphogenetic field, Turing reaction-diffusion, evolving cellular automata, developmental models, modeling biological structures

1. Introduction

One of the classical problems of morphogenesis, or the development of complex forms and patterns found in living organisms (Prusinkiewicz, 1993), is to explain how patterns of different animals (i.e., mammals, seashells, marine fishes) evolved resulting in a consolidated and stable pattern generation after generation. In 1952 Turing published a paper showing how patterns might grow from an initially nearly homogeneous state and how diffusion could drive to instability. A result of such instability is the emergence of patterns as a consequence of the breakdown of symmetry and homogeneity. Turing proposed that the temporal variation of the concentrations of two different chemicals, named by Turing as morphogens (the activator morphogen M_A and the inhibitor morphogen M_I), both diffusible but at different rates (D_A and D_I), can create patterns (D_A and D_I nonzero) on an initially homogeneous tissue by reacting in accordance with the nonlinear functions f and g:

$$\frac{\partial M_A}{\partial t} = f(M_A, M_I) + D_A \nabla^2 M_A,$$

$$\frac{\partial M_I}{\partial t} = g(M_A, M_I) + D_I \nabla^2 M_I.$$

In Turing's model, the activator morphogen M_A activates the production of itself and the production of the inhibitor M_I, whereas M_I inhibits the production of itself and decreases the activator M_A production. Twenty years after Turing's contribution, Gierer and Meinhardt (1972) developed a model of pattern formation based on a short-range activator and a long-range inhibitor which would promote the future development of simulation models by means of cellular automata. Cellular automata are discrete space and time models that have been used to model biological systems as a counterpart method to differential equations (Lahoz-Beltra, 1997; 1998). A cellular automaton in two dimensions consists of a regular grid of cells, and each of them can be in one of a finite number of possible states, being updated synchronously in discrete time steps according to a local, identical interaction rule (Toffoli and Margolus, 1987; Wolfram, 1984b). The state of each cell at the next time step $t+1$ is determined by its present state at time t and the states of its surrounding neighbors. Depending on the pattern of initial cell states and transition rules, thus on how neighboring states influence the state of a particular cell, patterns of cell states in the checkerboard evolve over time and can propagate, interact, store and compute information (Wolfram, 1983;

1984a). Based on such an approach, Young (1984) simulated Turing's reaction-diffusion model considering that cells lay out on a grid with two states, pigmented or nonpigmented, assuming that the pigmented cells produce activator M_A and inhibitor M_I morphogens diffusing at different rates across the grid. The results obtained by Young were similar to those obtained with continuous reaction-diffusion equations. Recently, the morphogens of vertebrate and invertebrate animals have been identified as proteins which exhibit diffusion and organize protein gradients (The and Perrimon, 2000). This is an important fact if we consider that computer simulation experiments using genetic algorithms (Goldberg, 1989), thus search algorithms based on the Darwinian natural selection, have demonstrated to be useful helping to find new clues and insights about the molecular evolution of proteins (i.e., enzymes (Lahoz-Beltra, 2001; Lahoz-Beltra *et al.*, 2002)). In the aforementioned papers we explored how electronic circuits modeling the catalytic function performed by enzymes and called 'electronic enzymes' evolved, leading to a metabolic ring similar to those present in organisms.

In this paper, we simulate the evolution of two hypothetical morphogens, or proteins, that diffuse across a grid. The grid models the zebra skin in an embryonic state, composed of pigmented and nonpigmented cells. The simulation experiments were carried out applying a genetic algorithm to Young cellular automata searching for proper values of the diffusion distance and the field value: the two main features that define a morphogen like activator or inhibitor in Young's model. In the biological realm, it is generally believed that zebras are dark animals, the white stripes being the areas where the pigmentation is inhibited. In agreement with Kipling (1908) and considering that predators such as lions cannot see colors, the zebra stripes may serve as camouflage. Such camouflage is an adaptation that prevents zebras from being seen by predators, confusing the zebras with most natural backgrounds. Murray (1981) showed that the chevrons at the base of zebra limbs result from the overlapping of two reaction-diffusion systems. Nevertheless, at present it is unknown how the zebra skin pattern is generated even when Turing's reaction-diffusion model could produce their characteristic stripe pattern. Our results show that on molecular and cellular levels and under the assumption of the absence of predators, recombination is the genetic mechanism that plays the key role in morphogen evolution. However, spot patterns appear more often than stripe patterns on the simulated skin of zebras. These results could support the view that on the population level Kipling's hypothesis or another alternative hypothesis could explain why zebras have stripe patterns. We propose that other biological features (i.e., genetic interactions, Kipling's hypothesis) different from those included in the present model (i.e., pigmented and nonpigmented cells, two morphogens with different diffusion rates) would promote those zebras with the stripe pattern against the zebras with spot patterns.

2. Model Description

2.1. Cellular Automata Approach to Turing's Reaction–Diffusion Model

The model proposed by Young (1984) assumes the animal skin in an embryonic state, formed by a uniform distribution of melanocytes or differentiated pigmented cells (black, state 1) and undifferentiated cells (white, state 0). Melanocytes produce the activator morphogen M_A which stimulates the transition from state 0 to 1 of nearby undifferentiated cells, as well as the inhibitor morphogen M_I promoting the opposite transition, thus from state 1 to 0, for nearby differentiated cells. When considered together, both morphogens define a morphogenetic field which is assumed to be circular and composed of two concentric rings (Fig. 1). The inner annulus is the activation region with radius R_1 and a large constant positive field value ω_1. The outer annulus is the inhibition area with radius R_2 and a constant small negative field value ω_2. As a consequence, the morphogenetic field results from a short-range activation and long-range inhibition areas defined around a melanocyte cell located at the origin or center of the morphogenetic field (Fig. 1). Once the pigmented cells are randomly distributed on a grid, the transition rules are applied to the pigmented and nonpigmented cells. In agreement with the model introduced by Young, each cell R on the grid in position (i, j) receives influences of the morphogens produced by all pigmented cells

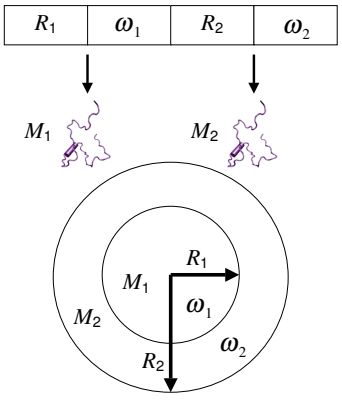

Fig. 1. Chromosome with genes modeling the diffusion distance R and the morphogenetic field ω values of two different morphogen molecules, or proteins, labeled as M_1 and M_2. The illustration assumes that in the morphogenetic field (obtained by decoding the chromosome) M_1 is the short-range activator and M_2 the long-range inhibitor.

R_p of the neighborhood in positions (i', j'). Thus, let R^* be the Euclidean distance between R and R_p:

$$R^* = |R - R_p| = \sqrt{(i - i')^2 + (j - j')^2}.$$

If the Euclidean distance R^* is less than or equal to radius R_1 ($R^* \leq R_1$), then the cell R located in position (i, j) receives the effect of the activator morphogen M_A, which is simulated by a positive field value ω_1. Otherwise, if R^* is greater than R_1 and less than or equal to R_2 ($R_1 < R^* \leq R_2$) then the cell R in position (i, j) would receive the effect of the inhibitor morphogen M_I, which is given by a negative field value ω_2. Finally, if for the cell R in position (i, j) we consider the composition of the effects of the morphogens produced by all nearby pigmented cells R_p in the neighborhood, then the future state of the cell R will be given by the sum of the field values. According to the model, the automaton transition rules for cells are given by the following rules:

Rule 1: If $\sum_p \omega(|R - R_p|) > 0$, then the state of cell R at time $t + 1$ is pigmented (state 1).

Rule 2: If $\sum_p \omega(|R - R_p|) = 0$, then the state of cell R at time $t + 1$ does not change and is equal to its state at time t (state 0 or 1).

Rule 3: If $\sum_p \omega(|R - R_p|) < 0$, then the state of cell R at time $t + 1$ is not pigmented (state 0).

2.2. Evolving Morphogen Features with a Genetic Algorithm

In this model, we define a population of chromosomes (Fig. 1) simulated as strings of real values. In the genetic algorithm terminology, genes were defined by four real values modeling the diffusion distance R and the morphogenetic field value ω of two different morphogen molecules, or proteins, labeled as M_1 and M_2. At each gene, and from left to right, the first gene position represents the diffusion distance R_1 and the second gene the field value ω_1 of the first morphogen molecule M_1, whereas the third gene represents the diffusion distance R_2 and the fourth gene the field value ω_2 of the second morphogen molecule M_2. Note that at the beginning of the simulation the two molecules are not defined as activator or inhibitor, both being candidates to be one or another type of morphogen during evolution. The current genetic algorithm uses one-point recombination and a population size of 60, testing recombination probability, as well as mutation probability values in different simulation experiments. Starting with a random population of chromosomes, reproduction, recombination and mutation were simulated, thus obtaining new generations of equal sizes. The initial population of chromosomes was obtained choosing R_1 and R_2, as well as ω_1 and ω_2, from a uniform distribution with $0 \leq R_1,\ R_2 \leq 10$ and $-5 \leq \omega_1,\ \omega_2 \leq 5$.

2.2.1. Reproduction

At each generation, the fitness f of each chromosome, thus the degree of the achievement of the M_1 and M_2 molecules in the morphogenic field,

$$f = \begin{cases} f_{spot} & \text{if } f_{spot} > f_{stripes}, \\ f_{stripes} & \text{if } f_{spot} < f_{stripes}, \end{cases}$$

was evaluated choosing the highest value of the following functions:

$$f_{spot} = C - \big(|\alpha_{ww} - N_{ww}| + |\alpha_{wb} - N_{wb}| + |\alpha_{bb} - N_{bb}|\big),$$

$$f_{stripes} = C - \big(|\beta_{ww} - N_{ww}| + |\beta_{wb} - N_{wb}| + |\beta_{bb} - N_{bb}|\big),$$

where C is 2500, thus the total number of contacts among cells. In the above functions, N_{ww}, N_{wb} and N_{bb} are the numbers of contacts between white-white, white-black and black-black cells, respectively, and α and β are the parameters of the fitness functions whose values were set up as follows: $\alpha_{ww} = 1430$, $\alpha_{wb} = 350$, $\alpha_{bb} = 720$, $\beta_{ww} = 720$, $\beta_{wb} = 350$ and $\beta_{bb} = 1430$. The functions f_{spot} and $f_{stripes}$ were proposed assuming that in the search space spot and stripe patterns are both stable solutions. The main reason to justify such an assumption is that both types of patterns are widely spread in animals, and other motives found in nature are spot and stripe variations. The parameters α and β were obtained experimentally based on Young's cellular automata model. In agreement with Young (1984), simulation experiments were performed considering the values reported by Young: the activator morphogen M_A was a molecule with $R = 2.30$, $\omega = 1$, and the inhibitor morphogen M_I a molecule with $R = 6.01$. Based on these values we generated 300 spot patterns ($\omega = -0.34$ for the inhibitor) and 300 stripe patterns ($\omega = -0.20$ for the inhibitor) obtaining the average values of N_{ww}, N_{wb} and N_{bb} in each kind of pattern. The mean values of N_{ww}, N_{wb} and N_{bb} were labeled as α_{ww}, α_{wb} and α_{bb} in the spot patterns and β_{ww}, β_{wb}, β_{bb} in the stripe patterns, respectively.

In order to obtain the value of f, the diffusion distances as well as the field values of the M_1 and M_2 morphogens were decoded from chromosomes. Thus, the simulation of Young's cellular automaton model was carried out based on R_1, R_2 and ω_1, ω_2 values coded by each chromosome. The simulation begins assuming 95% of white cells (undifferentiated cells) and 5% of black cells (melanocytes or differentiated pigmented cells) randomly distributed on a rectangular lattice, updating the state of the cells with transition rules until the resulting pattern no longer changes. With proper values of R_1,

R_2 and ω_1, ω_2, Young found that five iterations were enough for convergence to a stable pattern and that the final pattern is not sensitive to the initial distribution of melanocytes or black cells. However, in preliminary experiments we found unstable patterns as a consequence of the fact that we simulated how the Darwinian natural selection was able to find the proper activator and inhibitor diffusion distances and field values during evolution. In order to find the most suitable M_1 and M_2 molecules, we applied the following criterion: Once the cellular automaton evolves during five iterations, we test the stability of the final pattern, comparing the obtained pattern in the fifth iteration with the pattern obtained in the next iteration, thus in the sixth one. The reason to compare the fifth iteration only with the next one is that in preliminary experiments where we used more iterations we found similar results to those obtained with the present criterion. If during such an additional iteration the pattern is not stable, then there will be a change in the number of different kinds of cellular contacts (white-white, white-black and black-black) and, consequently, in the value of the fitness function f. In such a case we introduced the assumption that a pattern is unstable, and its fitness is equal to 0 if the difference N_{diff} of the total number of contacts among cells is above a threshold value ($N_{diff} > \theta$):

$$N_{diff} = |N_{ww}^{t+1} - N_{ww}^{t}| + |N_{wb}^{t+1} - N_{wb}^{t}| + |N_{bb}^{t+1} - N_{bb}^{t}|$$

being the iteration time t equal to 5. The present simulation experiments were carried out with a threshold value θ equal to 100. Otherwise, if $N_{diff} \leq \theta$, then the pattern is stable and $f > 0$. Note that the number of unstable patterns that pass the test increases as the threshold value θ does. In order to study the effect of the threshold, 250, 500 and 2500 values were studied in a different set of simulation experiments. Once the chromosomes are evaluated, we select the mating pool of the next generation using the roulette wheel algorithm of parents selection (Davis, 1991). This is a method for implementing reproduction, and thereby the Darwinian natural selection, by spinning a roulette wheel that assigns to each chromosome a slot whose arc size is proportional to its fitness value. Of course, other selection schemes are possible such as tournament selection, truncation selection, as well as linear and exponential ranking selection (Blickle and Thiele, 1995); however, the roulette wheel parents selection scheme bears a better resemblance to the Darwinian natural selection (Lahoz-Beltra, 2001).

2.2.2. Recombination

Once a new generation of offspring chromosomes is obtained, a single point crossover proceeds with pairs of mates randomly selected. Whether or not we are going to perform crossover on a current pair of parent chromosomes is decided on the basis of a Bernoulli trial regarding recombination as having a given probability (recombination probability). In the present model we assume that genes which define the features of the activator and inhibitor morphogens are linked and inherited together. Thus, and from left to right, the first and second genes define a first segment in the chromosome codifying the M_1 features and the third and fourth genes define a second chromosomic segment in which the M_2 features are codified. Therefore, since the first gene is linked with the second one and the third gene is linked with the fourth one, the crossover point is not randomly selected from a uniform distribution as is usual in a genetic algorithm. As a consequence, the crossover point is equal to 2 in the present simulations, just in the middle position of the chromosome. Finally, a single point crossover occurs when the segments of the two parent chromosomes i, j ($R_1^i, \omega_1^i, R_2^i, \omega_2^i$ and $R_1^j, \omega_1^j, R_2^j, \omega_2^j$) are swapped ($R_1^i, \omega_1^i, R_2^j, \omega_2^j$ and $R_1^j, \omega_1^j, R_2^i, \omega_2^i$).

2.2.3. Mutation

Mutation at a gene was simulated changing at random the value gene, choosing the mutated values of R or ω from a uniform distribution with a similar range to those defined to obtain the initial population of chromosomes ($0 \leq R \leq 10$ and $-5 \leq \omega \leq 5$). Once again whether or not to change a gene value on a chromosome is decided on the basis of a Bernoulli trial, mutation being a success with a given probability (mutation probability).

3. Genetic Algorithm Protocols

The goal of the following experiments was to find how a different arrangement of recombination and mutation operators performed in different protocols. Experiments were carried out applying recombination (single point crossover) and mutation to the chromosomes, setting the recombination and mutation probabilities to 0.25, 0.50, 0.75 and 0.05, 0.30, respectively. We tried out four protocols (Fig. 2) which were called R (only recombination), SGA (simple genetic algorithm), SDS (simulated DNA shuffling) and Wallace (a name arbitrarily chosen). The first protocol, R (Fig. 2(a)), evaluates a population of chromosomes before reproduction and once a new generation is obtained based on a single cycle of recombination. The protocol SGA (Fig. 2(b)) is a conventional genetic algorithm (Goldberg, 1989) which evaluates a population of chromosomes before reproduction and once a new generation is obtained based on a single cycle of recombination and mutation. The protocol SDS (Fig. 2(c)) was inspired by protein in vitro evolution experiments (Lahoz-Beltra, 2001; Lahoz-Beltra *et al.*, 2002). The protocol

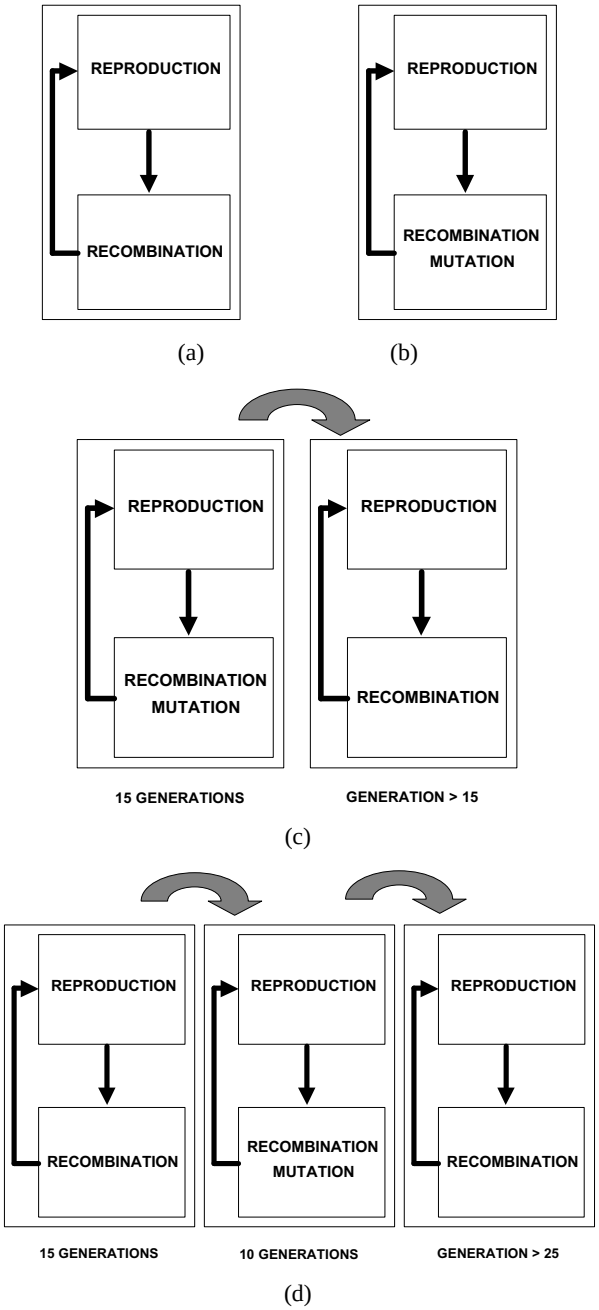

Fig. 2. Genetic algorithm protocols: (a) R; (b) SGA; (c) SDS and (d) Wallace.

involves a cycle of mutation and recombination through 15 generations as emulation of error-prone PCR or random nucleotide insertion and DNA reassembly by homologous recombination followed by repeated cycles of recombination, one per generation, in the absence of mutation. Finally, Wallace (Lahoz-Beltra, 2001) was defined (Fig. 2(d)) as follows: A simulation experiment starts with the first cycle that includes only recombination for 15 initial generations. In the second cycle, mutation is introduced together with recombination until the 25-th generation. The experiment concludes in the third cycle with recombination in the absence of mutation.

All simulation experiments were carried out during 40 generations repeating each experiment twenty times including the stability test (Section 2.2.1) into the protocols for the obtained patterns. The results were compared with the second set of experiments, where the protocols were similar except that the stability test for the obtained patterns was not carried out before reproduction, thus during chromosomes evaluation. In consequence, a total number of 920 trials or simulation experiments were performed based on the model described in Section 2.

4. Simulation Results

Computer simulation experiments were carried out using the population size, the morphogen diffusion distances and morphogenetic field values described in Section 2.2 as well as the recombination and mutation probabilities referred to Section 3. In Fig. 3, we show the performance graph (average fitness per generation) for each of the experiments performed under different protocols. Indeed, all protocols drive the population of chromosomes to a uniform population of similar genotypes with a maximum average fitness close to 2500. In the R protocol as well as the SGA one, and regardless of the recombination and mutation probabilities, the number of white (Fig. 4(a)) and black (Fig. 4(b)) patterns (the white pattern constitutes almost 90% of the sum of black and white zebras) is greater than the number of zebras with the spot pattern (Fig. 4(c)) and this last one is greater than those with the stripe pattern (Fig. 4(e)). We also obtained in the R protocol a few zebras bearing a thin stripe pattern (Fig. 4(f)). However, when the stability test was not carried out before reproduction, the number of zebras with the spot pattern was higher than those with white and black patterns, being again the lowest number of zebras with the stripe pattern. In the SDS protocol with the mutation probability equal to 0.05, and regardless of the recombination probability, the results were similar to those obtained with the R and SGA protocols, obtaining a few zebras bearing a thin stripe pattern. A similar result was obtained in the SDS protocol when patterns were evolved without applying the stability test, but zebras with the thin stripe pattern were not obtained. Increasing the mutation probability to 0.30 no matter whether the stability test is applied or not, the SDS protocol drives the population of chromosomes to results similar to those obtained with the R and SGA protocols in the absence of the stability test. Finally, the Wallace protocol and setting up the mutation probability of 0.05 lead to results that were similar to those obtained with the R protocol, as well as the SGA and SDS protocols with a similar mutation probability. In the absence of the stability test and setting up a mutation probability of 0.05 or

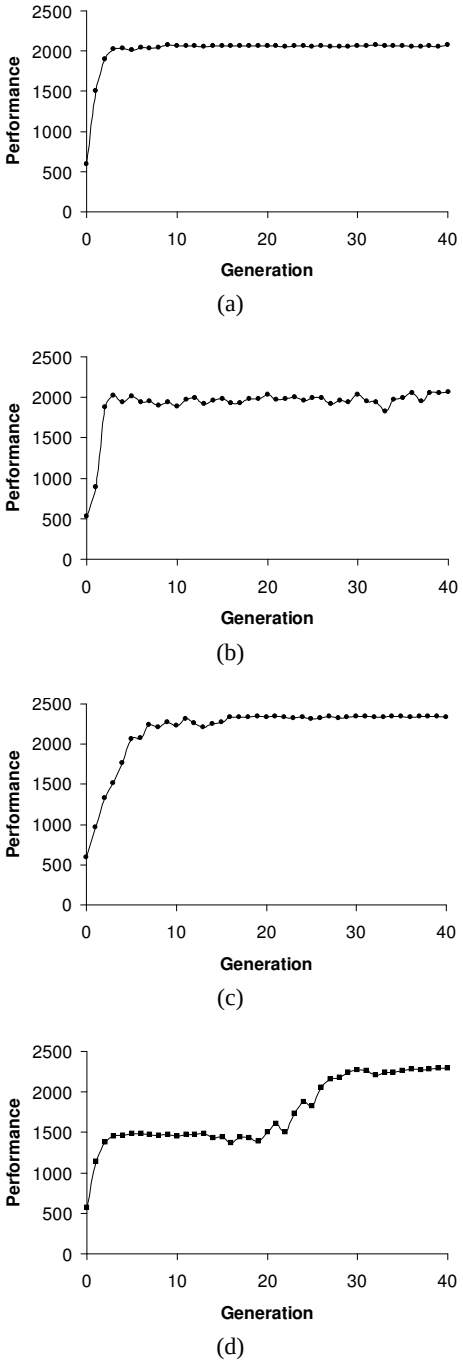

Fig. 3. Performance graph obtained under different protocols. (a) R; (b) SGA; (c) SDS; (d) Wallace.

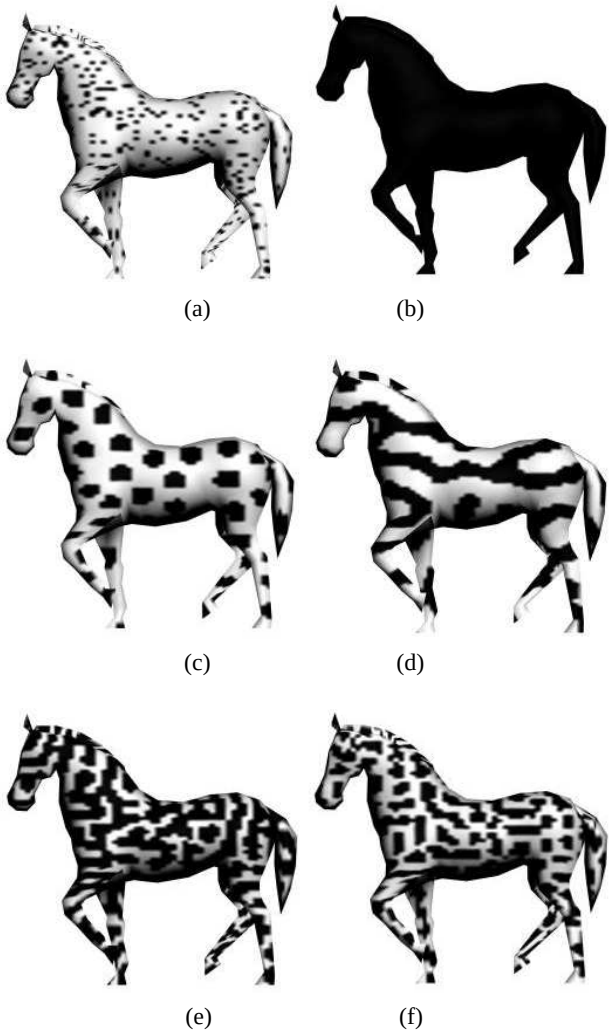

Fig. 4. Zebra skin patterns obtained after the evolution of the skin morphogenetic field: (a) white; (b) black; (c) spot pattern; (d) unstable stripe pattern; (e) stripe pattern; (f) thin stripe pattern.

0.30, the results obtained with the Wallace protocol were similar to those obtained with a mutation probability of 0.30 under the SDS protocol or the SGA protocol in the absence of the stability test. Therefore, our results indicate that the evolution of the morphogens leads to a higher number of white and black zebras, as well as zebras bearing the spot pattern, being always the lowest number of zebras with the stripe pattern.

In Figs. 5 and 6 we show two representative zebras bearing a spot pattern and stripe pattern, respectively. An example of the skin development of a zebra bearing an unstable thin stripe pattern is shown in Fig. 7. Such patterns were obtained with the activator-inhibitor model proposed by Young using the M_1 and M_2 features, thus their diffusion distances and field values. The morphogen features were obtained by decoding one of the chromosomes selected from a uniform population (all the chromosomes are the same) evolved under one genetic algorithm protocol with the maximum average fitness. Then, the morphogenetic field was updated during six iterations based on Young's transition rules, illustrating the changing states for each cell until the resulting pattern no longer changes. Therefore, our results indicate that the evolution of morphogens does not depend on the genetic algorithm

protocol and similar results are obtained in the presence or absence of mutation. Our findings support the view that recombination is a process that plays the key role in morphogen evolution. As a consequence, if mutation was present, morphogen evolution would not be sensitive to its presence during all generations (SGA protocol), during the initial generations (SDS protocol) or during the middle generations (Wallace protocol). In the biological realm, biologists have recently observed the diffusion of proteins in particular morphogens into the extracellular space (The and Perrimon, 2000) as well as how these molecules are organized into extracellular protein gradients (Sog, Wingless/Wnt, etc.), providing for the first time a confirmation of Turing's hypothesis. For instance, Srinivasan *et al.* (2002) found a protein gradient in developing fruit fly embryos, which is believed to trigger the division of the embryo into the nervous system and different types of epidermis within complex organisms like humans. In consequence, if morphogens are proteins, then such molecules could have evolved by the recombination of exons as has been suggested for the enzymes of eukaryotic organisms (Fersht, 1985; Price and Stevens, 1996).

5. Discussion

The simulation results are consistent with the general picture of pattern modeling and simulation based on Turing's reaction-diffusion scheme. In contrast to the stripe

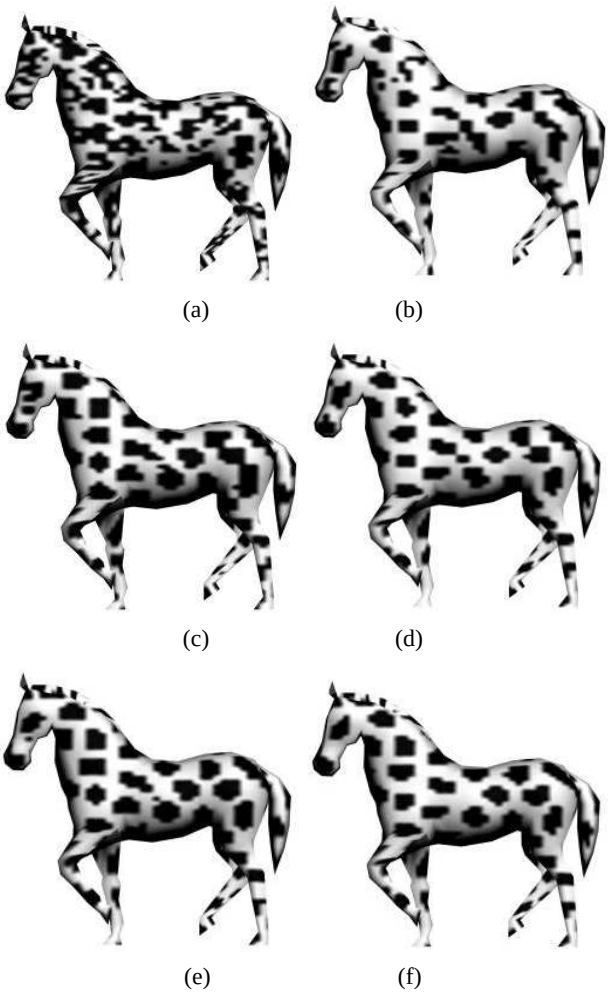

Fig. 5. Representative zebra showing a spot pattern obtained with the activator and inhibitor features obtained by decoding one of the chromosomes with maximum average fitness. The morphogenetic field was updated during six iterations based on Young's transition rules, illustrating the changing states for each cell until the resulting pattern no longer changes.

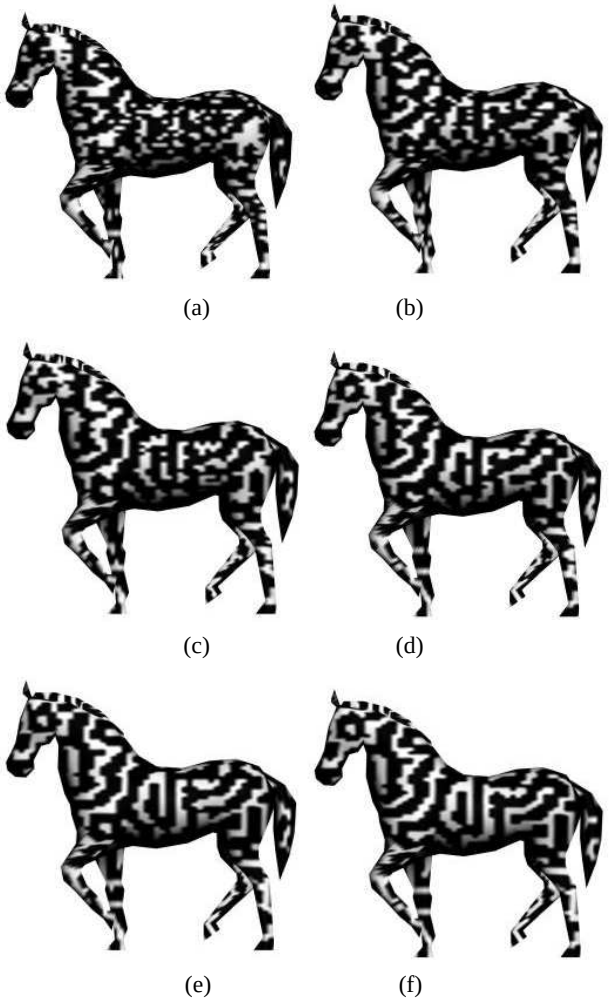

Fig. 6. Representative zebra showing a stripe pattern obtained with the activator and inhibitor features obtained by decoding one of the chromosomes with maximum average fitness. The morphogenetic field was updated during six iterations based on Young's transition rules, illustrating the changing states for each cell until the resulting pattern no longer changes.

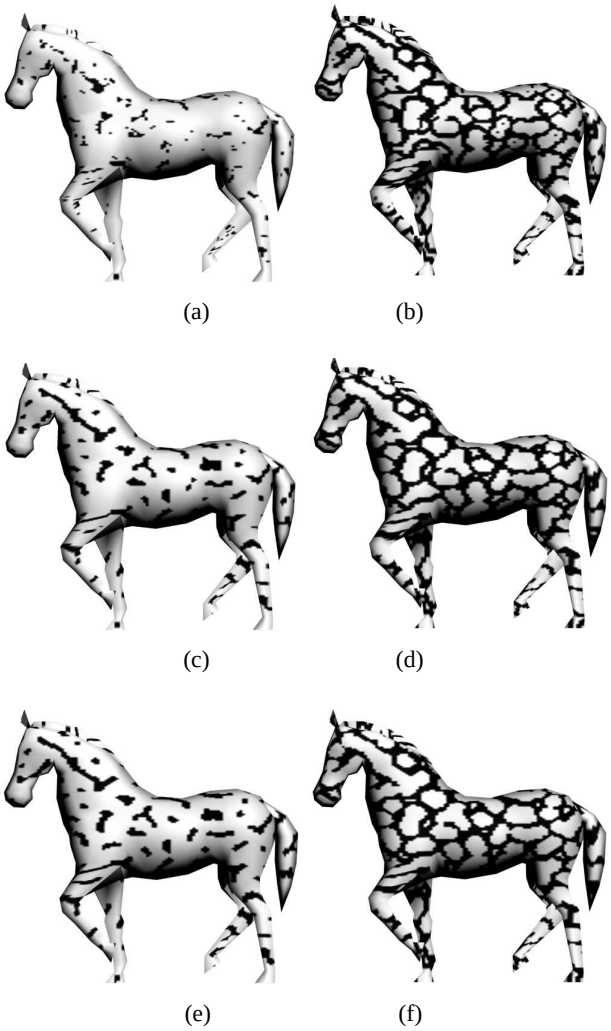

Fig. 7. Example of zebra showing an unstable thin stripe pattern which morphogenetic field was updated during six iterations.

pattern on fish skin which changes as the fish body size increases (Kondo and Asai, 1995), the case of zebras as well as other patterns of mammals (Murray, 1989) is easier to simulate. The reason for that is that once the pattern is formed in early developmental stages, the number of stripes does not change during their lifetime even if the body grows. The fact that the evolution of morphogens in the absence of predators (Kipling, 1908)—or other selection pressures not included in the proposed fitness function—leads to a higher number of zebras bearing the spot pattern than zebras with the stripe pattern is in agreement with previously published results. In general, striped patterns are more difficult to generate by reaction-diffusion models than spotted patterns, as has been observed by several authors (e.g., Murray, 1989). In two morphogen reaction-diffusion models stripe patterns tend to dissolve into spots, but they can be stabilized by adding further reactants, thus morphogens, or modifying the non-linearity of the system (Ball, 1999). Furthermore, the obtained pattern is related to the shape, in our model a squared lattice, where diffusion takes place (Varea *et al.*, 1999). It was also observed by Murray (1981) that if two reaction-diffusion systems that produce stripes meet together, then a chevron pattern emerges which is similar to the zebra pattern. In agreement with Painter *et al.* (1999), the applicability of classical Turing models to biological pattern formation—in our case the Young cellular automaton is an approximation to Turing's approach—is limited by the sensitivity of patterns to model parameters: in our case the proper values of R_1, R_2 and ω_1, ω_2 were found with a genetic algorithm. In two-dimensional models (i.e., the zebra pattern), the relative distance of the equilibrium level of the morphogen activator between two constraint terms which confine the variables within a finite range determines pattern selection (Shoji and Iwasa, 2003). Moreover, regular stripes are unlikely to arise reliably without tight parameter control (Painter *et al.*, 1999). In consequence, how did the zebra get its stripes? Bard (1977; 1981) proposed a mechanism for the production of zebra stripes surmising that while neural crest cells begin migration in the second week of gestation, zebra stripes are generated between the third and the fifth week depending upon the species. It is generally accepted that pattern formation during the early development is usually controlled by networks like genetic networks with complex genetic interactions or biochemical networks with coupled biochemical reactions (i.e., Drosophila embryo). In both cases a stochastic reaction-diffusion algorithm (Gillespie, 1976; 1977) seems to be the most suitable method to simulate pattern formation. This suggests that zebra stripe pattern formation should be simulated based on a cellular automaton approximation of a stochastic reaction-diffusion model, instead of Young's cellular automaton model used as an approximation of the classical Turing approach. In the population biology realm, Nelson and Shnerb (1998) applied a reaction-diffusion equation with a stochastic element to describe the development of a bacteria population. In the aforementioned model, when a relative small amplitude is introduced via the stochastic element, a very strong localization of the steady state solution occurs. Stochastic cellular automata models, thus cellular automata where the transition rules incorporate stochastic or probabilistic elements, have been widely used in the simulation of dynamical systems (for example, in plant ecology the model introduced by Inghe (1989) or the model of traffic flow developed by Schreckenberg *et al.* (1995)) and, in consequence, in reaction-difussion systems. For instance, Weimar (1997) introduced a cellular automaton with a look-up table of probabilistic rules for reaction-diffusion systems. Such a cellular automaton was constructed using the solutions of partial differential equations being

always more efficient than explicit numerical techniques and in many cases more efficient than better numerical techniques. Savill and Hogeweg (1997) developed a three-dimensional hybrid model, thus a model based on a cellular automaton with stochastic rules and partial differential equations, for the study of morphogenesis in simple cellular systems, such as the slime mold *Dictyostelium discoideum*. Based on a different approach, Rohlf and Bornholdt (2003) introduced a stochastic cellular automaton showing a model of pattern formation which is not based on the Turing instability. In the model, information is transmitted through soliton-like particles instead of a morphogen gradient, whose collective dynamics results in pattern formation. These models suggest for future work the interest to develop a stochastic cellular automaton version of Young's model, which could be applied as an alternative to stochastic reaction-diffusion equations in modeling the zebra stripe pattern.

Another possibility which was investigated by Kitano (1994; 1995) is to apply a genetic algorithm simulating the metabolism of cells, cell division and neurogenesis. In such a case, the rules which describe chemical reactions using differential equations, in our case the morphogens, were encoded into chromosomes. Nevertheless, the fact that pattern formation is ruled by a network (i.e., a genetic network) supports the application of a genetic algorithm to tune up the parameters of the model. For instance, Hamahashi and Kitano (1999) used a genetic algorithm to tune the parameters of a model describing the diffusion of proteins along the body of Drosophila. The aforementioned model was based on a reaction-diffusion system where pattern formation is regulated by maternal genes. However, in our model the application of a genetic algorithm to a cellular automaton raises some difficulties related to the fitness function. In the context of fine arts and in nature, the difficulties in applying genetic algorithms to cellular automata in pattern generation (i.e., the evaluation of stable patterns or the sort, synchronous or asynchronous, of cellular automaton updating) were studied in detail by Bentley (2002). In this study it is suggested that the asynchronous updating of cellular automata is more biologically plausible bearing a resemblance to human artistic updating of paintings.

One of the main results in our experiments is the key role of recombination as the main genetic mechanism during morphogen evolution. This result is in agreement with the results obtained by other authors in the field of computer simulation. For instance, Kerszberg (1996) developed a model of morphogen gradient in embryos assuming that evolution operated in part by shuffling DNA responsive elements (i.e., promoter sequences). In the context of genetic networks in developmental evolution (Gibson, 2002), the robustness or stability of such networks was studied simulating recombination among parameter solutions. Furthermore, in experiments evolving a cellular automaton, Werfel *et al.* (2000) found the advantage for crossover or recombination versus mutation alone, as well as other common features (i.e., identifiable building blocks that contribute to high-fitness solutions, metastable periods of fitness punctuated by rapid periods of innovation, etc.) which are applicable to other evolutionary systems based on cellular automata. In the biological realm, Xia and Levitt (2002) studied the effects of both mutation and recombination events on the evolution of protein stability. They found an opposite effect of both mechanisms which may be relevant to morphogens: protein sequences under mutation tend to be far from optimal, whereas under recombination the sequences tend towards the optimum.

At present cellular automata models are considered as an alternative to differential equations in the modeling and simulation of dynamical systems (for a detailed explanation, see Culik *et al.*, 1990; Lahoz-Beltra, 1998; Omohundro, 1984; Toffoli, 1984). In consequence, cellular automata can be used as an alternative to reaction-diffusion simulators which have been developed based on different reaction schemes. Such an approach could be used to develop a counterpart of simulators such as Ilya, which simulates a reaction-diffusion model based on the Brusselator scheme, or Xmorphia, which shows a beautiful simulation of the Gray-Scott reaction diffusion mechanism (Pearson, 1993).

The hybridation of genetic algorithms with cellular automata seems a promising field in biology, physics and computer science in the solution of theoretical and practical problems (Mitchell *et al.*, 1993; Corno *et al.*, 2000). In the biological realm, such a methodology could offer the possibility to simulate easily the evolution of morphogenetic fields, embryo development as well as the modeling and simulation of animal skin patterns.

Acknowledgments

We thank Jorge Perales Gravan for his support of the computer simulation experiments. We are grateful to Professor Alberto Perez de Vargas for his suggestions concerning the fitness function and to Chelo M. Mas for her helpful revision of the manuscript style. This work was supported by Laboratorio de Bioinformatica, Universidad Complutense.

References

Ball P. (1999): *The Self-Made Tapstry: Pattern Formation in Nature*. — Oxford: Oxford University Press.

Bard J.B.L. (1977): *A unity underlying the different zebra striping patterns*.- J. Zool. Vol. 183, pp. 527–539.

Bard J.B.L. (1981): *A model for generating aspects of zebra and other mammalian coat patterns*. — J. Theoret. Biol., Vol. 93, pp. 363–385.

Bentley K. (2002): *Evolving asynchronous adaptive systems for an exploration of aesthetic pattern formation*. — EASy M.Sc. thesis, Computer Science Department, University College of London.

Blickle T. and Thiele L. (1995): *A mathematical analysis of tournament selection*. — Proc. 6-th Int. Conf. *Genetic Algorithms*, Pittsburgh, USA, pp. 9–16.

Corno F., Sonza Reorda M. and Squillero G. (2000): *Evolving cellular automata for self-testing hardware*. — 3-rd Int. Conf. *Evolvable Systems: From Biology to Hardware (ICES2000)*, Edinburgh, UK, pp. 31–39.

Culik II K., Hurd L.P. and Yu S. (1990): *Computation theoretic aspects of cellular automata*. — Physica D, Vol. 45, pp. 357–378.

Davis L. (Ed.). (1991): *Handbook of Genetic Algorithms*. — New York: Van Nostrand Reinhold.

Fersht A. (1985): *Enzyme Structure and Mechanism*. — San Francisco, CA: W.H. Freeman.

Gibson G. (2002): *Developmental evolution: Getting robust about robustness*. — Current Biol., Vol. 12, pp. R347–R349.

Gierer A. and Meinhardt H. (1972): *A theory of biological pattern formation*. — Kybernetic, Vol. 12, pp. 30–39.

Gillespie D.T. (1976): *A general method for numerically simulating the stochastic time evolution of coupled chemical reactions*. — J. Comp. Phys., Vol. 22, pp. 403–434.

Gillespie D.T. (1977): *Exact stochastic simulation of coupled chemical reactions*. — J. Phys. Chem., Vol. 81, pp. 2340–2361.

Goldberg D.E. (1989): *Genetic Algorithms in Search, Optimization, and Machine Learning*. — Reading, MA: Addison-Wesley.

Hamahashi S. and Kitano H. (1999): *Parameter optimization in hierarchical structures*. — Proc. 5-th Europ. Conf. *Artificial Life (ECAL)*, Lausanne, Switzerland, pp. 461–471.

Inghe O. (1989): *Genet and ramet survivorship under different mortality regimes — A cellular automata model*. — J. Theor. Biol., Vol. 138, pp. 257–270.

Kerszberg M. (1996): *Accurate reading of morphogen concentrations by nuclear receptors: A formal model of complex transduction pathways*. — J. Theor. Biol., Vol. 183, pp. 95–104.

Kipling R. (1908): *Just So Stories*. — London: Macmillan.

Kitano H. (1994): *Evolution of metabolism for morphogenesis*, In: Artificial Life IV. Proc. Workshop *Artificial Life* (R. Brooks and P. Maes, Eds.). — Cambridge, MA: MIT Press, pp. 49–58.

Kitano H. (1995): *A simple model of neurogenesis and cell differentiation based on evolutionary large-scale chaos*. — Artificial Life, Vol. 2, pp. 79–99.

Kondo A. and Asai R. (1995): *A reaction-diffusion wave on the marine angelfish Pomacanthus*. — Nature, Vol. 376, pp. 765–768.

Lahoz-Beltra R. (1997): *Molecular automata assembly: principles and simulation of bacterial membrane construction*. — BioSyst., Vol. 44, pp. 209–229.

Lahoz-Beltra R. (1998): *Molecular automata modeling in structural biology*. — Adv. Struct. Biol., Vol. 5, pp. 85–101.

Lahoz-Beltra R. (2001): *Evolving hardware as model of enzyme evolution*. — BioSyst., Vol. 61, pp. 15–25.

Lahoz-Beltra R., Recio Rincón C. and Di Paola V. (2002): *Autómatas moleculares evolutivos: algoritmo SDS y sus aplicaciones*. — Actas del Primer Congreso Español de Algoritmos Evolutivos y Bioinspirados (AEB02), pp. 333–340.

Mitchell M., Hraber P.T. and Crutchfield J.T. (1993): *Revisiting the edge of chaos: Evolving cellular automata to perform computations*. — Complex Syst., Vol. 7, pp. 89–130.

Murray J.D. (1981): *A pre-patterns formation mechanism for animal coat markings*. — J. Theor. Biol., Vol. 88, pp. 161–199.

Murray J.D. (1989): *Mathematical Biology*. — New York: Springer.

Nelson D.R. and Shnerb N.M. (1998): *Non-Hermitian localization and population biology*. — Physical Review E, Vol. 58, pp. 1383–1403.

Omohundro S. (1984): *Modelling cellular automata with partial differential equations*. — Physica D, Vol. 10, pp. 128–134.

Painter K.J., Maini P.K. and Othmer H.G. (1999): *Stripe formation in juvenile Pomacanthus explained by a generalized Turing mechanism with chemotaxis*. — Proc. Nat. Acad. Sci. USA, Vol. 96, pp. 5549–5554.

Pearson J.E. (1993): *Complex patterns in a simple system*. — Science, Vol. 261, pp. 189–192.

Price N.E. and Stevens L. (1996): *Fundamentals of Enzymology*. — London: Oxford University Press.

Prusinkiewicz P. (1993): *Modeling and visualization of biological structures*. — Proc. Conf. *Graphics Interface'93*, Toronto, Canada, pp. 128–137.

Rohlf T. and Bornholdt S. (2003): *Self-organized pattern formation and noise-induced control from particle computation*. — Available at: http://www.arxiv.org/abs/cond-mat/0312366

Savill N.J. and Hogeweg P. (1997): *Modelling morphogenesis: From single cells to crawling slugs*. — J. Theor. Biol., Vol. 184, pp. 229–235.

Schreckenberg M., Schadschneider A., Nagel K. and Ito N. (1995): *Discrete stochastic models for traffic flow*. — Physical Rev. E, Vol. 51, pp. 2939.

Shoji H. and Iwasa Y. (2003): *Pattern selection and the direction of stripes in two-dimensional Turing systems for skin pattern formation of fishes*. — Forma, Vol. 18, pp. 3–18.

Srinivasan S., Rashka K.E. and Bier E. (2002): *Creation of a sog morphogen gradient in Drosophila embryo*. — Developmental Cell, Vol. 2, pp. 91–101.

The I. and Perrimon N. (2000): *Morphogen diffusion: The case of the Wingless protein*. — Nature Cell Biol., Vol. 2, pp. 79–82.

Toffoli T. (1984): *Cellular automata as an alternative to (rather than an approximation of) differential equations in modeling physics*. — Physica D, Vol. 10, pp. 117–127.

Toffoli T. and Margolus N. (1987): *Cellular Automata Machines*. — Cambridge, MA: The MIT Press.

Turing A. (1952): *The chemical basis of morphogenesis*. — Philos. Trans. Roy. Soc. London B, Vol. 237, pp. 37–52.

Varea A., Aragon J.L. and Barrui R.A. (1999): *Turing patterns on a sphere*. — Physical Review E, Vol. 60, No. 4, pp. 4588–4592.

Weimar J.R. (1997): *Cellular automata for reaction-diffusion systems*. — Parallel Comput., Vol. 23, pp. 1699–1715.

Werfel J., Mitchell M. Crutchfield J.P. (2000): *Resource sharing and coevolution in evolving cellular automata*. — IEEE Trans. Evolut. Comput., Vol. 4, pp. 388–393.

Wolfram S. (1983): *Statistical mechanics of cellular automata*. — Rev. Mod. Phys., Vol. 55, No. 3, pp. 601–644.

Wolfram S. (1984a): *Cellular automata as models of complexity*. — Nature, Vol. 311, pp. 419–424.

Wolfram S. (1984b): *Universality and complexity in cellular automata*. — Physica D, Vol. 10, pp. 1–35.

Xia Y. and Levitt M. (2002): *Roles of mutation and recombination in the evolution of protein thermodynamics*. — Proc. Nat. Acad. Sci., Vol. 99, pp. 10382–10387.

Young D.A. (1984): *A local activator-inhibitor model of vertebrate skin patterns*. — Math. Biosci., Vol. 72, pp. 51–58.