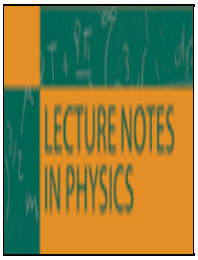
Biological Evolution and Statistical physics. (L assig M. & Valleriani A., eds.)

Chapter



Lecture Notes in Physics
Publisher: Springer Berlin / Heidelberg
ISSN: 1616-6361
Subject: Physics and Astronomy
Volume 585 / 2002
Chapter: p. 217

Multilevel processes in evolution and development: Computational models and biological insights

P. Hogeweg [A1]

[A1] Theoretical Biology and Bioinformatics Group, Utrecht University Padualaan 8, 3584CH Utrecht, the Netherlands. P.Hogeweg@bio.uu.nl

Abstract:

We argue that it is profitable (and necessary) to study biotic systems as multilevel systems in which the behavior of 'higher levels' is not only determined by the lower level processes but where the reverse is true as well: the higher level processes determine the structure of the lower level entities In this paper we will discuss two rather different examples of this phenomenon.

The first example discusses how the properties of (mutating) reproducing entities are, over evolutionary time, shaped by the dynamics of the larger scale spatial patterns which they generate. Also the dynamics of the evolutionary process, as represented in the shape of the phylogeny, is determined by these patterns. A novel type of 'punctuated evolution', i.e. periods of rapid change alternated by periods of stasis at the genotypic level is shown to be the result of unstable pattern formation. We will also show that the dynamics of the large scale spatial patterns, in its turn, shows features which result from the evolutionary dynamics of the micro entities.

The second example examines morphogenesis as multilevel process. Here, unlike the previous example, the basic model formulation incorporates several levels, and morphogenesis is the result of intricate coordination between these levels. The modeling of such coordination is achieved through modeling an evolutionary process. We will discuss both an example of an resulting morphogenetic process, and some properties of the evolutionary paths which leads to it. The re-occurrence of similar morphogenetic 'inn ovations' is one feature which our model evolutionary process shares with an ever increasing database of examples of biotic molecular phylogenies in which such re-inventions at the morphological level also seem to occur frequently.

Keywords:

PACS: 87.10.+e; 87.14.-g; 87.14.Cc; 87.14.Ee; 87.14.Gg; 87.15.-v; 87.15.Aa; 87.15.By; 87.15.Cc; 87.15.He; 87.15.Kg; 87.15.La; (7.15.Mi; 87.15.Nn; 87.15.Rn; 87.15.Tt; 87.15.Vv; 87.15.Ya; 87.16.-b; 87.16.Ac; 87.16.Dg; 87.16.Gj; 87.16.Ka; 87.16.Nn; 87.16.Qp; 87.16.Sr; 87.16.Tb; 87.16.Uv; 87.16.Xa; 87.16.Yc; 87.17.-d; 87.17.Aa; 87.17.Ee; 87.17.Jj; 87.17.Nn; 87.18.-h; 87.18.Bb; 87.18.Ed; 87.18.Hf; 81.18.La; 87.18.Pj; 87.18.Sn; 87.19.-j; 87.19,Bb; 87.19.Dd; 87.19.Ff; 87.19.Hh; 87.19.Jj; 87.19.La; 87.19.Nn; 87.19.Pp. 87.19.Rr; 87.19.St; 87.19.Tt; 87.19.Uv; 87.19.Xx; 87.23.-n; 87.23.Cc; 87.23.Ge; 87.23.Kg

1 Multilevel processes in evolution and development: computational models and biological insights

Paulien Hogeweg

Theoretical Biology an Bioinformatics Group, Utrecht University Padualaan 8, 3584CH Utrecht, the Netherlands.
Email P.Hogeweg@bio.uu.nl

Abstract. We argue that it is profitable (and necessary) to study biotic systems as multilevel systems in which the behavior of 'higher levels' is not only determined by the lower level processes but where the reverse is true as well: the higher level processes determine the structure of the lower level entities In this paper we will discuss two rather different examples of this phenomenon.

The first example discusses how the properties of (mutating) reproducing entities are, over evolutionary time, shaped by the dynamics of the larger scale spatial patterns which they generate. Also the dynamics of the evolutionary process, as represented in the shape of the phylogeny, is determined by these patterns. A novel type of 'punctuated evolution', i.e. periods of rapid change alternated by periods of stasis *at the genotypic level* is shown to be the result of unstable pattern formation. We will also show that the dynamics of the large scale spatial patterns, in its turn, shows features which result from the evolutionary dynamics of the micro entities.

The second example examines morphogenesis as multilevel process. Here, unlike the previous example, the basic model formulation incorporates several levels, and morphogenesis is the result of intricate coordination between these levels. The modeling of such coordination is achieved through modeling an evolutionary process. We will discuss both an example of an resulting morphogenetic process, and some properties of the evolutionary paths which leads to it. The re-occurrence of similar morphogenetic 'innovations' is one feature which our model evolutionary process shares with an ever increasing database of examples of biotic molecular phylogenies in which such re-inventions at the morphological level also seem to occur frequently.

1.1 Introduction

The extensive formal theory developed in population genetics has shaped our insight in the process of Darwinian evolution. Curiously most of these theoretical population genetics models fail to deal in a realistic manner with the process of mutation, they focus on selection only (e.g. [1]). More recently statistical physics approaches have been applied to biological evolution. Here the basic dynamics is formulated in microscopic equations, including mutations etc. and macroscopic properties, are derived.

"A theory should be as simple as possible, but no simpler" is an often quoted slogan of Einstein. In my opinion many of the evolutionary models, although yielding useful baseline insights, are in the 'too simple' category, because they ignore the multilevel nature of biotic systems. The aim of this paper is to develop models which are 'as simple as possible' but in which we can explore and exploit the interplay between levels in biotic systems Level refers here to a process at a particular space and time scale.

I will argue and demonstrate that 'clean' micro-macro transitions are not sufficient to describe biological evolution satisfactorily Instead I argue that because of the very process of evolution, a two-way information flow among levels occurs. In figure 1.1 of the next session this proposition is illustrated in a eco-evolutionary setting. Interacting, replicating entities living in space are prone to the generation large scale spatial temporal patterns. These macro-scale properties of the micro-scale interactions between the entities constitute the 'environment' in which these entities live. Thus, it is rather obvious that, when the entities are subjected to mutations and selection, the environment and hence the spatial temporal pattern generated by the entities determine the evolutionary fate of the mutants and ultimately the micro-scale interactions which define the system.

The aim of this paper is not only to argue that such two way interactions between levels do occur, but also to argue that it is a good research strategy to examine these two or multiple way processes explicitly: it allows us to make our models, and the analysis of the models, simpler.

An important reason for the usefulness is that the macro-scale spatial-temporal dynamics may be relatively well understood even when it is beyond present day analysis tools to derive the macro-scale dynamics from the micro-scale interactions. This is true for 'generic' patterns which occur for a large set of micro-scale

dynamics. In the example we discuss in section 1.2 spiral wave and turbulent patterns arise from a stochastic model for host-virus interactions. This is easily seen in movies of monte-carlo simulations of the system. Once the wave patterns are recognized, we can use their well studied properties to try to understand the evolutionary dynamics at the micro-scale level. We demonstrate that the evolutionary dynamics of the viruses can best be explained in terms of generic properties of spiral waves. On the other hand 'not so generic properties' of the spiral waves in the studied system, can best be explained in terms of the mutational dynamics of the viruses.

The entanglement of levels in the products of biological evolution, i.e. present day biotic systems, is mostly seen as simply an unwanted complication in both experimental and theoretical studies. Experimental and theoretical studies therefore try to isolate processes at one space/time scale, or to study simple one way micro-macro transitions. An example of the latter is the use of spatial pattern formation in e.g. Turing systems as theories of development. I think this approach is in the 'too simple category', as indeed was acknowledged by Turing himself when he allegedly said: "the stripes are simple, but what about the horse part?". In my opinion entanglement of levels is such a basic feature of biotic systems, that we should address it explicitly in our models. I think it is the very process we want to understand about biotic systems. In the second part of this paper I will explain a modeling approach by which we can explore and exploit the entanglement of various processes at various space and time scales. By using an evolutionary process. and by taking into account the entanglement of levels our models (and the parameter space to be explored) can in fact become simpler/smaller because the levels constrain each other. We will apply this approach to study biological development We will show that by using a simple evolutionary process we can derive interesting simple models for morphogenesis, i.e. we can at least begin to address "the horse part". In these models the interplay between differential adhesion, cell signaling and cell growth, cell death and cell differentiation, is shaped by evolution to produce intricate morphologies.

In short, in this paper I discuss how entanglement between levels arise through Darwinian evolution, and how a Darwinian evolutionary process can be used as tool to study (evolved) complex entangled systems. Moreover in both contexts I will study the evolution process itself - and how it is shaped by the entanglement of levels. I will stress in both context the biological insights we have obtained. The insights, of course, are what validates the modeling approach!

1.2 Multilevel evolution: pattern formation, phylogenetic trees and punctuated equilibria

1.2.1 Dynamics of meso-scale patterns shape of micro-scale replicators through evolution: a review

In explicit spatial ecological models of interacting replicators, pattern formation will occur in many circumstances. Minimally local replication will result in clumping of lineages and larger scale clumping of species is also often observed. Moreover many ecological interactions, as modeled in classical ODE models, give rise to oscillatory population densities. Examples are Host-Parasitoid, Predator-Prey, Hypercycles, In-transient competition etc . In space such oscillations lead to waves - which may organize in e.g. spiral waves. The formation of such spatial patterns is crucial for the qualitative outcome of the interactions. For example coexistence of mutualistic species and so called 'cheaters', which take extra benefits and do not reciprocate, is impossible in the classical ODE models, but is possible in spatial models over a large range of parameters. Such coexistence is also observed in nature, and may persist over geological time scales.

Here we discuss how such patterns formed by ecological interactions of replicating entities, influence the evolutionary 'fate' of these replicators. Thus we consider a feedback of the larger scale patterns to the micro-scale entities which generate these patterns (fig.1.1).

We can best analyze this feedback in cases where the dynamics of the the meso-scale patterns is generic and well understood in systems other than the eco-evolutionary system under consideration. This is pre-eminently the case for spiral waves, and the patterns formed by the complex Ginsberg-Landau equation, in which regions of spirals and regions of turbulence occur. We have demonstrated the feedback of spiral wave on the evolutionary dynamics of the replicators which generate the waves, in a variety of cases and a variety of aspects. For example we have shown:

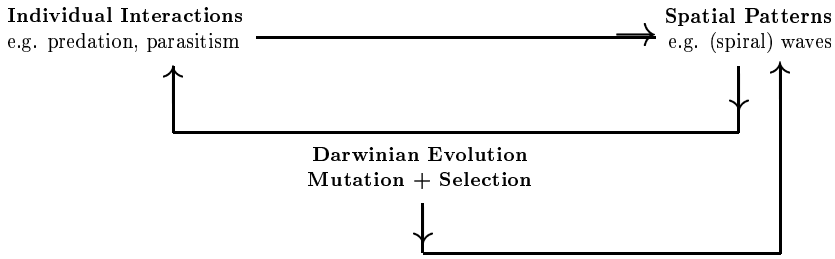


Fig. 1.1. Relation between local interactions and spatial pattern formation in eco-evolutionary models

- **Positive selection for fast decay**
  Boerlijst and Hogeweg [2] [3] have shown that very counter-intuitive selection pressures occur in an explicit spatial version of the classical ODE Hypercycle model of Eigen and Schuster [4]. The model is formulated as a stochastic cellular automata (CA) model. The states represent the presence of one molecule of a particular type at that position; state 0 represents an empty position. Surrounding molecules can replicate into an empty position with a probability depending on their type and the catalysis they get. Non-zero states have a probability to become zero, representing the decay of molecules. For Hypercycles of length 5 and longer stable spiral waves, mostly organized in pairs, form from initial random distribution of the molecules. It is well known that faster rotating spiral waves overtake the domain of slower rotating ones. Increase of decay rate increases the rotation speed and hence spirals with faster decaying molecules will overtake the domain of spirals consisting of species with greater longevity - a feature which is entirely incomprehensible from the standpoint of the individual.
- **Self-reinforcing spatial patterns**
  Savill et al. [5] have shown that selection pressures in an explicit spatial version of the classical Nicholson Bailey Host-Parasitoid model [1] are such that the spatial pattern in which the evolving parasitoids happen to find themselves are reinforced. The model is formulated as a Lattice Map model. At each Lattice site Host and Parasitoids interact as in the N-B model, i.e. the within site dynamics is an unstable spiral. The hosts diffuse randomly to neighboring sites, whereas the parasitoids have biased diffusion towards the Hosts: $\beta_{A->B} = cH_B/{H_{tot}}^{\mu}$, where H_{tot} is the total number of hosts surrounding the local parasitoid population in A , and $\beta_{A->B}$ the fraction thereof which will migrate to B. The parameter on which the mutation selection process operates is μ; The patterns which form in a large parameter region resemble those studied extensively in the Complex Ginsburg-Landau equation: regions of spiral waves and regions of turbulence occur in models where all hosts/parasitoids are identical. Selection pressure will be towards 'overexploitation (large μ) in turbulent regions, leading to μ values for which only turbulence occurs. Conversely selection pressure within the spiral regions is towards under-exploitation, but alternate between increasing and decreasing μ within a range for which spiral patterns are reinforced. Within a spiral selection is towards low μ, i.e. towards remaining in the spiral core (from which all offspring descend in the long run), whereas competition between spirals favor higher μ values, as this leads to faster rotation spirals. New spirals are formed at the border with turbulent regions and have high μ .
  From the standpoint of the parasitoids neither outcome seems 'optimal'. The outcome is determined by the spatial patterns and indeed reinforces them.
- **Long term information integration**
  Pagie and Hogeweg [6] [7] have shown that in explicit spatial models of co-evolving antagonistic populations long term evolution leads to entities which can cope with a full set of circumstances, although each generation experiences only a very small subset of these (i.e. experiences sparse fitness evaluation). We have even shown that sparse fitness evaluation leads to better performance of individual entities towards all circumstances than full fitness evaluation [7]. We studied this in stochastic CA models in which an external fitness criterion is given. One species represent instances of this fitness function (e.g. coordinates

[1] Host: $H_{t+1} = \lambda H_t \exp(-aP_t)$ Parasitoid: $P_{t+1} = bH_t(1 - \exp(-aP_t))$

of a function) and the other population should generate the value of the function at these coordinates. Information integration depends crucially on spatial pattern formation, although the patterns formed in these circumstances are not characterized (characterizable) independently, so that we cannot point at the dynamical features of the patterns which explain the effects seen at the level of the individuals, as in the cases mentioned above. However, as shown in [6] , mixing the populations after each step leads to so called red queen dynamics: no information integration occurs but the populations "have to run very fast to remain at the same place" i.e. the individuals only adapt to the current circumstances.

In this paper we extend these studies by looking at "tempo and mode" of an evolutionary process resulting from the interplay between evolution and pattern formation, rather than at selection pressures and the resulting the properties of the evolved individuals, as done above.

1.2.2 Interplay between pattern formation and the shape of phylogenetic trees

Reconstruction of phylogenetic trees on the basis of DNA sequence similarity is widely used to study the evolutionary history of species. Traditionally the main interest was "who was more closely related to whom". More recently there are "New uses for new phylogenies" [8], e.g. (1) assessment of population dynamics [9] from phylogenetic trees, using simple models of evolution (see [10] for an application to virus evolution) and (2) assessment of 'novelty' and 'reinvention' of morphological/behavioral features ([11] and many other examples in the same volume). In this paper we will use reconstructed phylogenetic trees of *simulated* evolutionary processes to investigate the phylogenetic signatures one might expect with respect to these latter uses. In this section we focus on the first type of issue, i.e. we study how the shape of phylogenetic trees is influenced by spatial patterns generated by the ecological interactions of the species under consideration.

We examine a virus host situation, in which the host becomes immune to viruses which have infected it. The study is inspired by data on influenza virus NP protein. Sequence data are available of human, pig and bird infecting strains from 1910 onwards. Phylogenetic reconstruction of the human and pig strains shows a very skewed, almost linear tree shape, with older strains at the bottom and younger strains at the top. Thus, although some strain diversification occurs, in the longer run it is only one of the strains whose offspring produce the next generation [12]. In contrast, the bird phylogeny is rake-shaped. The time of isolation and position in the tree are not related. The length of the tree from the common ancestor to present day strains is an order of magnitude less than that of the human and the pig trees [13]. It is well know that influenza epidemics sweep as waves over the world. We study a generic spatial model of host virus interaction, which is not particularly moulded to influenza. We want to study which general properties might lead to such different tree shapes.

A model of host virus interactions We model long lived hosts which become (temporarily) immune for the virus which has infected it. The virus is represented by a amino-acid sequence, which is subject to mutation. Specifically the assumptions of the model are:

- Immunity declines exponentially in time
  $I = m * exp(-z * t)$ $(m = 120; z = .04)$
- sensitivity: saturated function of protein distances
  $s = d/(K + d)$ $(K = 15; sizeprotein = 60)$
- probability of infection: Hill function of immunity and sensitivity
  $P = s^n/(I^n + s^n)$ $(n = 3)$ (see fig. 1.2)
- Decay probability of viral infectivity: (zet=.2 - .4)
  Host decay probability: (decay=.01 - .0025)
  Mutation rate (mut=.0001 - .0004)
  Field size (200x200 or 300x300 patches)

The populations are modeled in a multi-layer stochastic CA. State 0 represents empty in all layers; all other states represent a virus type or immune memory for a virus type or I, strength of immunity. I is coded as integer and exponential decay is simulated in terms of probability of unit decay of memory. Initial only 1 virus type is present.

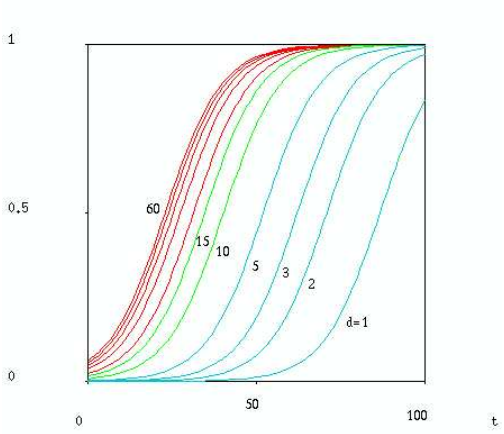


Fig. 1.2. Infectivity as a function of time after previous infection for indicated difference between current and previous infecting virus

Virus evolution and shape of phylogenetic trees Fig.1.3 shows the phylogenetic tree of an initial identical viral population evolved in two different hosts. The difference of the host is their longevity. In the long-lived host the phylogenetic pattern is similar to that reported for influenza in humans and pigs, whereas for a host which lives half as long phylogeny is rake-like, like observed of influenza in birds. Indeed figure 1.3 is strikingly similar to the similarly composed figure 6 in [13].

Fig. 1.3. Phylogeny of viruses evolved in two different hosts: the differences in evolutionary dynamics are caused by the presence/absence of patterns

The difference in shape of phylogeny coincides, and is indeed caused by (see next section), a difference in pattern formation. For the longer lived hosts, no stable spatial structures emerge. Infection spreads as target waves, which sometimes convert temporarily into spiral waves. New mutations can invade somewhat before the ancestral types and are therefore strongly selected. Temporarily several strains can coexist, in different regions, but regularly a new strain sweeps away all other existing strains, as none maintains a stable 'source'. For the shorter lived hosts, in contrast, pattern formation does occur, so that hosts are alternately infected by virus strains which have diverged significantly.

When the duration of infectivity is relatively long (e.g. $zet = .2$), two- or three-armed spirals develop of the different strains New mutants are not strongly selected because the other strains will always be more different from the parent strain than the mutant is. In spiral waves only mutants in the spiral core can replace their ancestor because in the long run all offspring originates from the core. However, mutants in the core tend to destabilize the core leading to slower rotation speeds (and therefore, in a slow process, to loss of spiral domain). However the de-stabilization will also increase selection on mutants of the other strains in that core, eventually leading to more divergence of the strains coexisting in the core, and thereby to larger rotation speeds. The result is that several strains per spiral domain and several different spiral domains, which alternately gain domain on the other, coexist. The strains are stably diverged (i.e. 'radiation' or 'speciation' has occurred) and few new mutants can establish themselves. The resulting phylogeny is 'rake-shaped'.

In the complex wave patterns which occur when infectivity is short ($zet = .4$), so that only a subset of the hosts are infected in an infection wave, a similar host longevity related difference in phylogeny occurs, although somewhat less strong. Also this case the organization of the waves is such that very few mutants are 'at the right moment at the right place' to be able to invade, and, if they are, their domain remains restricted.

Other parameter differences also lead to these two types of evolutionary dynamics. They even occur with identical parameter values, with as only difference the size of the 'world'. Putting the long-lived hosts and their viruses in a world of 300*300 patches, instead of the 200*200 in the simulation above, there is enough space to develop e.g. stable and multi-armed spirals, the strains can diverge, and the evolutionary change will stagnate, just as in the shorter lived ones described above. In the next section we see that it is indeed the occurrence of pattern formation which determines the evolutionary dynamics and the shape of the evolutionary tree.

Genotypic punctuated equilibria For intermediate parameter values the two evolutionary modes can alternate, as shown in fig. 1.4. This leads to periods of rapid evolutionary change of the viral strains alternated by periods of very little change but the concurrence of several significantly diverged strains. In the figure the lowest panel shows the (cumulative) number of mutations of the strains which are present in more than 200 individuals in the 'world', over time. The spatial patterns are shown over the shaded period spanning from a period of stasis, through a period of fast change to a period of stasis. Snapshots of the world, and a space-time plot of a horizontal section are shown. For clarity of presentation the type of the last infecting strain is shown: the thin viral waves are hard to see. The snapshot is taken at the last time of the corresponding space-time panel. The two-armed spiral with slowly changing strains persist to t=23000, i.e. just before the end of panel 2. The target wave patterns continue to ca. t=26000, when again two-armed spirals appear - now as coupled pair; they last to ca t=29000; after a short unstable period, with fast changing strains, one single two-armed spiral reappears which remains stable for a longer period.

Alternation of periods of fast and slow change have been observed at the phenotypic level in many evolutionary models, in long term evolutionary experiments and in the fossil record. It is known as 'punctuated equilibria' or 'epochal evolution'. The phenomenon is well studied in models with redundant genotypic coding, where it is a consequence of diffusion on of neutral networks and occasional shifts to other (fitter) neutral nets [14], [15] Thus the epochal evolution on the phenotypic level is accompanied with a constant rate of change at the genotypic level. In the fossil record punctuated equilibria may be caused by external environmental changes. However, neither of these scenario's is true in the present model. The epochal dynamics occurs here at the genotypic level as well as on the phenotypic level. It is caused by self induced 'environmental' conditions, i.e. spatial pattern formation, by which periods of 'radiation' and subsequent neutrality alternate with periods of strong selection.

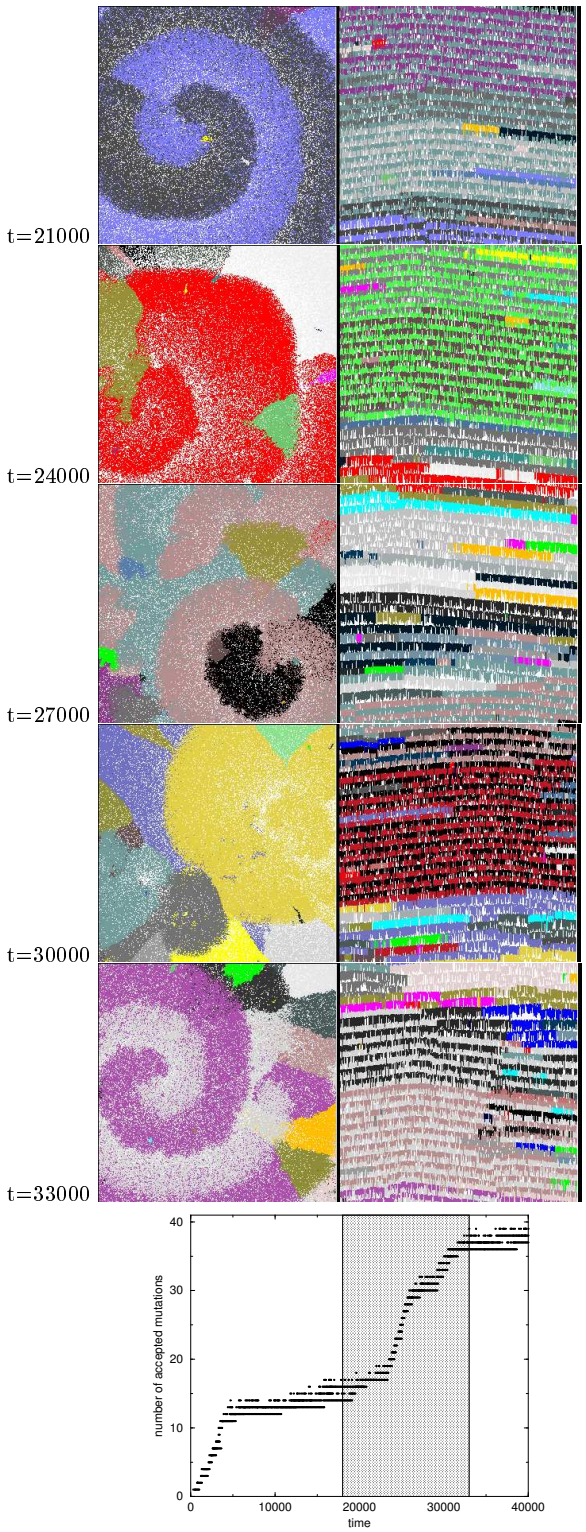


Fig. 1.4. Genotypic Punctuated equilibria due to unstable spatial pattern formation. The bottom panel shows the cumulative number of accepted point mutations in the dominant strains which coexist in the population over time: periods of rapid change and slow change alternate. The snapshots and space time plot above show that the periods of slow change correspond with periods in which the dominant species occur in stable two-armed spiral patterns, while during the period of fast change the spatio-temporal dynamics is chaotic. For more detail, see text

Mutation induced pattern dynamics The dynamics of the spatial patterns shows interesting features induced by the mutations as well. For example:

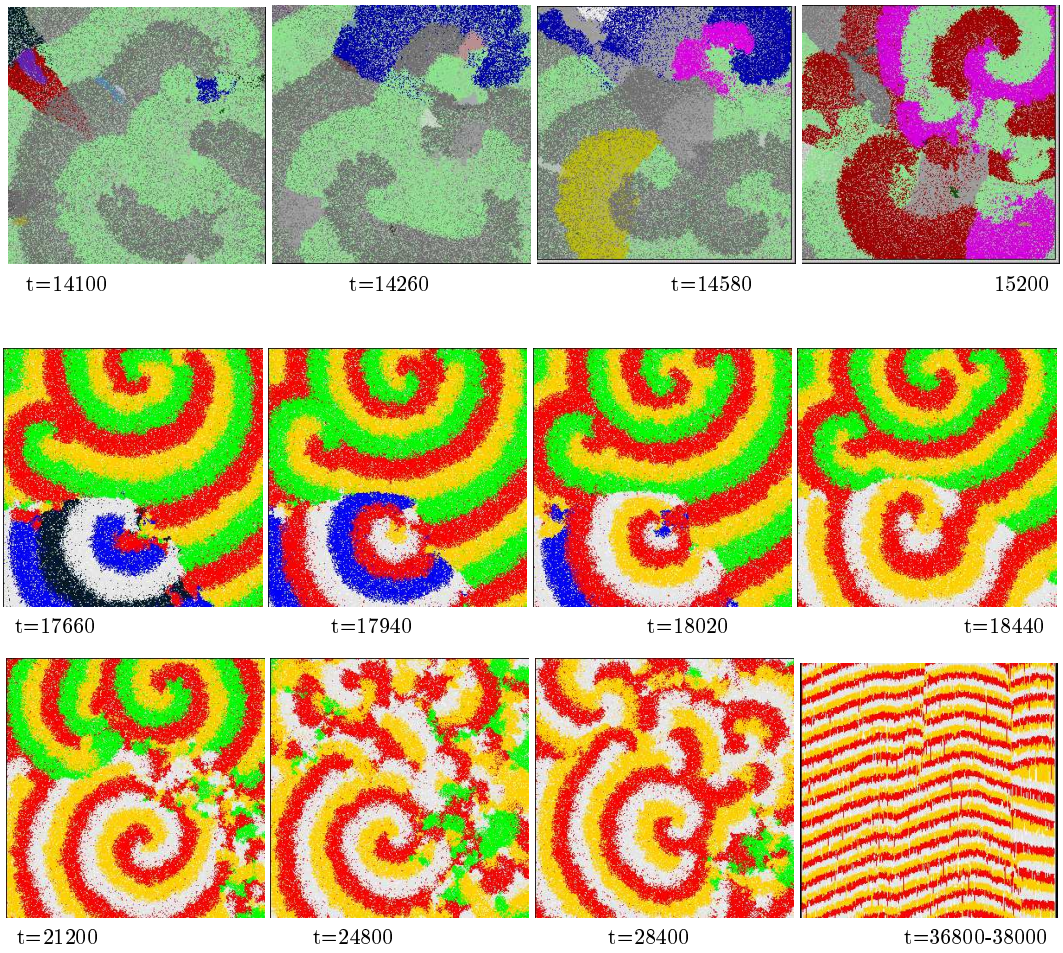


Fig. 1.5. Mutation induced features of spatial patterns. (a) transition from two-armed to three-armed spiral (b) 'infection' of weaker spiral by species of the stronger spiral. This infection induces new mutations, and a faster rotating spiral, which gains domain from the 'donor' spiral. Colors denote lineages, not individual strains for clarity of presentation (c) induction and suppression of chaotic waves (later in the same run as (b); same coloring)

- Mutations in a spiral arm produce inhomogeneities which, long after the mutants have died out, can generate new spiral cores "straddling' spiral arms of other cores (see e.g. fig. 1.5-t17940)
- Transition from single spirals to two-armed spirals to three-armed spirals is caused by mutations close to the core and interference by neighboring domains. An example is shown in fig. 1.5-a. Mutants cause destabilization of the core. The unstable core favors mutants. This way sufficient divergence of the strains can build up rapidly and stabilize the core with more members. Once it is stabilized it spreads over the entire worlds
- When one spiral domain is overtaking an other spiral domain the former may 'infect' the core of the latter: a member of the 'winning' spiral may become member of the loosing spiral. An example is shown in fig. 1.5-b The infection will breed a faster rotating spiral because an immigrating 'foreign' member is more dissimilar to the remaining 'native' strains than the replaced one. Moreover in a short period of core de-stabilization following the infection fixation of new mutations are favored. Therefore such infection

'saves' the spiral, which than starts to gain domain on the 'donor' spiral. Thus, 'horizontal transfer' of strains takes place, by which diversity in the world remains limited, not withstanding very long lived separate spiral domains.

- On the border between spiral domains which have at least two members which are very dissimilar (not withstanding the homogenizing effect mentioned above) the domain can start to intercalate and chaotic dynamics slowly spreads (see fig. 1.5-c). Phylogenetically this leads to a somewhat faster rate of accepted mutations, but the shape of the phylogeny remains rake-shaped. We studied the dynamics of the chaos - spiral 'competition' without the interference of new mutations: mutation is stopped after a small chaotic zone has developed. The chaos first spreads, like in the full evolutionary run. However, after it has spread over half of the world it is subsequently suppressed - leaving a multi-domain pattern of three-armed spirals, all composed of the same strains. The mechanism of chaos suppression is through the occurrence of spirals of the same members but in reversed order. These turn out to be more stable relative to the 'green' lineage (see fig. 1.5-c) and effectively annihilate that lineage. Collision of spirals in which the strains occur in reversed order, leads to regions which are less frequently 'visited' by viral infection waves, as clearly seen in the space time plot of fig. 1.5-c.

Thus summarizing: host-virus interactions lead to formation of wave patterns. These can be complex and chaotic, or can form (multi-armed) spirals, or target waves. The pattern formation leads to two different modes of evolution - which are seen in phylogenetic trees as 'narrow-and skewed' and 'rake-like'. They are similar to evolutionary modes verbally known as 'red queen dynamics' and 'radiation' respectively. In the former case there is a strong selection for being 'different' so as to (partially) escape the immunity of the hosts. Thus, many mutations are 'accepted', i.e. remain in the population. In the latter case strains diverge and waves organize themselves so that hosts are alternately infected by different (lineages of) strains. No (strong) positive selection on mutants occurs, because the diverged strains can infect the hosts better than the newly arising mutant. We have shown how the spiral patterns determine the evolutionary dynamics, and how 'genotypic punctuated equilibria' result from alternating periods of stable spirals and target waves. Moreover we have shown how the occurrence of mutations affects the spiral dynamics. Characterization of the global organization of the chaotic waves is a challenging task for the future.

1.3 Multilevel morphogenesis: coordination, evolution and re-inventions along phylogenetic trees.

In the previous section we have studied minimal conditions in which patterns generated at a larger scale can determine the structure of the interacting entities which form these patterns. It is shown that mutual interaction between 'levels of organization' is 'generic' in locally interacting systems subjected to mutation and selection (Darwinian evolution. In this section we study (and study how we can study) the products of long term multilevel evolution. We do this in the context of morphogenesis, i.e. the generation of a multicellular critter from a single cell. This process involves cell division, cell differentiation (i.e. alternative gene expression and therewith alternative function of the cells), differential growth and death of cells , and often (in animals) cell movement. Intricate coordination between intra- and inter-cellular dynamics is a prerequisite for the formation of 'well shaped' critters. Such a process does not only involve pattern formation, but also the reaction of the cells on the pattern formation.

A striking example of an interplay between pattern formation and morphogenesis in the full sense of the word is demonstrated in a recent study on the development of the slime mold *Dictyostelium discoideum* [16]. The last phase of the life cycle fruiting bodies are formed. This involves an intricate developmental process which is referred to in the experimental literature as the 'reverse fountain' because the top cells are 'pushed' down in the middle, while the bottom cells raise to the top at the periphery. We have shown how the *interaction* of waves of the signalling molecule cAMP (cyclic andenosine mono phosphate), (which are initiated by cells which produce cAMP oscillatory, and which are relayed by the other cells), differential adhesion between cells, and cell differentiation into a type of cells which produce an extracellular stiff slime layer and which do not produce or relay cAMP, generates such a pattern of cell movement and thus turns a 'blob' of cells into a fruiting body, i.e. a slender stalk with spores on top.

The *Dictyostelium* model successfully mimics a specific well studied morphogenetic process, and uses known biochemical processes. The question we examine here (and in [17],[18]) is how we can study the

'generic' behavior of the interplay between the various processes which play a role in morphogenesis. One should note, however, that here (and in evolved systems in general) the term 'generic' behavior should be understood in a special way. It is not so that from arbitrary initial conditions and for a large range of parameter values we should expect intricate morphologies to appear: this is not true for organisms either. Generic behavior should be understood in the sense of the common properties of the rare cases which do form interesting morphologies (see [19] for a discussion on such 'generic non generic properties').

We can study those 'rare' cases by formulating an evolutionary process, using an artificial fitness criterion. This fitness criterion should be chosen in such a way that maximizing it *enable* the phenomena of interest to occur *as side-effect*. The fitness criterion should not (and usually cannot) incorporate the phenomena of interest directly. Here we take 'cell differentiation' (i.e. number of gene-expression patterns and their Hamming distance) as fitness criterion to study morphogenesis.

1.3.1 A model for multilevel morphogenesis

An overview of the model is given in fig. 1.6. Two 'tricks' enable us to formulate a very *minimal* model of multilevel development. The first is using an evolutionary process to zoom in on interesting cases as explained above. The second is the use of a 'two scale CA model' as introduced by Glazier and Graner [20]. In this model a biological cell is represented as many (here ca 40) cellular automaton cells, with the same state (= cell identification number). The model can be seen as an extension of the large Q Potts model. The cellular automaton transition rules represent a surface energy minimization process which is conditional on properties of the cell, notably its volume, which is conserved. Thus, this larger scale feeds back on the micro-scale. Glazier and Graner [20] have shown that cell sorting is a generic property of this model when cell surface tensions differ.

The basic model can be easily extended/interfaced with other processes, like chemotaxis (as was done in the slime mold model mentioned above, but not here), cell growth (i.e. increase of target volume V), cell death (which occurs automatically for small λ), cell differentiation (cell identification refers to cell type which defines surface energy $J_{i,j}$). In the model studied here the biological cells contain a boolean gene-regulation network. Some of the nodes of the network define (through a bit-matching 'mask') the surface energies with neighboring cells. Two nodes define signaling molecules, which may (dependent on the gene regulation network) influence gene expression in neighboring cells. [2] Cell growth is modeled as a reaction on stretching of the cell. [21] has shown that stretching can indeed trigger cell growth in experiments.

The development starts with one large cell, representing a fertilized egg. The first 7 divisions are pre-scheduled for all cells simultaneously (such initial cell divisions are called 'cleavage' divisions in developmental biology). Further cell growth and division (when cell size is twice a reference size) and cell death is governed by the dynamics of the development.

Evolution shapes the gene-regulation networks. The development defines the genotype to fitness mapping: cell differentiation depends on cell movement, which co-determines the cell-neighborhood and the received signals.

The full model is defined by very few parameters. Apart from the parameters defining the evolutionary process, and the gene regulation network (number of nodes, number of connections, and mapping of nodes to surface energy and cell signaling) - which are all held constant in all experiments done so far, there are only four: the strength of the volume conservation λ, the cell growth threshold τ and the dissipation constant and 'temperature' of the Boltzmann equation which defines the probability of copying the state of a neighboring cell into the current cell given a change in surface energy. These were varied, but (over the range tried) did not systematically influence the results. The initial (randomly generated) gene regulation networks, and initial stochastic differences in the evolutionary process produce over evolutionary time very different morphemes It is the properties of these evolved morphogenetic processes which we study as 'generic rare cases' as discussed above.

[2] All nodes have 2 inputs coming from nodes in the same cell, of from neighboring cells (denoted by negative numbers). Because there are only 2 signalling nodes defined, negative inputs from other nodes are always 0. This defines a variable connection network.

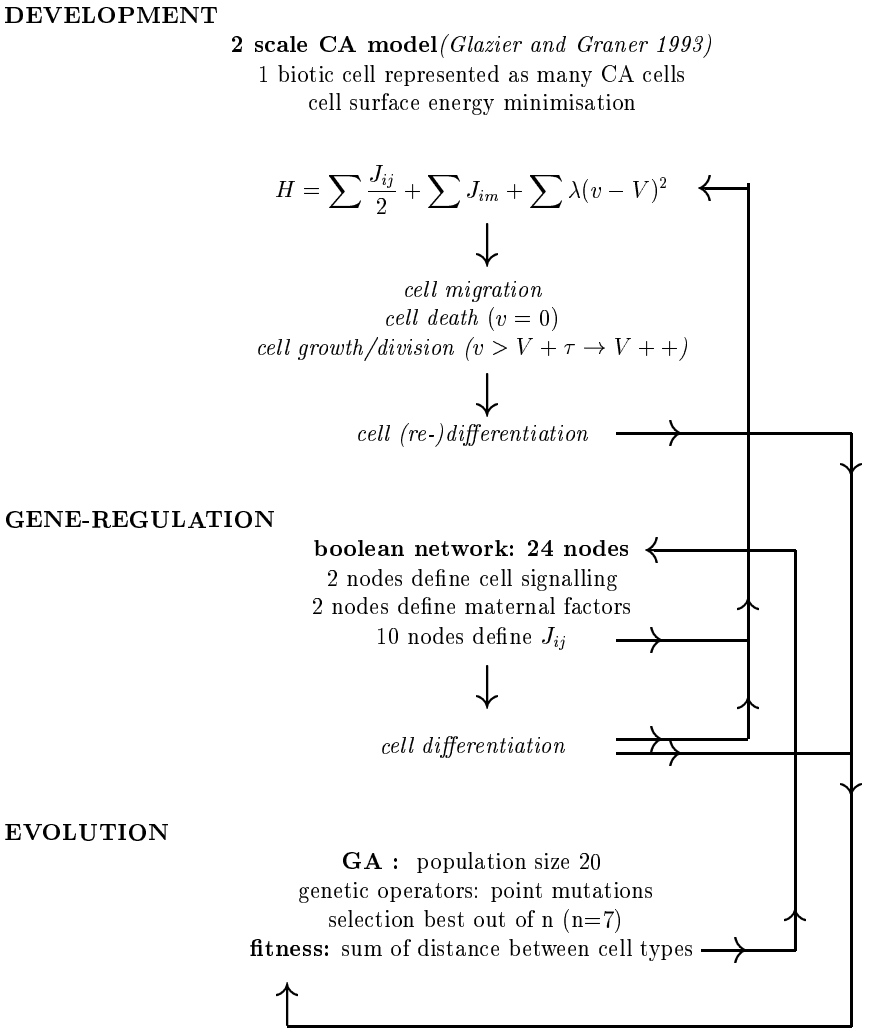


Fig. 1.6. Overview of the model: entanglement between gene regulation, development and evolution

1.3.2 Mechanisms of Morphogenesis

About 20 % of the evolutionary runs produce (more or less) intricate morphogenesis (a small sample is shown in fig. 1.7). One example is described in detail below (see fig. 1.8). Overall the resulting morphogenetic processes as have been described in [17] lead to the following conclusions:

- Morphogenesis results as 'sustained transient' from surface energy minimization and 'intrinsic conflict' which is maintained by cell differentiation, cell growth and cell death. Without continued 'interference', the initial, high energy state would change, through shape changes to, at the end a 'blob' like low energy shape. The development shown in fig. 1.8 without cell growth is an example of such 'transient to a blob'. Growth and cell division supply a continued interference which sustains zones where there is a conflict between energy minimization and internal cell changes which maintain higher energy states and intricate shapes.
- These intrinsic conflicts lead to automatic orchestration of adhesion, migration, differentiation, cell growth/division and death. It results in "pseudo-isomorphic outgrowth". Although the shapes do change during 'maturation' a 'critter' preserves its general appearance.

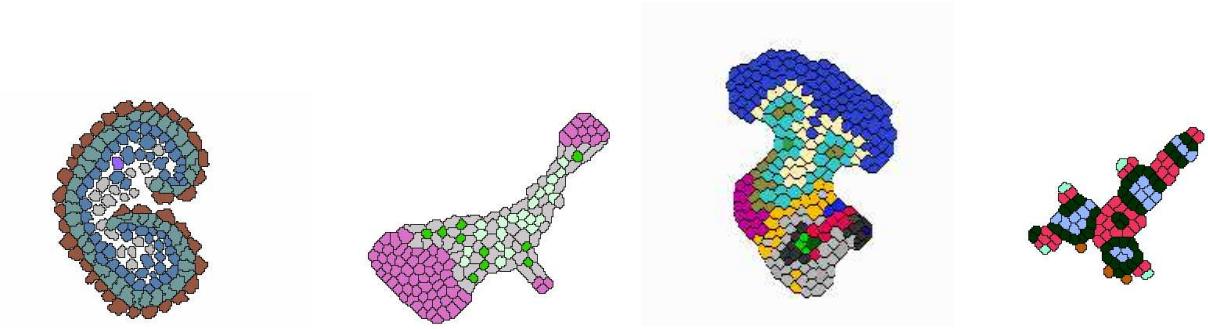


Fig. 1.7. Examples of morphogenetic mechanisms; (a) engulfing combined with apoptosis (b) elongation by budding (c) meristematic growth (d) dynamic growth-death and redifferentiation

- Many different morphemes result from few mechanisms. Mechanisms found are:
  - engulfing: one cell type surrounds an other cell type. An example is fig. 1.7-a
  - meristematic growth: a layer of dividing cells which differentiate into non-dividing (or rarely dividing) cells of several types. The zone is maintained because cell types depend on location. An example is the blue cap in fig. 1.7-c
  - elongation by 'budding': a small group of differentiated cells is pushed outwards because an other cell type on the one hand tends to engulf them, but on the other stick together more firmly than to the 'bud' (fig. 1.7-b) The situation is maintained because cells which do, nevertheless engulf, differentiate into bud-type cells.
  - elongation by 'convergence extension' like process. Two layers of cells maximize their border. Often a other cell types outside these layers contribute by intercalation in the boundary cell types and differentiating into boundary cell types themselves when they touch the other boundary cell type (see fig. 1.8)
  - elongation by intercalation of stably differentiated cell types (not shown)
  - dynamic re-differentiation, cell division and cell death (fig. 1.7-d).

1.3.3 A case study of the development and evolution of a morphotype

Here we examine a particular case of morphogenesis and its evolution. The morphogenesis is shown in fig. 1.8. It is an example in which morphogenesis is mainly caused by the 'convergence extension' mechanism mentioned above, and clearly demonstrates 'conflict induced morphogenesis'. As shown in the upper panel after the cleavage divisions a two-armed structure develops. However in the long run the arms 'retract' and finally only a blob consisting of a number of cell types remains, and is stable; many cells have died. This is how the development unfolds without cell growth and division (large τ), i.e. for the parameter setting used during the evolutionary runs (developmental time up to t=250000).

Gene regulation and cell differentiation The 'functional' gene regulation network is shown in the left (lower) panel of fig. 1.9 (network was drawn with daVinci graph drawing software). The functional gene regulation network was extracted from the full gene regulation network, by an iterative procedure which determines per node whether its state is invariant or dependent on one of its inputs only. If so, non-functional inputs are deleted, and invariant states are used as such in the next iteration. As seen in the figure, although the network is defined with 2 inputs per node, the majority of nodes have zero or one inputs. The structure of the network is very hierarchical, i.e. the terminology 'upstream' and 'downstream' genes makes sense. This in contrast to the situation in random boolean networks. There is some crosstalk between pathways. Genes 1 and 2 code for signalling molecules. The nodes -1 and -2 refer to the expression of the gene in neighboring cells.

Cell differentiation is initiated at the first cleavage division, when the state of gene 21 is set to 0 in one of the daughter cells for one time step. This signals causes stable differentiation between the two cell lineages: gene 7, which is auto-regulatory, preserves the signal. Further cell differentiation is through induction. Most

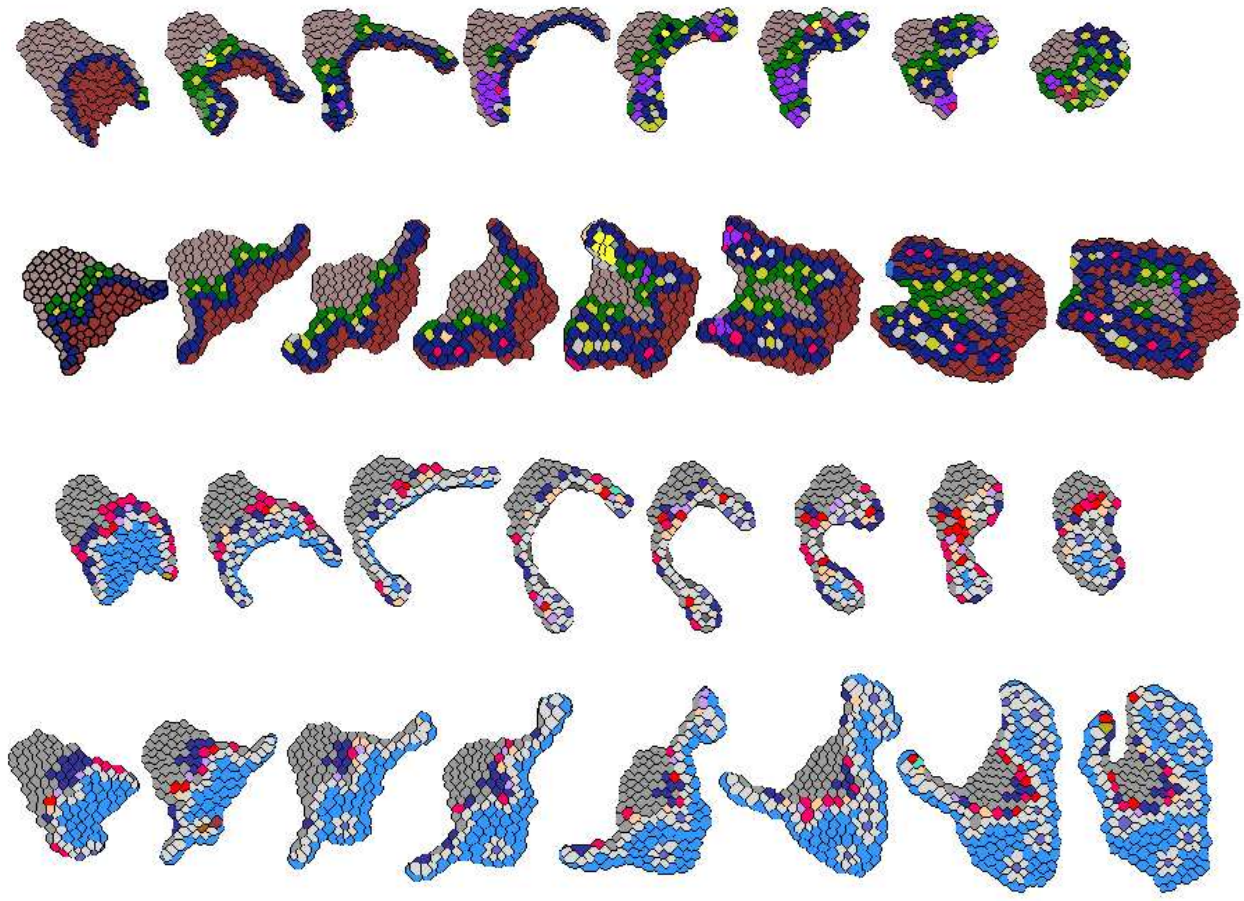


Fig. 1.8. Developmental histories of 'reinvented' morphologies; (a) development of ET2854 without cell growth frame=f=20,50,100,200,400,600,700,800; (1 frame equals 250 timesteps) (b) development of ET2854 with cell growth f= 35,70,140,200,273,400,440,480 (c) reinvention in an alternative evolutionary run restarted from ET=2488 The morpheme occurs at ET=1809 of the new run; Without cell growth f=20,50,100,200,280,300,600 With cell growth f=20,50,100,150,200,300,350,400

importantly the dark-brown cell lineage differentiates by having the light-brown cell lineage in its neighborhood (and becomes blue). Extensive further history dependent cell differentiation occurs, but is of minor importance for the formation of the arms.

Morphogenesis The mechanism of the development of the two-armed shape is as follows. The intermediate cell layer is very 'hydrophobic', i.e.the cell surface energy with the medium is very high (22) while that of the other cell types is 5 (dark brown) and 10 (light brown). Therefore the latter cell types engulf the middle layer. However when they do so they will touch each other, causing the dark-brown cells to differentiate into the blue cell type of the intermediate layer. This causes the extension of the tips of the arms. This process is aided by the fact that the dark-brown cells also intercalate into the intermediate layer: their surface energy with blue cells is only 7 whereas the surface energy among blue cells is 12. Also here re-differentiation occurs when they do intercalate and touch the light-brown cells. Thus the intermediate layer is extended by this process as well. Extension of the arms stops because of a 'shortage' of cells. Therefore the tip of the arms are 'closed off' by 'abnormal' intermediate cells (abnormal due to changing neighborhoods), which do have high surface energy toward the medium. This causes the arms to retract.

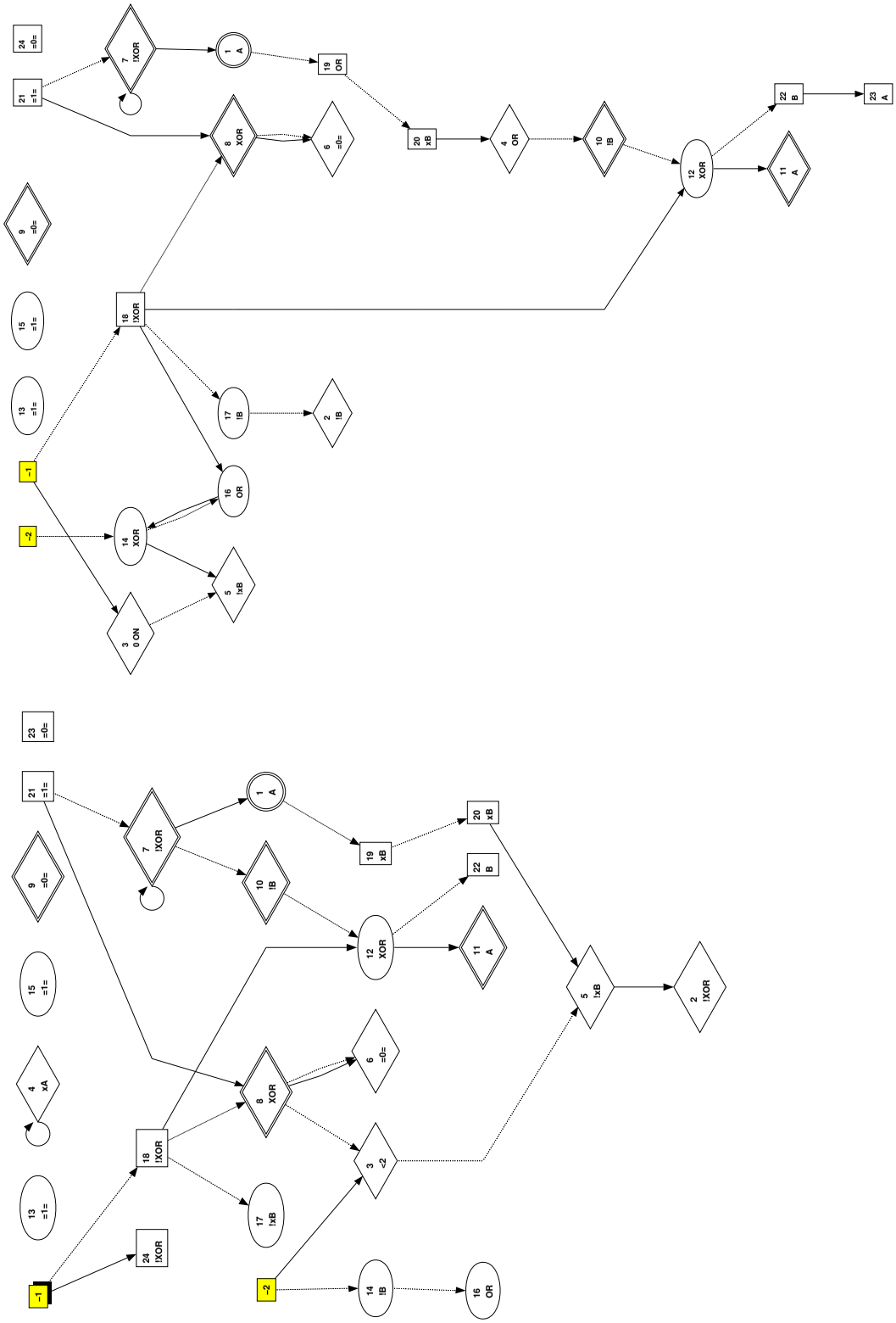


Fig. 1.9. Gene regulation networks of 2-armed morpheme as evolved in two parallel evolutionary histories, starting from a differentiated common ancestor

As mentioned, this is the development without cell growth and division. When cell growth is incorporated ($\tau = 4$) the morphogenesis of the 'critter' unfolds as shown in the second panel of fig. 1.8: relative to the non-growing case it turns itself 'inside out'. The mechanism is as follows. The dark-brown cells have low surface energy among themselves (6) and, as described above, are 'pulled' because of intercalation and engulfing tendency. This causes them to be stretched, and thereby they grow, and eventually divide. Therefore the 'shortage' of cells mentioned above does not arise at the side which engulfs most efficiently. The net result is to push the intermediate layer backwards. When finally the two arms meet each other at the back the engulfing stops, and the growth rate of the critter drops. The last depicted stage is stable.

Evolution From all the de novo initiated evolution runs, this particular morphogenesis was only observed once. However, during the evolutionary run similar critters arise repeatedly anew at wide intervals of time (data not shown). This is after full cell differentiation has evolved. The further evolution is along the 'neutral' path on which cell differentiation is maintained. The gene-regulation networks continue to change (at fixed rate), but the rate of change of the 'functional' gene regulation network slows down, but continues as well [18]. Within these constraints the morphogenesis as described remains a 'likely' mutant of the dominant evolutionary type, which forms rather undistinguished shapes (see fig. 1.10). In these morphemes one cell type engulfs all the other cell types. This is true in the same evolutionary run, but also when evolution is restarted (i.e. 'when the tape were played twice', [22]) after the full cell differentiation pattern has arisen. In fig. 1.8 an example is shown from such a parallel evolutionary history; it is 'critter' 1809 to arise in that run. Both without and with cell growth its development qualitatively mimics the one described above. Nevertheless the general shape of the functional gene regulation network has clearly changed (fig. 1.9), although closer inspection reveals quite some similarities: 9 of the 24 genes are identically regulated. However the surface energies of the main cell types are identical. The observed differences between the two morphogenetic processes are partly due to noise (no 2 critters are identical, also not when they have identical genomes because of randomness in cell movement) and partly due to differences in the history dependent cell types, which have slightly different surface energies, and arise in different circumstances. The latter differences are dominant over the former: it is possible to recognize the critters in a 'color-independent way' when the full morphogenesis is observed (the 're-evolved' critter sticks its arms out more during growth, and becomes thinner without growth). Note, however, that the morphogenetic process is quite sensitive to the ratio of the value of the Boltzmann 'temperature' T and the stretch induced growth parameter τ. Together they define the growth rate of the dark brown cells.

The phylogenetic tree of the two independent evolutionary histories is shown in fig. 1.10. The phylogenetic tree was calculated on the basis of the full gene regulation networks (using the neighbor-joining method). A sample of critters is shown. The evolutionary situation is comparable to that of independently evolving population on islands. Recent research (for a number of striking examples see [23] has shown that on islands repeatedly similar morphotypes and ecological adaptation evolve independently. This discovery has upset many traditional classifications, which assumed that the elaborate morphological or physiological adaptations surely would suggest a common origin. Within our experimental setup, the described two-armed critter would surely be judged as 'special'. And indeed, in a way it is, as it occurs only in one of the de novo runs. However, several samples of de novo evolution is not what we observe in our biosphere, rather we see alternative lineages evolved from a common ancestor. Our experiments suggest that given conserved basic cell differentiation pattern, the *propensity* to evolve certain seemingly complex adaptations is conserved as well. Such a 'pre-adaptation' to potential evolutionary change, is a 'heresy' current evolutionary theory. Our experiments suggest, however. that it is a 'generic' property of Darwinian evolution of critters evolved by Darwinian evolution when entangling of several levels of organization is not ruled out.

1.4 Conclusions

We have discussed models of multilevel evolution and of multilevel morphogenesis. In both cases we have seen that it is feasible to do so in quite simple models. In fact including several levels explicitly 'tunes' the processes at the various levels so as that it in fact modeling becomes easier. Moreover we have shown that two puzzling features of observed evolutionary processes can be explained in terms of the multilevel models presented.

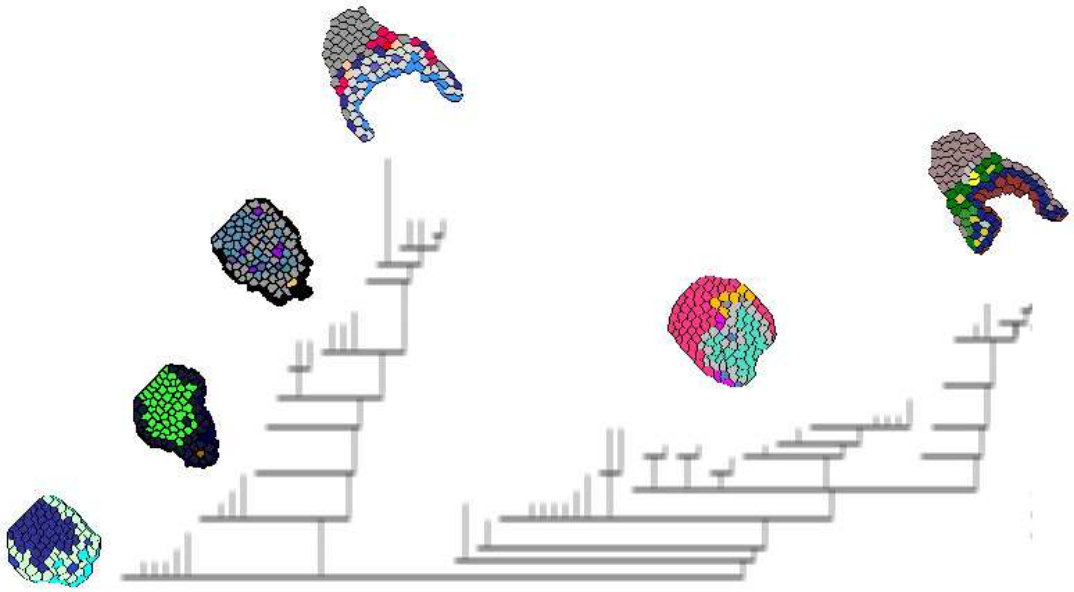


Fig. 1.10. Phylogenetic tree depicting evolutionary divergence and convergence from one common ancestor (left most picture).

In eco-evolutionary context we have seen that the same virus in different hosts, or even simply in a different sized world can evolve in very different modes. The modes of evolution depend on spatial pattern formation: a rake-like phylogeny, and radiation of strains occurs when wave patterns are such that hosts are infected by alternately by different strains. This is the case in multi-armed spirals. When this is not the case there is strong selection on escape of the immune system and a skewed, almost linear phylogeny arises.

In the context of morphogenesis we have seen that the 're-invention' of intricate morphologies appears to be a generic feature in multilevel evolution: some things are and remain 'easy' to evolve, given conservation of some basic features, e.g. (early) cell differentiation.

In conclusion, we have demonstrated that in a multilevel setting Darwinian evolution is an even more versatile mechanism than previously recognized. There is still much work to do to unravel its full potential.

1.5 Acknowledgements

I am much indebted to Roeland Merks for the great amount of work he did in setting up and programming the combined developmental and evolutionary model. I also thank my former students Maarten Boerlijst, Nick Savill, Ludo Pagie and Stan Mare'e for their contributions to studying evolution and development as multilevel processes. I thank Ben Hesper for long term inspiration and support.

References

1. D.L. Hartl and C.H. Taubes. Towards a theory of evolutionary adaptation. *Genetica*, 102-103:525–533, 1998.
2. M. C. Boerlijst and P. Hogeweg. Self-structuring and selection: spiral waves as a substrate for evolution. In C. G. Langton, editor, *Artificial Life II*, pages 255–276, Redwood City, CA, 1991. Addison-Wesley.
3. M. C. Boerlijst and P. Hogeweg. Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. *Physica D*, 48:17–28, 1991.
4. M. Eigen and P. Schuster. *The hypercycle: a principle of natural self-organization.* Springer, Berlin, Heidelberg, New York, 1979.
5. N. J. Savill, P. Rohani, and P. Hogeweg. Self-reinforcing spatial patterns enslave evolution in a host-parasitoid system [published erratum appears in J Theor Biol 1997 Oct 21;188(4):525-6]. *J. theor. Biol.*, 188:11–20, 1997.

6. L. W. P. Pagie. *Information Integration in evolutionary processes.* PhD thesis, Utrecht University, 1999.
7. L. Pagie and P. Hogeweg. Evolutionary consequences of coevolving targets. *Evol. Comput.*, 5:401–418, 1997.
8. P. H. Harvey, A. J. L. Brown, J. Maynard Smith, and S. Nee. *New Uses for New Phylogenies.* Oxford University Press, Oxford, 1996.
9. S. Nee, E. C. Holmes, A. Rambaut, and P. H. Harvey. Inferring population history from molecular phylogenies. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 349:25–31, 1995.
10. C. K. Ong, S. Nee, A. Rambaut, H. U. Bernard, and P. H. Harvey. Elucidating the population histories and transmission dynamics of papillomaviruses using phylogenetic trees. *J. Mol. Evol.*, 44:199–206, 1997.
11. T. Jackman, J. B. Losos, A. Larson, and K. de Wueiroz. Phylogenetic studies of convergent adaptive radiations in caribbeab anolis lizards. In T.J. Givnish and J. Sytsma, editors, *Molecular Evolution and Adaptive Radiation*, pages 535–557. Cambridge University Press, Cambridge, 2000.
12. W. M. Fitch, J. M. Leiter, X. Q. Li, and P. Palese. Positive Darwinian evolution in human influenza A viruses. *Proc. Natl. Acad. Sci. U.S.A.*, 88:4270–4274, 1991.
13. W. M. Fitch. Uses for evolutionary trees. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 349:93–102, 1995.
14. M. A. Huynen, P. F. Stadler, and W. Fontana. Smoothness within ruggedness: the role of neutrality in adaptation. *Proc. Natl. Acad. Sci. U.S.A.*, 93:397–401, 1996.
15. E. Van Nimwegen and J. P. Crutchfield. Metastable evolutionary dynamics: crossing fitness barriers or escaping via neutral paths? *Bull. Math. Biol.*, 62:799–848, 2000.
16. A. F. M. Marée. *From Pattern Formation to Morphogenesis: Multicellular Coordination in Dictyostelium discoideum.* PhD thesis, Utrecht University, 2000.
17. P. Hogeweg. Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation. *J. theor. Biol.*, 203:317–333, 2000.
18. P. Hogeweg. Shapes in the shadow: evolutionary dynamics of morphogenesis. *Artif. Life.*, 6:85–101, 2000.
19. P. Hogeweg. On searching generic properties of non-generic phenomena, an approach to bioinformatic patter analysis. In C. Adami, R. K. Belew, H. Kitano, and C.E. Taylor, editors, *Artificial Life VI*, pages 285–294. MIT Press, Cambridge, Mass., 1998.
20. J. A. Glazier and F. Graner. Simulation of the differential adhesion driven rearrangement of bio logical cells. *Phys. Rev. E*, 47:2128–2154, 1993.
21. C. S. Chen, M. Mrksich, S. Huang, G. M. Whitesides, and D. E. Ingber. Geometric control of cell life and death. *Science*, 276:1425–1428, 1997.
22. W. Fontana and L.W. Buss. What would be conserved 'if the tape were played twice'. *PNAS*, 91:757–761, 1994.
23. T.J. Givnish and J. Sytsma. *Molecular Evolution and Adaptive Radiation.* Cambridge University Press, Cambridge, 2000.