THE ROYAL SOCIETY

doi 10.1098/rspb.2000.1280

Robustness as an evolutionary principle

Stefan Bornholdt[1]* **and Kim Sneppen**[2]

[1]*Institut für Theoretische Physik, Universität Kiel, Leibnizstrasse 15, D-24098 Kiel, Germany*
[2]*Nordic Institute for Theoretical Physics, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*

We suggest simulating evolution of complex organisms using a model constrained solely by the requirement of robustness in its expression patterns. This scenario is illustrated by evolving discrete logical networks with epigenetic properties. Evidence for dynamical features in the evolved networks is found that can be related to biological observables.

Keywords: evolution; gene regulation; epistasis; simulation; networks

1. INTRODUCTION

A common concept in evolution is fitness and fitness landscapes (Wright 1982); often evolution is viewed as hill climbing, possibly with jumps between fitness maxima (Lande 1985; Newman *et al*. 1985). However, fitness landscapes implicitly assume that fitness varies over a well-defined metric in genomic space. This would be the case if single point mutations were a driving force. However, significant genome rearrangements have already been observed in the rather brief real-time evolution experiments of *Escherichia coli* cultures of Papadopoulos *et al*. (1999). Genomic rearrangements short-circuit the simple metric generated by one-point mutations, usually underlying the intuition of evolution on landscapes. As a consequence, the combinatorial distance for moving from a genome A to a genome B may easily be different from the distance of the opposite move, most simply exemplified by deletions and insertions. Thus, although fitness landscapes have a meaning for the small-scale adjustments associated with fine-tuning of binding constants, it is an unjustified concept for evolutionary changes on the scale of speciation events.

In this paper, we should like to draw a distinction between molecular neutrality, as suggested by Kimura (1983), and the possibility of neutrality on the scale of genetic regulation networks which is discussed here. Molecular neutrality deals with single base-pair mutations that do not influence the conformation and function of single molecules. Genetic network neutrality deals with architectural changes in the regulatory genetic network that do not influence its output pattern substantially. Dynamically, the one-point mutations in proteins or RNA (Gruener *et al*. 1996*a*,*b*; Reidys *et al*. 1997; Schuster 1997) proceed slowly, with a rate that reflects the probability that a mutation does not change molecular properties (Kimura 1983). We here similarly assume that genetic network rearrangement proceeds stochastically, with a rate determined by the probability that the networks do not change their expression pattern substantially. However, in contrast to molecular neutrality and random walks on neutral plateaux on the energy landscapes as studied by Gruener *et al*. (1996*a*,*b*), Reidys *et al*. (1997) and Schuster (1997), we evolve the network without any explicit energy function. Thus we do not have any absolute energy or fitness of our evolving network. This mimics the fact that the functional dynamics of genetic regulatory networks do not exclusively depend on their architecture. Often, other factors from the cellular environment (and beyond) interact with gene regulation. These factors can change on a much faster time-scale than the architecture of genetic networks and may be viewed as additional boundary conditions varying in time.

Abandoning fitness landscapes, we here instead discuss the possibility that evolution progresses through a process where genotypes and phenotypes subsequently set the frame at which the other may change. Of particular relevance for this view of evolution is the fact that one often observes different phenotypes for the same genotype. This viewpoint is in part supported by cell differentiation within one organism, together with epigenetics and the large class of organisms which undergo metamorphosis and thus exist in several phenotypes for the same genotype. A mechanism for evolution may thus be exposure of the same species to different environments. The species then faces a variable selection criterion, with the consequence that what is phenotypically neutral at some instant may not be phenotypically neutral at later instants. Thus, in contrast to the molecular neutrality where many RNA genotypes have the same phenotype (Gruener *et al*. 1996*a*,*b*; Reidys *et al*. 1997; Schuster 1997), we here use that, on higher level organization, more than one phenotype for each genotype may occur.

In general, evolutionary models with genotype–phenotype ambiguity are currently discussed as the basis of sympatric speciation events, where new species can emerge without the strict need for geographical separation (see Dieckmann & Doeberli (1999) and references cited therein). Proposed mechanisms range from the divergence of coexisting phenotypes (Kaneko & Yomo 1999) to the evolution of assortative mating conditional on a selectively neutral marker (Dieckmann & Doeberli 1999). In the present paper we do not discuss a full model of sympatric speciation, because we restrict ourselves to following a single germline. However, it is well conceivable that epigenetics on the level of genetic networks could contribute to sympatric speciation.

Here we consider a class of systems that exhibits epigenetics, which is represented by the logical networks, where nodes in the network take values on or off, as a function of the output of other specified nodes. This has

*Author for correspondence (bornhol@theo-physik.uni-kiel.de).

Received 10 May 2000 *Accepted* 21 August 2000

been suggested to model the regulatory gene circuits (Kauffman 1969, 1990; Somogyi & Sniegoski 1996; Thieffry & Thomas 1998), where specific genes may or may not be expressed as a function of other genes. In terms of these models it is natural to define genotypes in the form of the topology and rules of the nodes in the network. The phenotypes are similarly associated with the dynamical expression patterns of the network.

To define the rules under which phenotypes and genotypes set the frame for each other's development, a model for evolution should fulfill the requirement of robustness. Robustness is defined as the ability to function in the face of substantial change in the components (Savageau 1971; Hartwell 1997; Alon *et al*. 1999; Little *et al*. 1999). Robustness is an important ingredient in simple molecular networks and probably also an important feature of gene regulation on both a small and large scale. In terms of logical networks, robustness is implemented by constraining subsequent networks to have similar expression patterns.

This paper is organized as follows: first, we discuss dynamics on logical networks and numerically review the basic properties of attractors of random threshold networks and Boolean networks. Then we propose a minimal evolution model and investigate its statistical and structural implications for the evolved networks. Finally, biological implications and possible experimental approaches to the dynamics of real genetic networks are discussed.

2. DYNAMICS ON LOGICAL NETWORKS

Let us first discuss two prototype networks that exhibit epigenetics, Boolean networks (Kauffman 1969, 1990; Somogyi & Sniegoski 1996) and threshold networks (Kürten 1988*a*,*b*). These are both networks of logical functions and share similar dynamical properties. Here we briefly describe their definition and dynamical features. In both networks each node takes one of two discrete values, ± 1, that at each time-step is a discrete function of the value of some fixed set of other nodes specified by a wiring diagram. The links that provide input to node i are denoted by $\{w_{ij}\}$ with $w_{ij} = \pm 1$. A crucial structural parameter of the network so defined is its connectivity K, which we will define as the average number of incoming (non-zero) weights per node. The updating rule for the dynamics on the network remains to be specified. For the threshold network case it is additive:

$$\sigma_i = 1 \quad \text{if} \quad \sum_{j\in\{w_i\}} w_{ij}\sigma_j \geqslant 0, \tag{1}$$

$$\sigma_i = -1 \quad \text{if} \quad \sum_{j\in\{w_i\}} w_{ij}\sigma_j < 0. \tag{2}$$

In the Boolean network case the updating is a general Boolean function of the input variable

$$\sigma_i = B(\sigma_j\text{-values which provide input to } i). \tag{3}$$

Thus, the threshold networks form a hugely restricted set of the Boolean networks. Boolean networks include all nonlinear combinations of input nodes, including functions such as the 'exclusive-OR'. The threshold networks are well known as a type of neural network, where a certain number of input firings are necessary to induce firing in a given neuron (Kürten 1988*a*,*b*). Boolean networks are mostly discussed in connection with genetic networks, because in principle the specificity of protein binding enables the implementation of more detailed logical functions.

The basic property of logical networks is a dynamics of the state vector $\{\sigma_i\}$ characterized by transients that lead to subsequent attractors. The attractor length depends on the topology of the network. Below a critical connectivity, $K_c \approx 2$ (Kauffman 1969, 1990; Derrida & Pomeau 1986), the network decouples into many disconnected regions, i.e. the corresponding genome expression would become modular, with essentially independent gene activity. Above K_c any local damage will initiate an avalanche of activity that may propagate throughout most of the system. For any K above K_c the attractor period diverges exponentially with respect to system size N, and in some interval above K_c the period length also increases nearly exponentially with connectivity K (Bastola & Parisi 1996). Note that in Boolean networks the critical connectivity (or coordination number) equals two, compared to unity in usual random graphs (Erdös & Renyi 1960; Bollobas 1985), due to the Boolean logic. Criticality means that a change at a node in the network spreads marginally throughout the network. This picture is particularly simple for Boolean networks, where any change has probability 0.5 to propagate along any link for random Boolean rules, so that an average of two links have to leave each node to create the critical state. For neural threshold networks similar arguments apply.

3. STRUCTURAL EVOLUTION OF NETWORKS

Dynamics may occur on networks as defined by the rule above, but at least as important is the dynamics of network topology (Bornholdt & Sneppen 1998; Paczuski *et al*. 2000; Bornholdt & Rohlf 2000). In terms of network topology, an evolution means a change in the wiring $\{w_{ij}\}\rightarrow\{w'_{ij}\}$ that takes place on a much slower time-scale than the $\{\sigma_j\}$ updating. The evolution of such networks represents the extended degree of genetic network engineering that seems to be needed to account for the large differences in the structure of species genomes (Shapiro 1998), given the slow and steady speed of single protein evolution (Kimura 1983). The model will extend neutral evolution on the molecular scale (Kimura 1983) to neutral evolution on the regulatory level, and demonstrate that neutrality in itself enforces constraints on the evolved graphs.

We have, in an earlier publication, proposed to evolve Boolean networks with the sole constraint of continuity in expression pattern (Bornholdt & Sneppen 1998). Here we simplify this model by simple damage spreading testing:

The model evolves a new single network from an old network by accepting rewiring mutations with a rate determined by expression overlap.

This is a minimal constraint scenario with no outside fitness imposed. Further, the model tends to select for networks which have high overlap with neighbour mutant networks, thus securing robustness.

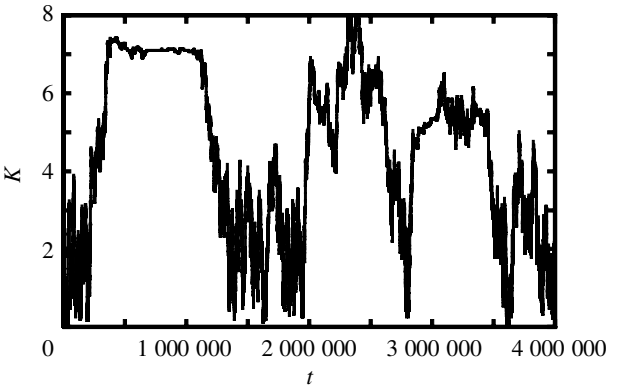

Figure 1. Long time evolution for the connectivity of a threshold network with $N = 32$ nodes. Connectivities are constrained to be below $K = 8$. One observes long periods of stasis interrupted by sudden changes, reminiscent of punctuated equilibrium.

Now let us formulate an operational version of the evolution in terms of threshold networks, because these have comparable structural and statistical features to the Boolean ones (Kürten 1988*a*,*b*). Consider a threshold network with N nodes. To each of these let us assign a logical variable $\sigma_i = -1$ or $+1$. The states $\{\sigma_i\}$ of the N nodes are simultaneously updated according to equation (1) where the links w_{ij} are specified by a matrix. The entry value of the connectivity matrix w_{ij} may take values -1 and $+1$ in case of a link between i and j, and the value zero if i is not connected to j.

The system that is evolved is the set of couplings w_{ij} in a single network. One evolutionary time-step of the network is as follows.

(i) Create a daughter network by (1) adding, (2) removing, or (3) adding and removing a weight in the coupling matrix w_{ij} at random, each option occurring with probability $p = 1/3$. This means turning $w_{ij} = 0$ to a randomly chosen ± 1 or vice versa.

(ii) Select a random input state $\{\sigma_i\}$. Iterate simultaneously both the mother and the daughter system from this state until either they have reached and completed the same attractor cycle, or until a time where $\{\sigma_i\}$ differs between the two networks. In case their dynamics are identical then replace the mother with the daughter network. In case their dynamics differ, keep the mother network.

Thus, the dynamics look for mutations which are phenotypically silent, i.e. these are neutrally inherited under at least some external conditions. Note that adding a link involves selecting a new w_{ij}, thus changing the rule on the same time-scale as the network connectivity. Iterating these steps represents an evolution which proceeds by checking overlap in expression pattern between networks. If there are many states $\{\sigma_i\}$ that give the same expression of the two networks, then transitions between them are fast. In contrast, if there are only very few states $\{\sigma_i\}$ which result in the same expression for the two networks, then the transition rate from one network to the other is small. If this is true for all its neighbours then the evolutionary process will be hugely slowed down. Interestingly, other than in existing concepts of selective neutrality (Gruener *et al*. 1996*a*,*b*; Reidys *et al*. 1997; Schuster 1997; Sibani & Pedersen 1999; Van Nimwegen *et al*. 1999), these transition rates are not constant in our model of regulatory neutrality. In particular, they are instead a function of the evolving connectivity K of the network.

In figure 1 the connectivity change with time for a threshold network of size $N = 32$ is shown. Time is counted as number of attempted mutations. To understand the evolution of the networks, we first remark that, because mutations are local, only the cluster at which the mutation took place is visible to the proposed phenotypic test. For subcritical networks, which usually consist of disconnected graphs, this means that acceptance rate below critical connectivity will not depend on system size. In the case of supercritical networks the dynamics are dominated by their giant component. The acceptance criterion, therefore, gets harder with increasing system size: transition probabilities of neutral evolution towards larger connectivity K decrease with K. Thus, most evolution will in practice be arrested slightly above critical conditions. We observe that the relatively large variance in K for small systems as shown in figure 1 is confined to a smaller interval for larger systems simulated over the same time. It is interesting to note that the effective critical connectivity of the evolved networks lies somewhere above K_c for a random network. This is a consequence of the evolved structural features of the network, a number of which will be described below.

One also observes that, especially for high connectivity, the system may stay for a long time at a particular network before an allowed mutation leads to punctuations of the stasis. The overall distribution of waiting times is approximately $1/t^{2\pm0.2}$. The wide variety of time-scales implied by the $1/t^2$ distribution reflects the different time-scales that are associated with networks of different connectivity K. Thus, any particular network will have a characteristic time-scale with exponentially distributed waiting time. The $1/t^2$ distribution originates from integration over this broad range of time-scales, reflecting that the probability of accepting a mutation decreases exponentially with K, whereas the probability per attempt to add a specific link equals the probability to remove it again.

One important feature of the evolution is the structure of the evolved networks, which can be quantified by the average length of attractors for the generated networks. This is shown in figure 2, where they are compared with attractor lengths for random networks at the same connectivity. One observes that the evolved networks have much shorter attractors than the random ones; thus our evolution scenario favours simplicity of expression.

To examine further the expression behaviour of the networks, let us consider the size of frozen components as introduced by Kauffman (1969, 1990) for Boolean networks. A frozen component is the set of nodes connected to a given attractor that does not change at any time when you iterate along the attractor, i.e. a frozen component represents genes which are anaesthetized under a given attractor/initial conditions. In figure 3 one sees that the frozen component for the evolved network typically involves half the system, and thus is much larger

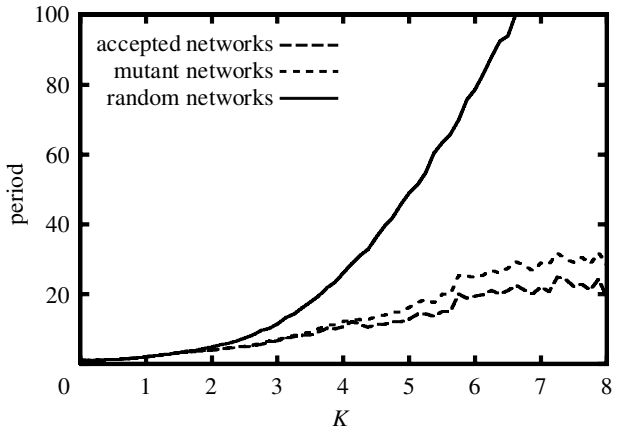

Figure 2. Average length of periodic attractors for evolved and for random networks. Also the periods of the unsuccessful mutations in the presence of newly chosen random initial conditions are shown, demonstrating that selection of networks is indeed operating in structure space and the specific input configuration in the event of selection does not play a major role.

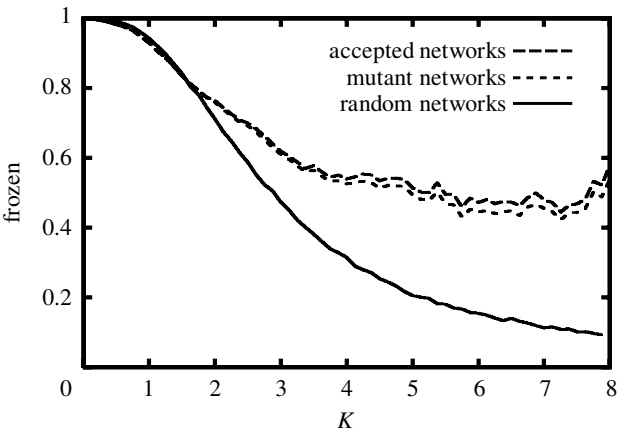

Figure 3. Average size of frozen components as a function of connectivity for evolved and random networks. The frozen component is the set of all nodes that do not switch during the attractor. One observes that the robustness constraint in evolution favours a larger frozen component.

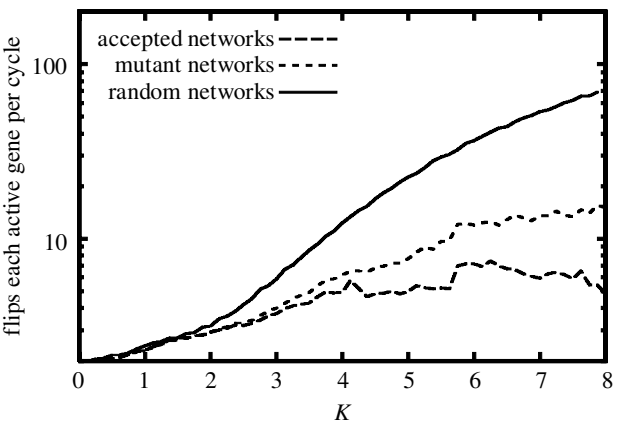

Figure 4. Average number of flips per node in the non-frozen part of the network, as a function of connectivity for evolved and for random networks. The evolved networks show a reduced activity in the non-frozen nodes resulting in simple expression patterns compared to those of random networks of the same connectivities. Notice that the number counts off–on and on–off transitions of the nodes as separate events.

than the typical frozen component associated with attractors of randomly generated threshold networks. Also we test frozen components for random one-mutant neighbours of the selected ones, and find that these networks also have huge frozen components.

Let us finally look at the active part of the network and the complexity of its expression pattern. Because a quite large fraction of the nodes may belong to the frozen component of the network, the remaining active part of the nodes may behave differently from the average dynamics of the whole network. One possible measure is the number of times each non-frozen node switches its state during the dynamical attractor. In figure 4 this quantity is shown for both random networks and evolved networks. One observes that the active part of the evolved networks exhibits a much simpler expression pattern than that of a random network of comparable connectivity.

Overall, implementing robustness as an evolutionary criterion has observable consequences for both the temporal evolution pattern and for confining possible genetic network architectures to those with simple expression patterns.

4. DISCUSSION

Some quantitative testing of the minimal evolution scenario is possible on the macro-evolutionary scale. Here, the intermittent evolution of the networks bears a resemblance to the punctuated equilibrium observed for species in the fossil record (Gould & Eldredge 1993). Quantitatively, the $1/t^2$ distribution of lifetimes for single networks that one finds for this model and for the earlier version (Bornholdt & Sneppen 1998), compares well with the similar scalings observed for the statistics of birth and death of individual species in the evolutionary record (Sneppen *et al*. 1995). In fact, the analogy can even be fine-grained into a sum of characteristic lifetimes, each associated with a given structural feature of the networks (Bornholdt & Sneppen 1998). A similar decomposition is known from the fossil record (Van Valen 1973), where groups of related species display Poisson-distributed lifetimes and, therefore, similar evolutionary stability.

A validation on the microlevel based on statistical properties of genetic regulatory circuits has to be based either on properties of genetic networks (Somogyi & Sniegoski 1996) or on evolution and mutation experiments of short-lived organisms such as *E. coli* (Papadopoulos *et al*. 1999). A key number is the estimated average connectivity K of $2 \rightarrow 3$ in the *E. coli* genome (Thieffry *et al*. 1998). Information on the overall organization of these genetic networks is obtained from gene knock-out experiments.

Quantitative support for a connected genome can be deduced from the experiments of Elena & Lenski (1999) on double mutants, which demonstrated that about 30–60% of these (dependent on interpretation) change their fitness in a cooperative manner. In terms of our networks, we accordingly should expect a coupled genetic expression for about half the pairs of genes. Although our evolved networks can give such correlations for the

connectivity estimate of 2–3 given by Thieffry *et al.* (1998), the uncertainty is still so large that random networks are also in accordance with the data. Further, one should keep in mind that the *E. coli* genome is large and not well represented by threshold dynamics of all nodes, and also that only between 45 and 178 of the 4290 genes of *E. coli* are likely to mediate regulatory functions (F. Blattner, http://www.genetics.wisc.edu/html/k12.html). Thus, most of the detected gene–gene correlations presumably involve genes that are not even regulatory, but instead metabolic, with more indirect effects on each other than in the case of the regulatory genes. Presumably, one would obtain stronger elements of both coupling and correlation, if one specialized in regulatory genes. Thus, one may wish for experiments where one- and two-point mutations are performed in regulatory genes only. A more direct test of our hypothesis of robustness in the form of damage control as a selection criterion may be obtained from careful analysis of the evolution of gene regulation in evolving *E. coli* cultures.

Another interesting observation is the simplicity of biological expression patterns. For example, as observed in yeast many genes are only active one or two times during the expression cycle (Cho *et al.* 1998); thus switching from off to on or on to off occurs for each gene in this system only a few times during expression. For random dynamical networks of comparable size, one would expect much higher activity. Thus, surprisingly simple expression patterns are observed in biological gene regulatory circuits. This bears resemblance to our model observation where simplicity of expression patterns emerges as a result of the evolutionary constraint of robustness.

5. SUMMARY

In this paper we have proposed a computer simulation of evolution operating on logical networks. The scenario mimics an evolution of gene regulatory circuits that is governed by the requirement of robustness only. The resulting dynamics evolve networks that have very large frozen components and short attractors. Thus, they evolve to an ordered structure that counteracts the increasing chaos when networks become densely connected. The evolved architecture is characterized by simplicity of expression pattern and increased robustness to permanent mutational fluctuations in the network architecture—features that are also seen in real molecular networks.

We thank Stanley Brown for valuable comments on the manuscript.

REFERENCES

Alon, U., Surette, M. G., Barkai, N. & Leibler, S. 1999 Robustness in bacterial chemotaxis. *Nature* **397**, 168–171.

Bastola, U. & Parisi, G. 1996 Closing probabilities in the Kauffman model: an annealed computation. *Physica* A **98**, 1–25.

Bollobas, B. 1985 *Random graphs*. Academic Press.

Bornholdt, S. & Rohlf, T. 2000 Topological evolution of dynamical networks: global criticality from local dynamics. *Phys. Rev. Lett.* **84**, 6114–6117.

Bornholdt, S. & Sneppen, K. 1998 Neutral mutations and punctuated equilibrium in evolving genetic networks. *Phys. Rev. Lett.* **81**, 236–239.

Cho, R. J. (and 10 others) 1998 A genome-wide transcriptional analysis of the mitotic cell cycle. *Mol. Cell* **2**, 65–73.

Derrida, B. & Pomeau, Y. 1986 Random networks of automata: a simple annealed approximation. *Europhys. Lett.* **1**, 45–49.

Dieckmann, U. & Doeberli, M. 1999 On the origin of species by sympatric speciation. *Nature* **400**, 354–357.

Elena, S. F. & Lenski, R. E. 1999 Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395–398.

Erdös, P. & Renyi, A. 1960 On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kut. Intez. Kozl.* **5**, 17–61.

Gould, S. J. & Eldredge, N. 1993 Punctuated equilibrium comes of age. *Nature* **366**, 223–227.

Gruener, W., Giegerich, R., Strothmann, D., Reidys, C., Weber, J., Hofacker, I. L., Stadler, P. F. & Schuster, P. 1996*a* Analysis of RNA sequence structure maps by exhaustive enumeration. 1. Neutral networks. *Monatshefte Chem.* **127**, 355–374.

Gruener, W., Giegerich, R., Strothmann, D., Reidys, C., Weber, J., Hofacker, I. L., Stadler, P. F. & Schuster, P. 1996*b* Analysis of RNA sequence structure maps by exhaustive enumeration. 2. Structure of neutral networks and shape space covering. *Monatsh. Chem.* **127**, 375–389.

Hartwell, L. 1997 Theoretical biology: a robust view of biochemical pathways. *Nature* **387**, 855–857.

Kaneko, K. & Yomo, T. 1999 Symbiotic speciation from a single genotype. *Proc. Natl Acad. Sci. USA*. (Submitted.)

Kauffman, S. A. 1969 Metabolic stability and epigenesis in randomly constructed genetic nets. *J. Theor. Biol.* **22**, 437–467.

Kauffman, S. A. 1990 Requirements for evolvability in complex systems: orderly dynamics and frozen components. *Physica* D **42**, 135–152.

Kimura, M. 1983 *The neutral theory of molecular evolution*. Cambridge University Press.

Kürten, K. E. 1988*a* Correspondence between neural threshold networks and Kauffman Boolean cellular automata. *J. Phys.* A **21**, L615–L619.

Kürten, K. E. 1988*b* Critical phenomena in model neural networks. *Phys. Lett.* A **129**, 157–160.

Lande, R. 1985 Expected time for random genetic drift of a population between stable phenotypic states. *Proc. Natl Acad. Sci. USA* **82**, 7641–7644.

Little, J., Shepley, D. P. & Wert, D. W. 1999 Robustness of a gene regulatory circuit. *EMBO J.* **18**, 4299–4307.

Newman, C. M., Cohen, J. E. & Kipnis, C. 1985 Neo-darwinian evolution implies punctuated equilibrium. *Nature* **315**, 400–402.

Newman, M. E. J. & Eble, G. J. 1999 Power spectra of extinction in the fossil record. *Proc. R. Soc. Lond.* B **266**, 1267–1270.

Paczuski, M., Bassler, K. & Corral, A. 2000 Self-organized networks of competing Boolean agents. *Phys. Rev. Lett.* **84**, 3185–3188.

Papadopoulos, D., Schneider, D., Meier-Eiss, J., Arber, W., Lenski, R. E. & Blot, M. 1999 Genomic evolution during a 10,000-generation experiment with bacteria. *Proc. Natl Acad. Sci. USA* **96**, 3807–3810.

Reidys, C. M., Stadler, P. F. & Schuster, P. 1997 Generic properties of combinatory maps—neutral networks of RNA secondary structures. *Bull. Math. Biol.* **59**, 339–397.

Savageau, M. A. 1971 Parameter sensitivity as a criterion for evaluating and comparing the performance of biochemical systems. *Nature* **229**, 542–544.

Schuster, P. 1997 Landscapes and molecular evolution. *Physica* D **107**, 351–365.

Shapiro, J. A. 1998 Genome system architecture and natural genetic engineering in evolution. In *Molecular strategies in biological evolution* (ed. L. Carporale), pp. 23–25. Annls NY Acad. Sci. Book Ser. no. 870. New York Academy of Sciences.

Sibani, P. & Pedersen, A. 1999 Evolution dynamics in terraced NK landscapes. *Europhys. Lett.* **48**, 34–52.

Sneppen, K., Bak, P., Flyvbjerg, H. & Jensen, M. H. 1995 Evolution as a self-organized critical phenomenon. *Proc. Natl Acad. Sci. USA* **92**, 5209–5213.

Sole, R. V., Manrubia, S. C., Benton, M. & Bak, P. 1997 Self-similarity of extinction statistics in the fossil record. *Nature* **388**, 764–767.

Somogyi, R. & Sniegoski, C. A. 1996 Modelling the complexity of genetic networks: understanding multigenic and pleiotropic regulation. *Complexity* **1**, 45–63.

Thieffry, D. & Thomas, R. 1998 Qualitative analysis of gene networks. *Pacific Symp. Biocomput.* **3**, 77–88. Available at http://www.smi.stanford.edu/projects/helix/psb98/thieffry.pdf.

Thieffry, D., Huerta, A. M., Perez-Rueda, E. & Collado-Vides, J. 1998 From specific gene regulation to global regulatory networks: a characterisation of *Escherichia coli* transcriptional network. *BioEssays* **20**, 433–440.

Van Nimwegen, E., Crutchfield, J. P. & Huynen, M. 1999 Neutral evolution of mutational robustness. *Proc. Natl Acad. Sci. USA* **96**, 9716–9720.

Van Valen, L. 1973 A new evolutionary law. *Evol. Theory* **1**, 1–30.

Wright, S. 1982 Character change, speciation, and the higher taxa. *Evolution* **36**, 427–443.