J. theor. Biol. (2002) **216,** 73–84
doi:10.1006/jtbi.2002.2530, available online at http://www.idealibrary.com on IDEAL®

Tangled Nature: A Model of Evolutionary Ecology

KIM CHRISTENSEN*, SIMONE A. DI COLLOBIANO*, MATT HALL†
AND HENRIK J. JENSEN†‡

**Blackett Laboratory, Imperial College, Prince Consort. Road, London SW7 2BW, U.K.*
and †Department of Mathematics, Imperial College, 180 Queen's Gate, London SW7 2BZ, U.K.

(*Received on* 5 *April* 2001, *Accepted in revised form on* 3 *January* 2002)

We discuss a simple model of co-evolution. In order to emphasize the effect of interaction between individuals, the entire population is subjected to the same physical environment. Species are emergent structures and extinction, origination and diversity are entirely a consequence of co-evolutionary interaction between individuals. For comparison, we consider both asexual and sexually reproducing populations. In either case, the system evolves through periods of hectic reorganization separated by periods of coherent stable coexistence.

Introduction

It is difficult in experiments and observations to bridge the gap between ecological and evolutionary time (Pimm, 1991). Nevertheless, since Darwin's publication of *The Origin of Species* (Darwin, 1859) it has been generally agreed that the intricate and complex ecologies surrounding us are the product of Natural Selection operating on vast numbers of successive generations. We know that the slow gradual effect of mutations and Natural Selection is the long-term mechanism underlying evolution in ecological systems, but we are often unable to answer questions concerning stability and the nature of the dynamical evolution (intermittent vs. gradual). It is also difficult to measure the degree of interrelatedness of an ecology (e.g. Bjørnstad *et al.*, 2001): who is interacting with whom and how strongly, and it is difficult to determine the species abundance in detail. What is especially difficult is to monitor the temporal variation in the aforementioned quantities over evolutionary time.

Can general principles be identified for the overall dynamical behaviour of evolution? Even if the characteristics of each individual species have to be considered in their proper specific contexts, perhaps general laws do operate at an overall level. Obviously, the answers to these questions are empirical, but indicators may be obtainable from deliberately simplified theoretical models. It is obviously important to consider carefully the type of simplification assumed. Simplified models often investigate mean field dynamics directly at species level and typically consider only a few coupled species with a focus on a single locus or qualitative character [e.g. May & Anderson, 1983; Seger, 1988; Gavrilets & Hastings, 1998 (and references therein);

---

‡Author to whom correspondence should be addressed.
E-mail: h.jensen@ic.ac.uk

0022-5193/02/$35.00/0

Doncaster *et al.*, 2000]. In order to capture the consequences of the complexity characteristic of ecology we believe it is important to treat species as emergent structures and to allow for the multitude of interactions each individual (and therefore species) is subject to. Our study also indicate the importance of including stochastic fluctuations in the dynamics, this is done by applying probabilistic dynamics at the level of individuals. We find that the very nature of the dynamics of the ecology is strongly influenced by the complexity of the system.

Temporal as well as spatial variations in the physical environment are known to play an important role in evolution. It is also often assumed that co-evolution with interaction between co-existing individuals, or species, may influence the evolutionary dynamics in a significant way (Kauffman, 1995; Bak & Sneppen, 1993). The relative importance of selective pressure of purely physical origins and co-evolutionary effects is not clear, and it seems difficult to resolve the issue solely by use of selected specific case studies. Moreover, seen from an ecological point of view, the biotic and the physical environment are coupled.

In the present paper, we present an individual-based mathematical model of an evolving ecology. An individual-based model has recently been studied by Taylor & Higgs (2000). Their model focus on pleitropy and epistasis whereas our model is designed to emphasize the ecological *interactions* between the individuals. Our model is purposely kept sufficiently simple to allow us to simulate evolutionary time-scales. For this reason, we identify individuals with genotype and neglect the difference between genotype and phenotype. Kaneko (2000) recently studied synpatric speciation in a model that explicitly investigate the effect of phenotype being different from genotype.

We attempt here to gain some insight into the possible effects of co-evolution through the study of a model in which variations in the physical environment are altogether neglected. Our model is not meant to be a realistic representation of biological evolution, but rather a theoretical approximation in which co-evolution is made to be the prominent driving force. We then demonstrate within this model that speciation *does* occur and we study in some detail the dynamical features of the evolution of the model as well as the nature of the ecology created by the co-evolutionary dynamics.

We are interested in the qualitative behaviour of a system in which the mutual interaction between co-existing individuals of different genetic composition determines the possibility of the individual to thrive. The model emphasizes the web of interactions between individuals of different genomic composition, to stress this aspect we will talk about the Tangled Nature model—or the TaNa model for short. We represent biotic factors in terms of the co-evolutionary effects on the fitness of individuals. The model is a simplification. No distinction is made between genotype and phenotype and the details of the reproductive mechanism are kept to a minimum. This simplification allows us to represent evolution in terms of the dynamics of the distribution of the population in genome space. We demonstrate that at a qualitative level the complex dynamics of the model resembles known aspects of long-term biological evolution such as speciation and intermittent behaviour. It is natural to relate these stable epochs to periods of evolutionary stable strategies (ESS), as introduced by Maynard Smith (1982). The stable periods of our model are, however, not perfectly stable as fluctuations caused by mutations can trigger a switch from one stable period to another. We, therefore, suggest calling these periods "quasi-Evolutionary Stable Strategies" or q-ESS. The overall effect of the evolutionary dynamics of the present model is *to increase the average duration of the q-ESS.*

Definition of the Model

INTERACTION

We now define the TaNa model in detail. We represent individuals in the same way as in models considered by, e.g. Kauffman (1995), Higgs & Derrida (1992), Gavrilets (1999), Eigen *et al.* (1988) and Wagner *et al.* (1998). An individual is represented by a vector $\mathbf{S}^\alpha = (S_1^\alpha, S_2^\alpha, \ldots, S_L^\alpha)$ in genotype space. Here, S_i^α may take the values ± 1. These may be interpreted as genes with two alleles, or a string of either pyrimidines or purines. Individuals are labelled by Greek letters $\alpha, \beta, \ldots = 1, 2, \ldots, N(t)$, where $N(t)$ is the total number of individuals at time t. When we refer, without reference to a specific individual, to one of the 2^L positions in genome space, we use roman superscripts $\mathbf{S}^a$, $\mathbf{S}^b$,... with $a, b, \ldots = 1, 2, \ldots, 2^L$. Note, many different individuals $\mathbf{S}^\alpha, \mathbf{S}^\beta, \ldots$, may reside on the same position, say $\mathbf{S}^a$ in genome space. Geometrically, the vector $\mathbf{S}^a$ represents one of the corners of the L dimensional hyper-cube $S = \{-1, 1\}^L$ (see Fig. 1). The ability of an individual α to reproduce is controlled by $H(\mathbf{S}^\alpha, t)$

$$H(\mathbf{S}^\alpha, t) = \frac{1}{N(t)} \sum_{\beta=1}^{N(t)} \sum_{i=1}^{L} J_i(\mathbf{S}^\alpha, \mathbf{S}^\beta) S_i^\alpha S_i^\beta - \mu N(t). \qquad (1)$$

The sum over individuals β in eqn (1) is more conveniently expressed as a sum over the positions $\mathbf{S}$ in the genome space $\mathscr{S}$, using the occupancy $n(\mathbf{S}, t)$ of the positions we obtain

$$H(\mathbf{S}^\alpha, t) = \frac{1}{N(t)} \sum_{\mathbf{S} \in \mathscr{S}} n(\mathbf{S}, t) \sum_{i=1}^{L} J_i(\mathbf{S}^\alpha, \mathbf{S}) S_i^\alpha S_i - \mu N(t). \qquad (2)$$

Two positions $\mathbf{S}^a$ and $\mathbf{S}^b$ in genome space are coupled with the fixed strength $\mathbf{J}^{ab} = \mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$. This coupling is non-zero with probability Θ, in which case we assume $\mathbf{J}^{ab} \neq \mathbf{J}^{ba}$ to be random and uniform on the interval $[-c, c]$, where c is a constant. The structure of the coupling in genome space is sketched in Fig. 1.

Some comments about the interaction matrix $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$ are appropriate. In our simplistic approach, a given genome is imagined to lead uniquely to a certain set of attributes (phenotype) of the individuals/organisms. The positions $\mathbf{S}^a$ and $\mathbf{S}^b$ represent blueprints for organisms that exist *in potentia*. The positions may very likely be unoccupied but, if we were to construct individuals according to the sequences $\mathbf{S}^a$ and $\mathbf{S}^b$ the two individuals would have some specific features. Anecdotally, we can imagine that $\mathbf{S}^a$ corresponds to rabbits and $\mathbf{S}^b$ represents foxes. The number $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$ now represents the potential influence of an individual constructed according to the genome sequence $\mathbf{S}^b$ on an individual constructed according to the genome sequence $\mathbf{S}^a$. In our toy example, $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$ represents the fact that the foxes will tend to eat the rabbits and thereby decrease the rabbits ability to survive and $\mathbf{J}(\mathbf{S}^b, \mathbf{S}^a)$ represents the fact that the availability of rabbits as a food source will help to sustain the foxes. Other examples could be parasitic or collaborative relationships. In order to emphasise co-evolutionary aspects we have *excluded* "self-interaction" among individuals located at the same positions $\mathbf{S}$ in genome space, that is we use $\mathbf{J}(\mathbf{S}, \mathbf{S}) = \mathbf{0}$.

In reality, the mutual influence between two individuals of a certain genotype (phenotype) is, of course, not a random quantity. The interaction may be collaborative, competitive or neutral. It is this aspect, we represent by ascribing a set of fixed randomly assigned coupling strengths between the positions in genome space.

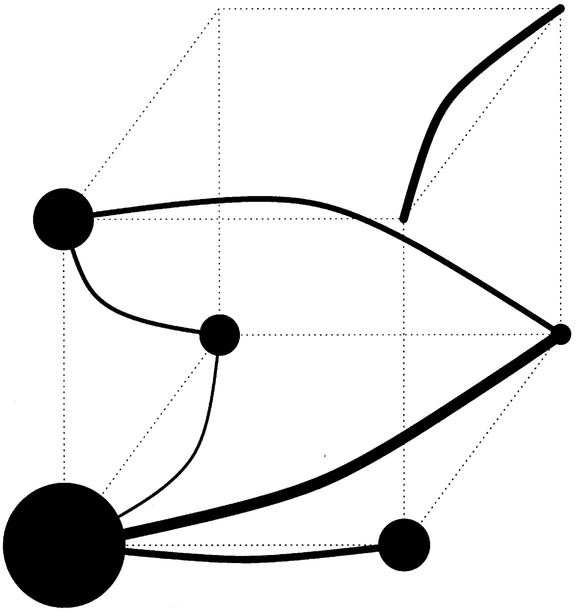


FIG. 1. A three-dimensional genome space. For $L = 3$, the sequence of genes uniquely defines a vertex of a cube. The number of edges (dotted lines) that must be traversed between two vertices defines their Hamming distance. Interactions between vertexes are shown as solid curves with thickness indicating the strength and circles placed at the vertexes have radii proportional to the occupation (number of individuals present) with the genome in question. Note that interactions are defined even for unoccupied vertices.

We stress that the segregation (or speciation), to be discussed below, is an effect of different couplings between different positions $\mathbf{S}^a$ and $\mathbf{S}^b$. When we assume $J_i(\mathbf{S}^a, \mathbf{S}^b) = J_0$ constant independent of $\mathbf{S}^a$ and $\mathbf{S}^b$, we find the population not to be concentrated around a subset of all positions in genome space, instead the population is smeared out through the genome space in a diffuse manner.

The conditions of the physical environment are simplistically described by the term $\mu N(t)$ in eqn (1), where μ determines the average sustainable total population size. An increase in μ corresponds to more harsh physical conditions. This is a simplification, though one should remember that the physical environment encountered by an organism is to some extent produced by the presence of other living organisms. Consider, for example, the environment experienced by the bacterial flora in the intestines. Here, one type of bacteria very much live in an environment strongly influenced by the presence of other types of bacteria. In this sense, some fluctuations in the environment may be thought of as included in the matrix $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$.

REPRODUCTION

Asexual reproduction consists of one individual being replaced by two copies, this event occurs for individuals $\mathbf{S}^\alpha$ with a probability per time unit proportional to

$$p_{off}(\mathbf{S}^\alpha, t) = \frac{\exp[H(\mathbf{S}^\alpha, t)]}{1 + \exp[H(\mathbf{S}^\alpha, t)]} \in [0, 1]. \qquad (3)$$

In the case of sexual reproduction, an individual $\mathbf{S}^\alpha$ is picked at random and paired with another randomly chosen individual $\mathbf{S}^\beta$ with Hamming distance $d = \frac{1}{2}\sum_{i=1}^{L} |S_i^\alpha - S_i^\beta| \leqslant d_{max}$ (allowing at most d_{max} pairs of genes to differ). The pair produces an offspring γ with a probability $\sqrt{p_{off}(\mathbf{S}^\alpha, t)\, p_{off}(\mathbf{S}^\beta, t)}$, with S_i^γ chosen at random from one of the two parent genes, either S_i^α or S_i^β. Choosing genes at random from the parents may be thought of as a process similar to recombination for $d_{max} \geqslant 1$. A maximum separation criterion was used by Higgs & Derrida (1992) to show that this criterion is sufficient to induce speciation in a population evolving in a flat fitness landscape. Gavrilets and collaborators (Gavrilets *et al.*, 1998, 2000; Gavrilets, 1999) used the separation criterion to study parapatric speciation in holy adaptive landscapes, i.e. fitness landscapes with a large degree of neutrality.

Mutation

Genes mutate with probability p_{mut}, represented by a change of sign $S_i^\gamma \to -S_i^\gamma$, during the reproduction process.

Annihilation

For simplicity, an individual is removed from the system with a constant probability p_{kill} per time step. This procedure is implemented both for asexual and sexually reproducing individuals.

Time Step

A time step consist of *one* annihilation attempt followed by one reproduction attempt. One generation consists of $N(t)/p_{kill}$ time steps, the average time taken to kill all currently living individuals.

Stability

At an average level of description, and neglecting mutations, the above dynamics is described by the following set of equations (one equation for each position in the genotype space):

$$\frac{\partial n(\mathbf{S}, t)}{\partial t} = [p_{off}(\mathbf{S}, t) - p_{kill}]\, n(\mathbf{S}, t) \qquad (4)$$

controlling the temporal evolution of the occupancy $n(\mathbf{S}, t)$ of the positions $\mathbf{S}$ in genotype space $\mathcal{S}$. Stationary solutions (i.e. those for which $\partial n/\partial t = 0$) demand either $n(\mathbf{S}, t) = 0$ or $p_{off}(\mathbf{S}, t) p_{kill}$. During the q-ESS, the system manage to find a configuration in genotype space for which all occupied positions satisfy the balance between production of offspring and decease. The fitness $p_{off}(\mathbf{S}^a, t)$ of individuals at a position $\mathbf{S}^a$ depends on the occupancy $n(\mathbf{S}^b, t)$ of all the sites $\mathbf{S}^b$ with which site $\mathbf{S}^a$ is connected through couplings $\mathbf{J}^{ab}$. Accordingly, a small perturbation in the occupancy at one position is able to disturb the balance between $p_{off}(\mathbf{S}, t)$ and p_{kill} on connected sites. In this way, an imbalance at one site can spread as a chain reaction through the system, possibly affecting a global reconfiguration of the genotypical composition of the population.

Dynamical Behaviour

We consider two different types of populations: (1) a purely asexual population and (2) a purely sexual population.

ASEXUAL REPRODUCTION

In this subsection, we discuss the model when all reproduction is assumed to be asexual.

INITIATION

Let us consider the initiation of the model. First, we place the entire population $N(0)$ at a randomly chosen location $\mathbf{S}^*$ in genome space. The H-function in eqn (2) will be given by $H(\mathbf{S}^*, 0) = -\mu N(0)$ since $n(\mathbf{S}) = 0$ for $\mathbf{S} \neq \mathbf{S}^*$ and $\mathbf{J}(\mathbf{S}^*, \mathbf{S}^*) = 0$. If no mutations can occur the population will remain confined at the location $\mathbf{S}^*$ and the size of the population $n(\mathbf{S}^*, t)$ will according to eqn (4) approach the value

$$N^* = \frac{1}{\mu} \ln\left(\frac{1 - p_{kill}}{p_{kill}}\right).$$

Mutations do occur, however, and the population will migrate away from the original location $\mathbf{S}^*$ into the surrounding region of genome space. In Fig. 2, we show a cladogram indicating the evolution of the first 110 generations. During this initial period, the newly invaded positions are only occupied for a few generations (indicated by the short horizontal lines in Fig. 2). After this period of rapid changes, a relatively stable configuration is achieved, and the occupied positions to the right in Fig. 2 indicate that the system has entered its first q-ESS.

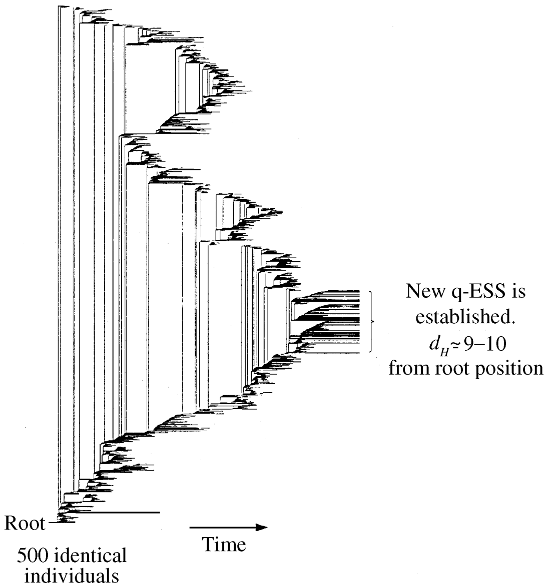


FIG. 2. The initial diversification from a single position in genome space. The system is initialized at time $t = 0$ with 500 identical individuals and allowed to develop autonomously. Time is plotted horizontally. Similar to ordinary cladograms different genotypes are located at different vertical positions. Vertical lines represent parentages. Horizontal lines starts at the time a genotype is created and stops when the genotype becomes extinct. The system mutates away from the initial location, which becomes extinct relatively quickly. After 34 branchings the system finds a stable configuration and enters the first q-ESS (see Figs 3 and 4).

We have also studied simulations started out from an initial population spread out over many randomly chosen positions in genome space. Most of these initially occupied positions rapidly become extinct. In this way, the diversity in genome space passes through a "bottleneck" before the population starts to migrate out into genome space from one or a few positions which were able to pass through the bottleneck. From then on, the evolution of the ecology behaves in the same way as when started out from one single position in genome space.

LONG TIME BEHAVIOUR

Now, we turn to a discussion of the nature of the long time dynamics of the model. The model consists of a variable number of co-evolving individuals all subject to the same physical environment. An individual's ability to thrive depends on its own genetic composition as well as the genetic composition of the other individuals present. The dynamical evolution, driven by mutations, will have to strike a balance between the multiplication of the individuals and the total carrying capacity of the environment. Different types of genotypical compositions of the population can achieve this balance.

One possibility consists of very numerous populations distributed on a relatively small number of isolated regions in genotype space corresponding to a small number of species (compared to the total number of genotypes for a given genome length). These configurations can be stable for very many generations and allow the species to co-exist quietly during coherent periods of little variation in the total size or composition of the population, see Fig. 3. In Fig. 4, we demonstrate that the occupancy of the positions in genome space fluctuates only very little during these stable periods. We call these epochs q-ESS, or quasievolutionary stable strategies (Maynard Smith, 1982). The q-ESS exhibit a degree of stability against mutation-induced changes, but fluctuations in the frequency distribution in genome space can abruptly destabilize such a configuration. We show, however, in Fig. 5 that the distribution of durations of the q-ESS, measured in numbers of generations, is very broad.

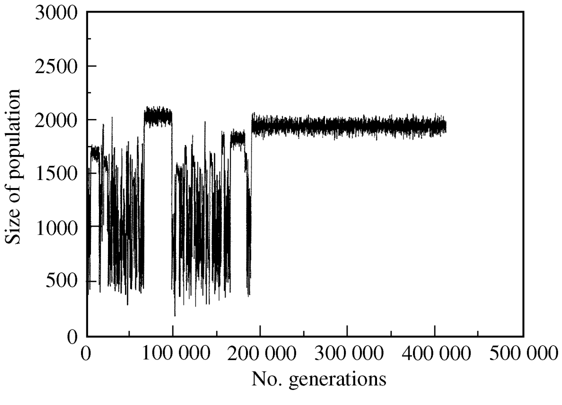


FIG. 3. Change of total population, $N(t)$, with time for a system with $L = 20$, $\mu = 0.1375$, $p_{kill} = 0.2$, $p_{mutate} = 0.01$, $c = 100$ and $\theta = 0.25$. Regions of high population and low relative fluctuations (q-ESS) are clearly distinct from regions of low population and high relative fluctuations (transition periods).

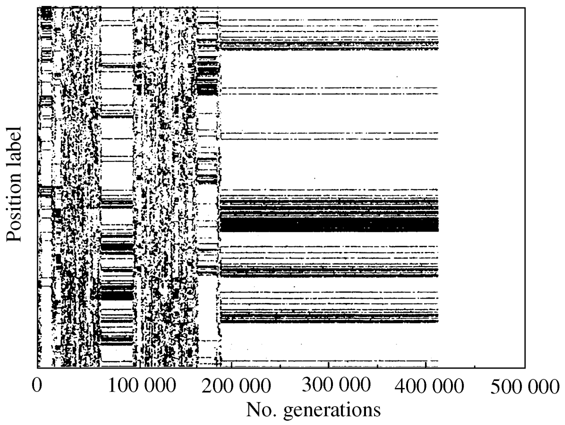


FIG. 4. Occupation of genome space vs. time for the same simulation as in Fig. 3. We arrange the positions in genome space in a convenient arbitrary way along the y-axis and place a dot for each occupied location at a given time. Periods of stability (q-ESS) interrupted by periods of hectic rearrangement are clearly visible.

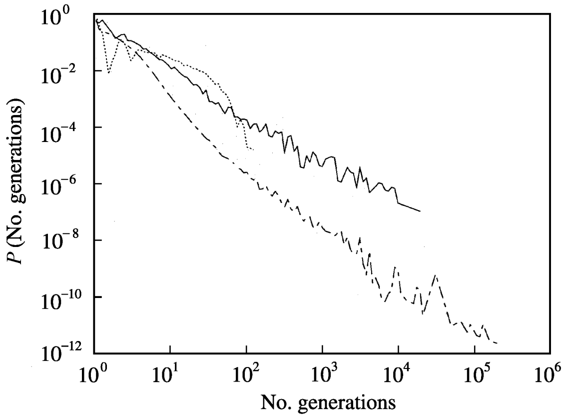


FIG. 5. Log–log plot of the distributions of lengths of q-ESS (solid line), transition periods (dotted line) and lifetimes of occupied locations in genome space (dot-dashed line). We observe power-law-like behaviour in both the q-ESS and the lifetimes of genome space locations, but the transition periods exhibit an abrupt cut-off at much shorter times. The lifetimes curve extends further than that of the q-ESS, indicating that locations may remain occupied from one q-ESS to another, surviving the transition.

TRANSITIONS

The q-ESS periods are separated by periods of hectic rearrangements of the genotypical composition of the entire population. During these periods of rapid change, the total number of individuals is smaller and populations located at specific positions in genome space undergo sequences of bifurcations as seen in Fig. 6, where we follow the evolution across the hectic transition period from one q-ESS to the next. The figure is a cladogram tracing out all the descendants originating from one root. One notices that most of the new positions spun off from the root die before the next q-ESS is reached. While new branches are created old ones die. The periods of rapid rearrangement in genome space are transition periods during which the system searches for a new stable configuration.

The nature of the transition from one q-ESS to the next is indicated in Fig. 7. This set of diagrams represents in a quantitative way the positions with the largest occupation together with the couplings in genome space.

Gene sequences in the diagrams are as follows:

$$
\begin{aligned}
\mathbf{S}^1 &= (+1,+1,+1,-1,-1,+1,-1,+1,-1,-1,+1,-1,+1,-1,+1,-1,-1,-1,+1,-1),\\
\mathbf{S}^2 &= (+1,+1,-1,-1,-1,+1,-1,+1,+1,+1,-1,-1,+1,-1,+1,-1,-1,-1,+1,-1),\\
\mathbf{S}^3 &= (+1,+1,-1,-1,-1,+1,-1,-1,-1,+1,+1,-1,-1,-1,-1,-1,-1,-1,-1,-1),\\
\mathbf{S}^4 &= (+1,+1,-1,-1,-1,+1,-1,+1,+1,+1,-1,-1,-1,-1,+1,+1,-1,-1,+1,-1),\\
\mathbf{S}^5 &= (+1,+1,-1,-1,-1,+1,-1,+1,+1,+1,-1,-1,-1,-1,+1,+1,-1,-1,+1,-1),\\
\mathbf{S}^6 &= (+1,+1,+1,-1,-1,+1,-1,-1,-1,-1,+1,-1,+1,+1,+1,-1,-1,-1,+1,-1),\\
\mathbf{S}^7 &= (+1,+1,-1,-1,-1,+1,-1,+1,+1,+1,+1,-1,+1,-1,+1,-1,+1,-1,+1,+1),\\
\mathbf{S}^8 &= (+1,+1,-1,-1,-1,+1,-1,+1,+1,+1,+1,-1,+1,-1,+1,-1,+1,-1,-1,+1),\\
\mathbf{S}^9 &= (-1,+1,+1,-1,-1,+1,-1,+1,-1,-1,+1,-1,+1,-1,-1,-1,-1,+1,+1,-1),\\
\mathbf{S}^{10} &= (-1,-1,-1,-1,-1,+1,-1,+1,-1,-1,-1,+1,+1,-1,-1,-1,-1,+1,+1,-1),\\
\mathbf{S}^{11} &= (-1,+1,-1,-1,-1,+1,-1,+1,+1,-1,+1,-1,-1,-1,+1,-1,+1,-1,-1,+1),\\
\mathbf{S}^{12} &= (-1,+1,-1,+1,-1,+1,-1,+1,+1,-1,+1,-1,+1,-1,+1,-1,+1,-1,-1,+1),\\
\mathbf{S}^{13} &= (-1,-1,-1,-1,-1,+1,-1,+1,-1,-1,-1,+1,+1,-1,-1,-1,-1,+1,+1,-1).
\end{aligned}
$$

During a transition between one q-ESS and the next, the systems behaviour becomes very hectic. Starting at (a), we see that a new mutant $\mathbf{S}^4$ has invaded the previous coherent configuration [originally similar to (b)] with a negative interaction with most of the existing flowers but a strong enough positive interaction with one of them for it to survive. This causes the coherent state to be destabilized. In (b), we see that although the new mutant does not survive for long, it has drastically reduced the population at $\mathbf{S}^1$, which in turn has a harmful effect on $\mathbf{S}^2$ and $\mathbf{S}^3$. In (c), we observe that two further new mutants, $\mathbf{S}^4$ and $\mathbf{S}^5$ have been able to invade, this is due to the reduced fitness of the original sites from the effect of the first invader. The new mutants are transient, they represent steps on an adaptive walk. The system is now in a situation where it is partly executing such a walk, and partly still in the previous coherent state. This continues into (d), where we can see that $\mathbf{S}^1$ and $\mathbf{S}^3$ are still holding on, and their complete first circles evince they are still reproducing. By (e), however, things have changed again. The adaptive walkers are now out-competing the originals, $\mathbf{S}^1$ has become extinct and S^3 has a very low population. We also observe the formation of a double flower ($\mathbf{S}^7$ and $\mathbf{S}^8$) which consists of two fit centres in adjacent locations in genome space. The adaptive walk continues for some time until a new q-ESS is found at (f).

We have studied the distribution of non-zero couplings $\mathbf{J}(\mathbf{S}^a, \mathbf{S}^b)$ between a given occupied position $\mathbf{S}^a$ and another occupied position $\mathbf{S}^b$. During the q-ESS this distribution is narrow and its average is smaller than during the transition periods, where the distribution broadens.

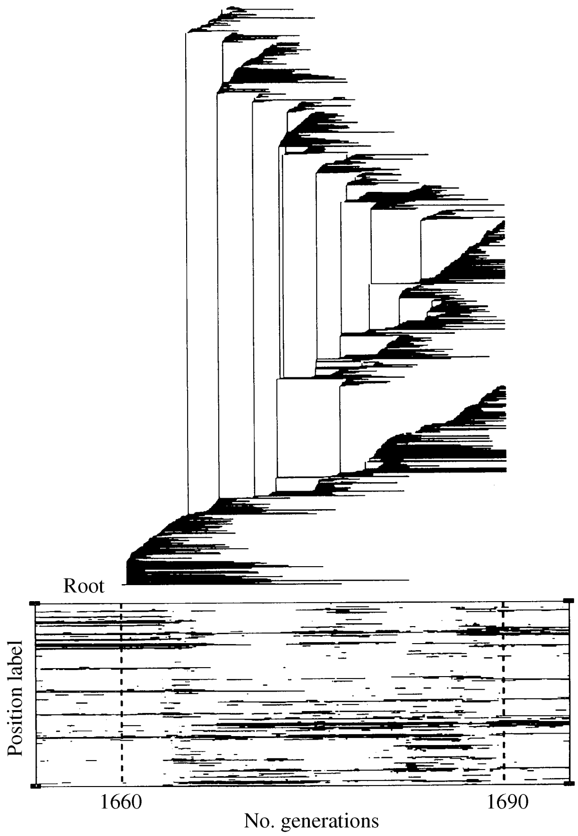


FIG. 6. The descendants of a fit location across a transition period. The lower portion of this figure is a slice from an occupation plot, similar to Fig. 4. We track the descendants of a single, fit location across the transiton and into the next q-ESS. After 12 branchings the descendants have found a new fit configuration that is stable enough to form part of the new q-ESS. The original location does not survive the transition and lineages of the other fit locations from the original state die out very rapidly.

EPOCH DISTRIBUTIONS

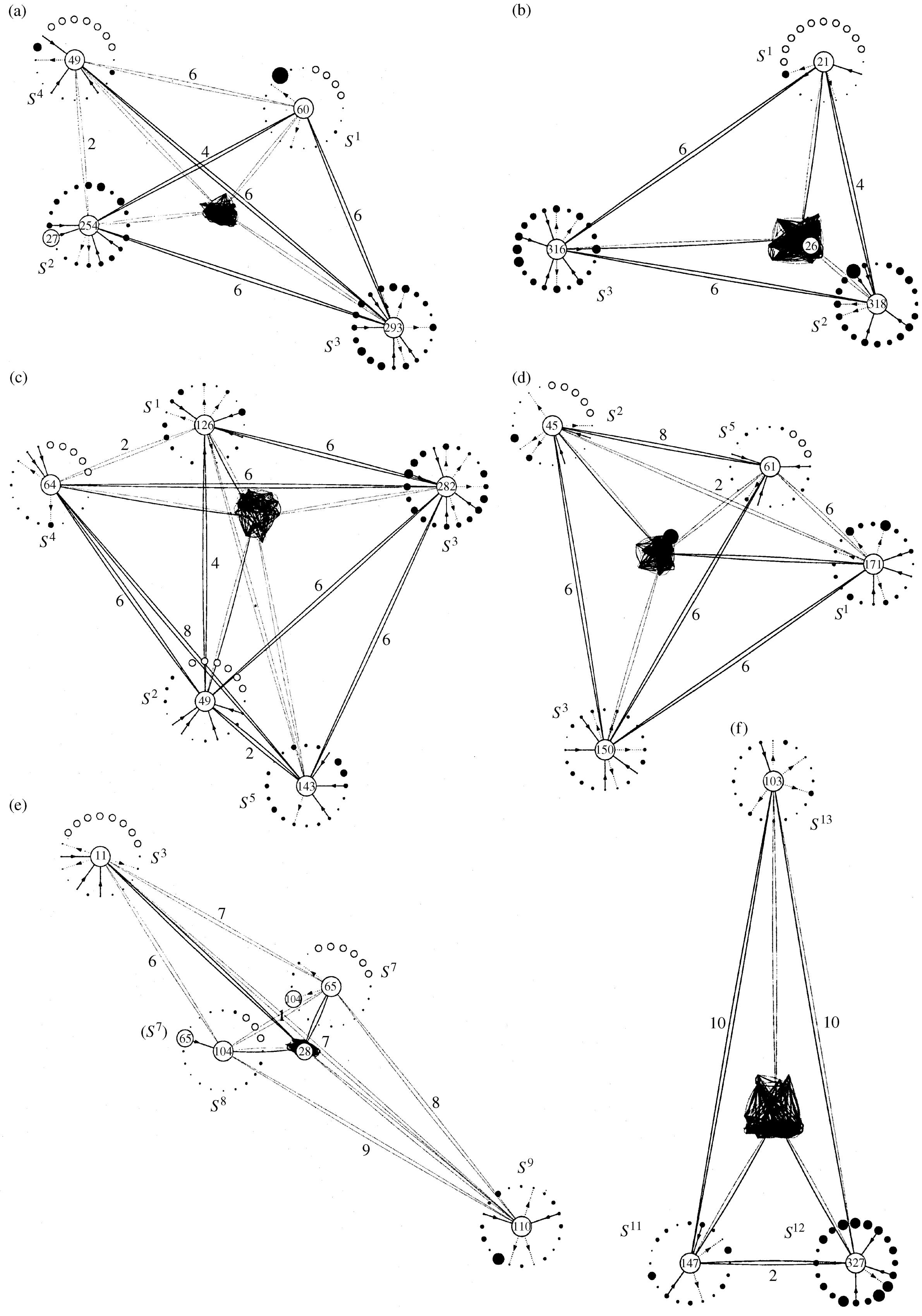
It is interesting to take a further look at Fig. 5. One notices that the distribution of lifetimes of occupied positions reaches as far out as the distribution of q-ESS durations. In fact, we observe in the simulations that positions sometimes are able to remain occupied across the transition from one q-ESS to the following, corresponding to a species that survives a mass extinction. Figure. 5 also shows clearly that the periods of hectic reconfiguration typically last for a significantly smaller number of generations than do the q-ESS periods. Finally, it is very interesting to note that both the lifetimes of individual positions and the distribution of q-ESS epoch lengths are power-law-like with exponents around -2.3 and -1.8, respectively. We mention that the distribution of q-ESS durations can be compared to the distribution of lifetimes of genera obtained from the fossil record. The latter has a shape similar to the distribution of q-ESS durations shown in Fig. 5. Power-law fits to the fossil record data leads to an exponent around 2. For a recent analysis of data from the fossil record, see Newman & Sibani (1999).

ADAPTATION LEVEL INCREASES

We now turn to a discussion of the overall long time effect of the dynamics of the TaNa model. How does the genomic composition of early configurations differ from those generated after hundreds of thousands of generations? In Fig. 8, we show the running average of the durations of the q-ESS as well as the transition epochs. One notices that there is no significant, trend in the duration of the hectic periods of rearrangement separating the consecutive q-ESS. The average duration of the q-ESS periods, however, slowly increases with time. This means that the entire ecology gradually becomes more stable. Or we may say that the ecology (represented by the distribution of the population through genome space) becomes increasingly better adapted; not adapted to some fixed external environment, but adapted in the sense that the ecology as a whole achieves collectively increasingly stable configurations among the total set of all possible ways of distributing a

→

FIG. 7. Flower diagrams of the configuration of a system during a transition from one coherent state to another. Flower diagrams visualize the interactions and genome space proximities in a system at a given time. Very fit locations have a complete set of nearest mutational neighbours (these make up a "flower"). Each occupied location is represented by a circle of radius proportional to its occupation (or a number, for very large locations) positive interaction strengths are solid lines, and negative interaction strengths are dotted lines. If a flower is out-competed by a new mutant, the q-ESS is disturbed and the system executes an adaptive walk, searching for a new q-ESS. Each diagram is separated by approximately 5–10 generations. See the main text for discussion. The Hamming distance between two highly occupied positions appears midway between them.


(a)
49
S^4
6
60
S^1
2
4
6
6
254
27
S^2
6
293
S^3
(b)
S^1
21
6
4
316
26
S^3
6
318
S^2
(c)
S^1
126
2
6
64
6
282
S^4
S^3
4
6
6
6
8
S^2
49
2
143
S^5
(d)
45
S^2
8
S^5
61
2
6
171
6
6
S^1
6
S^3
150
(e)
11
S^3
7
6
65
S^7
104
1
(S^7)
65
104
28
7
S^8
8
9
S^9
110
(f)
103
S^{13}
10
10
S^{11}
S^{12}
147
2
327

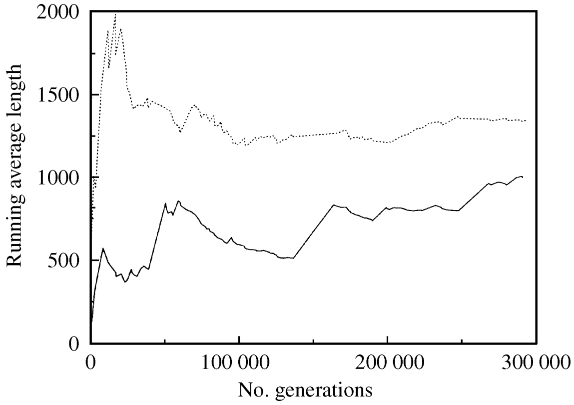


FIG. 8. Running averages of the duration of q-ESS (solid curve) and transition periods (dotted curve). At the end of a particular state, we evaluate the average length of states up until that point. We see that the average length of the transition periods settles down and fluctuates slightly around constant, whereas the average q-ESS length continues to increase.

population through genome space. Does this mean that eventually some maximally "fit" or adapted configuration is reached? Our simulations indicate, as expected, that the time to reach a stationary state increases exponentially with increasing genome length L. We will accordingly expect that for biologically relevant systems an ecology would never have the time to reach a final stationary state. Moreover, even if the system becomes stationary in the sense that the average duration of the q-ESS becomes time independent, switching between different equally well-adapted configurations is likely to continue forever. From the statistical mechanics of disordered systems, we do not expect the optimally adapted configuration to be unique. Hence, transitions between equally maximally adapted configurations may continue even in the mathematical limit of infinitely long time.

The increase of the average duration of the q-ESS can be viewed as an optimizing process. This is in accordance with the suggestion (Mayr, 1988) that the effect of biological evolution is to optimize some quantity. The identification of the quantity being optimized is still debated (Fogel & Beyer, 2000). Unfortunately, we cannot identify a specific mathematical function of the distribution $n(\mathbf{S}, t)$ in genome space which is optimized as an effect of the dynamics. However, it is very interesting to relate the average duration of the q-ESS to the extinction rate. Due to insufficient statistics we cannot, unfortunately, make a quantitative comparison. We note, qualitatively, that an increasing average duration of the q-ESS corresponds to a decreasing extinction rate. This is consistent with the Raup & Sepkoski (1982) analysis of the fossil record, which suggests that the extinction rate might have declined through the Phanerozoic.

Sexual reproduction

We now briefly discuss a model in which all individuals are assumed to reproduce sexually. More detail will be presented in a future communication.

LONG TIME BEHAVIOUR

In Fig. 9, we show the occupancy in genome space at the top and the temporal behaviour of the total number of individuals together at the bottom. We have assumed that the maximum number of genes in which parents can differ, $d_{max} = 2$, and in Fig. 9, we plot only the species occupancy, that is we have coarse-grained genome space with a resolution of d_{max}. This is done in the following way: in each time step, we identify the position with the largest population, we lump this position together with all positions within a distance d_{max}. Next, we find the position with the second largest population, and lump this location together with all positions within distance d_{max}. We continue this until all occupied positions have been considered. The locations in genome space are labelled in a convenient, but arbitrary, way. For each time step, we place a dot along the y-axis for each occupied (coarse grained) positions in genome space. Finally, along the x-axis, we convert time steps into time measured in generations. We observe that, similar to the asexual case, the model evolves through a set of q-ESS phases separated by short transition periods. We also emphasize that well-established species can be identified as isolated occupied locations in genome space separated by a distance larger than d_{max}. However, clusters of a diameter larger than d_{max} of adjacent occupied genotypes can also occur. This corresponds to a cluster of subspecies. Even configurations corresponding to ring species are sometimes observed.

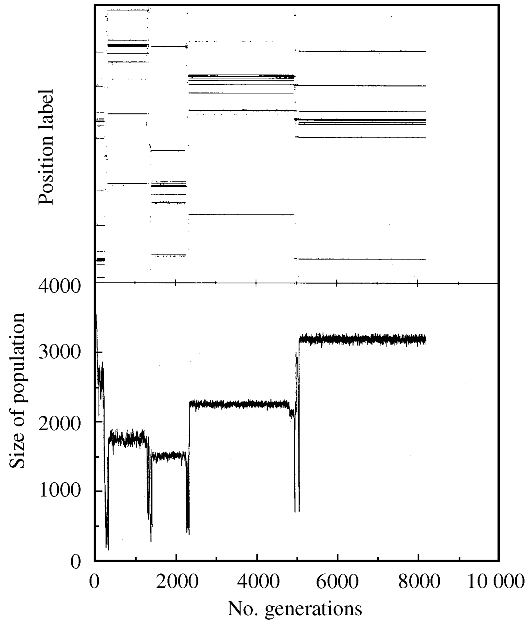


FIG. 9. The total number of individuals (below) and the occupation plot of the species (above). The horizontal axis is the generational time. The different plateaus in the population size correspond to rearrangements of the population in genome space. Parameters are $L = 20$, $\mu = 0.07$, $p_{kill} = 0.2$, $p_{mutate} = 0.01$, $c = 100$ and $\theta = 0.25$.

For a recent observational study of ring species see Irwin *et al.* (2001).

LIFETIME STATISTICS

In Fig. 10, we show the distribution of lifetimes of occupations of individual multiple occupied positions in genome space. A slow power-law-like decay is observed. Note the similarity with the distribution found in the asexual case and with the distributions reported from the fossil record, see e.g. Newman & Sibani (1999).

Discussion and Conclusion

We have discussed a simple, very general mathematical metaphor [see Gavrilets (1999) for a very interesting discussion of the significance to evolutionary theory of mathematical metaphors] by which we can study the long time (of order 10^5 or 10^6 generations) behaviour of an ecology.

Both asexual and sexually reproducing populations evolve through a set of relatively stable configurations, the q-ESS, separated by short transition periods of hectic reorganization of the genomic composition of the ecology. The population segregates in genome space into well separated clusters of highly occupied positions. Speciation events occur when a position or a tight cluster of positions undergoes successive bifurcations in genome space. This type of behaviour is observed for a broad range of control parameters.

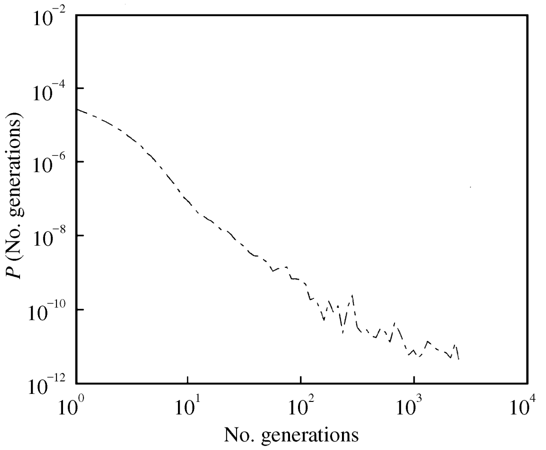


FIG. 10. Distribution of position's lifetime (log–log plot) for the sexual case. Same simulation as in Fig. 9.

The co-evolutionary dynamics produce a highly tangled interdependent population of species. The evolution gradually increases the robustness of the entire ecology against fluctuations in the genomic and physical environment. In agreement with analysis of the fossil record, we find that the average duration of the q-ESS increases slowly with time.

We have benefited greatly from discussions with A. Burt, A.M. Lerio, R.R.B. Azevedo and, at a very early stage, from conversations with C. Godfray. M.H. and S.A.C. are both supported by research studentships from the EPSRC and K. C. acknowledges the support from a computational EPSRC grant.

REFERENCES

BAK, P. & SNEPPEN, K. (1993). Punctuated equilibrium and criticality in a simple model of evolution. *Phys. Rev. Lett.* **71,** 4083–4086.

BJØRNSTAD, O. N., SAIT, S. M., STENSETH, N. C., THOMPSON, D. J. & BEGON, M. (2001). The impact of specialized enemies on the dimensionality of host dynamics. *Nature* **409,** 1001–1006.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*, 1951 Edn. John Murray. Oxford: Oxford University Press.

Doncaster, C. P., Pound, G. E. & Cox, S. J. (2000). The ecological cost of sex. *Nature* **404,** 281–285

Eigen, M., McCaskill, J. & Schuster, P. (1988). Molecular quasi-species. *J. Phys. Chem.* **92,** 6881–6891.

Fogel, D. B. & Beyer, H.-G. (2000). Do evolutionary processes minimize expected losses? *J. theor. Biol.* **207,** 117–123.

Gavrilets, S. (1999). A dynamical theory of speciation on holey adaptive landscapes. *Am. Nat.* **154,** 1–22.

Gavrilets, S. & Hastings, A. (1998). Coevolutionary chase in two-species systems with applications to mimicry. *J. theor. Biol.* **191,** 415–427.

Gavrilets, S., Li, H. & Vose, M. D. (1998). Rapid parapatric speciation on holey adaptive landscapes. *Proc. R. Soc. London* **265,** 1483–1489.

Gavrilets, S., Li, H. & Vose, M. D. (2000). Patterns of parapatric speciation. *Evolution* **54,** 1126–1134.

Higgs, P. G. & Derrida, B. (1992). Genetic distance and species formation in evolving populations. *J. Mol. Evol.* **35,** 454–465.

Irwin, D. E., Bensch, S. & Proce, T. D. (2001). Speciation in a ring. *Nature* **409,** 333–337.

Kaneko, K. (2000). Sympatric speciation: compliance with phenotype diversification from a single genotype. *Proc. R. Soc. London B* **267,** 2367–2373.

Kauffman, S. (1995). *Coevolution,* In: At Home in the Universe. Chapter 10. London: Viking.

May, R. M. & Anderson, R. M. (1983). Epidemiology and genetics in the coevolution of parasites and hosts. *Proc. R. Soc. London* **219,** 281–313.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Maynard Smith, J. (1995). *The Theory of Evolution*. Cambridge: Cambridge University Press.

Mayr, E. (1988). *Toward a New Philosophy of Biology: Observations of an Evolutionist*. Harvard, MA: Belknap.

Newman, M. E. J. & Sibani, P. (1999). Extinction, diversity and survivorship of taxa in the fossil record. *Proc. R. Soc. London B* **266,** 1593–1599.

Pimm, S. L. (1991). *The Balance of Nature*. University of Chicago Press.

Raup, D. M. & Sepkoski Jr, J. J. (1982). Mass extinction in marine fossil record. *Science* **215,** 1501–1503.

Seger, J. (1988). Dynamics of some simple host–parasite models with more than two genotypes in each space. *Philos. Trans. R. Soc. London* **319,** 541–555.

Taylor, C. F. & Higgs, P. G. (2000). A population genetics model for multiple quantitative traits exhibiting pleiotropy and epistasis. *J. theor. Biol.* **203,** 419–437.

Wagner, H., Baake, E. & Gerische, T. (1998). Ising quantum chain and sequence evolution. *J. Stat. Phys.* **92,** 1017–1052.