Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Theoretical Biology 235 (2005) 393–401
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Stochastic evolutionary dynamics on two levels

Arne Traulsen[a,*], Anirvan M. Sengupta[b], Martin A. Nowak[c]

[a]Institut für Theoretische Physik und Astrophysik, Christian-Albrechts Universität, Olshausenstr. 40, D-24098 Kiel, Germany
[b]Department of Physics and Astronomy, BioMaPS Institute, Rutgers, The State University of New Jersey, 136 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA
[c]Program for Evolutionary Dynamics, Department of Mathematics, Department of Organismic and Evolutionary Biology, Harvard University, One Brattle Square, Cambridge, MA 02138, USA

Received 12 October 2004; received in revised form 21 January 2005; accepted 25 January 2005
Available online 13 March 2005
Communicated by Lewis Wolpert

Abstract

We consider a population that is subdivided into groups. Individuals reproduce proportional to their fitness. When a group reaches a certain size it has a probability to split into two groups while another group is eliminated. In this stochastic process, the number of groups is constant, while the total population size fluctuates between well-defined bounds. We calculate the fixation probability of newly introduced mutants under constant selection. We show that the described population structure acts as a suppressor of selection compared to an unstructured population of the same size. The maximum suppression of selection is obtained, when the number of groups equals the number of individuals per group. We also study opposing selection on two or more levels by analysing the evolutionary dynamics of hierarchically embedded Moran processes.

Keywords: Group selection; Moran process

1. Introduction

There is much interest in studying selection on multiple levels (Michod, 1999; Michod and Roze, 2001; Gould and Lloyd, 1999; Keller, 1999; Swenson et al., 2000; Kerr and Godfrey-Smith, 2002; Boyd and Richerson, 2002). Multi-level selection arises whenever reproducing units form aggregates that also reproduce. For example, genes assemble in cells, cells form multicellular organisms, animals form social groups. Multi-level selection is related to kin selection (Hamilton, 1964), spatial selection (Nowak and May, 1992, 1993; Lindgren and Nordahl, 1994; Killingback and Doebeli, 1996; Nakamaru et al., 1997, 1998; Mitteldorf and Wilson, 2000; Irwin and Taylor, 2001; Hauert and Doebeli, 2004) and to the long standing debate of group selection (Wynne-Edwards, 1962; Maynard Smith, 1964; Williams, 1966; Wilson, 1975; Maynard Smith, 1976; Slatkin and Wade, 1978; Wilson, 1979; Goodnight, 1997; Sober and Wilson, 1998).

Wynne-Edwards first argued that animals might limit their breeding in order to prevent the population from over-exploiting its resources (Wynne-Edwards, 1962). This could lead to a higher survival probability of populations that produce less offspring. Williams criticized this overuse of group selection, emphasizing that in general individual selection is stronger than group selection (Williams, 1966). This criticism led to a denial of group selection for the following years.

During this debate, a wealth of models of group selection has been developed. In early models, new groups have been formed as a random sample from a

*Corresponding author. Tel.: +49 431 880 3899; fax: +49 431 880 4094.
E-mail addresses: traulsen@theo-physik.uni-kiel.de (A. Traulsen), anirvans@physics.rutgers.edu (A.M. Sengupta), martin_nowak@harvard.edu (M.A. Nowak).

0022-5193/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2005.01.019

migrant pool in which individuals from all groups are mixed in each generation (Maynard Smith, 1964; Levins, 1970; Levin and Kilmer, 1974). In this framework, group selection is usually less effective than individual selection. However, empirical work has shown that artificial group selection can be highly effective when new groups are derived from a single parent group (Wade, 1977, 1980; Craig and Muir, 1996; Muir, 1996; Goodnight, 1997).

Wade (1977) carried out experiments with the red flour beetle, where population sizes are mainly governed by cannibalism rates. In his experiments, he considered the development of several subpopulations, each starting with the same number of adults. After a certain period, he selected the largest populations to contribute to the next generation. In contrast to earlier models, he did not derive the new subpopulations from a migrant pool, but from a single group. Wade's experiments showed that group selection can be much more effective than previously thought. A comparison between different group formation mechanisms provided a theoretical explanation for the effectiveness of group selection in these experiments (Slatkin and Wade, 1978). More recently, additional experiments confirmed the effectiveness of artificial group selection (Craig and Muir, 1996; Muir, 1996; Swenson et al., 2000).

Group selection can also be relevant for parasite evolution. In general, the optimum level of virulence maximizes the parasite's basic reproductive ratio, R_0 (Anderson and May, 1979, 1981, 1982). Evolutionary dynamics within an infected host, however, can favor mutants that do not maximize R_0 (Nowak et al., 1991; Bonhoeffer and Nowak, 1994; Levin and Bull, 1994; Nowak and May, 1994; May and Nowak, 1994, 1995; Frank, 1998; Maynard Smith, 1998b). Hence, there is a conflict of selection on two levels. The parasite population in any one host can be seen as a group. Group selection can favor the cooperation of parasites in a host and thereby lead to virulence levels which are optimal for the parasite.

Wilson (1975) presented an elegant calculation demonstrating how altruistic behavior can evolve in subdivided populations. In Wilson's model, however, altruistic traits can only evolve if cooperators are more likely to form groups with other cooperators than with defectors. Therefore, in Wilson's original model the key mechanism that leads to cooperation is assortative interaction of cooperators rather than group selection per se (Charlesworth, 1979; Bergstrom, 2002; Sober and Wilson, 1998; Wilson and Dugatkin, 1997). Such assortative interaction can also be linked to kin selection (Hamilton, 1963, 1964; Wilson, 1979; Maynard Smith, 1976, 1998a).

Michod (1983), Szathmáry and Demeter (1987) and Maynard Smith and Szathmáry (1995) use group selection arguments in the context of the origin of life.

Here we study a model that is motivated by the idea of (proto-)cell division. A population of finite size is subdivided into groups (cells). Individuals are chosen for reproduction at random from the whole population but proportional to their fitness. If a group (cell) reaches a certain maximum size it is split into two groups. If this happens another group is eliminated. Thus the total number of groups remains constant. Our formulation is an extension of the classical Moran process (Moran, 1962; Ewens, 2004) toward studying selection on two levels. We analyse a hierarchy of two (or more) Moran processes at the individual and group level. In our model, new groups are generated by splitting existing groups. In this way, we follow the approach of Slatkin and Wade (1978) and Leigh (1983) where new groups are derived from a single parent group. Our model is also inspired by the celebrated "stochastic corrector" (Szathmáry, 1986; Szathmáry and Demeter, 1987; Grey et al., 1995) but very different in detail. We aim to formulate the simplest possible approach for studying stochastic evolutionary dynamics on multiple levels. We obtain analytic results in some limits.

The paper is organized as follows: First, we recall the Moran process for population dynamics within finite populations. Then we introduce a structured population which works as a suppressor of selection. We give heuristic arguments for the average group size of this model and derive a rough approximation for the fixation probability in the limit of large population size. Another analytical approximation is presented which is based on a hierarchy of two Moran processes. We discuss mechanisms in which selection on the two levels of evolutionary dynamics work in the same direction and mechanisms in which they oppose each other. We extend the model to selection on more than two levels.

2. The Moran process

Let us briefly review selection dynamics in a finite population as described by the Moran process (Moran, 1962). There are two types of individuals, A and B, with relative fitnesses r and 1, respectively. At each time step, an individual is randomly selected for reproduction with a probability that is proportional to its fitness. The offspring replaces a randomly chosen individual. The total population size, n, remains constant. The resulting birth–death process on the states $i = 0, \ldots, n$ is described by the tri-diagonal transition matrix

$$P_{i,i+1} = \frac{ri}{ri + n - i}\,\frac{n - i}{n}, \tag{1}$$

$$P_{i,i-1} = \frac{n - i}{ri + n - i}\,\frac{i}{n}, \tag{2}$$

$$P_{i,i} = 1 - P_{i,i+1} - P_{i,i-1} \tag{3}$$

All other entries are 0. There are two absorbing states, 0 and n. The probability to end in state n starting from state 1 denotes the probability that a single individual of type A will invade and take over the whole group. This fixation probability is given by (Moran, 1962; Karlin and Taylor,1975)

$$\rho_A = \frac{1-1/r}{1-1/r^n}. \tag{4}$$

For neutral mutants with $r=1$ we have $\rho = 1/n$. The probability that an advantageous mutant with $r>1$ fixates in a finite population converges to $1-1/r$ for large group sizes. Disadvantageous mutants can also reach fixation in finite populations, but this probability converges to zero as n increases.

3. A simple stochastic process of selection on two levels

Let us now consider a finite population that is subdivided into m groups. Each group has a maximum size of n. As before, A individuals have relative fitness r, while B individuals have fitness 1.

In each time step, a randomly chosen individual from the whole population is selected for reproduction, proportional to fitness. The offspring is added to that individual's group. If the new group size (which was augmented by one) is still less than or equal to n, then nothing else happens and the step ends. If, however, the new group size exceeds n, then the group is divided into two groups. The members of the parent group are randomly distributed over the two daughter groups. A randomly chosen group is eliminated. Levins (1970) called such a process "colonization selection" , because selection occurs when new groups are formed.

In this stochastic process, the total number of groups is constant and given by m. The total population size can fluctuate between a minimum value, m, and a maximum value, mn: there are m groups, each group contains at least 1 and at most n individuals.

Note that $n=1$ implies that the total population size is strictly constant. In this case, our process is equivalent to a Moran process with population size m.

When a group is split, each individual has a probability of $\frac{1}{2}$ to end up in the first of the two offspring groups. This "binomial sampling without replacement" has a probability of $1/2^{n-1}$ to lead to the situation where one of the groups remains empty. In order to avoid splitting a group into itself and an empty group, we repeat the group division process until each of the two groups is non-empty.

3.1. Average population size

What is the average population size of this process? For the case of neutral mutants, $r=1$, we can derive a rough approximation in the limit of large populations. For $m \gg n \gg 1$, we can assume that a large number of groups of each accessible size exists. In a single step of the process, the number of groups of a certain size can only change by ± 1 or ± 2: If no group splits, one group changes its size. Hence, the number of groups of two sizes is changed. If a group splits, the number of groups of size n is decreased by 1. In addition, a second group is eliminated and two daughter groups of different size are created. In general, this changes the number of groups at four different group sizes by ± 1. However, if both daughter groups have the same size, the number of groups of size $n/2$ can increase by two. Hence, in the limit $m \gg n \gg 1$ where a large number of groups of every size exists, we can assume that the fraction of groups of size x, $\rho(x)$, changes approximately continuously. The fraction of groups of size x decays by group elimination at rate $a\rho(x)$, where the elimination rate is determined self-consistently (see below). The rate at which a group of size x increases its size by 1 is given by bx/N, where N is the total population size which is assumed to be roughly constant. The rate of change of frequency of groups of size x is given by $-b\rho(x)\,x/N + b\rho(x-1)\,(x-1)/N$, the first term coming from the process $x \to x+1$, and the second from $x-1 \to x$. This yields for the rate of change of $\rho(x)$

$$\frac{\partial}{\partial t}\rho(x) = -a\rho(x) - b\frac{x}{N}\rho(x) + b\frac{x-1}{N}\rho(x-1). \tag{5}$$

If the maximum group size is large, $n \gg 1$, we can simplify this equation further,

$$\frac{\partial}{\partial t}\rho(x) = -a\rho(x) - \frac{\partial}{\partial x} b\frac{x}{N}\rho(x). \tag{6}$$

An additional assumption on the boundaries has to be made. A group that is divided at size n generates two daughter groups. These groups have a binomial size distribution around $n/2$. Since we are interested in the average group size divided by the carrying capacity, $\bar{x}/n$, we can neglect the fluctuation of the order $\pm\sqrt{n}$ around the mean value for large n. Hence, we can assume as a rough approximation that each group splits into exactly two groups of size $n/2$.

For the stationary solution $(\partial/\partial t)\rho(x) = 0$, we have $\rho(x) \propto x^{-2}$. As the group sizes are always between $n/2$ and n for large n, a has to be chosen in such a way that $\int_{n/2}^{n} \mathrm{d}x \rho(x) = 1$. For self-consistency, we need $a = b/N$. We find for the average group size

$$\bar{x} = \frac{\int_{n/2}^{n} x\rho(x)\,\mathrm{d}x}{\int_{n/2}^{n} \rho(x)\,\mathrm{d}x} = n \ln 2 \approx 0.693n. \tag{7}$$

Numerical simulations of large populations support this finding. For $n=m=100$ we find $\bar{x}/n = 0.691 \pm 0.017$. For $n=m=1000$ we have $\bar{x}/n = 0.693 \pm 0.009$.

3.2. Fixation probability

Fig. 1 shows the fixation probability of a single A mutant as function of the relative fitness r. We can compare the fixation probability of the group process with the corresponding fixation probability of the Moran process describing an unstructured population of the same average size, which was found numerically as 67.15 ± 0.02 for $n = m = 10$. We observe that the group process acts as a suppressor of selection (Nowak et al., 2003; Lieberman et al., 2005). The fixation probability of the Moran process, Eq. (4), defines a certain balance between selection and drift. In our group division process, advantageous mutants have a lower probability of fixation than in a simple Moran process, whereas disadvantageous mutants have a higher probability of fixation. Thus, the group division process reduces the intensity of selection and augments random drift.

For large m and n, we can argue that fixation probability becomes very small in contrast to the usual Moran process, details can be found in the Appendix. In a Moran process, the fixation probability of an advantageous mutant, $r > 1$, converges to $\rho = 1 - 1/r$ for large population size, $n \to \infty$ (see Eq. (4). Conversely, the fixation probability of a disadvantageous mutant, $r < 1$, converges to zero in a Moran process of large population size, $n \to \infty$. For our group division process, we have shown in the appendix that the fixation probability of any mutant converges to zero for large m and n.

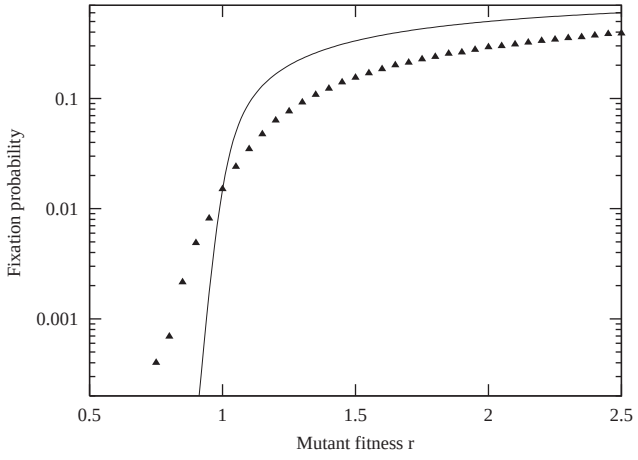

Fig. 1. Dependence of the fixation probability of a single mutant on its fitness in the group selection model. The triangles show the fixation probability of the introduced process. As the population size varies, we compute the average population size $\bar{N}$ numerically in our model in order to compare the fixation probability with the Moran process. Inserting $\bar{N}$ into Eq. (4), we find the fixation probability of the corresponding Moran process given by the full line. As the corresponding Moran process has a higher fixation probability for advantageous mutants, $r > 1$ and a lower fixation probability for disadvantageous mutants $r < 1$, we can conclude that the population structure is acting as an suppressor of selection ($\bar{N} = 67.15 \pm 0.02$, $m = 10$, $n = 10$, averages are over 10^5 realizations).

4. An extended process

So far we have only used three parameters, the number of groups, m, the maximum group size, n, and the relative fitness r. We now introduce a fourth parameter: upon exceeding the maximum group size, n, a group is divided with probability p. With probability $1 - p$, the group is not divided, but a randomly chosen individual of this group is eliminated. The process described in the previous section is the special case of $p = 1$: a group is always divided when exceeding the maximum size.

In the limit of very small p, most groups will be at the maximum size, n, and consist either of all A or all B individuals. The stochastic population dynamic is given by a recursion of two Moran processes. Within each group there is a Moran process with population size n. Among groups there is a Moran process with population size m.

4.1. Fixation probability in the extended process

What is the fixation probability of a single A mutant in the limit of very small p ? First the mutant has to reach fixation in its group. This happens with probability $\rho(r, n)$ given by Eq. (4). Subsequently the group has to out-compete all other groups. This happens with probability $\rho(r, m)$. Thus the overall fixation probability is given by

$$\phi(r, n, m) = \rho(r, n)\rho(r, m) = \frac{1 - 1/r}{1 - 1/r^n} \frac{1 - 1/r}{1 - 1/r^m}. \qquad (8)$$

This fixation probability has to be compared with the probability $\rho(r, nm)$ that a single A mutant reaches fixation in a large, unstructured population of size mn. For $r > 1$ we have $\phi(r, n, m) \leqslant \rho(r, mn)$. For $r < 1$ we have $\phi(r, n, m) \geqslant \rho(r, mn)$. Thus, disadvantageous mutants are more likely to be fixed in the group division process than in a corresponding Moran process describing an unstructured population of the same size. Advantageous mutations are less likely to be fixed. Hence, we have proven that the group division process, in the limit of small p acts as a suppressor of selection.

In Fig. 2, we show that Eq. (8) is a perfect approximation for the fixation probability that is observed in numerical simulations using a splitting probability of $p = 0.001$.

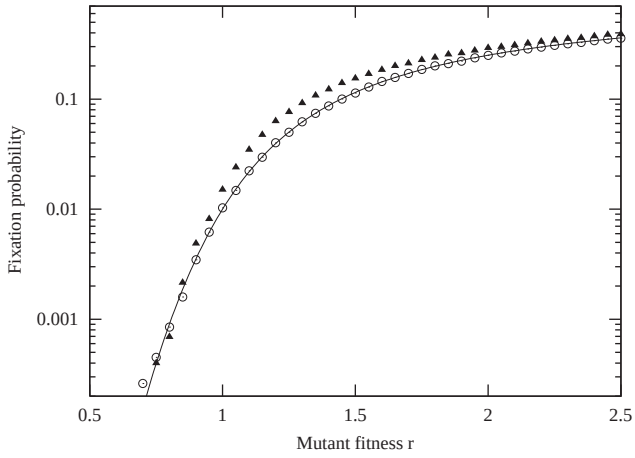

Fig. 2. Dependence of the fixation probability of a single mutant on its fitness in the extended process, where groups are split with a very small probability $p \ll 1$. The circles show the fixation probability in the extended process for $p = 0.001$. The line shows the corresponding analytical approximation given by Eq. (8), which agrees perfectly with the simulations. For comparison, the triangles show the process with splitting probability $p = 1$ ($m = 10$, $n = 10$, averages are over 10^5 realizations).

4.2. The maximum suppression of selection is obtained for $n = m$

Which population structure leads to the most effective suppression of selection? In order to answer this question, let us analyse a population with fixed maximum size, $N = mn$, in the limit of small p. For $r > 1$ we are looking for the value n which minimizes the function $\phi(r, n, N/n)$. For $r < 1$ we are looking for the value n which maximizes the function $\phi(r, n, N/n)$. In both cases, we find that the extremum is given by $n = \sqrt{N}$ and therefore by $n = m$. Thus the maximum suppression of selection is achieved when the group size equals the number of groups.

In Fig. 3, we show the perfect agreement between numerical simulations using $p = 0.001$ and our theoretical prediction that the fixation probability of advantageous mutants has a minimum at $n = m$. For $p = 1$ the minimum can be found at slightly larger values of the group size n.

Taking the limit of small p allows us to separate two different time-scales: (i) there is a fast time-scale for individual selection and (ii) a slow time-scale for group selection. Alternative mechanisms for introducing different time-scales in such contexts can be found in Slatkin (1981, 1977); Wakeley and Takahashi (2004). A slower time-scale for group selection arises, when the life time of groups is longer than the life time of individuals (Keller, 1999).

In the model so far, groups divide faster if they contain individuals that reproduce faster. Thus the "fitness" of the group is determined by the fitness of its individuals. This is a case of "soft group selection", where individual and group selection work in the same direction (Mayr, 1997).

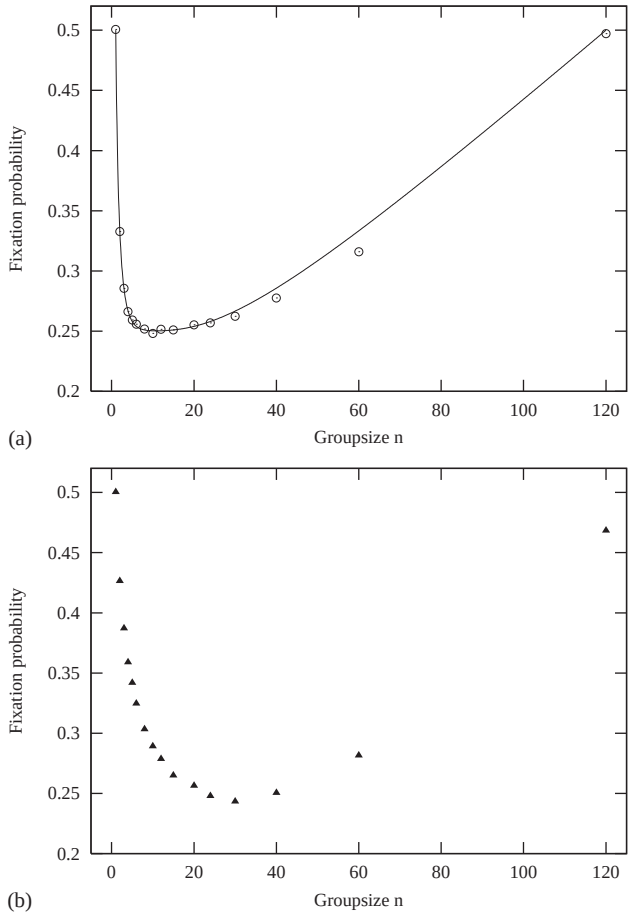

Fig. 3. Fixation probability for different population structures. A population with carrying capacity $N = 120$ is divided into $m = N/n$ groups of size n. The fixation probability is minimal for $m = n$ for small group splitting probabilities, $p \ll 1$. Panel (a) Simulation results for $p = 0.001$ (circles) agree well with the theory for constant group sizes Eq. (8) (full line). Panel (b) For $p = 1$ (triangles) the situation is qualitatively different as the population size varies. A minimum of the fixation probability is observed for larger group sizes than in the case $p \ll 1$. The corresponding fixation probability for a mixed group is approximately $1 - 1/r = 0.5$. Any group structure leads to a suppression of selection ($r = 2$, averages are over 10^4 realizations).

5. Opposing selection on two levels

Let us now study a case where group and individual selection work in opposite directions. This is the case when the benefit of the individual and the group contradict each other. For example, birds giving warning calls to their group increase the fitness of their group, but reduce their own survival probability. Each individual is tempted not to give warning calls, but this reduces the survival probability of the group. Another

example is the evolution of parasites. If a parasite maximizes its own reproductive ratio, it can reduce the fitness of the host too fast and thereby decrease the probability to spread to other hosts. Cooperation between parasites within a host can lead to virulence levels that are optimal for the parasite.

As before, the population is subdivided into m groups of size n. Within each group, selection dynamics are characterized by a Moran process. Occasionally, a group divides into two groups. In this case, a randomly chosen group is eliminated. Again, we assume that individual selection dynamics are much faster than group selection dynamics. Thus, in all likelihood any mixed group has reached fixation for one of the two types, before the group is divided. As before, the individual fitnesses of types A and B within each group are given by r and 1, respectively. But let us now assume that all-A groups divide at rate R whereas all-B groups divide at rate 1. Once again we have a recursive arrangement of two Moran processes.

The fixation probability of a single A mutant is

$$\phi(r,R,n,m)=\frac{1-1/r}{1-1/r^n}\,\frac{1-1/R}{1-1/R^m}. \tag{9}$$

For example, if $r>1$ and $R<1$, then individual selection favors A, while group selection favors B. If $\phi>1/(nm)$ then A is an advantageous mutant, once selection on both levels is combined. The following observations are of interest.

1. *Limit of group selection*: For given n and m there exists a critical value r_c. If $r\leqslant r_c$ then no value of R exists such that A is an advantageous mutant. The critical value r_c is implicitly given by $\rho(r_c,n)=1/(nm)$.
2. *Limit of individual selection*: for given n and m there exists a critical value R_c. If $R\leqslant R_c$ then no value of r exists such that A is an advantageous mutant. The critical value R_c is implicitly given by $\rho(R_c,m)=1/(nm)$.

For example, if $n=10$ and $m=100$ then $r_c\approx 0.5014$ and $R_c\approx 0.9643$. For example, for $r=0.5$ no R can make the mutant advantageous overall. In the case of $r=0.6$, the mutant becomes advantageous when R is larger than 1.3273. The critical values for $n=m=10$ are $r_c=R_c\approx 0.6785$.

Fig. 4 illustrates three different parameter regions for fixed n and m. If $\phi(r,R,n,m)>1/(nm)$ then A is favored by combined selection on the two levels. If $\phi(1/r,1/R,n,m)>1/(nm)$ then B is favored by combined selection on the two levels. These two regions are mutually exclusive. There is, however, a third region where combined selection on the two levels opposes both the fixation of a single A mutant in a B population and the fixation of a single B mutant in an A population.

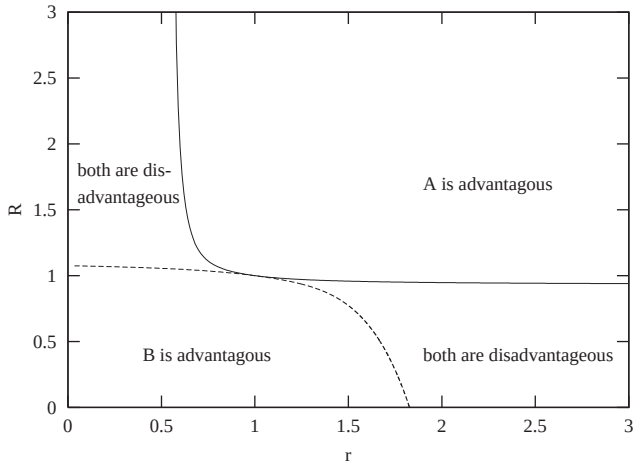

Fig. 4. Comparison of the fixation probability of the two types with a neutral mutant for a fixed population structure. For $r<1$ and $R<1$, the B individuals are always advantageous, as they are advantageous on both levels. Similarly, for $r>1$ and $R>1$, A individuals are always advantageous. For opposing selection on two levels, i.e. $r>1$, $R<1$ and vice versa, the overall advantage of the individuals is determined by the population structure. Both types can be disadvantageous, for $n\to\infty$ and $m\to\infty$ this is always the case if selection works in opposite directions on the two levels. Here, increasing r maintains the advantage of B much longer than increasing R, as fixation of groups has a higher impact on the total fixation probability than fixation within groups due to $m>n$ ($m=50$, $n=10$).

This third region is defined by $\phi(r,R,n,m)<1/(nm)$ and $\phi(1/r,1/R,n,m)<1/(nm)$. There is no parameter region that would simultaneously favor the fixation in both directions.

6. Multi-level selection

The results of the previous section can be generalized to simultaneous selection on more than two levels. Consider a recursive hierarchy of k Moran processes with respective population sizes $n_1,n_2,\ldots,n_k$. Hence, there are groups of groups of groups and so on. Denote by r_i the relative fitness of A on level i. Assume that the time-scales can be neatly separated as before. The fixation probability of a single A mutant in a population of B individuals is

$$\phi(r_1,\ldots,r_k;n_1,\ldots,n_k)=\prod_{i=1}^{k}\frac{1-1/r_i}{1-1/r_i^{n_i}} \tag{10}$$

The corresponding results of the previous section can be derived here. For a given configuration of the population, $(n_1,\ldots,n_k)$, there exists a critical selection coefficient, r_i^*, on level i such that no choice of selection coefficients on other levels can make A advantageous over all.

We can also generalize the result on the maximum suppression of selection. Consider a population of fixed size N with fixed selection coefficients, $r_i = r > 1$ for every i, which is to be distributed into a hierarchy of Moran processes on k levels. The fixation probability $\phi(r_1, ..., r_k; n_1, ..., n_k)$ has a minimum for $n_i = n = N^{1/k}$ for every i, i.e. the population size has to be equal on each level to realize the maximum suppression of selection. This can be shown straightforward by analysing the extremum of $\phi(r_1, \ldots, r_k; n_1, \ldots, n_k)$ with the side condition $\prod_{i=1}^{k} n_i = N$.

7. Discussion

We have introduced a model for stochastic, multilevel selection based on a hierarchy of Moran processes. In this model, a population of size N is sub-divided into m groups. There are two different types, A and B, with fitness values r and 1. Individuals reproduce proportional to their fitness. The off-spring is added to the same group. Whenever a group reaches a certain size, n, it splits into two groups with probability p. If this happens then another group is eliminated. The total number, m, of groups is constant. The total population size can fluctuate, but is bounded between the minimum value m and the maximum value mn.

The fixation probability of the Moran process, given by Eq. (4), denotes a certain balance between drift and selection. If a population structure with the same total population size increases the fixation probability of disadvantageous mutants and reduces the fixation probability of advantageous mutants, then it is a suppressor of selection. Our model acts as suppressor of selection. In the limit of small p, the suppression of selection is most effective, when the number of groups equals the group size, $m = n$.

We also discuss the limits of group selection and individual selection when selection works in opposite directions on different levels. Suppose that A and B individuals reproduce at rates r and 1 inside each groups, but all-A groups and all-B groups divide at relative rates R and 1, respectively. Group division is much slower than individual reproduction. Hence, groups become fixed for one or the other type before they divide. In this framework, we have calculated the fixation probability of a single A mutant in a population of B. We find, for a given m and n, there exists a critical value r_c. If $r < r_c$ then no value R can make A an advantageous mutant over all. This is the limit of group selection. Likewise, there exists a critical value R_c. If $R < R_c$ there is no value r that can make A an advantageous mutant over all. This is the limit of individual selection. These results can be extended to selection on more than two levels.

Another very intuitive observation is the following. Consider multiple levels of selection. If A is advantageous on at least one level then there exists a population structure which can make A advantageous over all.

A lot of previous models for multilevel selection describe systems in which groups are formed from a global pool of individuals, as the haystack model of Maynard Smith (1964) or the famous trait group model of Wilson (1975). While these works concentrate on the evolution of altruistic traits, we consider constant fitness here. Wilson (1975) addresses the question when an altruistic trait is advantageous in a population structured by groups. It is tacitly assumed that only advantageous traits will survive. However, in a finite population it is possible that disadvantageous traits fixate due to random drift. Our approach based on the Moran process allows to analyse this effect in population structured by groups. We find that the population structure itself shows interesting phenomena: It acts as a suppressor of selection, reducing the probability that advantageous mutants fixate in the population.

Slatkin and Wade (1978) also considered a model with constant fitness. They analysed group formation from a global pool of individuals as well as group formation from a single group. In contrast to our model with two types, they worked on a quantitative trait. They computed the expected change of these traits and the expected change of the variance for different group formation scenarios. However, an advantageous mutants entering a population with lower fitness will always take over the population in this model. In contrast to our model where one type is ultimatively eliminated, a finite variance from genetic drift will always be present. However, a direct comparison of both models is intricated, as our approach based on Markov processes cannot easily be transferred to more types.

In this paper, we have calculated the effect of multilevel selection in a system where the offspring of individuals are added to the same group. Once the group reaches a certain size it can divide into two groups. This population structure describes, for example, the evolutionary dynamics of proto-cells (Szathmáry, 1986; Szathmáry and Demeter, 1987; Maynard Smith and Szathmáry, 1995; Grey et al., 1995): genes replicate within proto-cells. Proto-cells divide if they reach a certain size. The genes are randomly distributed over the daughter cells. Another application is groups of animals or early human ancestors. Offspring are added to the group. Larger groups have a higher probability of splitting into two groups. In both cases, selection on different levels can oppose each other. For example, a trait may be beneficial for the competition of individuals within a group but disadvantageous to the group. Here we have shown how to calculate whether the trait is advantageous over all. We do not consider migration, mutation or frequency-dependent selection. All of these

phenomena are extremely important and have been analysed by many authors in the context of group selection (Wilson, 1975; Slatkin and Wade, 1978; Slatkin, 1981; Leigh, 1983). We wish to study these phenomena in the framework of our model in forthcoming papers.

In conclusion, our work provides a simple analytical approach to stochastic multi-level selection that may serve as a useful basis for further investigations.

Acknowledgements

A. T. gratefully acknowledges support by the Studienstiftung des deutschen Volkes (German National Academic Foundation). The Program for Evolutionary Dynamics is supported by Jeffrey Epstein.

Appendix A. Fixation probability of the basic process with $p = 1$

If groups are always split when they reach their carrying capacity, the fixation probability becomes very small in contrast to the usual Moran process. To see that, let us consider the early phase of the dynamics, once a mutant of type A is introduced in a population of type B. Since n is large, for a certain time, mutants would be a small minority in the group and the over all growth rate of the group would not be seriously affected by mutants. During this period, a group that is not eliminated takes a certain time to grow to its carrying capacity n from $n/2$, which is the typical group size after the splitting process. Consider a group with a small fraction of mutants. As the growth rates of the mutants is r times higher than the growth rate of the rest, the mutants will, typically, grow by a factor 2^r when the rest just doubled. We could estimate the number of doubling cycles k, required for the mutant population size in the group to be comparable with the number of residents, B. Assuming that the group has γn residents (where γ is order one), we find for k the equation

$$2^{rk} \approx 2^k \gamma n \quad \text{or} \quad k \approx \frac{1}{r-1}\frac{\ln(\gamma n)}{\ln 2}. \tag{A.1}$$

If $\ln n \gg r - 1$, we have $k \gg 1$. For a numbers of cycles much less than k one could ask, what is the probability of extinction of the mutant? For the fixation probability we have $\rho \leqslant 1 - p_{ext}(t)$, where $p_{ext}(t)$ is the mutant extinction probability after t cycles from the introduction of a single mutant. Thus if we could provide a lower bound on $p_{ext}(t)$, we could bound ρ.

In the model, the only way to become extinct is, if all the groups carrying mutants get eliminated. In the regime of interest, $t \ll k$, all one needs to do is to keep track of which groups do or do not have mutants.

The probability to get a group without any mutants after the splitting of a group with y mutants is 2^{-y}, where y is the number of mutants when the carrying capacity is reached. For the purpose of estimating a lower bound, we consider a modified process, where each group containing mutants always generates two groups containing mutants, upon reaching the maximum size n. The original process has a greater extinction probability than the modified process for the same time period. The modified process turns out to be a branching process with constant average number (Feller, 1968). Every group that reaches size n generates two offspring groups. Both have a probability $\frac{1}{2}$ of growing to n. The extinction probability of such a process in t cycles is roughly $1 - c/t + O(1/t^2)$ (in fact, numerically, one finds $c \approx 4$). After k cycles, a different dynamics takes over and groups with mutants grow faster. In this phase, the extinction of the mutants is very unlikely, if r is sufficiently large. The fixation probability can be bounded from above by a term scaling as c/k, which vanishes for large n. The argument given so far holds for $m \gg k \gg 1$. However, if we have $k \gg m \gg 1$, most groups will contain some mutants after some time, but the mutants do not have a significant influence on the fitness of their groups, since the mutants are a small fraction only. Hence, all groups have approximately the same fitness and the group dynamics looks like the conventional Moran process with fitness $r = 1$. Hence, the fixation probability would be $1/m$, which converges to zero for large m. Figs. 1–4

References

Anderson, R.M., May, R.M., 1979. Population biology of infectious diseases I. Nature 280, 361–367.

Anderson, R.M., May, R.M., 1981. The population dynamics of microparasites and their invertebrate hosts. Philos. Trans. R. Soc. London 291, 451–524.

Anderson, R.M., May, R.M., 1982. Coevolution of hosts and parasites. Parasitology 85, 411–426.

Bergstrom, T.C., 2002. Evolution of social behavior: individual and group selection. J. Econ. Persp. 16, 67–88.

Bonhoeffer, S., Nowak, M.A., 1994. Mutation and the evolution of virulence. Proc. R. Soc. London B 258, 133–140.

Boyd, R., Richerson, P.J., 2002. Group beneficial norms can spread rapidly in a structured population. J. Theor. Biol. 215, 287–296.

Charlesworth, B., 1979. A note on the evolution of altruism in structured demes. Am. Nat. 113, 601–605.

Craig, D.M., Muir, W.M., 1996. Group selection for adaptation to multiple-hen cages: beak related mortality, feathering, and body weight responses. Poultry Sci. 75, 294–302.

Ewens, W.J., 2004. Mathematical Population Genetics. Springer, New York.

Feller, W., 1968. An Introduction to Probability Theory and its Applications, vol. I. Wiley, New York.

Frank, S.A., 1998. Foundations of Social Evolution. Princeton University Press, Princeton.

Goodnight, C.J., 1997. Experimental studies of group selection: what do they tell us about group selection in nature? Am. Nat. 150, S59–S79.

Gould, S.J., Lloyd, E.A., 1999. Individuality and adaption across levels of selection: how shall we name and generalize the unit of Darwinism? Proc. Natl Acad. Sci. 96, 11904–11909.
Grey, D., Hutson, V., Szathmáry, E., 1995. A re-examination of the stochastic corrector model. Proc. Biol. Sci. 262, 29–35.
Hamilton, W.D., 1963. The evolution of altruistic behavior. Am. Nat. 97, 354–356.
Hamilton, W.D., 1964. The genetical evolution of social behavior I and II. J. Theor. Biol. 7, 1–52.
Hauert, C., Doebeli, M., 2004. Spatial structure often inhibits the evolution of cooperation in the snowdrift game. Nature 428, 643–646.
Irwin, A.J., Taylor, P.D., 2001. Evolution of altruism in stepping-stone populations with overlapping generations. Theor. Popul. Biol. 60, 315–325.
Karlin, S., Taylor, H.M.A., 1975. A First Course in Stochastic Processes. Second ed. Academic, London.
Keller, L. (Ed.), 1999. Levels of Selection in Evolution. Princeton University Press, Princeton.
Kerr, B., Godfrey-Smith, P., 2002. Individualist and multi-level perspectives on selection in structured populations. Biol. Philos. 17, 477–517.
Killingback, T., Doebeli, M., 1996. Spatial evolutionary game theory: Hawks and doves revisited. Proc. R. Soc. London B 263, 1135–1144.
Leigh, R.G., 1983. When does the good of the group override the advantage of the individual. Proc. Natl Acad. Sci. 80, 2985–2989.
Levin, B.R., Bull, J.J., 1994. Short-sighted evolution and the virulence of pathogenic microorganisms. Trends Microbiol. 2, 76–81.
Levin, B.R., Kilmer, W.L., 1974. Interdemic selection and the evolution of altruism. Evolution 28 (4), 527–545.
Levins, R., 1970. Extinction. Some Mathematical Questions in Biology, vol. 2, American Mathematical Society, Providence, RI, pp. 75–108.
Lieberman, E., Hauert, C., Nowak, M.A., 2005. Evolutionary graph theory. Nature 433, 312–316.
Lindgren, K., Nordahl, M.G., 1994. Evolutionary dynamics of spatial games. Physica D 75, 292–309.
May, R.M., Nowak, M.A., 1994. Superinfection, metapopulation dynamics, and the evolution of diversity. J. Theor. Biol. 170, 95–114.
May, R.M., Nowak, M.A., 1995. Coinfection and the evolution of parasite virulence. Proc. R. Soc. London B 261, 209–215.
Maynard Smith, J., 1964. Group selection and kin selection. Nature 201, 1145–1147.
Maynard Smith, J., 1976. Group selection. Q. Rev. Biol. 51, 277–283.
Maynard Smith, J., 1998a. The origin of altruism. Nature 393, 639–640.
Maynard Smith, J., 1998b. Evolutionary Genetics. Second ed. Oxford University Press, Oxford.
Maynard Smith, J., Szathmáry, E., 1995. The Major Transitions in Evolution. W.H. Freeman, Oxford.
Mayr, E., 1997. The objects of selection. Proc. Natl Acad. Sci. 94, 2091–2094.
Michod, R.E., 1983. Population biology of the first replicators: the origin of genotype, phenotype and organism. Am. Zool. 23, 5–14.
Michod, R.E., 1999. Darwinian Dynamics. Princeton University Press, Princeton.
Michod, R.E., Roze, D., 2001. Cooperation and conflict in the evolution of multicellularity. Heredity 86, 1–7.
Mitteldorf, J., Wilson, D.S., 2000. Population viscosity and the evolution of altruism. J. Theor. Biol. 204, 481–496.
Moran, P.A.P., 1962. The Statistical Processes of Evolutionary Theory. Clarendon, Oxford.
Muir, W.M., 1996. Group selection for adaptation to multiple-hen cages: selection program and direct responses. Poultry Sci. 75, 447–458.
Nakamaru, M., Matsuda, H., Iwasa, Y., 1997. The evolution of cooperation in a lattice-structured population. J. Theor. Biol. 184, 65–81.
Nakamaru, M., Nogami, H., Iwasa, Y., 1998. Score-dependent fertility model for the evolution of cooperation in a lattice. J. Theor. Biol. 194, 101–124.
Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature 359, 826–829.
Nowak, M.A., May, R.M., 1993. The spatial dilemmas of evolution. Int. J. Bifur. and Chaos 3, 35–78.
Nowak, M.A., May, R.M., 1994. Superinfection and the evolution of parasite virulence. Proc. R. Soc. London B 255, 81–89.
Nowak, M.A., Michor, F., Iwasa, Y., 2003. The linear process of somatic evolution. Proc. Natl Acad. Sci. 100 (25), 14966–14969.
Nowak, M.A., Anderson, R.M., McLean, A.R., Wolfs, T., Goudsmit, J., May, R.M., 1991. Antigenic diversity thresholds and the development of AIDS. Science 254, 963–969.
Slatkin, M., 1977. Gene flow and genetic drift in species subject to frequent local extinction. Theor. Popul. Biol. 12, 253–262.
Slatkin, M., 1981. Fixation probabilities and fixation times in a subdivided population. Evolution 35, 477–488.
Slatkin, M., Wade, M.J., 1978. Group selection on a quantitative character. Proc. Natl Acad. Sci. 75, 3531–3534.
Sober, E., Wilson, D.S., 1998. Unto Others: The Evolution and Psychology of Unselfish Behavior. Harvard University Press, Cambridge.
Swenson, W., Wilson, D.S., Elias, R., 2000. Artificial ecosystem selection. Proc. Natl Acad. Sci. 97, 9110–9114.
Szathmáry, E., 1986. The eukaryotic cell as an information integrator. Endocytobiol. Cell Res. 3, 113–132.
Szathmáry, E., Demeter, L., 1987. Group selection of early replicators and the origin of life. J. Theor. Biol. 128, 463–486.
Wade, M.J., 1977. An experimental study of group selection. Evolution 31, 134–153.
Wade, M.J., 1980. An experimental study of kin selection. Evolution 34, 844–855.
Wakeley, J., Takahashi, T., 2004. The many-demes limit for selection and drift in a subdivided population. Theor. Popul. Biol. 66, 83–91.
Williams, G.C., 1966. Adaption and Natural Selection: A Critique of some Current Evolutionary Thought. Princeton University Press, Princeton.
Wilson, D.S., 1975. A theory of group selection. Proc. Natl Acad. Sci. 72, 143–146.
Wilson, D.S., 1979. Structured demes and trait-group variation. Am. Natl 113, 606–610.
Wilson, D.S., Dugatkin, L.A., 1997. Group selection and assortative interactions. Am. Natl 149, 336–351.
Wynne-Edwards, V.C., 1962. Animal Dispersion in Relation to Social Behavior. Oliver & Boyd, Edinburgh.