



ELSEVIER

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Positive effects of multiple gene control on the spread of altruism by group selection

Tomáš Kulich^{a,*}, Jaroslav Flegr^b^a Department of Computer Science, Faculty of Mathematics, Physics and Informatics, Comenius University, Mlynská dolina, 84248 Bratislava, Slovak Republic^b Department of Philosophy and History of Science, Charles University, Viničná 7, CZ-128 44 Praha 2, Czech Republic

ARTICLE INFO

Article history:

Received 21 July 2010

Received in revised form

10 May 2011

Accepted 11 May 2011

Available online 23 May 2011

Keywords:

Epistasis

Context-dependent influence

Altruistic behavior

Frozen evolution

Simulation

ABSTRACT

The origin of altruistic behavior has long been a challenge for students of evolutionary biology. The populations with altruistic individuals do better than those without altruists; however, the altruists within a population do worse than the non-altruists and their prevalence in the population decreases due to individual selection. Under certain conditions, the strength of group selection, i.e., the selection driven by competition between populations, can surpass the strength of individual selection; however, such conditions seem to be relatively strict and probably do not hold in many natural systems where the altruistic behavior was observed. It was suggested recently that chances for altruistic behavior to spread highly increase when it is controlled not by a single gene but by multiple independent genes substitutable in their effects on the phenotype of the individual. Here we confirm the original verbal model by numerical modeling of the spread of altruistic/selfish alleles in a metapopulation consisting of partly isolated groups of organisms (demes) interconnected by migration. We have shown that altruistic behavior coded by multiple substitutable genes can stably coexist with selfish behavior, even under relatively high mutation and migration rates, i.e., under such conditions where altruistic behavior coded by a single gene is quickly outcompeted in a metapopulation.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

The problem of evolution and persistence of altruistic behavior has long been a challenge for theoretical and evolutionary biology. It can be relatively easily explained in systems where either reciprocity (Paolucci et al., 2006) or preferential help to genetically related individual, i.e., kin selection (Agrawal, 2001; Hamilton, 1964; Ratnieks and Helanterä, 2009) occurs. The origin and persistence of altruistic behavior in systems without reciprocity and preferential help to genetically related organisms, i.e., the evolution by the form of group selection called interdemic selection, can also be explained, but the explanation is much more difficult (Ono et al., 2003; Fletcher and Doebeli, 2009; van Baalen and Rand, 1998). For review of old models, see Alexander and Borgia (1978), for survey of recent opinions, see Lehmann and Keller (2006) and Leigh (2010) and to read about historical aspects of development of the group selection approach, see Okasha (2005), Wilson and Wilson (2007) and Borrello (2005).

According to classical models, a behavioral pattern that provides an advantage to a group and at the same time places its carrier at

a disadvantage has a low chance of spreading and enduring in nature. Groups in which the altruistic trait spreads would do better than those in which this trait is lacking and the average fitness of their members would be greater; however, selfish individuals who do not exhibit this trait and do not behave altruistically, but only enjoy the advantages provided by the presence of altruists, would have the greatest fitness within these groups. It has been shown that under certain conditions, the strength of group selection can surpass that of individual selection, especially in populations with a certain structure and certain population dynamics (Ono et al., 2003; Bowles, 2006; Fletcher and Doebeli, 2009; van Baalen and Rand, 1998). However, most of these analyzes have shown that under usual conditions, the spreading of an altruistic behavioral pattern in systems without reciprocity or kin selection is rather rare.

The chances for altruistic behavior to spread may considerably increase when complex gene interactions (Peck, 2004) or genes with pleiotropic effects (Foster et al., 2004) are responsible for the altruistic behavior. For example, the individuals behave altruistically when heterozygous in a particular altruistic gene while behaving selfishly when homozygous in such a gene or alternatively altruistic behavior is coded by multiple independent genes and the probability of altruistic behavior is a non-monotone function of the number of altruistic alleles in the genome (being the highest when this number approaches some intermediate

* Corresponding author.

E-mail addresses: kulich@dcs.fmph.uniba.sk (T. Kulich), flegr@cesnet.cz (J. Flegr).

value). It was, however, recently suggested (Goodnight, 2005) that the evolution of an altruistic trait is strongly influenced not only by the ratio of individual to group selection, but also by the ratio of the heritability at the group level to the heritability at the individual level. In situations where this ratio is higher than one, the altruism can evolve even if selection favoring selfish behavior is stronger than selection favoring altruism.

In the present study, we tested a verbal model (Flegr, 2008) based on the frozen plasticity theory (Flegr, 2010). This model suggests an increased probability of persistence of the altruistic behavioral patterns, when they are coded by several substitutable genes, in comparison to their persistence in case, when they are coded by single gene with large effect. This is due to the fact that such genetic architecture (several substitutable genes) of an altruistic trait decreases the individual level heritability of altruism/selfishness (average amount of variability in the altruism/selfishness trait of offspring that can be explained by genotype of their parents), while leaving group level heritability of altruism/selfishness (average amount of variability of interpopulation variability in frequencies of altruistic alleles in daughter populations that can be explained by frequencies of altruistic alleles in maternal populations) intact. The present study starts with the description of the model.

2. Model

We consider the *fitness* in the classical meaning of the word, i.e., if two individuals have the fitnesses equal to a and b , then the ratio of the expected number of their descendants is a/b . Especially, when the fitnesses of two individuals are 1 and $1+c$, we can say that the second individual has an *advantage* c over the first one. It is easy to see that if the fitnesses of two individuals are a and b and $a \leq b$, then the second individual has an advantage $(b-a)/a$ over the first one.

Our model consists of a *metapopulation* of $n \cdot m$ individuals. They are structured into n demes, of an average size of m . We monitor the metapopulation's behavior in N generations. Each generation consists of three phases: natural selection, migration and mutation.

The phase of natural selection results in the replacement of all the individuals by their descendants. Similarly as in Hamilton (1975) and Rousset (2004), this happens "at once" and the size of the metapopulation is preserved. In a simplified way, for each descendant (which we want to create), we randomly choose a deme, into which it will be put, and two individuals from the chosen deme as a parents. Finally, for every locus in the genome we randomly choose one of two parents, from whom the allele will be copied. Naturally, the random choices (except for the last one) that we make are not taken uniformly. The probability of an individual/deme to be chosen corresponds to their fitnesses. Once all descendants are created by this process, the whole metapopulation is updated by removing old individuals and inserting new descendants into demes.

More specifically, inside the metapopulation, two kinds of natural selection take place, intrademic and interademic. Under intrademic selection, selfish individuals have an advantage α over altruistic individuals, while under interademic selection, altruistic demes have an advantage β over selfish demes. More formally, the meanings of α , β are as follows: The probability of a new individual becoming the member of the i -th deme is proportional to $m_i + \beta a_i$, where m_i is the size of the i -th deme and a_i is the number of altruists of the i -th deme. The parents of this new individual are two randomly chosen individuals from the i -th deme. The probability of a random individual becoming a parent is proportional to its fitness which is equal to 1 for an altruist individual and to $1+\alpha$ for a selfish individual. The phase of

natural selection is ended by the extinction of any deme whose size is less than or equal to 2. Its place is taken by the deme of the largest size, which randomly splits in to two new demes, with each individual being put randomly with a probability of 1/2 into the first or the second new deme.

Altruism is controlled by g non-linked loci of each individual of a sexually reproducing haploid hermaphrodite organism. Each locus has two alleles: an altruistic and a selfish one. The phenotype of the individual depends on whether or not the number of altruistic alleles in the genome is at least t (threshold). In quantitative genetics such a system is called a threshold character (Falconer, 1989), and the underlying number of altruistic alleles in the genome can be a breeding value (or something linearly correlated to it) of the individual.

Mutations occur randomly and for all generations, all individuals and all loci. The probability of an allele being switched from altruistic to selfish or vice versa is μ . The individuals can migrate between demes and the probability of an individual leaving its deme for another (randomly chosen) deme (i.e., migration rate) is η .

So far, the model has the following parameters: advantages of selfish individuals and altruistic demes (α, β), number of demes (n), average deme size (m), mutation rate (μ), migration rate (η), altruism controlling mechanism (g and t), number of generations used for simulation (N) and the initial rate of altruists. By p_a let us denote the average rate of altruists during the whole evolution. We shall say that the *metapopulation is altruistic* if $p_a \geq \varepsilon$, where ε is a constant. By R let us denote such an advantage of the altruistic demes, which leads to a 1/2 probability of an altruistic metapopulation occurring. More formally, R is the advantage of the altruistic demes β such that

$$\mathbb{P}(p_a \geq \varepsilon) = \frac{1}{2}.$$

Naturally, R should be seen as a function of all the parameters from the model including ε . For practical usage, we shall always treat R as a function of just one parameter, other parameters will be fixed and their values will be clear from the context. Under weak selection we can assume that R , as a function of a selfish individual's advantage α , is proportional to α ($R \sim \alpha$). Therefore, rather than R , we shall analyze the fraction $r := R/\alpha$ for low values of α , in our case $\alpha = 0.05$. This value was chosen as a compromise between a too big α where $R \sim \alpha$ holds no more and a too small α where the effects of both intrademic and interademic selection are too subtle comparing to random fluctuations of the system which enormously complicate the numerical analysis.

Finally, let us note that in our model we use threshold twice: the first time for determining whether the individual is altruistic. We find this reasonable and well-motivated. For example, in cooperatively breeding vertebrates and social insects a young adult can be either a non-reproductive helper (worker) or a reproductive individual. Therefore we assume discrete phenotypes, i.e., an individual is either an altruist or a non-altruist, although one can alternatively assume a linear relationship between phenotype and genotype such that an individual becomes more altruistic in proportion to g . This would mean, that if individual has x altruistic alleles (out of g loci) he becomes x/g altruist. Appendix D shows some results using this (linear) model. Generally, the effect of genetic architecture on spreading of altruistic alleles is weaker in the linear model, however, the results obtained with both models are qualitatively the same.

The second use of the threshold is to determine, whether the whole population is altruistic or not. Without this we would not be able to obtain several important results, for example those in Fig. 1. More detailed discussion on this topic can also be found in Appendix A.

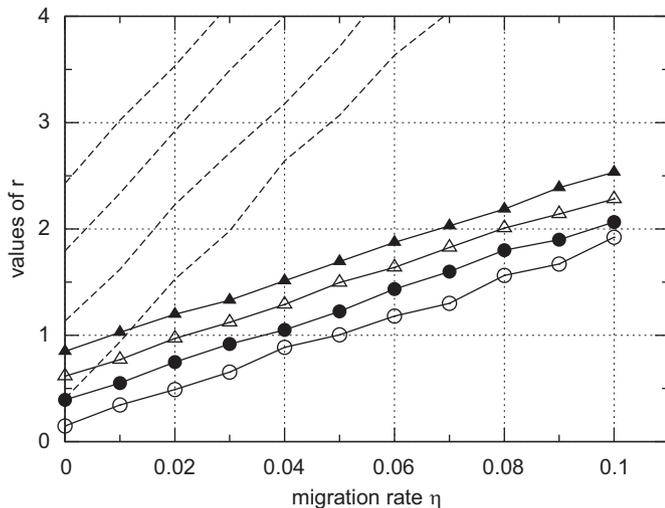


Fig. 1. Values of r as a dependence on migration rate η . Values of other parameters: number of demes $n=10$, average deme size $m=10$, $\varepsilon=1/2$. Mutation rate: $\mu=0.001$ (empty circles), $\mu=0.007$ (full circles), $\mu=0.013$ (empty triangles), $\mu=0.019$ (full triangles). Dashed lines show results for the same values of μ but for a different structure of the metapopulation: $n=5$, $m=20$. The markings are approximately the size of the 90% confidence intervals calculated by the maximum likelihood method. In this case we naturally put $g=1$, $t=1$.

3. Simulation

The model described above was straightforwardly simulated by a program written in the C++ language. From $N=10\,000$ generations we calculated the average rate of altruists p_a . If p_a is sufficient (i.e., $p_a \geq \varepsilon$), we decrease a little the value of β , and vice versa. By repeating this procedure, we get β oscillating around the searched value R . However, these oscillations and also the oscillations of averaged values of β are too big for calculating R with sufficient accuracy. Therefore, we use another approach: Instead of averaging the values of β , we attempt to find a sigmoidal model that would “explain” the measured values of β with some likelihood. We search the space of all possible models for the *best* model—such model that best explains the measured values. This approach is well known as the maximum likelihood method and is described in detail, for example, in Cramér (1946). The final value of R is then deduced from the best model. For more details see Appendix A.

4. Comparison of the model's behavior with the known theoretical results

The known theoretical results obtained by analytical computation generally refer to a weak selection (even if it is not stated explicitly by some authors, their assumptions only apply to weak selections). Therefore, they describe function r where $g=1$ and $t=1$ (where r is defined above).

Our first finding is that r does not significantly depend on the initial rate of altruists. It clearly follows from the fact that there is enough time for alleles to mutate in both ways, and since both α and β are low, there is also enough time for altruism to spread as a result of the genetic drift. However, in all our further experiments with weak selection we initialize the metapopulation in such a way that every allele is altruistic with probability $1/2$. This frequency should correspond to the equilibrium state in all of our experiments and therefore we expect the results not to be dependent on number of generations N .

Further results are summarized in Fig. 1. In accordance with Traulsen and Nowak (2006), Kimura (1983) and Lehmann

et al. (2007) we get linear dependency of r on η and in accordance with Traulsen and Nowak (2006) and Lehmann et al. (2007) we also get linear dependency of r on μ . In discordance with Traulsen and Nowak (2006) and Lehmann et al. (2007) the slope of the dependencies (i.e., $\partial R/\partial \eta$) changes with different parameters n , m . The slope change is roughly in agreement with Kimura (1983), although there is an inaccuracy of about 30% in its actual quantity. Furthermore, the dependencies of r on n and μ are not analyzed in Kimura (1983).

The disagreements mentioned above are a consequence of specific assumptions, which are necessary for obtaining any theoretical results. This makes the results valid only in the domain of the corresponding approximations. Once we test the results in a more general scenario, inaccuracies naturally arise. As an example, the stochasticity of the system is neglected (Kimura, 1983), only demes containing either only altruistic or only selfish individuals are reproduced (Traulsen and Nowak, 2006) or the authors neglect of strong correlations between dependent random variables (Silva and Fontanari, 1999). This shows that solving this problem analytically is very hard and resists numerous different approaches, no matter how inventive they are in using a wide range of mathematical tools.

5. Weak selection and multiple loci control

By the weak selection we mean a regime where the fitness of all individuals is close to one. On the other hand, the strong selection means that the fitness can significantly vary between individuals. Naturally, these terms are not strictly defined and there is a continuous scale between the weak and strong selection.

The results from the simulation for some combinations of the parameters are presented in Tables 1 and 2. All r values in Table 1 are very close to each other. It is not accidental as we explain below.

Let us assume a metapopulation with the same rate p of the altruistic alleles in each deme and on each locus. The rate of the altruists (p_a) is then function of t , g and p . Especially for $t=1$, it can be calculated as follows:

$$p_a := 1 - (1-p)^g \quad (1)$$

(this is the case for models 1–7). For odd g and $t = \lceil g/2 \rceil$, the frequency of the altruists p_a equals p (the case for models 4 and 8). We can say that certain allele is *active* if its change (from altruistic to selfish or vice versa) changes the phenotype of the individual. It clearly happens if (and only if) there are exactly $t-1$ altruistic alleles on the other loci. Probability of such event will be denoted as $P=P(p,g,t)$. Let us estimate the advantage that is conferred by a selfish/altruistic allele to the individual/deme, respectively. With a probability $1-P$, this allele does not influence anything and therefore, does not confer any advantage to anyone. With a probability P , the allele confers advantage α to the individual (if the allele is selfish) or advantage β to the deme (if the allele is altruistic). We can regard the average advantages of the allele as the product of the probability of the allele being active multiplied by the advantage conferred to the individual/deme. We conclude that the selfish allele confers average

Table 1
Models' parameters: $n=20$, $m=30$, $\mu=0.001$, $\eta=0.1$.

Model 1	$g=1$, $t=1$, $\varepsilon=1/2$	$r=6.7 \pm 0.1$
Model 2	$g=3$, $t=1$, $\varepsilon=7/8$	$r=6.7 \pm 0.1$
Model 3	$g=5$, $t=1$, $\varepsilon=31/32$	$r=6.7 \pm 0.2$
Model 4	$g=5$, $t=3$, $\varepsilon=1/2$	$r=6.5 \pm 0.1$

advantage of $P \cdot \alpha$ to the individual and altruistic allele confers advantage $P \cdot \beta$ to the deme it inhabits. As the assumption about the equal distribution of the altruistic alleles holds, we can analyze the evolution of alleles of a certain locus independently on the evolution of other alleles of other loci. Let us now compare models 1 and 3. In model 1, we found a ratio $r_1 = \beta/\alpha$ which leads to a rate of altruists (and therefore as well to a rate of altruistic alleles) 50%. If the same values α and β are used in model 3, then the ratio of the actual advantages the allele confers in this model is

$$r_3 = \frac{P \cdot \beta}{P \cdot \alpha} = \frac{\beta}{\alpha} = r_1.$$

Since we are in the regime of low values α and β , we only expect the ratio of α to β (and not the absolute values of α and β) to be important for the rate of the altruistic alleles. This ratio does not change ($r_1 = r_3$) and therefore as in model 1, we also expect 50% of altruistic alleles on average also in model 3. By using Eq. (1), we get the final rate of altruists 31/32, which is exactly the value of ε in model 3. Using a similar approach, we can justify the similarity of all other results in Table 1. In Table 2, we work with lower values of η . This causes that the altruistic alleles cease to be evenly distributed and start grouping in some demes that will become more altruistic. It means that the results derived in Table 1 also cease to hold.

Summary: Under weak selection and strong migration (10% in our models), multiple loci control does not significantly affect the β/α ratio needed for altruistic alleles to occur. On the other hand, multiple loci control does significantly affect the final rate of altruists (Table 1).

Under weak selection and not so strong migration, multiple loci control also helps to spread altruism by changing the β/α ratio (Table 2).

To understand these counterintuitive results, it could help to compare models 5 and 7 from Table 2. Let us start with model 5 and assume a purely altruistic metapopulation where one selfish migrant emerges in one of the demes. It spreads in the deme which consequently shrinks and finally vanishes. Until the extinction it spreads selfish migrants and potentially infect other demes by selfishness. It may happen that the selfish migrant or its descendants will not reproduce (although they have advantage α over the rest of the deme). Naturally, the advantage α reduces the probability from something alike to happen. In comparison,

this probability is still quite high in model 7. Since $t=1$, even the descendants of a selfish migrant are with high probability altruists and therefore advantage α does not help them (so significantly) to spread. Although the selfish migrant brings some selfish alleles to the new deme, their spreading is influenced mostly by random drift and they do not gain significant advantage from the selfish behavior.

6. Strong selection and multiple loci control

In Ono et al. (2003) authors showed that under sufficiently low migration and mutation rates, the metapopulation can exist in two different semi-stable states. They are called S state, with almost all individuals being selfish, and A state, with almost all individuals being altruistic. Transitions between these two states are denoted as A–S and S–A.

The results of our simulation are in full accordance with existence of S and A states described in Ono et al. (2003). Although the authors of Ono et al. (2003) used extremely low values of μ and η , we observed the S and A states also when higher values of μ and η together with higher advantages α and β were set. Note that naturally there is a strong correlation between existence of S and A states and the fact that R depends on the initial frequency of altruists.

In our simulations, we focus on A–S transition, i.e., we initialize all alleles in the metapopulation as altruistic. A similar shape of R as a function of β is observed for S–A transitions, but higher values of β are necessary for S–A transitions.

We will start with a brief explanation of the behavior of R as a function of α under one-locus control of altruism. For small α function R is similar to linear dependence, but for bigger α , this no longer holds, and the function gets concavely or convexly shaped. Finally the R tends to asymptote either in horizontal or vertical direction. The horizontal asymptote is shown in Fig. 2 on the left. It means that a relatively small advantage for altruists is sufficient for compensating for a much bigger advantage of selfish individuals. The vertical asymptote is shown in Fig. 2 on the right. The interpretation is that a particular advantage of selfish individuals cannot be outweighed by any advantage of altruists, no matter how high.

We briefly describe the mechanism how different asymptotes occur. Let us suppose $\alpha \rightarrow \infty$, a relatively small β and a purely altruistic metapopulation where one selfish mutant emerges. This mutant spreads over its deme very quickly. The deme starts shrinking and finally goes extinct. If it succeeds in producing enough selfish migrants that spread, the selfishness will thrive. A similar effect occurs when $\beta \rightarrow \infty$ and α is relatively small. This time the selfish mutant spreads not do quickly in its own deme.

Table 2
Models' parameters: $n=20$, $m=20$, $\mu=0.0005$, $\eta=0.01$.

Model 5	$g=1$, $t=1$, $\varepsilon=1/2$	$r=0.49 \pm 0.01$
Model 6	$g=3$, $t=1$, $\varepsilon=7/8$	$r=0.44 \pm 0.01$
Model 7	$g=5$, $t=1$, $\varepsilon=31/32$	$r=0.38 \pm 0.02$
Model 8	$g=5$, $t=3$, $\varepsilon=1/2$	$r=0.48 \pm 0.01$

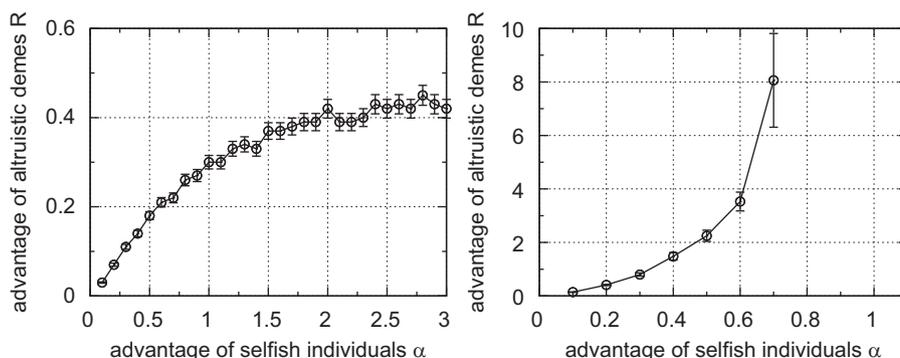


Fig. 2. Values of R as a function of α . The horizontal asymptote can be observed on the left, the vertical asymptote on the right. Parameters of the model: The number of demes $n=10$, average deme size $m=10$ (left) or $m=30$ (right), mutation rate $\mu=0.001$, migration rate $\eta=0.01$.

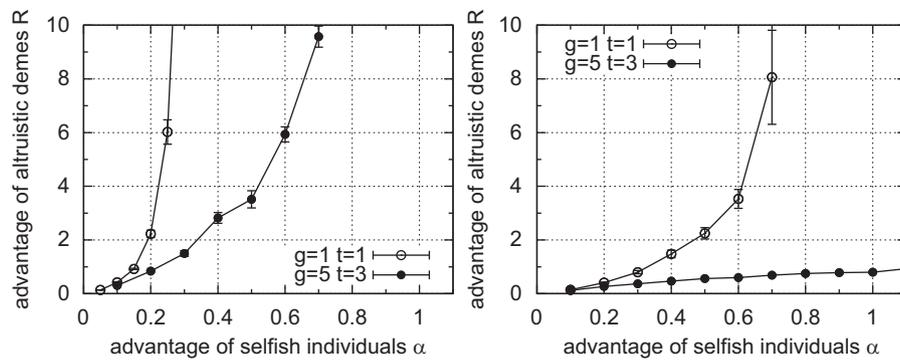


Fig. 3. Values of R as a function of α . Left: average deme size $m=55$, vertical asymptote also occurs for $g=1, t=1$ and for $g=5, t=3$. Right: average deme size $m=30$, vertical asymptote for $g=1, t=1$ (empty circles), but horizontal asymptote for $g=5, t=3$ (full circles). The other parameters are equal to those given in Fig. 2: $n=20$, $\eta=0.01$, $\mu=0.001$.

The deme is shrinking proportionately slowly because of a lower number of altruists in it. Once there are no altruists in the deme, it almost immediately goes extinct. Once again, the main factor that influences the spread of the selfishness is the number of selfish migrants that are produced by the infected deme until it goes extinct. The number of selfish individuals is significantly influenced by the quantity $m \cdot \beta$ which says how many migrants are produced in one generation by an average sized deme. It is also influenced by values β (for horizontal asymptotes) and α (for vertical asymptotes), since these values predict how fast the deme will be shrinking. Contrary to this, no matter how big the values α (for horizontal asymptotes) or β (for vertical asymptotes) are, a selfish individual cannot spread over its deme faster than in one generation and also a deme consisting of only selfish individuals needs at least one generation to extinct. Therefore, it does not matter how high these quantities are, once they are high enough, their change does not influence the number of selfish migrants. The asymptotic behavior is therefore a natural consequence of the systems being insensitive to a change of some parameters.

Let us now introduce the results for multiple loci control. We put $g=5$ and $t=3$. Let us remind that under the weak selection, there was no difference between $g=5, t=3$ and $g=1, t=1$. Differences now occur as an effect of strong selection and existence of semi-stable S and A states. Values of R , as a function of α , are presented in Fig. 3. The figure on the left demonstrates that under five-loci control, the observed values of R are significantly lower than under one-locus control. Also, the placement of the vertical asymptote is different. For one-locus control, its x coordinate is almost three times as high as that for five-locus control. It means that under five-loci control altruism can occur, even for such α that makes it completely impossible under one-locus control. If we reduce the average deme size to $m=30$, an interesting phenomenon occurs: for one-locus control, we still have vertical asymptote, while for five-loci control, we obtain horizontal asymptote. It is possible due to the fact that $m \cdot \eta = 0.3$ which is roughly close to one.

We conclude, that under five-loci control, metapopulation can be altruistic even for such low values of β that would lead to complete annihilation of altruism under one-locus control.

7. Discussion

Although quite unfavorable conditions were chosen for altruism to spread (unlike others we used higher mutation and migration rates), the advantages of altruistic demes needed for altruism to spread are rather low. Even for one-locus control, these values are lower than predicted by other theoretical studies. This is specific for the discussed model where altruistic demes

thrive and out produce other altruistic demes while selfish demes shrink and therefore produce fewer migrants. When intrademic selfishness only results in higher probability of deme extinction (but until extinction the size of the deme is constant, as is the case for example in Traulsen and Nowak (2006), which presents a model that appears to be less realistic than ours), the values of R are significantly higher.

The model for multiple loci control of altruism used in this work is absolutely symmetric—an altruistic allele has exactly the same phenotype effects as any other altruistic allele present in the genome. In reality, the situation will be different, with the phenotype effects of two altruistic alleles on two different loci varying from one another. Preliminary results of simulations of this scenario indicate that even this way of controlling altruism is quite efficient for spreading it.

Another important property of this model is that having one more altruistic allele always makes the individual at least as altruistic as it was before.

If some negative dependencies were considered (i.e., having one more “altruistic” allele could result in a more selfish individual), or when the gene (or alleles) interactions of dominance/recessiveness type were involved (Matessi and Karlin, 1984), then the spreading of altruism would be even much easier. The important message of this paper is that even quite simple gene architecture of multigene traits may have a strong impact on the spreading of altruism.

Our results are in qualitative agreement with the general conclusion of Goodnight (2005) as well as the specific suggestion of Flegr (2008). In situations where the altruistic trait is controlled not by one but by several genes, the altruism can evolve and spread even if it is selected against on the intrapopulation level. In sexual species, any behavioral trait, including the altruistic behavior, is usually determined by the greater number of genes and many of these genes have (due to epistasis) a context-dependent influence on the particular trait. Consequently, the heritability of most traits is low. Under such conditions, altruists emerge from the population as if by chance in families that are completely unrelated and have different phenotypes, i.e., individuals with both selfish and altruistic behavior, with a probability that is determined only by the proportion of particular alleles in the entire population. Thus populations can compete for the greatest average fitness of their members; those that have the greatest proportion of the relevant alleles, resulting in the greatest number of altruists being formed (emerging by chance), shall win in this competition. Our results show that group and interspecies selection can occur in nature in favor of altruistic traits (because the percentage proportion of altruistic alleles in the population is inherited from one generation to the next) and its results cannot be canceled out by individual selection because the

trait itself, altruistic behavior, is not inherited. The complex genetic architecture of traits in sexual organisms can highly increase the chance for persistence of altruistic behavior in local population and its spreading in structured metapopulations.

Acknowledgments

The first author was supported by VEGA Grant no. 1/0406/09. The second author was supported by Grant no. 0021620828 from the Ministry of Education, Youth and Sports of the Czech Republic.

Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version at doi:[10.1016/j.jtbi.2011.05.017](https://doi.org/10.1016/j.jtbi.2011.05.017).

References

- Agrawal, A.F., 2001. Kin recognition and the evolution of altruism. *Proc. R. Soc. Lond. B* 268, 1099–1104.
- Alexander, R., Borgia, G., 1978. Group selection, altruism, and the levels of organization of life. *Annu. Rev. Ecol. Syst.* 9, 449–474.
- Borrello, M.E., 2005. The rise, fall and resurrection of group selection. *Endeavor* 29, 43–47.
- Bowles, S., 2006. Group competition, reproductive leveling, and the evolution of human altruism. *Science* 314, 1569–1572.
- Cramér, H., 1946. *Mathematical Methods of Statistics*. Princeton University Press.
- Falconer, D.S., 1989. *Introduction to Quantitative Genetics*, third ed. Longman, New York.
- Flegr, J., 2008. Frozen evolution. Or, that's not the way it is, Mr. Darwin. Farewell to selfish gene. Charles University of Prague, Faculty of Science Press, Prague.
- Flegr, J., 2010. Elastic, not plastic species: frozen plasticity theory and the origin of adaptive evolution in sexually reproducing organisms. *Biology Direct* 5, 2.
- Fletcher, J., Doebeli, M., 2009. A simple and general explanation for the evolution of altruism. *Proc. R. Soc. Lond. B* 276, 13–19.
- Foster, K., Shaulsky, G., Strassmann, J., Queller, D., Thompson, C., 2004. Pleiotropy as a mechanism to stabilize cooperation. *Nature* 431, 693–696.
- Goodnight, C., 2005. Multilevel selection: the evolution of cooperation in non-kin groups. *Popul. Ecol.* 47, 3–12.
- Hamilton, W.D., 1964. The genetical evolution of social behavior. *J. Theor. Biol.* 7, 1–52.
- Hamilton, W.D., 1975. In: Fox, R. (Ed.), *Biosocial Anthropology*. Malaby Press, London, pp. 133–157.
- Kimura, M., 1983. Diffusion model of intergroup selection, with special reference to evolution of an altruistic character. *Proc. Natl. Acad. Sci. USA* 80, 6317–6321.
- Lehmann, L., Keller, L., 2006. The evolution of cooperation and altruism—a general framework and a classification of models. *J. Evol. Biol.* 19, 1365–1376.
- Lehmann, L., Keller, L., West, S., Roze, D., 2007. Group selection and kin selection: two concepts but one process. *Proc. Natl. Acad. Sci. USA* 104, 6736–6739.
- Leigh, E., 2010. The group selection controversy. *J. Evol. Biol.* 23, 6–19.
- Matessi, C., Karlin, S., 1984. On the evolution of altruism by kin selection. *Proc. Natl. Acad. Sci. USA* 81, 1754–1758.
- Okasha, S., 2005. Maynard smith on the levels of selection question. *Biol. Philos.* 20, 989–1010.
- Ono, S., Misawa, K., Tsuji, K., 2003. Effect of group selection on the evolution of altruistic behavior. *J. Theor. Biol.* 220, 55–66.
- Paolucci, M., Conte, R., Tosto, E.A.G.D., 2006. A model of social organization and the evolution of food sharing in vampire bats. *Adapt. Behav.* 41 (3), 223–238.
- Peck, J.R., 2004. Sex causes altruism, altruism causes sex, maybe. *Proc. R. Soc. Lond.* 271, 993–1000.
- Ratnieks, F.L.W., Helanterä, H., 2009. The evolution of extreme altruism and inequality in insect societies. *Proc. R. Soc. Lond. B* 364, 3169–3179.
- Rousset, F., 2004. *Genetic Structure and Selection in Subdivided Populations*. Princeton University Press, Princeton.
- Silva, A., Fontanari, J., 1999. Stochastic group selection model for the evolution of altruism. *Physica A* 268, 257–268.
- Traulsen, A., Nowak, M.A., 2006. Evolution of cooperation by multilevel selection. *Proc. Natl. Acad. Sci. USA* 103, 10952–10955.
- van Baalen, M., Rand, D., 1998. The unit of selection in viscous populations and the evolution of altruism. *J. Theor. Biol.* 193, 631–648.
- Wilson, D., Wilson, E., 2007. Rethinking the theoretical foundation of sociobiology. *Q. Rev. Biol.* 82, 327–348.

8 Supplementary materials

Appendix 1 - Measurements

Here we would like to explain in more detail how the values of R were obtained. Let $\text{Alt}(\beta)$ be the probability of obtaining altruistic metapopulation (as it was defined in the paper) with β denoting the advantage of altruists and all other parameters fixed to some values. With this denotation, the function R satisfies:

$$R = \beta \text{ is equivalent to } \text{Alt}(\beta) = 1/2.$$

In Figure 4 we see values of β recorded while obtaining values for the first row of Table 1. We start the whole process with some $\beta = \beta_0$. We run one experiment and see if the metapopulation is altruistic - if so, β is decreased a little, if not, β is increased. By repetition of this process many times, we get β oscillating around the searched value, where the probability of obtaining altruistic metapopulation is near 1/2.

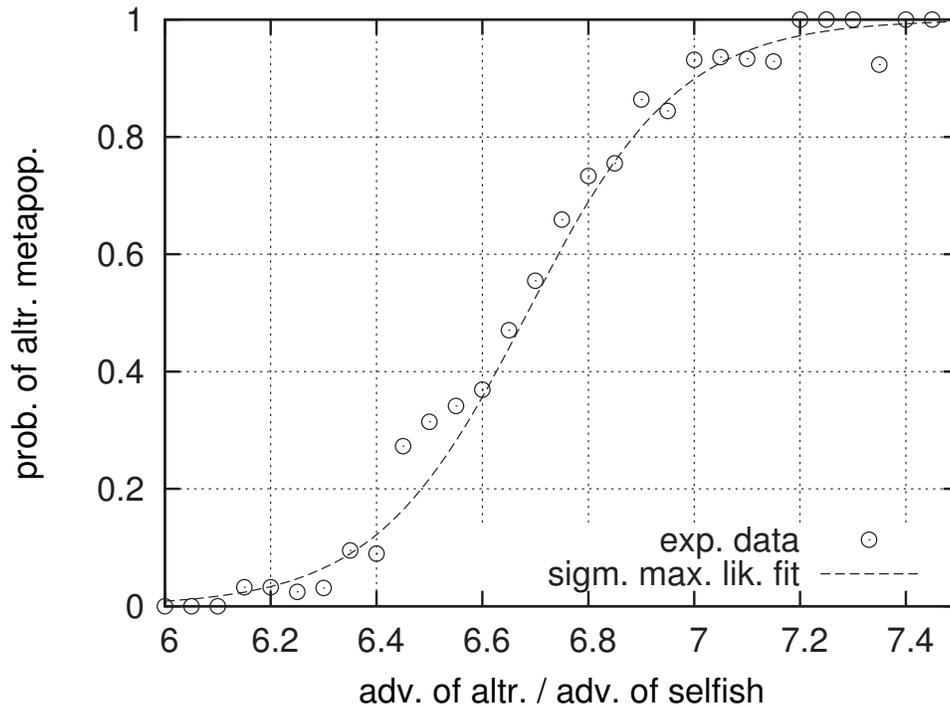


Figure 4: Visualized values from one measurement.

Visualization of this process is shown in the Figure 4. For the purposes of the visualization, the interval between values 6 and 7.5 is divided to 30 bins. In each bin, we show (circles)

the percentage of altruistic metapopulation that is obtained with β belonging to this bin. It is natural to expect that $\text{Alt}(x)$ would have sigmoidal shape - for low values of x its value is near 0, with bigger x the probability is increasing and finally for high x the probability is near 1. Therefore we suppose that $\text{Alt}(x) = 1 - 1/(1 + e^{\gamma(x-x_0)})$ for some γ, x_0 and then we find such γ, x_0 that maximizes the probability that the set of data obtained from random experiment based on our approximation of $\text{Alt}(x)$, is the same as the one actually measured. Finally, we find point where our approximation of $\text{Alt}(x)$ equals 1/2. This is very easy to do, since the solution of the equation

$$\text{Alt}(x) = 1 - \frac{1}{(1 + e^{\gamma(x-x_0)})} = 1/2$$

is $x = x_0$.

This approach is generally known as *maximum likelihood method*, since it maximizes the likelihood of obtaining the set of data actually measured. It gives us the sigmoidal fit, which is also plotted in the Figure 4.

Finally, we would like to argue, why such complicated analysis is necessary in this case.

1. The first reason is that we wanted to be compatible with the existing studies. Consider for instance the claims derived by Hamilton or Traulsen and Nowak. These say that altruism will prosper in the metapopulation (this has to be understood as a probabilistic claim, since everything is stochastic), if certain inequality holds. If one wants to test this hypothesis, it has to be defined what it means that altruism prospers. Therefore, we chose the most simple among many possible definitions and we set the threshold for ratio of altruists in the population (defining the population to be altruistic). Our understanding of the problem enables us to present interesting results, such as those in Figure 1.

2. Let us assume this hypothetical scenario: We compare two architectures. In the architecture A, 99.0% of population is altruistic, in the architecture B, the same quantity is 99.9%. The interpretation of these numbers is not easy. Either way, there is a lot of altruists in both populations and it is quite difficult to tell if the increase of altruistic ratio by 0.9% is insignificant (because it is just 0.9%) or very significant (because it reduces the ratio of selfish individuals to 1/10). Our interpretation is more intuitive. Let us note that β/α ratio can be seen as an effectiveness with which an individual helps its deme. For example, $\beta/\alpha = 3$ means that if individual gives up X resources (which it can alternatively use for its own breeding) for the sake of the whole deme, it helps the deme with 3X resources. With this in mind we can state the following: In A the altruism will favor, if the β/α ratio is at least 6 (everything the individual does for the sake of its deme helps the deme 6 times more than it hurts the individual). On the other hand, in B, the same ratio is reduced to 2. We consider such statements more eloquent than just saying that altruists ratio increases from 99.0 to 99.9 percent. Of course, there is one drawback - we must somehow define, what means that the metapopulation is altruistic. Luckily, there exists a canonical way how to do it. We present this definition in the next paragraph.

3. In our model, mutations cause many complications. Let us consider the case where just one allele out of five can make the individual altruistic. Clearly, even if there is no natural selection at all, this would make almost all individuals altruistic. Quite naturally, this is not what we want - it is misleading to say that 5-loci control is better for spreading of altruism just because the mutation pressure does its job. Therefore, to compare different architectures as objectively as possible, we always set the threshold exactly to the same value as corresponds to equilibrium frequency of altruists if only mutations were present.

Appendix 2 - Correspondence with known results

In the paper we have shown how r depends on η . These observations are in great accordance with [3] (for low mutation rate we have dependency that is very close to linear). Now we present how r depends on number of demes n . The figure 5 shows dependencies for three different values m (average number of individuals in deme). According to several papers such as [24], [3] we would expect these functions to be of a type a/x or $a + b/x$. Therefore we fitted the data-points with functions of a type $a + b/x$ with a, b constant.

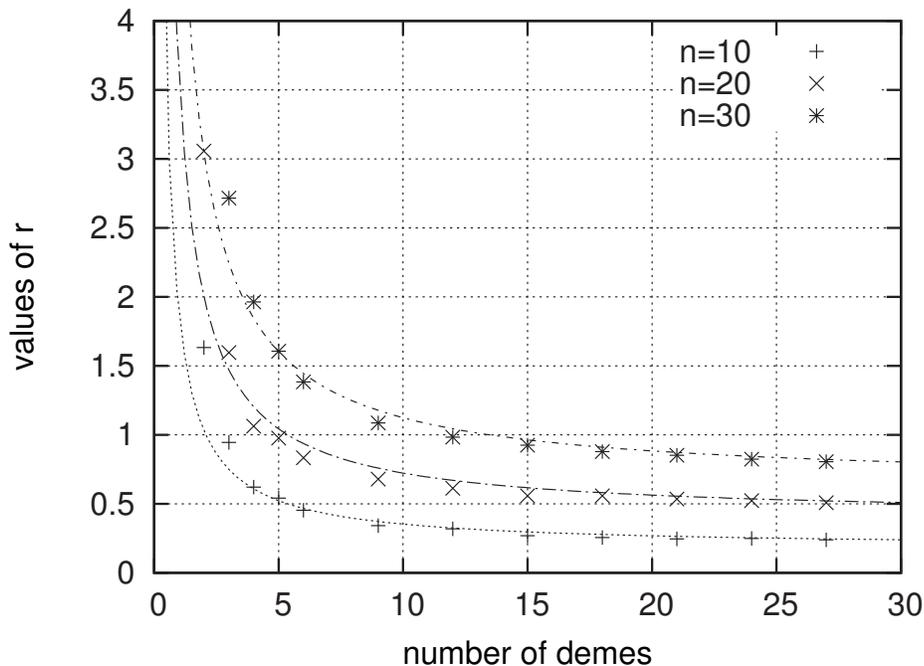


Figure 5: Dependence of r on n . Values of parameters: $\mu = 0.001$, $\eta = 0.01$.

Appendix 3 - Robustness of the results

All results were obtained by evolving the metapopulation for $N = 10000$ generations. Naturally, we would demand the results not to be (significantly) dependent on N . To test this we re-calculate results from Tables 1 and 2. In these experiments the metapopulation was evolving for $2N$ generations but the altruist-frequency data were collected only from the second half of the time. Results obtained with this process are shown in Tables 3 and 4 in its last column. As we can see, the results are almost identical.

Table 3: Models' parameters: $n = 20, m = 30, \mu = 0.001, \eta = 0.1$

Model 1	$g = 1, t = 1, \varepsilon = 1/2$	$r = 6.7 \pm 0.1$	6.7 ± 0.1
Model 2	$g = 3, t = 1, \varepsilon = 7/8$	$r = 6.7 \pm 0.1$	6.7 ± 0.1
Model 3	$g = 5, t = 1, \varepsilon = 31/32$	$r = 6.7 \pm 0.2$	6.8 ± 0.2
Model 4	$g = 5, t = 3, \varepsilon = 1/2$	$r = 6.5 \pm 0.1$	6.6 ± 0.1

Table 4: Models' parameters: $n = 20, m = 20, \mu = 0.0005, \eta = 0.01$

Model 5	$g = 1, t = 1, \varepsilon = 1/2$	$r = 0.49 \pm 0.01$	0.49 ± 0.01
Model 6	$g = 3, t = 1, \varepsilon = 7/8$	$r = 0.44 \pm 0.01$	0.43 ± 0.01
Model 7	$g = 5, t = 1, \varepsilon = 31/32$	$r = 0.38 \pm 0.02$	0.38 ± 0.02
Model 8	$g = 5, t = 3, \varepsilon = 1/2$	$r = 0.48 \pm 0.01$	0.49 ± 0.01

Appendix 4 - Non-threshold model for individual altruism

In this section we present results for linear model of individual altruism. This means that if an individual has x altruistic alleles he becomes x/g altruistic which means, that its disadvantage, as well as advantage for its deme, are only $\alpha^{x/g}$ and $\beta^{x/g}$. For weak selection the results can be easily summarized. Now we do not need any t and we set $\varepsilon = 1/2$. For parameters used in models 1-4 we obtain $r = 6.2 \pm 0.2$. For models 5-8 we obtain $r = 0.47 \pm 0.02$. Note that these values are very close to those from the paper and therefore we can conclude that threshold character is not very important here. For strong selection we present the same results as in Figure 3 in the paper with additional data for linear model.

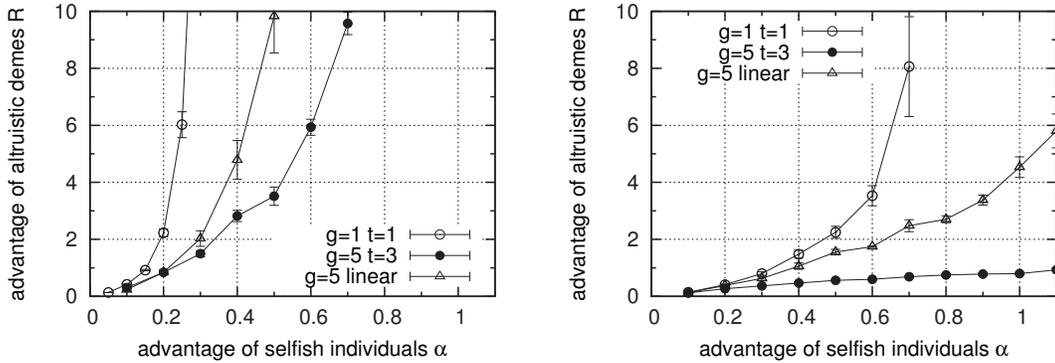


Figure 6: The same figure as Figure 3 in the paper with added dependency for linear control (triangles).