

1

2 **With a little help from my friends: Cooperation can accelerate**

3 **crossing of adaptive valleys**

4

5

6 Uri Obolski<sup>1,†</sup>, Ohad Lewin-Epstein<sup>1,†</sup>, Eran Even-Tov<sup>2</sup>, Yoav Ram<sup>1</sup>, Lilach Hadany<sup>1,\*</sup>

7 <sup>1</sup>Department of Molecular Biology and Ecology of Plants, Tel-Aviv University, Israel. <sup>2</sup>Department of

8 Molecular Microbiology and Biotechnology, Tel-Aviv University, Israel

9 <sup>†</sup> The 1st (Obolski U) and the 2nd (Lewin-Epstein O) authors contributed equally to this work

10

11

12 Running head: Cooperation accelerates adaptive peak shifts.

13

14

15 \*Corresponding author:

16 Lilach Hadany, Department of Molecular Biology and Ecology of Plants, Tel-Aviv

17 University, Israel 69978. Phone: +(972)-52-7455677, fax: +(972)36406886, eMail:

18 [lilach.hadany@gmail.com](mailto:lilach.hadany@gmail.com)

19

20 **KEY WORDS:** Peak shift, adaptive landscape, cooperation, rugged fitness landscape,

21 mathematical models/simulations, altruism.

## Abstract

Natural selection favors changes that lead to genotypes possessing high fitness. A conflict arises when several mutations are required for adaptation, but each mutation is separately deleterious. The process of a population evolving from a genotype encoding for a local fitness maximum to a higher fitness genotype is termed an adaptive peak shift.

Here we suggest cooperative behavior as a factor that can facilitate adaptive peak shifts. We model cooperation in a *public goods* scenario, wherein each individual contributes resources that are later equally redistributed among all cooperating individuals. We use mathematical modeling and stochastic simulations to study the effect of cooperation on peak shifts in well-mixed populations and structured ones. Our results show that cooperation can accelerate the rate of complex adaptation. Furthermore, we show that cooperation increases the population diversity throughout the peak shift process, thus increasing the robustness of the population to drastic environmental changes.

Our work could help explain adaptive valley crossing in natural populations and suggest that the long term evolution of a species depends on its social behavior.

## 41 Introduction

42 Adaptive landscapes, introduced by Sewall Wright in the 1930s (Wright 1932), are a  
 43 useful metaphor for the relationship between genotype and fitness. Under this  
 44 analogy, fitness is portrayed as a function of genotype, varying between different  
 45 allele combinations. Complex traits — which depend on two or more loci — can  
 46 produce rugged adaptive landscapes due to fitness interactions between loci. The  
 47 simplest instance of a rugged fitness landscape consists of two loci in which  
 48 mutations are jointly beneficial but separately deleterious. Adaptive valley crossing,  
 49 also known as an adaptive peak shift, has long been an evolutionary conundrum:  
 50 how can a population evolve to a higher fitness optimum if it has to "cross" a less fit  
 51 genotype on the way? Two main stages are needed for such an evolutionary process  
 52 (Michalakis and Slatkin 1996). First, the fitter genotype must appear in the  
 53 population. This could happen as a result of sequential mutations in the same  
 54 lineage, by recombinant offspring of mutant parents or by migration from another  
 55 population. Second, the fitter genotype will have to spread in the population.  
 56 However, these two stages have opposing optimal conditions (Weinreich et al.  
 57 2005). Mild selection facilitates the first stage of the process: single mutants are  
 58 more likely to survive in a mild selection regime, increasing the rate of appearance of  
 59 the fitter genotype by recombination or by acquisition of an additional mutation. In  
 60 contrast, the second stage is impeded when selection is mild because the fixation  
 61 probability of a rare genotype decreases (Eshel 1981).  
 62 In contrast to Wright's theory, some researchers have suggested that adaptive  
 63 landscapes, which can be extended to various topologies in multidimensional  
 64 genotype spaces (Kingman 1978; Kauffman and Weinberger 1989; Neidhart et al.

2014), might most commonly be single peaked (Bennett 1983; Whitlock 1995, 1997; Gavrillets 2004). However, these theories cannot account for all adaptive landscape topologies, and do not exempt evolutionary biologists from a characterization of evolution on a given rugged adaptive landscape.

Sewall Wright himself was the first to offer a theoretical solution to the adaptive peak shift problem (Wright 1932). His solution, the *Shifting Balance Theory*, was based on a subdivision of the population into small demes in which random genetic drift can increase the frequency of single deleterious mutants. Then, the second mutation can appear on the background of the single mutant. Finally, migration and natural selection can allow the new genotype to expand to other demes. In spite of its novelty and intuitive nature, the *Shifting Balance Theory* has been criticized for the limited range of realistic peak shift scenarios it explains (Coyne et al. 1997, 2000). Additionally, considerable amounts of research focused on finding unique conditions which can facilitate adaptive valley crossing: Dividing the population into smaller subpopulations connected by migration was shown to increase the rate of adaptive valley crossing, even without Wright's assumptions of increase and decrease in deme sizes as a function of the beneficial genotypes inhabiting them (Bitbol and Schwab 2014); Furthermore, dividing the population into only two populations connected by migration, but changing the selection pressure on each population, was also found to substantially reduce the waiting time for a peak shift (Hadany 2003); Mutation or recombination rates that increase with low fitness were shown to facilitate peak shifts, as this entails that less fit individuals can more rapidly adapt and traverse the fitness valley (Hadany and Beker 2003; Ram and Hadany 2014); Finally, assortative mating was also found to increase the rate of adaptive peak shifts in diploid

89 populations, as it increases the mating between individuals of the advantageous  
90 genotype, thus preventing recombination from breaking advantageous allele  
91 combinations (Williams and Sarkar 1994). Further theoretical research by Weissman  
92 et al. has established the rate of adaptive valley crossing for sexual (Weissman et al.  
93 2010) and asexual (Weissman et al. 2009) populations under different ranges of  
94 evolutionary forces such as selection, mutation, recombination and population size.  
95 Here we consider an additional factor that can affect the process of crossing  
96 adaptive valleys: cooperative behavior. We show that for a population subdivided  
97 into demes and connected by migration, cooperation between individuals within the  
98 same deme can considerably increase the rate of adaptive peak shifts.  
99 We focus on a *public goods* form of cooperation (Kagel and Roth 1995): all  
100 individuals within a deme contribute some resources (thus contributing fitness) to  
101 other deme members, and all deme members receive an equal amount of the  
102 redistributed resources. This cooperative behavior reduces the fitness difference  
103 between different genotypes and therefore effectively "smooths" the landscape to  
104 some degree. As a result, less fit mutants are more likely to survive with  
105 cooperation, increasing the rate of appearance of multiple mutants. Cooperation has  
106 an opposite effect on the fixation of the fittest genotype. Precisely because  
107 cooperation smooths the adaptive landscape, it reduces the relative advantage of  
108 the fittest genotype, and with it, its fixation probability. Nevertheless, we show that  
109 for intermediate levels of cooperation, the increase in the rate of appearance of the  
110 fittest genotype outweighs the decrease in its fixation probability, and altogether  
111 shortens the total adaptation time. Furthermore, we find that smoothing the  
112 adaptive landscape serves the cooperative population in another sense: it increases

the population diversity. This increase in diversity is beneficial in evolutionary terms, as it can help populations to overcome environmental changes, parasites etc. (Clarke 1979).

Overall, our results show that cooperation affects adaptive peak shifts substantially and might be an important and overlooked component of complex adaptation.

## Model

We model a population of sexually reproducing haploid individuals, containing two bi-allelic loci. The  $ab$  genotype is the wild type with fitness 1; the fitness of the single mutant genotypes,  $Ab$  and  $aB$ , is  $1 - s$ ; the fitness of the double mutant  $AB$  is  $1 + sH$ .  $s$  and  $H$  are the selection coefficients of the single mutant and the relative advantage of the double mutant, respectively ( $s > 0, H \geq 1$ ). We assume equal forward and backward mutation rates for both loci (defined in units of mutations per generation per locus), and denote them by  $\mu$ . Recombination occurs with rate  $r$  per generation per loci pair.

We model a population composed of  $n$  demes connected by migration, each of size  $k$ . Thus the total size of the population is  $N = n \cdot k$ . Each generation, individuals of the same deme cooperate. We model cooperation using a *public goods game* (Kagel and Roth 1995): Each individual in the deme contributes a constant fraction  $c$  of its fitness to the deme ( $0 \leq c \leq 1$ ), further referred to as the 'cooperation level'; the contributions are multiplied by a constant  $b$  and summed ( $b \geq 1$ ); finally, this sum is equally redistributed between the deme members. Hence,  $c$  is the cost of cooperative behavior whereas  $b$  determines the fold-increase of contributed resources due to cooperation. After cooperating, individuals migrate to other demes

with probability  $m$ . Setting  $m = 1 - 1/n$  defines a population in which offspring are uniformly distributed among all demes in every generation; setting  $m = 0$  determines each deme to be an effectively isolated population. For simplicity, we assume that migration to each deme is equiprobable. Finally, mating occurs between individuals of the newly formed demes. The offspring generation replaces the parent generation, so that population size remains constant and generations do not overlap.

**Table 1.** Parameters used in the analytical model and the stochastic simulation.

Parameter	Description
$n$	Number of demes
$k$	Size of each deme
$\mu$	Mutation rate
$r$	Recombination rate
$s$	Selection coefficient
$H$	Double mutant advantage
$c$	Fraction of resources contributed
$b$	Contributed resources multiplier
$m$	Migration rate

Fitness is determined by an individual's genotype and the effect of cooperation. We denote  $\omega_i$  as the initial fitness of individual  $i$ , determined by his genotype  $(ab, Ab, aB, AB)$ , and derive the fitness  $\omega_{i,D}$  of individual  $i$  in deme  $D$ , after considering cooperation:

152

$$(eq. 1) \omega_{i,D} = \omega_i \cdot (1 - c) + \frac{b}{k} \sum_{j \in D} \omega_j \cdot c$$

153

154

155 Within this framework, we first analyzed the expected waiting time to a peak shift  
 156 when each deme contains two individuals ( $k = 2$ ) and the population is fully mixed  
 157 ( $m = 1 - 1 / n$ ) using Branching processes analysis (Eshel 1981) (see  
 158 Supplementary Information 1). Then, we developed a Wright-Fisher stochastic  
 159 simulation to account for general deme sizes ( $k$ ) and migration rates ( $m$ ).  
 160 Simulations include the effects of natural selection, migration, mating,  
 161 recombination, mutation and drift (see Supplementary Information 2).  
 162 The simulations comprise three stages. In the first stage a population inhabited by  
 163 wild types evolves towards a mutation-selection balance. In the second stage, we  
 164 simulate the population until a double mutant appears for the first time. This allows  
 165 us to estimate the expected time for the appearance of a double mutant. In the third  
 166 stage, the double mutant either goes extinct or fixates in the population (determined  
 167 by reaching a frequency of 0.99). From this stage we can estimate the probability  
 168 that a double mutant will fixate (see Supplementary Information 2). Combining the  
 169 two measures (expected first appearance time and fixation probability) we can  
 170 estimate the expected waiting time for the appearance of a double mutant that  
 171 fixates (Hadany 2003). We compared simulation results to the analytical  
 172 approximation for a fully mixed population ( $k = 2, m = 1 - 1/n$ ); they were in  
 173 close agreement (Supplementary Information 1c).



We denote the adaptation time in a population with cooperation level  $c$  by  $\tau_c$ . The relative difference between the adaptation time of a cooperative population with cooperation level  $c > 0$  ( $\tau_c$ ) and a non-cooperative population ( $\tau_0$ ) is denoted by

$$\rho(c) := \frac{\tau_0 - \tau_c}{\tau_0}.$$

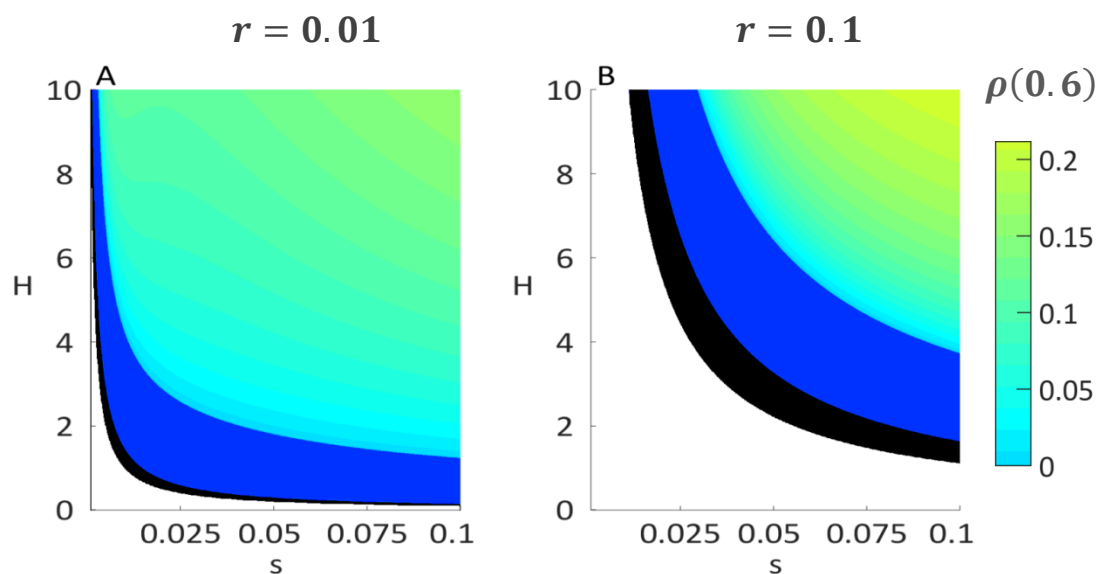
## Results

### A fully mixed population

We start with the special case where demes only have two individuals ( $k = 2$ ) and the population is fully mixed by migration in every generation ( $m = 1 - 1/n$ ). We estimate the relative difference in adaptation time due to cooperation,  $\rho(c)$ . Fig. 1 presents  $\rho(c = 0.6)$ , derived from our approximation (see Model and Supplementary Information 1), as a function of the selection coefficient,  $s$ , and the double mutant advantage,  $H$ . Higher recombination rates restrict the possibility of a peak shift, as recombination tends to break beneficial gene combinations (Eshel and Feldman 1970). This is consistent for cooperative behavior as well (Fig. 1, compare A and B). Cooperation has contrasting effects over the two main stages of the peak shift. First, it reduces the disadvantage of the single mutants ( $Ab$ ,  $aB$ ), therefore increasing their survival. As a result, the waiting time for the first double mutant ( $AB$ ) shortens in comparison to a non-cooperative population (Fig. 1). On the other hand, cooperation reduces the benefit of the double mutant, thus decreasing its survival probability. When selection against the single mutant ( $s$ ) or the advantage of the double mutant ( $H$ ) are high (upper right corners in Fig. 1), the decrease in waiting time for the double mutant's first appearance is more pronounced than the

decrease in its fixation probability, resulting in faster adaptation and thus a higher  $\rho$  value.

Note that some conditions allow for a peak shift only in a non-cooperative population, whereas a cooperative population is not expected to cross the adaptive valley (Fig. 1, black areas). The public goods cooperation, as modeled here, cannot expand the range of conditions allowing for a peak shift because cooperative behavior effectively decreases the double mutant advantage (see a formal proof for a fully mixed population in Supplementary Information 1b).



**Figure 1. Cooperation affects adaptation time in a mixed population.** Results are based on analytical approximations (Supplementary Information 1a). White areas represent parameters where both cooperatives and non-cooperatives are not expected to achieve a peak shift; black areas are parameters for which only non-cooperatives are expected to achieve a peak shift; blue areas represent parameters for which both cooperatives and non-cooperatives achieve a peak shift, but non-cooperatives accomplish the process faster. Teal to green areas are parameters for which cooperatives achieve peak shift faster than non-cooperatives, and the colors represent the relative difference in the expected time to peak shift due to cooperation ( $\rho$ ). Panels A and B present the

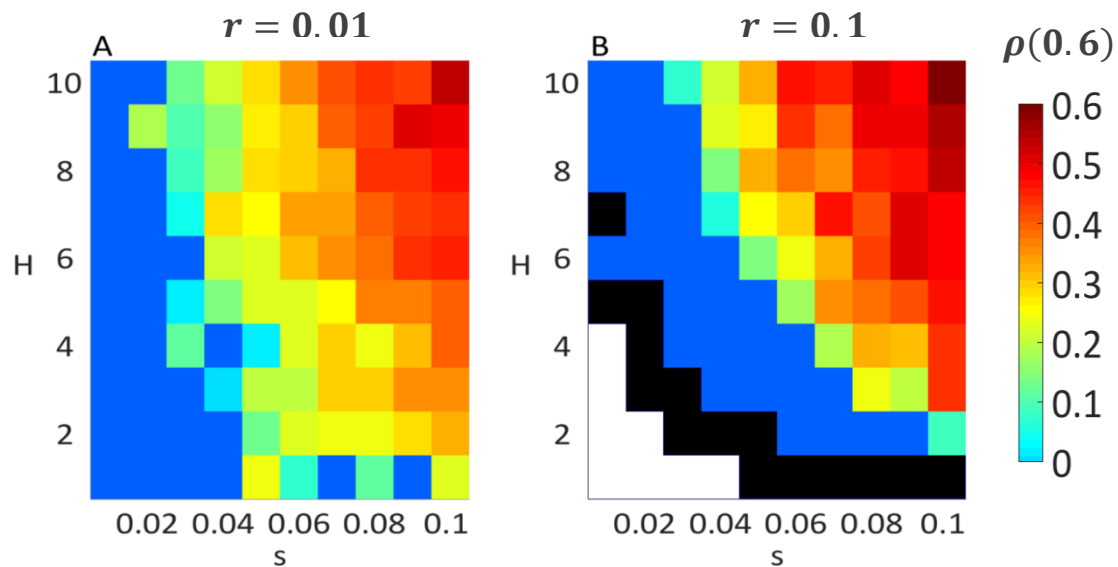
results for recombination rates  $r = 0.01, 0.1$ , respectively. Additional parameters are:  $N = 10,000$ ,  $\mu = 10^{-5}$ ,  $c = 0.6$ ,  $b = 1.2$ .

### **A subdivided population**

Next we analyzed the rate of adaptation in populations divided into demes containing more than two individuals ( $k > 2$ ). The division to larger demes changes the frequencies of single and double mutants, rendering our approximation no longer compatible for the multi-level selection between and within demes. Thus, all further analysis is based on stochastic simulations (see Model).

First, we examine the effects of selection ( $s$  and  $H$ ) on the adaptation time with and without cooperation. Fig. 2 shows the relative difference in adaptation time due to cooperation,  $\rho$ , for populations divided to demes that contain 10 individuals. Similarly to Fig. 1,  $\rho$  is a function of the selection coefficient,  $s$ , on the horizontal axis, and the double mutant advantage,  $H$ , on the vertical axis. The division of the population to demes retains the properties exemplified for a mixed population; high recombination values narrow the range of conditions leading to a peak shift (Fig. 2 compare A and B), and cooperation cannot extend the parameter range enabling the population to cross the adaptive valley (Fig. 2B, black area).

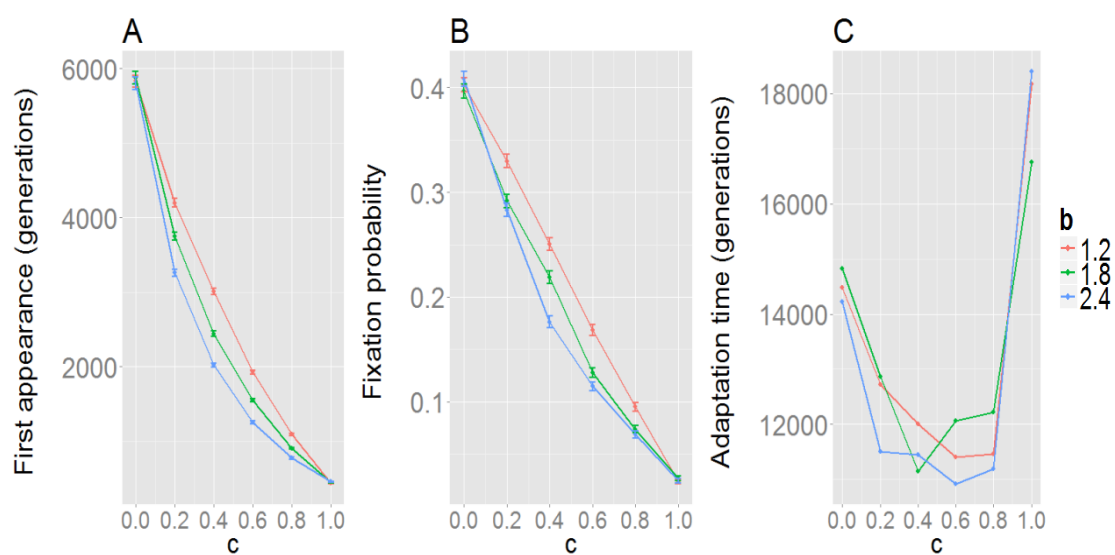
However, cooperation in a subdivided population accelerates complex adaptation even more than in a fully-mixed population; i.e., the subdivided population attains higher  $\rho$  values (note the red-yellow hues of Fig. 2 in comparison to Fig. 1; both figures have the same color scales).



**Figure 2. Cooperation affects adaptation time in a subdivided population.** Results are based on stochastic simulations (Supplementary Information 2). The x-axis represents the selection coefficient,  $s$ , and the y-axis represents the double mutant advantage,  $H$ . White areas represent parameters where both cooperatives and non-cooperatives are not expected to achieve a peak shift; black areas are parameters for which only non-cooperatives are expected to shift a peak; blue areas represent parameters for which both cooperatives and non-cooperatives achieve a peak shift, but non-cooperatives do so faster. Teal to red areas are parameters for which cooperatives achieve a peak shift faster than non-cooperatives, and the color represents the average relative difference in the expected time for a peak shift due to cooperation,  $\rho$ , averaged over  $\geq 1200$  simulations per parameter set. Panels A and B represent results for low and high recombination rates:  $r = 0.01, 0.1$ , respectively. Additional parameter values are  $n = 1,000$ ,  $k = 10$ ,  $\mu = 10^{-5}$ ,  $c = 0.6$ ,  $b = 1.2$ ,  $m = 0.01$ .

We investigated the effects of the cooperation levels,  $c$ , and the cooperation benefit,  $b$ , on the fixation of double mutants. In Fig. 3 we break down the dynamics of the double mutant fixation to the waiting for the appearance of a double mutant (Fig. 3A), the fixation probability of a double mutant (Fig. 3B), and the overall time to adaptation (Fig. 3C). For higher benefit produced from contributed resources,  $b$

values, the curves of first appearance and fixation probability become steeper (Fig. 3A and B). Higher  $b$  values effectively diminish the influence of each genotype's fitness and increase the influence of the pooled fitness. Thus, increasing  $b$  "smooths" the adaptive landscape and decreases the waiting time for the appearance of a double mutant (Fig. 3A). On the other hand, increasing  $b$  reduces the advantage of the double mutants and therefore decreases its fixation probability (Fig. 3B). Although high cooperation levels,  $c$  values, substantially shorten the waiting time for appearance of a double mutant, they also reduce its fixation probability. Therefore, intermediate cooperation levels minimize the adaptation time by striking a balance between shortening the waiting time for appearance of a double mutant and decreasing its fixation probability (Fig. 3C). Of course, when the population is non-cooperative ( $c = 0$ ),  $b$  doesn't affect the results. With full cooperation ( $c = 1$ ), all individuals within a deme have the same fitness regardless of their genotype, thus yielding the same relative fitness for all  $b$  values (eq. 1).



**Figure 3. Effect of cooperation on complex adaptation in a subdivided population.** The figure shows simulation results for (A) first appearance of a double mutant; (B) fixation probability of a double

mutant after appearance; and **(C)** total adaptation time. Markers show averages of  $\geq 5,000$  simulations; bars show standard error of the mean. Colors indicate different  $b$  values (red: 1.2, green: 1.8, blue: 2.4). Additional parameter values:  $r = 0.01$ ,  $n = 1,000$ ,  $k = 10$ ,  $\mu = 10^{-5}$ ,  $m = 0.01$ ,  $s = 0.05$ ,  $H = 5$ .

### **Population diversity**

Another interesting facet of the evolutionary process of peak shifts is the population diversity. A genetically diverse population is more robust to environmental changes that change genotypes' fitness, and is thus less likely to go extinct due to such changes (Clarke 1979). In order to measure the genetic diversity we use Shannon's Index, normalized by the number of genotypes:

$$D = \frac{-\sum_i p_i \ln(p_i)}{\ln(4)}, \text{ where } i \in \{ab, Ab, aB, AB\}$$

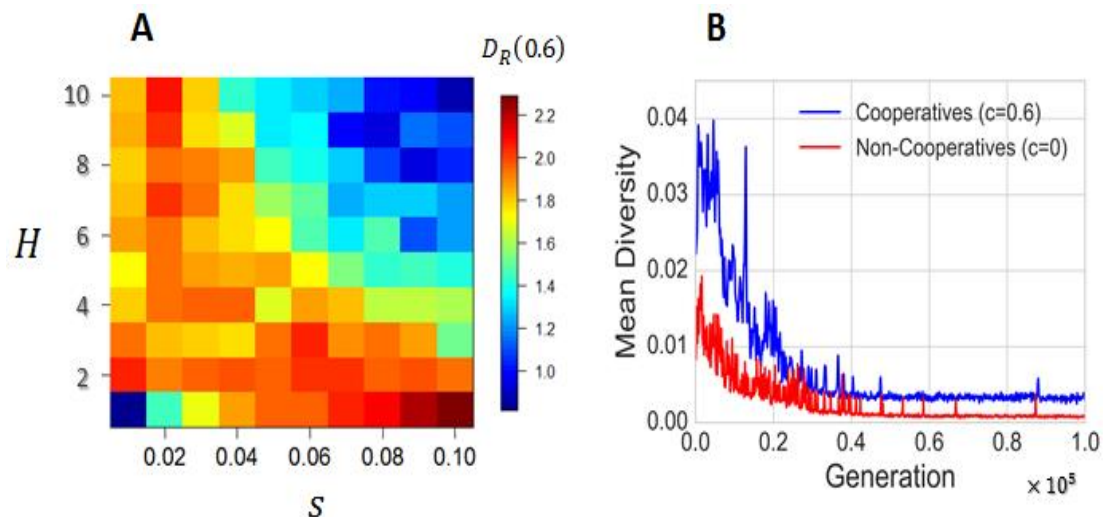
where  $p_i$  is the frequency of genotype  $i$  in the entire population.  $D$  ranges from zero to one, indicating only one genotype exists or that all genotypes are found in the population in equal frequencies, respectively.

We define the relative increase in diversity between cooperative and non-cooperative populations to be  $D_R(c) = \frac{D(c) - D(0)}{D(0)}$ . For diversity analysis, we simulated the peak shift dynamics for 100,000 generations and recorded the diversity every 100 generations. In Fig. 4 we show the difference in diversity between cooperative and non-cooperative populations,  $D_R(c)$ , against varying  $s$  and  $H$  values. For all examined selection coefficients,  $s$  and  $H$ , cooperation increases the diversity relative

296 to a non-cooperative population, usually by more than two-fold ( $D_R(0.6) > 1$ , Fig.  
 297 4A).

298 During a peak shift, populations reach mutation-selection balance under two  
 299 selective regimes: First before the successful double mutant appears and second  
 300 after it fixates. The population diversity in the second mutation selection balance is  
 301 lower than in the first, since the selective disadvantage of the single mutant  
 302 compared to the double mutant is higher than compared to the wild type.

303 Cooperation reduces the effective disadvantage of single mutants as well as the  
 304 advantage of double mutants, and therefore increases diversity both before and  
 305 after the peak shift (Fig. 4; Supplementary Information 3). However, when selection  
 306 is very weak, this effect diminishes (see Fig. 4, bottom-left corner). When selection is  
 307 strong, cooperators achieve a peak shift faster and spend more time in the second  
 308 mutation-selection balance, relative to non-cooperators (Supplementary Information  
 309 3). Hence, strong selection diminishes the diversity advantage of cooperators (Fig.  
 310 4A). Overall, we see that a cooperative population retains, on average, higher  
 311 diversity than a non-cooperative population during the entire peak shift process (Fig.  
 312 4B, Supplementary Information 3).



**Figure 4. Effect of cooperation on genetic diversity during a peak shift. (A)** The relative increase in diversity,  $D_R(0.6)$ , as a function of the selection coefficient,  $s$ , and the double mutant advantage,  $H$ . Cooperating populations are more diverse throughout the parameter range. Data is averaged over  $\geq 200$  simulations per parameter set. **(B)** The average genetic diversity (taken over  $\geq 200$  simulations) in a cooperative population (blue line) is higher than that of a non-cooperative population (red line) for selection coefficient  $s = 0.05$  and double mutant relative advantage  $H = 5$  at almost every time point. Additional parameter values are:  $n = 1,000$ ,  $k = 10$ ,  $\mu = 10^{-5}$ ,  $r = 0.01$ ,  $c = 0.6$ ,  $b = 1.2$ ,  $m = 0.01$ .

Although there is a marked increase in diversity for a cooperative population, the dynamics of the double mutant's fixation are similar in cooperative and non-cooperative populations. A peak shift can occur either by a rapid takeover of double mutants occupying entire demes, or by a gradual increase of the number of double mutants in each deme. We find that the maximal variance of double mutants among demes during the peak shift is similar with or without cooperation. However, the duration of the double mutant takeover in a cooperative population appears to be more variable than in a non-cooperative one (See Supplementary Information 4),



suggesting that a metapopulation composed of cooperative populations would be highly variable over a long period of time.

## **Discussion**

In this study we have shown that cooperation can be an important factor in the evolution of complex genotypes. We have found that in a *public goods* scenario, cooperative behavior can accelerate a peak shift, relative to non-cooperative behavior. However, in our model, cooperators can only achieve a peak shift under conditions enabling so for non-cooperators. Our results indicate that the adaptive advantage of cooperative behavior increases with the strength of selection, and that the range of conditions where cooperation is beneficial is expanded for low recombination rates. The effect of cooperation on fixation time is usually not monotonous, and intermediate values of cooperation (i.e. not a full investment of one's resources in cooperative behavior) might be optimal for achieving peak shifts. Additionally, we examined how cooperation affects genetic diversity. Cooperation smooths the adaptive landscape by decreasing selection intensity, and therefore increasing genetic diversity before, during and after a peak shift (Fig. 4 and Supplementary information 3).

Cooperation is a widespread biological phenomenon, with a vast body of theoretical literature supporting its evolutionary feasibility (Hamilton 1964; Trivers 1971; Axelrod and Hamilton 1981; Nowak 2006; West et al. 2007b). The focus of our work is the effect of cooperation on the dynamics of complex adaptation, rather than the conditions leading to the emergence of cooperative behavior. Importantly, we note

that our results hold for population with low relatedness (in a fully mixed population with migration rate  $m \approx 1$ ) as well as populations with high relatedness ( $m \ll 1$ ). Our model assumes that migration is equal between all demes (equivalent to spatial homogeneity between the demes). This assumption can be violated if migration is fitness-associated at the individual level. In this case, less fit individuals may be inclined to migrate more often in order to improve their offspring's genotypes (Gueijman et al. 2013) and cooperative populations would have lower effective migration rates than non-cooperative ones. However, we do not expect this to have a qualitative effect, because the advantage of cooperative populations does not depend on the level of relatedness.

Recently, Komarova has shown that peak shifts in asexual populations in a spatially heterogeneous environment can be facilitated by cooperation, when genes affecting cooperation also determine the fitness, and cooperators directly compete with non-cooperators (Komarova 2014). However, we model the loci determining the fitness as independent of cooperation, and the results are relevant for peak shifts of genes that are not directly affected by cooperation. Furthermore, we show that this can occur for sexual organisms, and even without explicit spatial constraints (*e.g.* fully mixed populations; Fig. 1).

Interestingly, cooperative behavior in our model does not change the maximal between-deme variance of the double mutant distribution (Supplementary information 4). This implies that the advantage of cooperation does not stem solely from increased genetic drift. Rather than increasing genetic drift, or bypassing the adaptive peak, like Wright and Fisher have suggested (Wright 1932; Bennett 1983), cooperation directly changes the intensity of selection. Nevertheless, cooperation

changes the variance of the double mutants' duration of fixation (Supplementary information 4). If different runs of our simulations are viewed as possible outcomes of independent populations under the same conditions, then the time needed for a double mutant to spread in these cooperative populations varies substantially. That some populations lag behind in crossing adaptive valleys might result in competition between populations or even lead to speciation. The added advantage due to high genetic diversity in cooperative populations can also influence their survival probability. Populations might encounter new parasites, predators, or abiotic environmental changes, against which some of the intermediate genotypes might have an advantage (Benton 2009). Maintaining intermediate genotypes could in such cases be a substantial advantage of cooperation.

Multicellular, sexually reproducing organisms are an obvious fit to our assumptions, if recombination rates between the relevant loci are low and selection is not too weak. Our model can also be relevant to bacteria, for example, which can face a peak shift challenge when developing antibiotic resistance (Salverda et al. 2011; de Visser and Krug 2014). Furthermore, some mutations that confer antibiotic resistance carry a fitness cost, but can be compensated by additional mutations that are beneficial in the presence of antibiotics and slightly deleterious in its absence (Andersson and Hughes 2010). Bacteria carrying both resistance and compensation mutations in an environment currently without antibiotics would need to cross an adaptive valley to become non-resistant and uncompensated. Although bacteria do not reproduce sexually, they can perform some horizontal gene transfer (Ochman et al. 2000; Thomas and Nielsen 2005), as befitting our model. Bacterial cooperation is also documented: bacteria often aggregate to produce biofilms or molecules that

can be considered as *public goods* (Kreft 2004; West et al. 2007a; Nadell et al. 2008).

Our results suggest that cooperative bacteria may enjoy an additional benefit of crossing adaptive valleys faster and having increased genetic diversity. Such knowledge on bacterial population dynamics might be used to devise strategies to fight antibiotic resistance, as other evolutionary processes are used for predictions of efficient treatment strategies (Obolski and Hadany 2012; Obolski et al. 2015; Perron et al. 2015; Caudill and Wares 2016; Obolski et al. 2016).

To conclude, we suggest a possible interplay between evolutionary forces and social behavior: Cooperative behavior can hasten the appearance of complex genotypes with increased fitness, which in turn might play a role in maintaining cooperative populations.

#### **Acknowledgements**

This project was supported by the Israeli Science Foundation 1568/13 (LH), the Minerva Center for Lab Evolution (LH) and by a fellowship from the Manna Program in Food Safety and Security (UO).

#### **References**

- Andersson, D. I. and D. Hughes. 2010. Antibiotic resistance and its cost: is it possible to reverse resistance? *Nature Reviews Microbiology* 8:260-271.
- Axelrod, R. and W. D. Hamilton. 1981. The evolution of cooperation. *Science* 211:1390-1396.
- Bennett, J. H. 1983. Natural selection, heredity, and eugenics: Including selected correspondence of RA Fisher with Leonard Darwin and others.

426 Benton, M. J. 2009. The Red Queen and the Court Jester: species diversity and the  
427 role of biotic and abiotic factors through time. *Science* 323:728-732.

428 Bitbol, A.-F. and D. J. Schwab. 2014. Quantifying the role of population subdivision in  
429 evolution on rugged fitness landscapes. *PLoS computational biology*  
430 10:e1003778.

431 Caudill, L. and J. R. Wares. 2016. The role of mathematical modeling in designing and  
432 evaluating antimicrobial stewardship programs. *Current Treatment Options*  
433 *in Infectious Diseases* 8:124-138.

434 Clarke, B. 1979. The evolution of genetic diversity. *Proceedings of the Royal Society*  
435 *of London B: Biological Sciences* 205:453-474.

436 Coyne, J. A., N. H. Barton, and M. Turelli. 1997. Perspective: a critique of Sewall  
437 Wright's shifting balance theory of evolution. *Evolution*:643-671.

438 Coyne, J. A., N. H. Barton, and M. Turelli. 2000. Is Wright's shifting balance process  
439 important in evolution? *Evolution* 54:306-317.

440 de Visser, J. A. G. and J. Krug. 2014. Empirical fitness landscapes and the  
441 predictability of evolution. *Nature Reviews Genetics* 15:480-490.

442 Eshel, I. 1981. On the survival probability of a slightly advantageous mutant gene  
443 with a general distribution of progeny size—a branching process model.  
444 *Journal of mathematical biology* 12:355-362.

445 Eshel, I. and M. W. Feldman. 1970. On the evolutionary effect of recombination.  
446 *Theoretical population biology* 1:88-100.

447 Gavrillets, S .2004 .Fitness landscapes and the origin of species (MPB-41). Princeton  
448 University Press Princeton, NJ.

449 Gueijman, A., A. Ayali, Y. Ram, and L. Hadany. 2013. Dispersing away from bad  
450 genotypes: the evolution of Fitness-Associated Dispersal (FAD) in  
451 homogeneous environments. *BMC evolutionary biology* 13:125.

452 Hadany, L. 2003. Adaptive peak shifts in a heterogenous environment. *Theoretical*  
453 *population biology* 63:41-51.

454 Hadany, L. and T. Beker. 2003. Fitness-associated recombination on rugged adaptive  
455 landscapes. *Journal of evolutionary biology* 16:862-870.

456 Hamilton, W. 1964. The genetical evolution of social behaviour. I.

457 Kagel, J. H. and A. E. Roth. 1995. *The handbook of experimental economics*.  
458 Princeton university press Princeton, NJ.

459 Kauffman, S. A. and E. D. Weinberger. 1989. The NK model of rugged fitness  
460 landscapes and its application to maturation of the immune response.  
461 *Journal of theoretical biology* 141:211-245.

462 Kingman, J. 1978. A simple model for the balance between selection and mutation.  
463 *Journal of Applied Probability*:1-12.

464 Komarova, N. L. 2014. Spatial interactions and cooperation can change the speed of  
465 evolution of complex phenotypes. *Proceedings of the National Academy of*  
466 *Sciences* 111:10789-10795.

467 Kreft, J.-U. 2004. Biofilms promote altruism. *Microbiology* 150:2751-2760.

468 Michalakis, Y. and M. Slatkin. 1996. Interaction of selection and recombination in the  
469 fixation of negative-epistatic genes. *Genetical research* 67:257-269.

470 Nadell, C. D., J. B. Xavier, S. A. Levin, and K. R. Foster. 2008. The evolution of quorum  
471 sensing in bacterial biofilms. *PLoS biology* 6:e14.

472 Neidhart, J., I. G. Szendro, and J. Krug. 2014. Adaptation in tunably rugged fitness  
473 landscapes: the rough Mount Fuji model. *Genetics* 198:699-721.

474 Nowak, M. A. 2006. Five rules for the evolution of cooperation. *science* 314:1560-  
475 1563.

476 Obolski, U., E. Dellus-Gur, G. Y. Stein, and L. Hadany. 2016. Antibiotic cross-resistance  
477 in the lab and resistance co-occurrence in the clinic: Discrepancies and  
478 implications in *E. coli*. *Infection, Genetics and Evolution* 40:155-161.

479 Obolski, U. and L. Hadany. 2012. Implications of stress-induced genetic variation for  
480 minimizing multidrug resistance in bacteria. *BMC medicine* 10:89.

481 Obolski, U., G. Y. Stein, and L. Hadany. 2015. Antibiotic Restriction Might Facilitate  
482 the Emergence of Multi-drug Resistance. *PLoS Comput Biol* 11:e1004340.

483 Ochman, H., J. G. Lawrence, and E. A. Groisman. 2000. Lateral gene transfer and the  
484 nature of bacterial innovation. *Nature* 405:299-304.

485 Perron, G. G., R. F. Inglis, P. S. Pennings, and S. Cobey. 2015. Fighting microbial drug  
486 resistance: a primer on the role of evolutionary biology in public health.  
487 *Evolutionary applications* 8:211-222.

488 Ram, Y. and L. Hadany. 2014. Stress-induced mutagenesis and complex adaptation.  
489 *Proceedings of the Royal Society B: Biological Sciences* 281:20141025.

490 Salverda, M. L., E. Dellus, F. A. Gorter, A. J. Debets, J. Van Der Oost, R. F. Hoekstra, D.  
491 S. Tawfik, and J. A. G. de Visser. 2011. Initial mutations direct alternative  
492 pathways of protein evolution. *PLoS Genet* 7:e1001321.

493 Thomas, C. M. and K. M. Nielsen. 2005. Mechanisms of, and barriers to, horizontal  
494 gene transfer between bacteria. *Nature reviews microbiology* 3:711-721.

495 Trivers, R. L. 1971. The evolution of reciprocal altruism. Quarterly review of  
496 biology:35-57.

497 Weinreich, D. M., L. Chao, and P. Phillips. 2005. Rapid evolutionary escape by large  
498 populations from local fitness peaks is likely in nature. Evolution 59:1175-  
499 1182.

500 Weissman, D. B., M. M. Desai, D. S. Fisher, and M. W. Feldman. 2009. The rate at  
501 which asexual populations cross fitness valleys. Theoretical population  
502 biology 75:286-300.

503 Weissman, D. B., M. W. Feldman, and D. S. Fisher. 2010. The rate of fitness-valley  
504 crossing in sexual populations. Genetics 186:1389-1410.

505 West, S. A., S. P. Diggle, A. Buckling, A. Gardner, and A. S. Griffin. 2007a. The social  
506 lives of microbes. Annual Review of Ecology, Evolution, and Systematics:53-  
507 77.

508 West, S. A., A. S. Griffin, and A. Gardner. 2007b. Evolutionary explanations for  
509 cooperation. Current Biology 17:R661-R672.

510 Whitlock, M. C. 1995. Variance-induced peak shifts. Evolution:252-259.

511 Whitlock, M. C. 1997. Founder effects and peak shifts without genetic drift: adaptive  
512 peak shifts occur easily when environments fluctuate slightly.  
513 Evolution:1044-1048

514 Williams, S. M. and S. Sarkar. 1994. Assortative mating and the adaptive landscape.  
515 Evolution:868-875.

516 Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in  
517 evolution. Pp. 356-366. Proceedings of the sixth international congress on  
518 genetics.



## Supplementary Information

### Supplementary Information 1

#### 1a. Analytic approximation

Using Branching processes (Harris 1948) we analyze the probability of a peak shift for the case of  $k = 2$  and  $m = 1 - \frac{1}{n}$ . We define  $\omega_{x,y}$  to be the fitness of genotype  $x$ , when genotype  $y$  is its partner. Since the fitness of  $Ab$  and  $aB$  is the same, and so is their contribution to their partners we will denote  $sm$  (single mutant) as either  $Ab$  or  $aB$ .

$$\omega_{sm_{ab}} = (1 - s) \cdot (1 - c) + \frac{1}{2} \cdot c \cdot (1 - s + 1) \cdot b$$

$$\omega_{AB_{ab}} = (1 + sH) \cdot (1 - c) + \frac{1}{2} \cdot c \cdot (1 + sH + 1) \cdot b$$

$$\omega_{AB_{sm}} = (1 + sH) \cdot (1 - c) + \frac{1}{2} \cdot c \cdot (1 + sH + 1 - s) \cdot b$$

$$\omega_{ab_{ab}} = (1 - c) + c \cdot b$$

$$\omega_{AB_{AB}} = (1 + sH) \cdot (1 - c) + c \cdot b \cdot (1 + sH)$$

$$\omega_{sm_{sm}} = (1 - s) \cdot (1 - c) + c \cdot b \cdot (1 - s)$$

Since  $ab$  is the wild type and is the most common genotype, and since the population is fully mixed, most cooperating couples are of two  $ab$  individuals. Therefore the fitness of most individuals is  $\omega_{ab_{ab}}$ .

We normalize the fitness of all individuals relative to the most common fitness  $\omega_{ab_{ab}}$  and define:

$$\omega_{sm_{ab}}' = \frac{\omega_{sm_{ab}}}{\omega_{ab_{ab}}}$$

$$\omega_{AB_{ab}}' = \frac{\omega_{AB_{ab}}}{\omega_{ab_{ab}}}$$

$$\omega_{AB_{sm}}' = \frac{\omega_{AB_{sm}}}{\omega_{ab_{ab}}}$$

532 By neglecting terms of the order of  $\mu^2$  we can say that the partner of an  $sm$   
 533 individual, at the mutation-selection balance, would be  $ab$ . Now we can calculate  $s'$ ,  
 534 the effective selection coefficient when taking cooperation into account:

$$535 \quad s' = 1 - \omega_{sm_{ab}}'.$$

536 We can approximate the probability  $q$  that an  $AB$  individual would be formed in the  
 537 next generation, given no other  $AB$  individuals exist in the current generation:

$$q = \left(\frac{\mu}{s'}\right)^2 \cdot (r + 2 \cdot s' + s'^2)$$

538 Hence, the probability that the first  $AB$  individual would appear in the population at  
 539 a certain generation, given that it had not appeared earlier is:

$$1 - (1 - q)^N$$

540 which can be approximated by  $q \cdot N$ , where  $N = n \cdot k$  is the size of the population.

541 The expected time for appearance of the first  $AB$  individual ( $T_{first}$ ), which is  
 542 geometrically distributed, is:

$$E[T_{first}] \approx \frac{1}{q \cdot N}$$

543 The approximation of the average fitness, marked by  $\bar{\omega}$ , is:

$$\bar{\omega} = 1 - 2 \cdot \left(\frac{\mu}{s'}\right) + 2 \cdot \left(\frac{\mu}{s'}\right) \cdot (1 - s') = 1 - 2\mu$$

544 The progeny of an  $AB$  individual, marked by  $p$ , is:

$$p = \frac{\left((1 - r) \cdot \left(1 - 2 \cdot \left(\frac{\mu}{s'}\right)\right) + 2 \cdot \left(\frac{\mu}{s'}\right)\right) \cdot \omega_{AB_{ab}}'}{1 - 2\mu}$$

545 Therefore, if we assume that the number of offspring is Poisson distributed and its  
 546 mean is slightly above one (Eshel 1981; Hadany 2003), the probability that an  $AB$

547 genotype would fixate in the population, rather than go extinct, denoted by  $\pi$ , is  
 548 approximated by:

$$\pi \approx \frac{2 \cdot (p - 1)}{p}$$

549 Thus we can approximate the expected time for appearance of an  $AB$  individual that  
 550 will go to fixation by:

$$E[T_{tot}] \approx \frac{E[T_{first}]}{\pi}$$

551

# 552 1b. Expected progeny analysis for partners:

553 If the fitness of an  $AB$  individual is  $1 + sH$ , and  $\mu \ll s$ , then the expected number of  
 554 the  $AB$ 's offspring which are themselves  $AB$  is approximated by:

555 (eq.s1)  $\alpha = (1 - r)(1 + sH)$

556 The fitness of an  $AB$  when its partner is  $ab$ :

557 (eq.s2)  $\omega_{AB_{ab}} = (1 - c)(1 + sH) + \frac{1}{2}cb(2 + sH)$

558 The fitness of a  $ab$  when its partner is  $ab$ :

559 (eq.s3)  $\omega_{ab_{ab}} = (1 - c) + cb$

560 We normalize the new fitness

561 (eq.s4)  $\omega'_{AB_{ab}} = \frac{\omega_{AB_{ab}}}{\omega_{ab_{ab}}} = \frac{2(1-c)(1+sH)+cb(2+sH)}{2(1-c)+2cb} = 1 + \frac{2(1-c)sh+cbsH}{2(1-c+cb)}$

562 Using (eq.s4) and (eq.s1) we get:

563 (eq.s5)  $\alpha = (1 - r) \cdot \left(1 + \frac{2(1-c)sh+cbsH}{2(1-c+cb)}\right)$

564 In order for  $AB$  to fixate, we require that  $\alpha > 1$ .

565 Since  $0 \leq c \leq 1$  and  $b \geq 1$  we get that the denominator in (eq.s5)  $2(1 - c + cb)$  is

566 always positive and therefore we can multiply (eq.s5) and get:

$$(eq.s6) (1 - r) + \frac{(2(1-c)sH + cbsH) \cdot (1-r)}{2(1-c+cb)} > 1$$

$$\Leftrightarrow (2(1-c)sH + cbsH) \cdot (1-r) > 2r(1-c+cb)$$

$$\Leftrightarrow sH \cdot (2-2c+cb)(1-r) > 2r(1-c+cb)$$

Since  $r < 1$  and  $c \leq 1$ , we have that  $(2-2c+cb)(1-r) > 0$  and therefore

$$(eq.s7) sH > \frac{2r(1-c+cb)}{(1-r)(2-2c+cb)} = \frac{r}{1-r} \cdot \frac{2-2c+2cb}{2-2c+cb} = \frac{r}{1-r} \cdot \left(1 + \frac{cb}{2-2c+cb}\right)$$

Differentiating the right hand side of (eq.s7) with respect to  $c$  yields:

$$(eq.s8) \frac{d\left(\frac{r}{1-r} \cdot \left(1 + \frac{cb}{2-2c+cb}\right)\right)}{dc} = \frac{r}{1-r} \cdot \frac{b(2-2c+cb) - cb(b-2)}{(2-2c+cb)^2} = \frac{r}{1-r} \cdot \frac{2b}{(2-2c+cb)^2} > 0,$$

$$\forall 0 \leq c \leq 1, b \geq 1$$

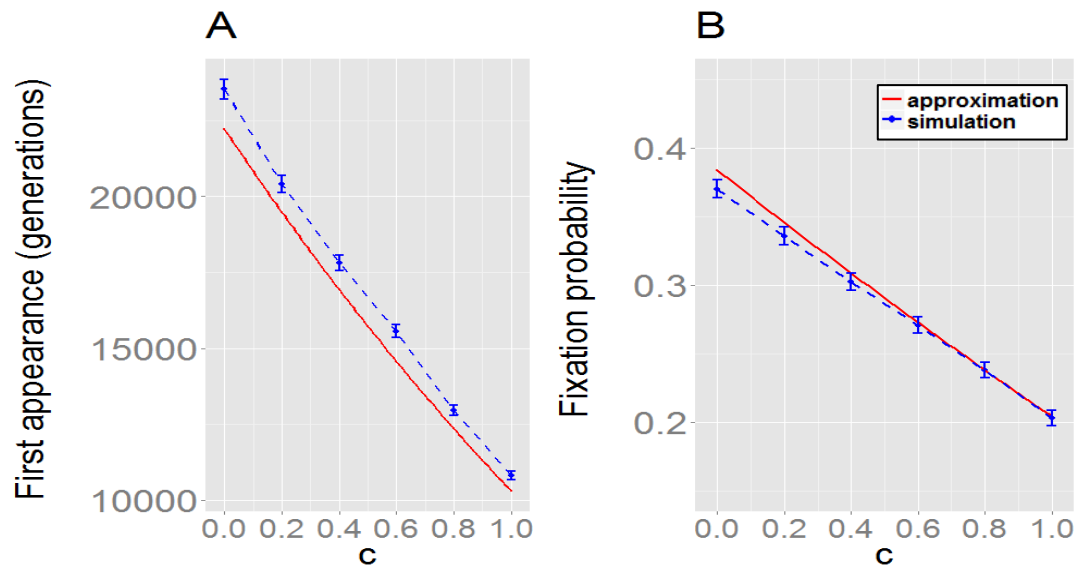
Therefore  $\frac{r}{1-r} \cdot \left(1 + \frac{cb}{2-2c+cb}\right)$  is monotonically increasing in  $c$  for all  $b \geq 1$ , and

$$\left(1 + \frac{cb}{2-2c+cb}\right) = 1 \text{ only when } c = 0.$$

**Thus, any increase in  $c$  restricts the parameter range enabling  $AB$ 's fixation.**

### 1c. Comparison between analytical approximation and simulation results

In order to verify our approximations we used stochastic simulations. Fig. S1 shows the results of the simulation (blue line) and the approximation (red line), for various cooperation levels. This is shown for the time of first appearance of the double mutant (Fig. 2A) and the double mutant's fixation probability (Fig. 2B), as a function of the cooperation level ( $c$ ). When  $c = 0$ , we are reduced to the results presented, and verified against simulations, in (Hadany 2003). We can see that our approximation is accurate when  $c$  increases, but it does remain slightly biased. Note that both the waiting time for appearance of the double mutant and its fixation probability decrease with  $c$ .



**Figure S1. Comparison between simulation and analytical approximation in a fully mixed population.** We plot the effect of cooperation on the first appearance of a double mutant and its fixation. As the level of cooperation increases (x-axis) both the waiting time for appearance of a double mutant (A) and the fixation probability of a double mutant (B) decrease. We can see that the analytical approximation and the simulation results are in close agreement. Other parameters are:

$$s = 0.05, H = 5, n = 5,000, k = 2, m = 1 - \frac{1}{n} \mu = 10^{-5}, r = 0.01, b = 1.2.$$

### Supplementary Information 2 – Simulation design

The simulation is composed of a genotype frequencies vector, representing the structure and composition of the population, and of several functions, each representing a different process: selection, migration, mating and recombination, mutation and drift. One generation ends after all functions have been applied and the population is replaced.

We initialize the simulation with a population comprised solely of wild type individuals. Every generation, we apply the selection function and calculate the

fitness separately for each deme. We calculate the 'donation pool' of every deme and then derive the new fitness of each individual by (eq. 1) in the main text:

$$\omega_{i,D} = \omega_i \cdot (1 - c) + \frac{b}{k} \sum_{j \in D} \omega_j \cdot c$$

After that we apply the migration function to the fitness-weighted frequencies. The frequency of genotype  $i$  in deme  $D$ , after selection and migration (marked as  $i'_D$ ) is:

$$(eq.s9) i'_D = (1 - m) \cdot i_D \omega_{i,D} + \sum_{d \neq D}^n i_d \omega_{i,d} \cdot \frac{m}{n-1}$$

After selection and migration, we apply the mating-recombination function. We calculate the frequencies of the genotypes in the next generation, within each deme, according to the frequencies in the parent generation, and the recombination rate,  $r$ . Finally, we apply mutation.

Note that we do not normalize the results after applying the selection function. Thus, demes with higher mean fitness would export more individuals.

The last phase in each generation is drift, simulated by drawing  $k$  random numbers from a multinomial distribution based on the relative genotype frequency in each deme, and repopulating the demes. We then transform the quantities of every genotype back into frequencies. After applying all the functions, and initiating the new frequency vector, the simulation proceeds to the next generation.

We start each simulation with running 3,000 generations, beginning with a population composed solely of wild type individuals. At this stage we define the fitness of genotype AB to be 0 in order to reach mutation selection balance while no double mutant is yet formed.

The parameters we use are  $n, k, r, \mu, s, h, b, c, m$ , as defined in Table 1 in the main text. The frequency vector has  $4 \cdot n$  coordinates, where 4 coordinates describe the frequencies of the different genotypes in a specific deme. For instance, for the population vector  $V$ , the frequencies of genotypes  $ab, Ab, aB, AB$  in deme  $j$  are  $[V(4j), V(4j + 1), V(4j + 2), V(4j + 3)]$ .

## 1. Fitness

For deme  $k$  we calculate the pool of fitness donations using matrix product:

$$pool_k = (ab \quad Ab \quad aB \quad AB)_k \begin{pmatrix} 1 \cdot c \cdot b \\ (1-s) \cdot c \cdot b \\ (1-s) \cdot c \cdot b \\ (1+sh) \cdot c \cdot b \end{pmatrix}$$

After calculating the pool we can calculate the new fitness of every individual, and by that, the frequency of every genotype after applying selection. The new frequencies of genotypes in deme  $k$  would be:

$$\begin{pmatrix} ab' \\ Ab' \\ aB' \\ AB' \end{pmatrix}_k = (ab \quad Ab \quad aB \quad AB)_k \overbrace{\begin{pmatrix} 1 \cdot (1-c) + pool_k \\ (1-s) \cdot (1-c) + pool_k \\ (1-s) \cdot (1-c) + pool_k \\ (1+sh) \cdot (1-c) + pool_k \end{pmatrix}}^{\text{New fitness vector}}$$

Where the vector  $(ab \quad Ab \quad aB \quad AB)_k$  represents the frequencies of the genotypes in generation  $N$ , and the vector  $(ab' \quad Ab' \quad aB' \quad AB')_k$  represents the frequencies of the genotypes after selection.

## 2. Migration

According to eq.s9, we can define a migration matrix that would simulate the change in the population composition due to migration. For instance, if there are 3 demes in the population, the matrix would be defined as follows:

$$\begin{pmatrix} 1-m & 0 & 0 & 0 & m/2 & 0 & 0 & 0 & m/2 & 0 & 0 & 0 \\ 0 & 1-m & 0 & 0 & 0 & m/2 & 0 & 0 & 0 & m/2 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & m/2 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & m/2 \\ m/2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & m/2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & m/2 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & m/2 \\ m/2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & m/2 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & m/2 & 0 & 0 & 0 & m/2 & 0 & 0 & 0 & 1-m & 0 \\ 0 & 0 & 0 & m/2 & 0 & 0 & 0 & m/2 & 0 & 0 & 0 & 1-m \end{pmatrix}$$

### 3. Mating and recombination

Let  $i \in \{ab, Ab, aB, AB\}$  be the frequency of genotype  $i$  in a specific deme in generation  $N$ , and  $i'$  the frequency of genotype  $i$  in that same deme in generation  $N + 1$ . By this we get:

$$\begin{pmatrix} ab' \\ Ab' \\ aB' \\ AB' \end{pmatrix} = \begin{pmatrix} ab^2 + ab \cdot aB + ab \cdot aB + ab \cdot AB \cdot (1-r) + Ab \cdot aB \cdot r \\ ab \cdot Ab + Ab^2 + ab \cdot AB \cdot r + Ab \cdot AB + Ab \cdot aB \cdot (1-r) \\ ab \cdot aB + ab \cdot AB \cdot r + Ab \cdot aB \cdot (1-r) + aB^2 + aB \cdot AB \\ ab \cdot AB \cdot (1-r) + Ab \cdot aB \cdot r + Ab \cdot AB + aB \cdot AB + AB^2 \end{pmatrix}$$

### 4. Mutation function

The frequency of each genotype after mutation takes place can be written as:

$$\begin{pmatrix} ab' \\ Ab' \\ aB' \\ AB' \end{pmatrix} = \overbrace{\begin{pmatrix} (1-\mu)^2 & \mu \cdot (1-\mu) & \mu \cdot (1-\mu) & \mu^2 \\ \mu \cdot (1-\mu) & (1-\mu)^2 & \mu^2 & \mu \cdot (1-\mu) \\ \mu \cdot (1-\mu) & \mu^2 & (1-\mu)^2 & \mu \cdot (1-\mu) \\ \mu^2 & \mu \cdot (1-\mu) & \mu \cdot (1-\mu) & (1-\mu)^2 \end{pmatrix}}^{\text{mutation matrix}} \begin{pmatrix} ab \\ Ab \\ aB \\ AB \end{pmatrix}$$



In this notation,  $i \in \{ab, Ab, aB, AB\}$  represent the frequency of genotype  $i$  in a specific deme before mutation and  $i'$  represents the frequency of genotype  $i$  in that same deme after mutation takes place. Since the mutation occurrences are independent variables, each deme is multiplied by the same matrix. Moreover, the probability of mutation occurrence is homogenous in time; therefore the same matrix is valid for the entire simulation.

Simulations were performed using Python 3.3. When multiple parameter sets were examined, assignments with random selection of the desired parameter sets were sent to a computer cluster until the pre-determined number of runs was satisfied for each parameter set. Hence the statements of 'at least ## simulations' found in the main text.

### **Supplementary Information 3 – Diversity analysis**

In order to see the change in genetic diversity as a function of generations, we plotted the average graph for several parameter sets. For each parameter set, we chose simulations that ended with fixation of double mutants (limited by 100,000 generations) and calculated the average of two time points:

- $T_1$  – The generation of the first appearance of a double mutant that will eventually fixate (in contrast with appearance of double mutants that appear and then extinct).
- $T_2$  – The generation in which the double mutants reached a frequency of 0.99 for the first time.

These two time points divide the process to three phases:

- 679 1. MSB1 – First mutation-selection balance. The time until the first  
680 appearance of a successful double mutant.
- 681 2. Takeover – The phase between the first appearance of a successful  
682 double mutant and its fixation.
- 683 3. MSB2 – Second mutation-selection balance, after the fixation of the  
684 double mutants.

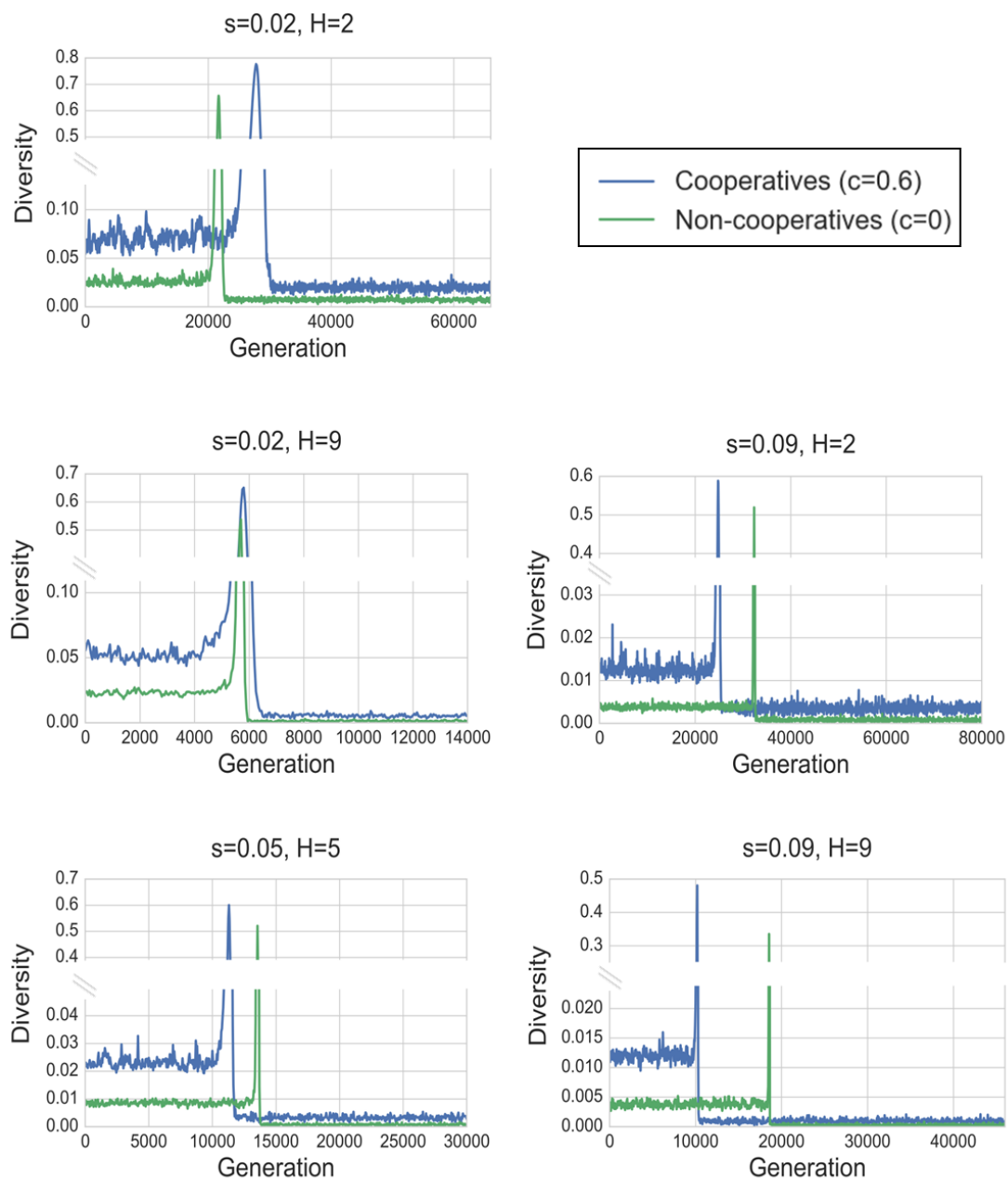
685 For simulation  $j$  let  $t_1^j, t_2^j$  be the time of the first appearance of a successful double  
686 mutant and the time of its fixation, respectively. We normalize each generation  $i$  in  
687 simulation  $j$  to its relative location within its phase in the following manner:

$$i' = \begin{cases} \frac{i}{t_1^j} \cdot T_1 & i \leq t_1^j \\ T_1 + \frac{i - t_1^j}{t_2^j - t_1^j} \cdot (T_2 - T_1) & t_1^j < i \leq t_2^j \\ T_1 + T_2 + \frac{i - t_2^j}{100000 - t_2^j} \cdot (100000 - T_2) & t_2^j < i \leq 100,000 \end{cases}$$

688

689 After normalization, we smoothed the diversity of the simulations with a 10  
690 generation moving average and plot the results. In the plots presented below, we  
691 show that cooperatives attain higher diversity during the first and the second  
692 mutation-selection balances. In addition, during the takeover cooperatives reach  
693 higher levels of diversity. We can see that occasionally the diversity in a non-  
694 cooperative population is higher than in a cooperative population. This happens only  
695 for relatively short periods and mostly when the two populations are not well  
696 synchronized: For example if the cooperative population has already crossed the  
697 valley and is in the second mutation-selection balance (after fixation) while the non-

698 cooperative population is still in the first stage (see  $s=0.09$ ,  $H=9$  plot; generations  
699  $\sim 10,000$ - $20,000$ ). Nevertheless, as shown in the Fig. 4 in the main text, the average  
700 diversity in the entire process is almost always higher for cooperative populations.



701  
702 **Figure S2. Mean diversity is higher for cooperative population.** For each phase (MSB1, Takeover,  
703 MSB2) the diversity in a cooperative population is higher than in a non-cooperative population. Each  
704 plot is based on at least 180 simulations. Note that the y-axes are broken.  
705 Parameters are:  $n = 1,000$ ,  $k = 10$ ,  $\mu = 10^{-5}$ ,  $r = 0.01$ ,  $c = 0.6$ ,  $b = 1.2$

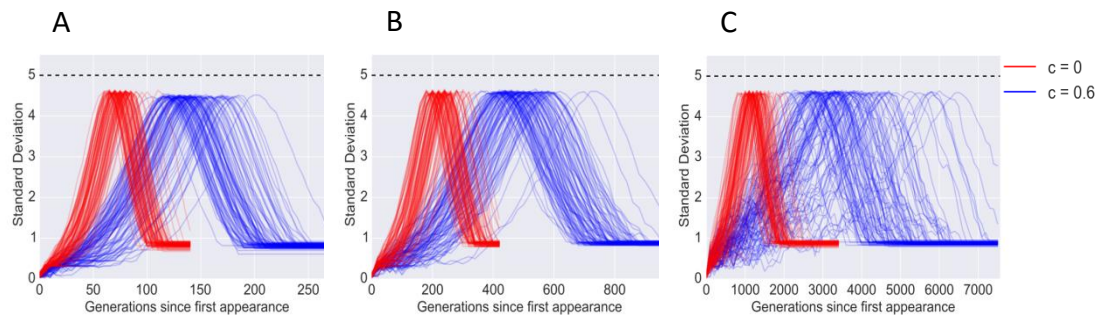
#### **Supplementary Information 4 – SD of double mutants between demes**

In this section we examined the variation in the number of double mutants between demes during the fixation process. A peak shift can occur either by a rapid takeover of double mutants occupying entire demes, or by a gradual increase of the number of double mutants in each deme. The spread of double mutants in the population can be expressed by the variance of the number of double mutants between the demes:

$$\sigma_{AB}^2 = \sum_{D=1}^n \frac{(AB_D - \overline{AB})^2}{n}$$

Where  $AB_D$  is the number of double mutants in deme D, and  $\overline{AB}$  is the average number of double mutants per deme  $\left(\frac{\sum_{D=1}^n AB_D}{n}\right)$ .

A population composed of demes with a similar number of double mutants would have low variance, whereas high variance is expected when the population is composed of demes fully inhabited by double mutants and demes completely devoid of them. We find that the cooperation level (c) does not influence the maximum variance attained during a peak shift (compare red to blue curves in the figure), indicating that the peak of the distribution of double mutants among demes during the fixation process is similar with or without cooperation. However, the dynamics of the double mutant spread in a cooperative population are more variable than in a non-cooperative one.



**Figure S3. The change in variance of double mutants across demes during a peak shift.** Standard deviation of the number of double mutants between demes ( $\sigma_{DM}$ ) is plotted for 100 simulation runs of a cooperative population ( $c = 0.6$ , blue) and a non-cooperative population ( $c = 0$ , red), under several sets of selection coefficients:  $s = 0.1, h = 10$  (A),  $s = 0.05, h = 5$  (B) and  $s = 0.02, h = 2$  (C). Each simulation begins from the first successful double mutant (a double mutant that fixates in the population) until 99% of the population are double mutants (yielding  $\sigma_{DM} < 1$ ). The dashed line represents the theoretical boundary of the  $\sigma_{DM}$  value, attained when half of the demes are inhabited solely with double mutants, and the other half is empty of double mutants. The qualitative dynamics are very similar for the different selection coefficients (compare A, B and C) though the scale is different. Other parameters:  $n = 1,000$ ,  $k = 10$ ,  $\mu = 10^{-5}$ ,  $r = 0.01$ ,  $b = 1.2$ .

## References

- Eshel, I. 1981. On the survival probability of a slightly advantageous mutant gene with a general distribution of progeny size—a branching process model. *Journal of mathematical biology* 12:355-362.
- Hadany, L. 2003. Adaptive peak shifts in a heterogenous environment. *Theoretical population biology* 63:41-51.
- Harris, T. E. 1948. Branching processes. *The Annals of Mathematical Statistics*:474-494.