JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12547

Overdominance interacts with linkage to determine the rate of adaptation to a new optimum

J. DRAGHI & M. C. WHITLOCK

Department of Zoology, University of British Columbia, Vancouver, BC, Canada

Keywords:

evolutionary constraint;
evolvability;
genetics of adaptation;
overdominance;
stabilizing selection.

Abstract

Overdominance, or a fitness advantage of a heterozygote over both homozygotes, can occur commonly with adaptation to a new optimum phenotype. We model how such overdominant polymorphisms can reduce the evolvability of diploid populations, uncovering a novel form of epistatic constraint on adaptation. The fitness load caused by overdominant polymorphisms can most readily be ameliorated by evolution at tightly linked loci; therefore, traits controlled by multiple loosely linked loci are predicted to be strongly constrained. The degree of constraint is also sensitive to the shape of the relationship between phenotype and fitness, and the constraint caused by overdominance can be strong enough to overcome the effects of clonal interference on the rate of adaptation for a trait. These results point to novel influences on evolvability that are specific to diploids and interact with genetic architecture, and they predict a source of stochastic variability in eukaryotic evolution experiments or cases of rapid evolution in nature.

Introduction

The phenomenon of overdominance, or a fitness advantage of heterozygotic individuals over either homozygote, is well known as the explanation for the high frequency of sickle-cell anaemia in humans (Allison, 1954). Beyond this classic example, there are only a few well-documented cases of overdominance for fitness, for example warfarin resistance in rats (Partridge, 1979; but see Kohn *et al.*, 2000), fecundity in domestic sheep (Gemmell & Slate, 2006), insecticide resistance in mosquitos (Labbé *et al.*, 2007) and male horn size in feral sheep (Johnston *et al.*, 2013). Although overdominance played an important part in earlier debates about genetic variation and the putative intrinsic superiority of heterozygotes, these controversies were largely resolved without a major role for overdominance (Crow, 2008). Genomic approaches have largely failed to detect widespread signatures of overdominance, confirming the view that loci with overdominant polymorphisms are rare (Hedrick, 2012).

Although overdominant mutations are sometimes observed in the laboratory with new mutations (e.g. Peters *et al.*, 2003), the apparent rarity of long-term overdominant polymorphisms has been invoked to argue against the validity of these observations (Fry, 2004). Recent theoretical work with Fisher's geometric model has challenged this idea and suggested that overdominance may arise frequently during adaptation in diploids (Manna *et al.*, 2011; Sellis *et al.*, 2011). In these models, overdominance arises when a large-effect mutation with additive phenotypic effects is strongly selected as a heterozygote, but overshoots a phenotypic optimum as a homozygote. Rather than invoking complementary functions of heterozygous alleles or opposing selection on a pleiotropic locus, this framework requires only stabilizing selection and the existence of mutations of large effect.

These theoretical observations can be reconciled with the empirical rarity of overdominant polymorphisms if the heterozygote advantage caused by phenotypic overshoot can be ultimately resolved by further adaptation. However, transient overdominance still may play a significant role in shaping the rate of adaption, or evolvability, of a population. A single overdominant polymorphism at one locus may slow further adaptation at other loci because of the epistatic nature of fitness under stabilizing selection, even for traits determined by additively interacting genes.

Correspondence: Jeremy Draghi, Department of Zoology, University of British Columbia, Vancouver, BC V6T 1Z4, Canada.
Tel.: 1 604 822 6879; fax: 1 604 822 2416; e-mail: jdraghi@gmail.com

The question of how an initial adaptive change might constrain and direct further evolution has received considerable recent attention in theoretical models (Weinreich *et al.*, 2005; Franke *et al.*, 2011), laboratory evolution experiments (Kvitek & Sherlock, 2011; Rokyta *et al.*, 2011; Salverda *et al.*, 2011; Kryazhimskiy *et al.*, 2012; Chou *et al.*, 2014), reconstructions of ancestral molecular phenotypes (Poelwijk *et al.*, 2007; Bridgham *et al.*, 2009; Lunzer *et al.*, 2010; Schenk *et al.*, 2013) and studies of rapid evolution in the wild (Bloom *et al.*, 2010). These studies have begun to uncover the causal factors that determine the rate and predictability of mutation-limited adaptive processes. Although much of this work has focused on haploid models and systems, pressing issues – such as the rapid evolution of eukaryotic disease vectors (Labbé *et al.*, 2007), as well as adaptation in threatened (Gonzalez *et al.*, 2013) or invasive species (Gilchrist & Lee, 2007) – highlight the need to understand the roots of diploid evolvability in new environments.

Previous analyses of the subsequent evolution of overdominant loci have considered modifiers of dominance (Otto & Bourguet, 1999) or duplication of alleles with complementary functions (Spofford, 1969; Otto & Yong, 2002; Proulx & Phillips, 2006). Here, we focus on the simple case, inspired by Sellis *et al.* (2011), of evolution at multiple loci with additive effects on the same trait. We find that overdominance during adaptation causes an epistatic constraint on the rate of subsequent adaptation and that the degree of this constraint depends strongly on the linkage among the sites determining the trait, as well as the shape of the fitness function near the optimum. We illustrate how this differential constraint arises from the epistasis caused by overdominance, which greatly restricts the potential for further adaptation at unlinked loci. As a consequence, recombination may slow or hasten the pace of adaptation, depending on whether mutations have large or small effects on the phenotype.

Materials and methods

We consider a model in which several loci contribute additive phenotypic effects to a single phenotypic trait under stabilizing selection. Under this model, overdominance for fitness – defined as the fitness superiority of a heterozygote to both homozygotes – implies that the homozygote trait values are on opposite sides of the optimum (Fig. 1). With the further assumption of a symmetrical fitness function, the range of allele effects leading to overdominance can easily be delineated. Scale the phenotype such that the optimal value is 0, then define z_{off} as the distance between the phenotype of the current genotype, labelled z_{aa}, and the optimum: $z_{\text{off}} = |0 - z_{aa}|$. If α_A is the effect of a single copy of the mutant allele A, then A will be beneficial if

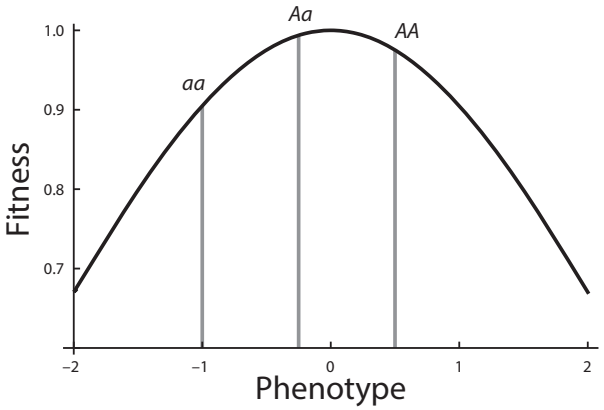

Fig. 1 Stabilizing selection on an additive trait can result in overdominance for fitness. If the genotype starts at *aa* (with mean phenotype 1 unit below the optimum for this trait), a mutation that increases the phenotype 0.75 units per copy has a higher fitness as a heterozygote with the ancestral allele because it closer to the optimum. The homozygote for the new allele overshoots the optimum and is in this case farther from the optimum than the heterozygote. This results in overdominance at this locus.

$|z_{aa} + \alpha_A| < z_{\text{off}}$. If A is favoured when rare, then A and a will be overdominant if $2z_{\text{off}} > |\alpha_A| > 2z_{\text{off}}/3$.

Fitnesses, w, are calculated from phenotypes, z, using two functions: the Gaussian function:

$$w_G(z) = e^{-\frac{z^2}{2\omega}} \tag{1}$$

with strength of selection determined by $1/\omega$ and a triangular function:

$$w_t(z) = \begin{cases} 1 - c|z| \text{ for } \frac{-1}{c} \le z \le \frac{1}{c} \\ 0 \text{ for } z < -\frac{1}{c} \text{ or } z > \frac{1}{c}. \end{cases} \tag{2}$$

Throughout, c is set to $1-\exp(-1/2\omega)$ so that both functions are equal at $z_0 = -1$. For convenience, define z_{xy} as the phenotype $z_0 + \alpha_x + \alpha_y$. This triangular function shows no curvature to the relationship between fitness and phenotype, except at the optimum itself, where the curvature is undefined. We use this function to explore how the (typically unknown) curvature of the fitness function affects the relationship between overdominance and evolutionary constraint.

Within this framework, we used three types of numerical methods to examine overdominance in the Wright–Fisher model. First, we performed individual-based simulations of either two or ten loci, each with a continuum of possible alleles and, for ten loci, a range of possible linkage patterns. For systems of two loci, each with two alleles, we performed forward iteration of a set of recursions (eqns 3a–d) to calculate a deterministic prediction of the equilibrium alleles frequencies obtained for a specified starting configuration. To calculate invasion probabilities, we combined these same equations with multinomial sampling to generate stochastic trajectories of allele frequencies.

Individual-based simulations

Simulations tracking N individual genotypes were used for the results in Figs 2, 3 and S1. This code simulated the Wright–Fisher model: generations were nonoverlapping and selection was based on differential fertility. Mutation numbers were Poisson-distributed with rate U per zygote. Populations were classified based on the properties of the first expected substitution: the first novel allele to exceed a frequency of 0.05 was analysed by calculating the relative fitnesses of its homozygote, the homozygote of the ancestral allele at that locus, and their heterozygote. These relative fitnesses were computed by sampling N genetic backgrounds (without linkage) from the population at this time t and measuring the marginal fitness of each allele at the putatively overdominant locus in the context of those backgrounds. A pair of alleles were classified as overdominant if their heterozygote was more fit than both homozygotes in the context of the population state at generation t.

Deterministic dynamics in two-locus, two-allele populations

Consider loci, A and B, with the ancestral state *aabb*. We assume without loss of generality that the effects of these ancestral alleles (α_a and α_b) are zero, such that the genotype *aabb* has initial phenotype z_0. We also assume random mating. For recombination probability r, the frequency of each haplotype in the next generation is given by the equations:

$$p'_{ab} = p_{ab}^2 w(z_{aabb}) + p_{Ab}p_{ab}w(z_{Aabb}) + p_{ab}p_{aB}w(z_{aaBb}) + (p_{ab}p_{AB}(1-r) + p_{aB}p_{Ab}r)w(z_{AaBb}) \quad (3a)$$

$$p'_{Ab} = p_{Ab}^2 w(z_{AAbb}) + p_{Ab}p_{ab}w(z_{Aabb}) + p_{AB}p_{Ab}w(z_{AABb}) + (p_{ab}p_{AB}r + p_{aB}p_{Ab}(1-r))w(z_{AaBb}) \quad (3b)$$

$$p'_{aB} = p_{aB}^2 w(z_{aaBB}) + p_{AB}p_{aB}w(z_{AaBB}) + p_{ab}p_{aB}w(z_{aaBb}) + (p_{ab}p_{AB}r + p_{aB}p_{Ab}(1-r))w(z_{AaBb}) \quad (3c)$$

$$p'_{AB} = p_{AB}^2 w(z_{AABB}) + p_{AB}p_{Ab}w(z_{AABb}) + p_{AB}p_{aB}w(z_{AaBB}) + (p_{ab}p_{AB}(1-r) + p_{aB}p_{Ab}r)w(z_{AaBb}). \quad (3d)$$

Our goal is to predict the equilibrium frequencies of the four haplotypes if alleles A and a constitute an overdominant polymorphism, allele b is resident, and allele B is a rare mutant. We therefore begin with p_{Ab} set to the equilibrium frequency for an overdominant polymorphism (Fisher, 1922):

$$\hat{p}_A = \frac{w(z_{Aa}) - w(z_{aa})}{2w(z_{Aa}) - w(z_{AA}) - w(z_{aa})}. \quad (4)$$

p_{ab} is then set to $1 - p_{Ab}$. To introduce B, we define a population size ($N = 10\,000$ is used throughout), set both p_{AB} and p_{aB} to frequencies of $1/2N$, and subtract $1/2N$ from both p_{ab} and p_{Ab}. Introducing the

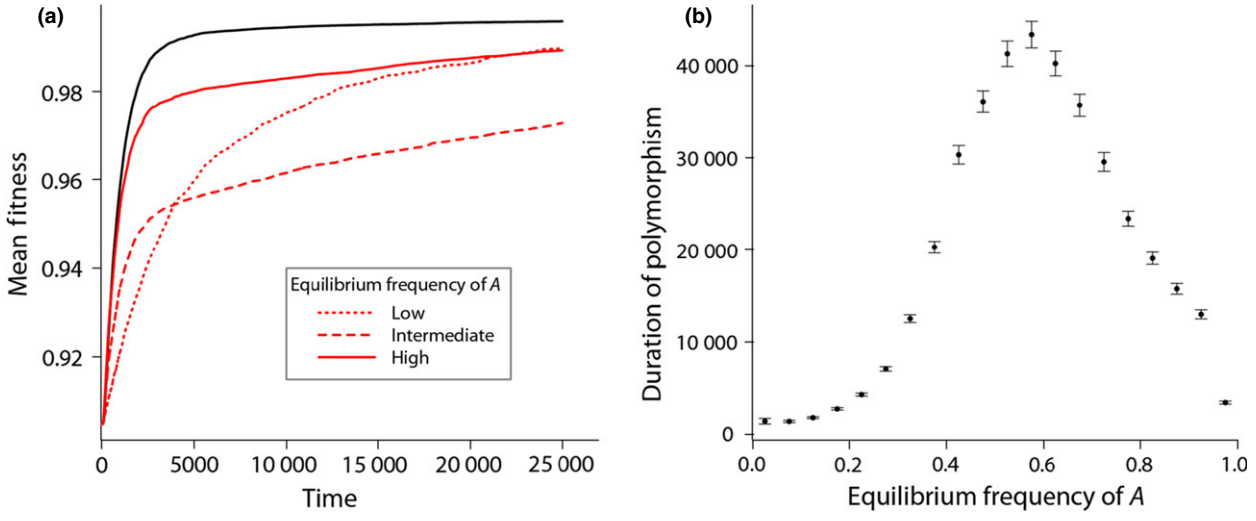

Fig. 2 (a) Mean fitness approaches its maximum value more quickly for populations in which the first invading mutation is predicted to fix (*black*) or reach a polymorphic equilibrium due to overdominance (*red*). Overdominant cases are divided based on the expected equilibrium frequency of the invading mutation: 'low' corresponds to frequencies between 0.05 (the detection threshold) and 0.33; 'intermediate' to 0.33–0.67; and 'high', to 0.67–1. (b) The predicted equilibrium frequencies of overdominant polymorphisms vs. their observed persistence times. Simulations run until either the ancestral allele or the invading mutation are lost from the population, for the subset of simulations in which the first invading mutation is predicted to cause overdominance. Bars show confidence intervals. Genotypes are comprised of ten unlinked loci with an initial phenotype of −1; selection is Gaussian with an optimum at zero and $\omega = 5$. Population size is 10 000 with a genomic mutation rate of 10^{-6} and normally distributed mutation effects with mean 0 and a standard deviation 0.75.

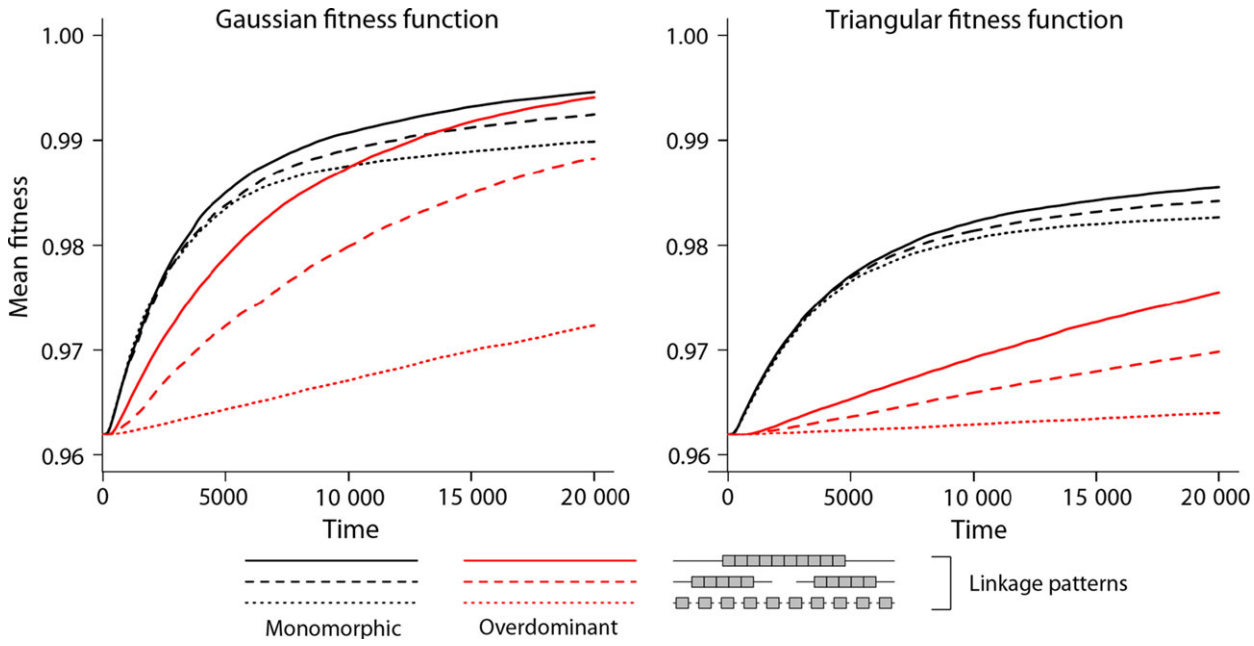

Fig. 3 Adaptation in populations which are either polymorphic for an overdominant allele pair (*red*) or fixed for alleles with equal population mean fitness (*black*). The trait is determined by one of three genetic architectures: ten fully linked loci (*solid lines*), two sets of five loci each with complete linkage within sets and free recombination between (*dashed*) or ten unlinked loci (*dotted*). In all scenarios, all loci but one, labelled *A*, are fixed for alleles of zero effect. In overdominant populations, *A* has two alleles, *A* and *a*, with effects 0.9 and 0, respectively, with the starting frequency of *A* binomially distributed around the equilibrium $p_A = 0.607$. In monomorphic populations, locus A is fixed for the allele *A* with an effect of approximately 0.189. Selection is either (*left side*) Gaussian and $\omega = 5$ or (*right side*) triangular with a slope of 0.0952. In each, the initial phenotype is set to −1. Population size is 10 000 with a genomic mutation rate of 10^{-6} and normally distributed mutation effects with mean 0 and standard deviation 0.75.

mutation on both backgrounds at once simplifies the calculation when the question is whether any mutation of a given allele could invade. Equations 3a–d are then iterated until the change in the frequency of both *A* and *B* was $< 10^{-10}$. An allele was classified as present at equilibrium if its frequency was $> 1/(2N)$ (or greater than its initial frequency, $2/(2N)$, for the *B* allele), and fixed if its frequency was greater than $(2N - 1)/(2N)$.

Stochastic invasion in two-locus, two-allele populations

The stochastic calculations of invasion probabilities (used for Fig. 4) are also determined by forward iteration, with the key difference that, each generation, the genotypes of the *N* breeding adults are drawn from a multinomial distribution. The expected frequencies of each parental genotype are calculated from eqns 3a–d with the additional assumption of random mating. This process simulates genetic drift as the stochastic and nonselective survival of *N* individuals from an unlimited pool of gametes.

Each replicate simulation begins by setting the frequency of *A* to its equilibrium, given by eqn 4, followed by evolution in the absence of *B* for one hundred generations; this 'burn in' period allows the frequency of *A* to vary among replicates at the time of the introduction of the *B* mutant. Then, a single *B* mutant is introduced in a randomly chosen gamete. The probability that this mutant will be linked to *A* is therefore equal to the current frequency of *A*. If *B* persists for at least 10 000 generations, then we record a successful invasion. Invasion probability is therefore the number of replicates in which the allele persisted for 10 000 generations divided by the total.

We can derive an approximate value for the probability that the *B* allele invades from a single-copy mutation from classic results. We calculate its selective coefficient when introduced and assume that the selective value of the new allele does not change much when it is rare, and therefore most at risk of stochastic loss. We calculate the probability of invasion as one minus the probability of loss. The selection coefficient for the *B* mutant, s_B, when p_B approaches zero, can be calculated as the difference in relative fitness between the *B* and *b* alleles. When the recombination rate $r = 0$, this calculation compares the mutant haplotype (*AB* or *aB*) to the corresponding resident haplotype:

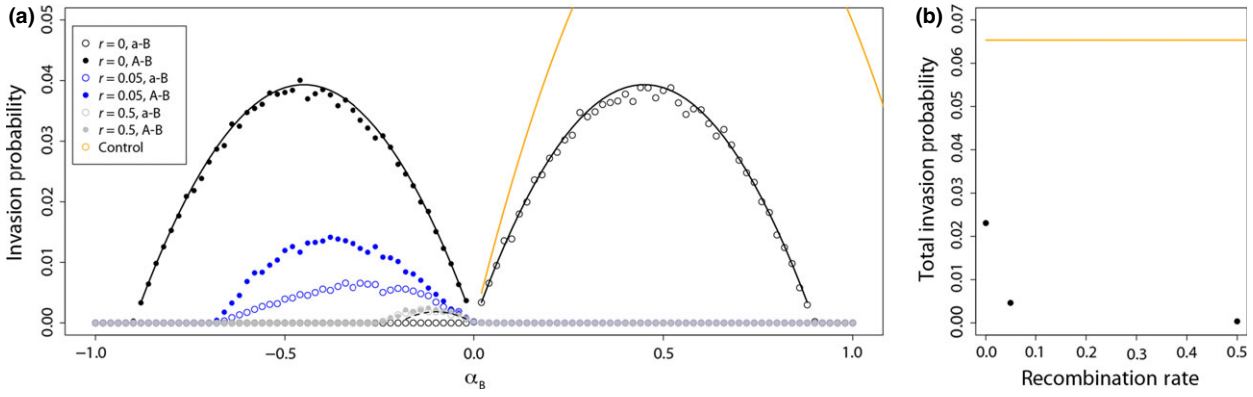

Fig. 4 The probability of invasion of an allele (a) at locus B into a population with overdominance at A selected according to a Gaussian fitness function for free (*grey*), intermediate (*blue*) and no recombination (*black*) between the loci. When r is small, the probability of invasion depends on whether the mutation arose in linkage with the a (*open symbols*) or the A (*filled symbols*) allele. Lines depict the expected invasion probability (eqns 5–7; see Methods) for free (*dashed*) and no (*solid*) recombination, as well as the scenario in which the mutant arose in a monomorphic population with the same initial mean fitness (*orange line*). Total invasion probability (b) is calculated as the integral of the invasion probability, over the range $\alpha_B = -1$ to $+1$, of B linked with a, times the frequency of a alleles at the overdominant equilibrium, plus the integral for B linked with A times the equilibrium frequency of A. Tight linkage to the previous overdominant locus greatly increases the total probability of invasion over the range of possible beneficial mutations. Each point is the mean of 50 000 stochastic simulations (see Methods) with a population size of 10 000. The initial phenotype is -1, the optimal trait value is at 0, and selection is Gaussian with $\omega = 5$, before the introduction of a new allele at the A locus with effect $\alpha_A = 0.9$, resulting in an equilibrium frequency of $p_A = \sim 0.607$.

$$s_B \approx \frac{\hat{p}_A w(z_{AABb}) + (1-\hat{p}_A) w(z_{AaBb})}{\hat{p}_A w(z_{AAbb}) + (1-\hat{p}_A) w(z_{Aabb})} - 1 \text{ when } AB \quad (5a)$$

and

$$s_B \approx \frac{\hat{p}_A w(z_{AaBb}) + (1-\hat{p}_A) w(z_{aaBb})}{\hat{p}_A w(z_{Aabb}) + (1-\hat{p}_A) w(z_{aabb})} - 1 \text{ when } aB. \quad (5b)$$

when $r = 0.5$, a different approximation can be derived by assuming that $p_B = 0$ and ignoring linkage disequilibrium:

$$s_B \approx \frac{\hat{p}_A^2 w(z_{AABb}) + 2\hat{p}_A(1-\hat{p}_A) w(z_{AaBb}) + (1-\hat{p}_A)^2 w(z_{aaBb})}{\hat{p}_A^2 w(z_{AAbb}) + 2\hat{p}_A(1-\hat{p}_A) w(z_{Aabb}) + (1-\hat{p}_A)^2 w(z_{aabb})} - 1. \quad (6)$$

Haldane (1927) used a branching process to approximate the probability of the stochastic loss when rare of a beneficial mutation ($s > 0$) with a fixed s:

$$p_{\text{loss}}(s) = 1 - 2s. \quad (7)$$

In the simple case Haldane considered, fixation is calculated as the converse of extinction. In the cases that we consider, the alternative to loss is not necessarily fixation, but it could be a quasi-stable equilibrium at an intermediate allele frequency. Therefore, we use eqn 7 in conjunction with the s_B values calculated by the above equations to calculate the probability that a given allele, beneficial when rare, is not lost stochastically when rare. We call this the invasion probability of that allele.

Results

Invasion of an overdominant mutation slows subsequent adaptation

To measure how overdominant polymorphisms shape subsequent adaptation, we performed replicate individual-based simulations of adapting diploid populations and classified each replicate based on the phenotypic effect of the first beneficial mutation to reach a frequency of 0.05. If this phenotypic effect was large enough to lead to overdominance (an effect greater than two-third of the distance to the optimum; see Methods), we predicted its equilibrium allele frequency in the absence of further mutation (eqn 4). Mean fitness in populations with initial overdominant mutations increases more slowly than in populations without predicted overdominance, indicating that overdominant mutations restrict the rate of adaptation (Fig. 2a). The degree of the reduction of evolutionary rate also appears to be most severe for polymorphisms of intermediate equilibrium frequency; these polymorphisms typically persist for tens of thousands of generations (Fig. 2b). Examination of the dynamics in individual simulations suggests that the gradual improvement in the ensemble mean fitness of populations with overdominant polymorphisms is primarily caused by beneficial mutations which eliminate the polymorphism, although adaptive changes that retain the original overdominant alleles are also seen (Fig. S1).

To control for the effects of a population's mean fitness on its subsequent adaptation, we measured the rate of adaptation in populations which were either polymorphic for an overdominant polymorphism or fixed for an adaptive mutation with an equivalent effect on mean fitness. These simulations suggest that an overdominant polymorphism reduces evolvability and reveals two additional patterns. First, genetic architecture determines the deficit of evolvability. For a fixed population size and genomic rate of mutation, the pattern of linkage among loci shapes the rate of adaptation; specifically, systems with linked loci adapt much more quickly (Fig. 3a). Second, the linearity of the fitness function interacts with the genetic architecture to determine evolvability: constraint in systems of unlinked loci adapting with a triangular fitness function is more severe than under Gaussian selection (Fig. 3b). We explore the reasons for these patterns in the next section.

Role of linkage in evolving out of overdominance

To better understand how linkage shapes evolution in populations with overdominant polymorphisms, we measured the invasion probability of mutants arising in polymorphic populations in stochastic simulations (Fig. 4). Simple approximations (see eqns 5–7) in the Methods) closely match the simulated invasion probabilities for fully linked and unlinked loci; intermediate levels of recombination show a more complex pattern. No *B* allele can be favoured under all possible genetic backgrounds. If *B* is favoured with one homozygote at the A locus, it will be selected against with the other homozygote. With free recombination, *B* will only invade when it causes the phenotype of the more fit homozygote at the A locus to be closer to the optimum. However, with no recombination, both positive and negative alleles can invade, conditioned on them arising on the genetic background which makes them favourable. When the mutation arises on the right background, it is consistently favoured by selection and the effective selection coefficient for the new *B* allele is higher than it could be with free recombination. Consequently, a mutant is overall more likely to invade when it occurs in tight linkage with the locus causing the overdominant polymorphism (Fig. 4b).

This effect of linkage can be easily understood with the example in Fig. 5. In our model of a single trait, the two homozygotes *aa* and *AA* must lie on opposite sides of the phenotypic optimum to cause overdominance. Therefore, any mutation *B* with additive effects cannot be beneficial with both *aa* and *AA* homozygotes; the mutation must move at least one of these genotypes further away from the optimum. Furthermore, in the worst-case scenario depicted in Fig. 5, any mutation *B* will also displace the heterozygote *Aa* from the optimal phenotype. This epistatic constraint, created by stabilizing selection on the trait and a symmetrical pair of overdominant alleles ($z_{aa} = -z_{AA}$), can be resolved by the combination of Gaussian selection and a tightly linked mutation at a locus B. If a negative effect *B* allele is tightly linked to A, then *B* improves the fitness of *AA* homozygotes and worsens the fitness of *Aa* heterozygotes. With the 'curvature' found near the optimum under Gaussian selection – namely, an accelerating loss of fitness with increasing deviation – the detrimental effect of *B* with *Aa* is smaller than the beneficial effect with *AA* (compare Fig. 5 upper and lower left). Linkage therefore maximizes the association between the mutant allele and the background in which it is most favoured, whereas the changing slope of the Gaussian function discounts the deleterious effect on the heterozygote background. Numerical exploration of the full range of *A* and *B* allelic effects confirms that the shape of the Gaussian fitness function, in comparison with the linear, triangular function, allows further adaptation for any value of *A* and promotes outcomes that fix one or both loci rather than maintain variation in both (see 'Deterministic dynamics in two-locus, two-allele populations' in Methods; Fig. S2).

Although these results suggest that populations with unlinked genomes might evolve more slowly, on average, than populations with strongly linked genes, strong linkage is typically thought to reduce the rate of adaptation due to competition between beneficial mutations in separate lineages – called the Hill–Robertson effect (Felsenstein, 1974) or clonal interference. However, strong epistasis may reverse this prediction (Pepper, 2003). To test whether the constraint caused by overdominance could outweigh the effects of clonal interference, we compared the mean rate of adaptation for sets of ten unlinked or fully linked loci, across several mutation rates in populations of 10 000 individuals. For large-effect mutations (standard deviation of mutation effect equal to 0.75 phenotype units), fully linked architectures adapt more quickly over a range of mutation rates (Fig. 6). When mutations are small (standard deviation mutational effect 0.05), this pattern is reversed; for mutations of intermediate size (standard deviation of mutational effect 0.15), neither architecture is universally more evolvable.

Discussion

Although overdominance was once seen as common and an important source of genetic variation (Lewontin 1974), more recent views have suggested that overdominant polymorphisms are too rare to make a substantial contribution to variation (e.g. Charlesworth & Charlesworth, 1987; Hedrick, 2012). In this study, we show that overdominance can be important for another reason not previously considered – it may delay adaptation to a new optimum for a phenotypic trait. For one form of overdominance – that created by overshoot of

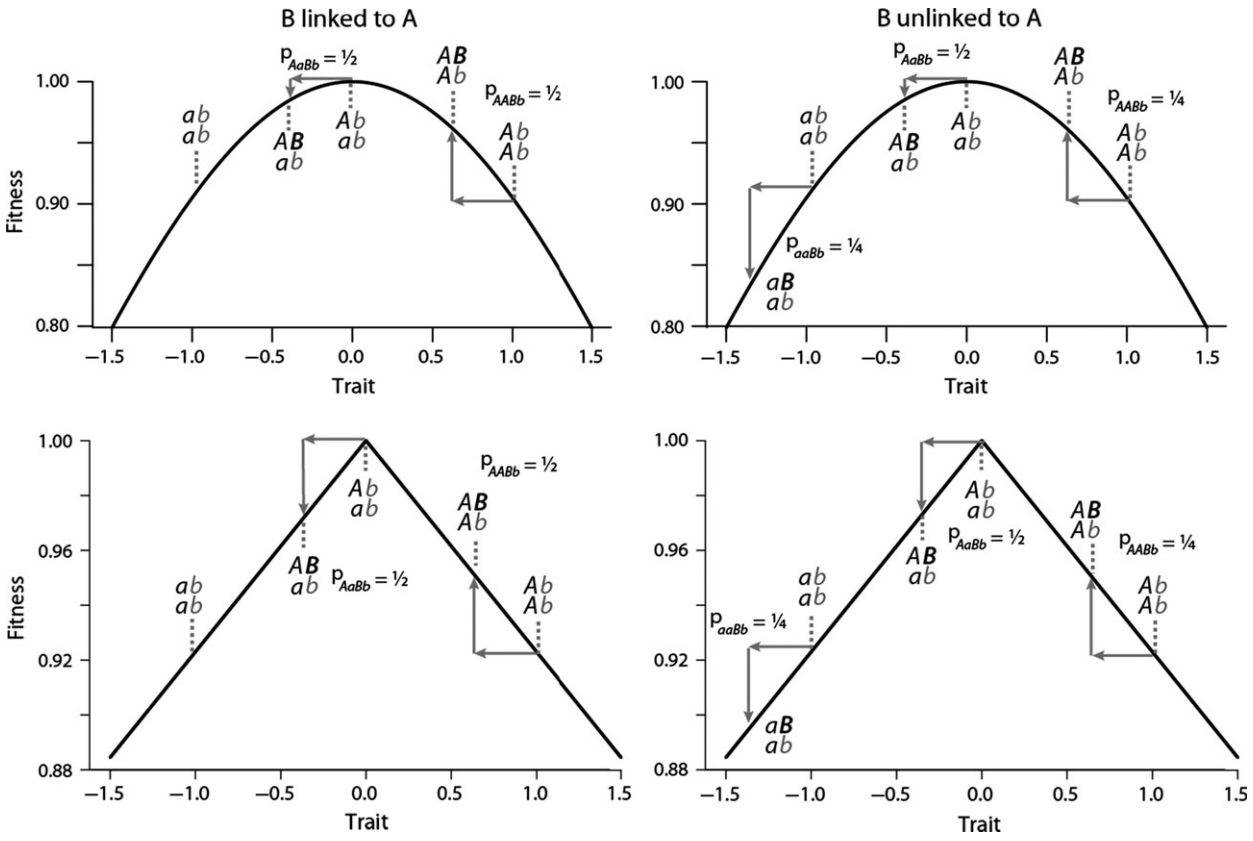

Fig. 5 Illustrations of the phenotype and fitness effects of a mutation at a locus *B* in the context of overdominance at locus A. For simplicity, the illustrated polymorphism is the symmetrical case in which $p_A = p_a = 0.5$. *Left*: When B is fully linked to A and the *B* allele appears on an *A* background, *B* occurs with the *AA* homozygote at a frequency p_A, and with the *Aa* heterozygote at a frequency $(1 - p_A)$. Nonlinearity in the fitness function (*upper left*) can lead to a positive selection coefficient of the *B* mutant, even when the *B* mutant displaces the heterozygote from the optimum; in contrast, the mutant is neutral at best when fitness falls off linearly from the optimum (*lower left*). *Right*: When B is unlinked with A, its selection coefficient is decreased by two differences relative to the fully linked scenario. First, the mutant occurs less often with the background on which it is favoured; in this case, the homozygote *AA*. Second, the *B* mutant now sometimes occurs with the other homozygote, *aa*, with which it is strongly disfavoured. Selection against this mutant is exacerbated under the triangular fitness function (*bottom right*); the deleterious effect of *B* with *Aa* is now equal to its benefit with *AA*, and *B* occurs in combination with *Aa* twice as often as it does with *AA*.

the optimum with stabilizing selection – the rate of adaptation after the emergence of overdominance can be substantially slower than in the absence of overdominance. This restriction on evolvability is shaped by the genetic architecture underlying the trait as well as the slope of the phenotype–fitness relationship near the optimum; the influence of both factors can be readily understood as the consequences of epistasis between new mutations and the overdominant polymorphism.

Previous work in a one-locus, multiple-trait model described how overdominance could cause differences in the evolvability of diploids in comparison to haploids (Sellis *et al.*, 2011); more recent work with this model highlights how overdominance can lead to stochastic differences between parallel populations (Venkataram *et al.*, 2013). This approach is particularly valuable for understanding how evolvability might determine the benefits of diploidy (Zeyl *et al.*, 2003; Gerstein *et al.*, 2011), but may be hampered by the unpredictable effects of ploidy on selection coefficients (Gerstein, 2013). In contrast, we focus on evolvability differences within diploids to clarify the role of genetic architecture – specifically, the pattern of linkage among the loci that contribute to a trait. Our results show that strong linkage significantly speeds adaptation out of overdominance, although the time required also depends on the allele frequencies of the overdominant polymorphism and the change in slope of the fitness function near the optimum. Although recombination can speed adaptation by bringing together beneficial mutations, this benefit can be outweighed by the advantage of linked loci in evolving solutions to overdominance.

Several models have looked at linkage between an overdominant (Strobeck *et al.*, 1976) or frequency-

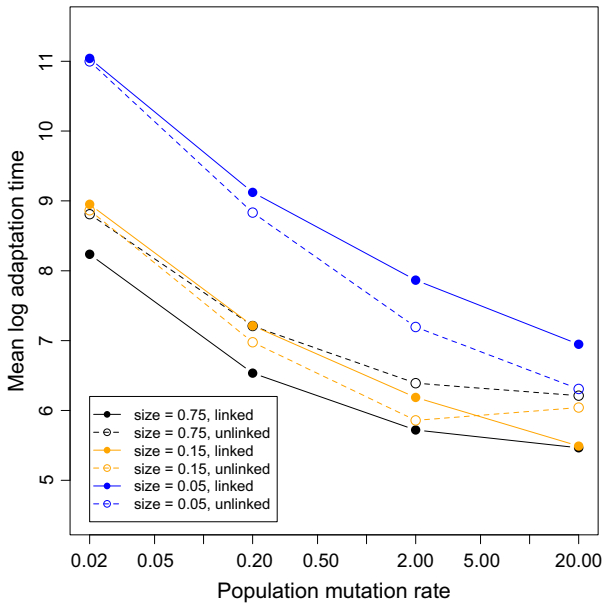

Fig. 6 Mean of the natural log of the time to adapt to a mean population fitness exceeding 0.995. Individual-based simulations with ten loci, either fully linked or freely recombining, evolve from a initial phenotype $z_0 = -1$ towards an optimum of zero under Gaussian selection with $\omega = 5$. Gaussian mutations, with means of zero and standard deviations as indicated by the 'size' parameter in the legend, arise in populations of size 10 000. Confidence intervals, not displayed, are comparable or smaller than plotting symbols.

dependent locus (Peck, 1993; Hodgson & Otto, 2012) and another locus under directional selection, finding that such a situation can select for modifiers of increased recombination. Although these results appear contrary to ours, the key difference is that the fitness effects of the two loci in these models combine multiplicatively, but in the case of overdominance caused by stabilizing selection, fitness is determined by epistatic interactions between loci. We also find that sets of linked genes can evolve more quickly than an equivalent set of unlinked genes under high mutation rates, which contradicts the view that recombination generally speeds adaptation. This difference may be explained by the observation that, in our model, the beneficial mutations with the highest selective coefficients when rare are also those that lead to overdominance. Adaptation in populations with a high input of mutations may therefore be especially vulnerable to constraint through overdominance.

Our finding that evolvability in overdominant populations is sensitive to the shape of the fitness function connects this model with efforts to understand Wright's symmetric viability model (two loci with additive effects under stabilizing selection; reviewed in Bürger 2000). Several studies have demonstrated that the stability of the polymorphic equilibria of this model is sensitive to the shape of the fitness function (Nagylaki, 1989; Willensdorfer & Bürger, 2003). One interesting result is that linkage in Wright's model enhances the possibility of stable polymorphism at two loci (Bürger & Gimelfarb, 1999), whereas in our treatment, tight linkage hastens adaptive loss of an overdominant polymorphism. This disparity suggests that different conclusions about the maintenance of variation under stabilizing selection may be reached by modelling the progress of adaptation when new mutants invade resident polymorphisms (the more realistic case that we consider here), rather than focusing on pairs of alleles with random effects.

Not all loci which are overdominant for fitness will fit our model. Our results pertain to the kind of overdominance for fitness that results from genotypes on opposite sides of the optimum with stabilizing selection on a single phenotypic trait. For example, selection on ram's horns shows stabilizing selection, and as a result, a large-effect mutation for horn length demonstrates overdominance for fitness (Gemmell & Slate, 2006). Other cases of overdominance do not so clearly fit this model of stabilizing selection on a single trait. The overdominance in acetylcholinesterase (*ace-1*) in *Culex pipiens* involves an allele that is susceptible to pesticide and another which is resistant but with lower function. Here, fitness is determined by a balance of the number of susceptible and resistant alleles, rather than a single phenotypic axis like enzyme activity (Labbé *et al.* 2007). Similarly, some apparent cases of overdominance are actually directionally dominant at multiple individual loci (a phenomenon known as apparent overdominance or associative overdominance). Different mechanisms of overdominance (or apparent overdominance) may behave differently in evolutionary terms, and the relative frequency of different types of overdominance in nature is currently not known.

We have shown that a particular form of overdominance – caused during adaptation to a new optimum with stabilizing selection – can cause substantial changes to the evolutionary trajectory of a population. The transience of the overdominance may cause it to leave little trace in the genome (although we expect fixed mutations to aggregate slightly in the genome because of the linkage effects discussed above). It would in many cases be difficult to reconstruct the evolutionary trajectories of a population, because many alleles that were important for the process may have been lost as the population evolved out of overdominant states. Evolution to an optimum may take substantially longer with overdominance that without, but subsequent reconstructions of the process would not see the reason why. In diploids, overdominance may be a moderately important cause of lowered evolutionary potential of populations responding to new environments.

Acknowledgments

We are grateful to the members of the Whitlock and Otto laboratory group for comments on an earlier version of this manuscript and especially to Sally Otto for her extensive suggestions. We also thank two anonymous reviewers for helpful comments. This work was supported by a Discovery grant from the Natural Science and Engineering Research Council (Canada).

References

Allison, A.C. 1954. Protection afforded by sickle-cell trait against subtertian malarial infection. *Br. Med. J.* **1**: 290–294.

Bloom, J.D., Gong, L.I. & Baltimore, D. 2010. Permissive secondary mutations enable the evolution of influenza oseltamivir resistance. *Science* **328**: 1272–1275.

Bridgham, J.T., Ortlund, E.A. & Thornton, J.W. 2009. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* **461**: 515–519.

Bürger, R. & Gimelfarb, A. 1999. Genetic variation maintained in multilocus models of additive quantitative traits under stabilizing selection. *Genetics* **152**: 807–820.

Bürger, R. 2000. *The Mathematical Theory of Selection, Recombination, and Mutation*. John Wiley & Sons, West Sussex, England.

Charlesworth, D. & Charlesworth, B. 1987. Inbreeding depression and its evolutionary consequences. *Annu. Rev. Ecol. Syst.* **18**: 237–268.

Chou, H.H., Delaney, N.F., Draghi, J.A. & Marx, C.J. 2014. Mapping the fitness landscape of gene expression uncovers the cause of antagonism and sign epistasis between adaptive mutations. *PLoS Genet.* **10**: e1004149.

Crow, J.F. 2008. Mid-century controversies in population genetics. *Annu. Rev. Genet.* **42**: 1–16.

Felsenstein, J. 1974. The evolutionary advantage of recombination. *Genetics* **78**: 737–756.

Fisher, R.A. 1922. On the dominance ratio. *Proc. R. Soc. Edinb.* **42**: 321–341.

Franke, J., Klözer, A., de Visser, J.A.G.M. & Krug, J. 2011. Evolutionary accessibility of mutational pathways. *PLoS Comput. Biol.* **7**: e1002134.

Fry, J.D. 2004. How common are overdominant mutations? *Genetics* **167**: 1031–1032.

Gemmell, N.J. & Slate, J. 2006. Heterozygote advantage for fecundity. *PLoS ONE* **1**: e125.

Gerstein, A.C. 2013. Mutational effects depend on ploidy level: all else is not equal. *Biol. Lett.* **9**: 20120614.

Gerstein, A.C., Cleathero, L.A., Mandegar, M.A. & Otto, S.P. 2011. Haploids adapt faster than diploids across a range of environments. *J. Evol. Biol.* **24**: 531–540.

Gilchrist, G.W. & Lee, C.E. 2007. All stressed out and nowhere to go: does evolvability limit adaptation in invasive species? *Genetica* **129**: 127–132.

Gonzalez, A., Ronce, O., Ferriere, R. & Hochberg, M.E. 2013. Evolutionary rescue: an emerging focus at the intersection between ecology and evolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **368**: 20120404.

Haldane, J.B.S. 1927. A mathematical theory of natural and artificial selection, part V: selection and mutation. *Mathematical Proceedings of the Cambridge Philosophical Society*, **23**: 838–844.

Hedrick, P.W. 2012. What is the evidence for heterozygote advantage selection? *Trends Ecol. Evol.* **27**: 698–704.

Hodgson, E.E. & Otto, S.P. 2012. The red queen coupled with directional selection favours the evolution of sex. *J. Evol. Biol.* **25**: 797–802.

Johnston, S.E., Gratten, J., Berenos, C., Pilkington, J.G., Clutton-Brock, T.H., Pemberton, J.M. & Slate, J. 2013. Life history trade-offs at a single locus maintain sexually selected genetic variation. *Nature* **502**: 93–95.

Kohn, M.H., Pelz, H.J. & Wayne, R.K. 2000. Natural selection mapping of the warfarin-resistance gene. *Proc. Natl. Acad. Sci. USA* **97**: 7911–7915.

Kryazhimskiy, S., Rice, D.P. & Desai, M.M. 2012. Population subdivision and adaptation in asexual populations of *Saccharomyces cerevisiae*. *Evolution* **66**: 1931–1941.

Kvitek, D.J. & Sherlock, G. 2011. Reciprocal sign epistasis between frequently experimentally evolved adaptive mutations causes a rugged fitness landscape. *PLoS Genet.* **7**: e1002056.

Labbé, P., Berticat, C., Berthomieu, A., Unal, S., Bernard, C. Weill, M. & Lenormand, T. 2007. Forty years of erratic insecticide resistance evolution in the mosquito *Culex pipiens*. *PLoS Genet.* **3**: e205.

Lewontin, R.C. 1974. *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

Lunzer, M., Golding, G.B. & Dean, A.M. 2010. Pervasive cryptic epistasis in molecular evolution. *PLoS Genet.* **6**: e1001162.

Manna, F., Martin, G. & Lenormand, T. 2011. Fitness landscapes: an alternative theory for the dominance of mutation. *Genetics* **189**: 923–937.

Nagylaki, T. 1989. The maintenance of genetic variability in two-locus models of stabilizing selection. *Genetics* **122**: 235–248.

Otto, S.P. & Bourguet, D. 1999. Balanced polymorphisms and the evolution of dominance. *Am. Nat.* **153**: 561–574.

Otto, S.P. & Yong, P. 2002. The evolution of gene duplicates. *Adv. Genet.* **46**: 451–483.

Partridge, G.G. 1979. Relative fitness of genotypes in a population of *Rattus norvegicus* polymorphic for warfarin resistance. *Heredity* **43**: 239–246.

Peck, J.R. 1993. Frequency-dependent selection, beneficial mutations, and the evolution of sex. *Proc. R. Soc. Lond. B Biol. Sci.* **254**: 87–92.

Pepper, J.W. 2003. The evolution of evolvability in genetic linkage patterns. *Biosystems* **69**: 115–126.

Peters, A.D., Halligan, D.L., Whitlock, M.C. & Keightley, P.D. 2003. Dominance and overdominance of mildly deleterious induced mutations for fitness traits in *Caenorhabditis elegans*. *Genetics* **165**: 589–599.

Poelwijk, F.J., Kiviet, D.J., Weinreich, D.M. & Tans, S.J. 2007. Empirical fitness landscapes reveal accessible evolutionary paths. *Nature* **445**: 383–386.

Proulx, S.R. & Phillips, P.C. 2006. Allelic divergence precedes and promotes gene duplication. *Evolution* **60**: 881–892.

Rokyta, D.R., Joyce, P., Caudle, S.B., Miller, C., Beisel, C.J. & Wichman, H.A. 2011. Epistasis between beneficial mutations and the phenotype-to-fitness map for a ssDNA virus. *PLoS Genet.* **7**: e1002075.

Salverda, M.L., Dellus, E., Gorter, F.A., Debets, A.J., Van Der Oost, J., Hoekstra, R.F. *et al.* 2011. Initial mutations direct alternative pathways of protein evolution. *PLoS Genet.* **7**: e1001321.

Schenk, M.F., Szendro, I.G., Salverda, M.L., Krug, J. & de Visser, J.A.G. 2013. Patterns of epistasis between beneficial mutations in an antibiotic resistance gene. *Mol. Biol. Evol.* **30**: 1779–1787.

Sellis, D., Callahan, B.J., Petrov, D.A. & Messer, P.W. 2011. Heterozygote advantage as a natural consequence of adaptation in diploids. *Proc. Natl. Acad. Sci. USA* **108**: 20666–20671.

Spofford, J.B. 1969. Heterosis and the evolution of duplications. *Am. Nat.* **103**: 407–432.

Strobeck, C., Smith, J.M. & Charlesworth, B. 1976. The effects of hitchhiking on a gene for recombination. *Genetics* **82**: 547–558.

Venkataram, S., Sellis, D. & Petrov, D.A. 2013. Ploidy and the predictability of evolution in Fisher's geometric model. *bioRxiv* doi: 10.1101/001016.

Weinreich, D.M., Watson, R.A. & Chao, L. 2005. Perspective: sign epistasis and genetic constraint on evolutionary trajectories. *Evolution* **59**: 1165–1174.

Willensdorfer, M. & Bürger, R. 2003. The two-locus model of Gaussian stabilizing selection. *Theor. Popul. Biol.* **64**: 101–117.

Zeyl, C., Vanderford, T. & Carter, M. 2003. An evolutionary advantage of haploidy in large yeast populations. *Science* **299**: 555–558.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Representative examples of mean fitness over time for populations invaded by a low-, intermediate-, or high-frequency overdominant mutation (top, middle, and bottom plots, respectively).

Figure S2 Invasion of an overdominant locus *A* by alleles at a locus *B*.

Data deposited at Dryad doi:10.5061/dryad.6f2j8

Received 3 June 2014; revised 14 October 2014; accepted 31 October 2014