Revisiting Dominance in Population Genetics

Chenlu Di[1], Kirk E. Lohmueller [1,2,3,*]

[1]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA, USA
[2]Interdepartmental Program in Bioinformatics, University of California, Los Angeles, CA, USA
[3]Department of Human Genetics, David Geffen School of Medicine, Los Angeles, CA, USA

*Corresponding author: Email: klohmueller@g.ucla.edu.

Accepted: June 24, 2024

Abstract

Dominance refers to the effect of a heterozygous genotype relative to that of the two homozygous genotypes. The degree of dominance of mutations for fitness can have a profound impact on how deleterious and beneficial mutations change in frequency over time as well as on the patterns of linked neutral genetic variation surrounding such selected alleles. Since dominance is such a fundamental concept, it has received immense attention throughout the history of population genetics. Early work from Fisher, Wright, and Haldane focused on understanding the conceptual basis for why dominance exists. More recent work has attempted to test these theories and conceptual models by estimating dominance effects of mutations. However, estimating dominance coefficients has been notoriously challenging and has only been done in a few species in a limited number of studies. In this review, we first describe some of the early theoretical and conceptual models for understanding the mechanisms for the existence of dominance. Second, we discuss several approaches used to estimate dominance coefficients and summarize estimates of dominance coefficients. We note trends that have been observed across species, types of mutations, and functional categories of genes. By comparing estimates of dominance coefficients for different types of genes, we test several hypotheses for the existence of dominance. Lastly, we discuss how dominance influences the dynamics of beneficial and deleterious mutations in populations and how the degree of dominance of deleterious mutations influences the impact of inbreeding on fitness.

Key words: population genetics, dominance, deleterious mutations, inference, natural selection.

Significance

Dominance refers to the phenotype of the heterozygous genotype relative to that of the two homozygous genotypes. It is a foundational quantity in population genetics because dominance affects how natural selection changes the frequencies of mutations. Despite intense study over 100 years, the dominance effects of mutations in different organisms remain mostly unknown. Further, the reasons for why some mutations are recessive are not fully understood. In this review, we describe conceptual models for the existence of dominance and discuss some methods that have been used to estimate dominance coefficients.

Introduction

Mutations may affect the fitness of individuals who carry them. In diploid organisms, mutations can be carried in heterozygous or homozygous genotypes. Dominance describes the effect of a mutant heterozygote relative to the two homozygotes. Some mutations only have a phenotypic effect when homozygous; these are known as recessive mutations. In contrast, dominant mutations have the

© The Author(s) 2024. Published by Oxford University Press on behalf of Society for Molecular Biology and Evolution.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

same effect when homozygous and heterozygous. Historically, there has been tremendous interest in explaining the evolutionary and physiological basis for the existence of dominance. Dominance has an even deeper impact on many evolutionary processes and strongly influences the dynamics of both deleterious and beneficial mutations in populations (Crow 1948), thereby influencing inbreeding depression (Lande and Schemske 1985; Kyriazis et al. 2021; Robinson et al. 2022, 2023), introgression (Harris and Nielsen 2016; Whitlock et al. 2000; Kim et al. 2018), and detection of selection (Teshima and Przeworski 2006), which we will review in later sections. In addition to these processes, dominance also has broad evolutionary implications, for example, on maintaining genetic diversity (Charlesworth and Hughes 2000), the evolution of sex chromosomes and dosage compensation (Charlesworth et al 2018), the evolution of species with different mating systems (Ronfort and Glemin 2013), and much more. Due to the lack of knowledge of dominance parameters, for convenience, many population genetic models assume that mutations act in an additive manner (Eyre-Walker et al. 2006; Keightley and Eyre-Walker 2007; Boyko et al. 2008; Huber et al. 2017; Kim et al. 2017; Tataru et al. 2017; Booker and Keightley 2018; Castellano et al. 2019; Tataru and Bataillon 2019; Huang et al. 2021). However, this assumption almost certainly does not hold for all mutations (Simmons and Crow 1977; Balick et al. 2015; Huber et al. 2018).

Here, we review population genetic aspects of dominance. We begin by defining dominance from a population genetic perspective. We then revisit the historical and more recent explanations of dominance and recessiveness. We review methods for estimating dominance coefficients and gather estimates of dominance coefficients from the literature from different methods and species. Next, we discuss how these estimates support or conflict with models for dominance. Finally, we review how dominance influences the dynamics of selected mutations and the evolution of populations, especially focusing on inbreeding depression, introgression, and adaptation.

Defining Dominance

The concept of dominance is as old as genetics itself. In Mendel's pea plant experiment, homozygous purple flowers were crossed with homozygous white flowers, and all the heterozygous offspring had purple flowers. In this case, the allele determining the purple flower phenotype is dominant (Mendel 1865) while the allele determining the white flower phenotype is recessive. Since these early days of genetics, dominance has been described in the context of Mendelian genetic traits, quantitative genetics, and population genetics. Here, we focus primarily on population genetic aspects of dominance.

In population genetics, the dominance coefficient (h) quantifies the fitness of heterozygotes relative to that of

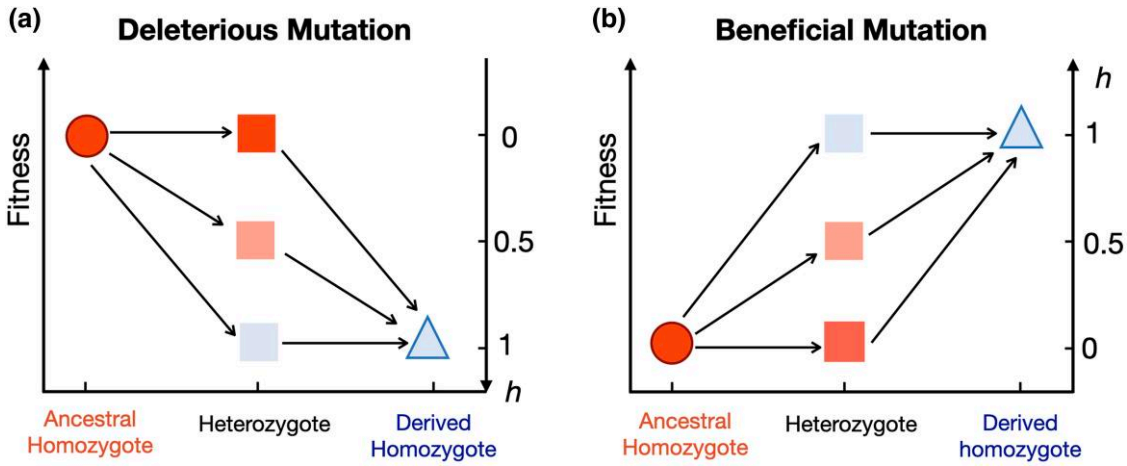

Fig. 1. Dominance refers to the fitness of the heterozygous genotype compared with that of the homozygous genotypes. The left y-axis indicates the fitness of different genotypes (x-axis). a) For deleterious mutations, the ancestral homozygote (circle) has the highest fitness and the derived homozygote has the lowest fitness (triangle). The fitness of the ancestral homozygous genotype is 1, the derived (mutant) homozygote is $1-s$, and the heterozygote is $1-hs$. If the fitness of the heterozygote is the same as that of the ancestral homozygote (top square), $h = 0$ and the deleterious mutation is recessive. If the fitness of the heterozygote is the average of the ancestral and the derived homozygotes (middle square), $h = ½$, and the deleterious mutation is additive. If the fitness of the heterozygote is the same as the derived homozygote, $h = 1$, and the deleterious mutation is dominant (bottom square). b) For beneficial mutations, the ancestral homozygote has the lowest fitness (bottom left circle) and the derived homozygote has the highest fitness (upper right triangle). If the fitness of the heterozygote is the same as the ancestral homozygote (bottom circle), $h = 0$, the beneficial mutation is recessive (bottom square). If the fitness of the heterozygote is the average of the derived and ancestral homozygotes, the mutation is additive (middle square). If the fitness of the heterozygote is as high as that of the derived homozygote, the mutation is dominant and $h = 1$ (upper square).

the homozygous genotypes (Fig. 1). Throughout this review, we will use the term "ancestral" to denote the original allele without the mutation. Such an allele would often be termed "wild-type" in molecular genetic studies. We use the term "derived" to denote the new allele that arose via mutation. Using ancestral and derived in place of wild and mutant types shows the direction of evolution and promotes an understanding of the fitness effects of derived mutations.

First focusing on deleterious mutations, the selection coefficient (s) represents the decrease in fitness in a derived homozygote (triangle in Fig. 1a, bottom right, fitness of $1 - s$) relative to the ancestral homozygote (circle in Fig. 1a, top left, fitness of 1). The dominance coefficient (h) of this derived mutation is the proportion of decreased fitness in heterozygotes relative to derived-type homozygotes. Thus, the fitness of heterozygotes can be written as $1 - hs$ (squares in Fig. 1a). For example, if the mutation is completely dominant, the decreased fitness in heterozygotes is the same as that in the derived homozygotes (bottom blue square in Fig. 1a) and thus $h = 1$. On the other hand, if $h = 0$, there is no decrease in fitness in the heterozygotes, and thus, the mutation is completely recessive (top orange square in Fig. 1a). If the mutation is additive, the decrease in fitness in heterozygotes relative to derived homozygotes is 0.5, that is, $h = 0.5$ (middle light orange square in Fig. 1a). While Fig. 1 focuses on completely recessive ($h = 0$) or completely dominant ($h = 1$) cases, mutations also could be partially recessive ($0 < h < 0.5$) or partially dominant ($0.5 < h < 1$). The same concepts apply to beneficial mutations (Fig. 1b), except that the derived homozygote has the highest fitness equal to $1 + s$. Heterozygotes here have fitness $1 + hs$. So, if one copy of the beneficial derived allele is just as beneficial as two copies, the derived allele is dominant and $h = 1$ (blue square in Fig. 1b). If being homozygous for the derived allele is required to enjoy an increase in fitness, then the derived beneficial mutation is recessive and $h = 0$ (bottom orange square in Fig. 1b).

In most cases, the fitness of a heterozygote ranges between that of the homozygous genotypes, and thus, h ranges between zero and one. However, if the fitness of heterozygotes is higher than both of the homozygous genotypes, then this is called overdominance. Here, h is negative. On the contrary, the fitness of heterozygotes can be lower than both of their homozygous parents. This scenario is called underdominance and h is above 1.

Historical Explanations for Dominance

Fisher's "Dominance Modifier" Model

Several theoretical and conceptual models have been put forth over the last 100 years to explain why dominance exists. One of the early explanations came from R.A. Fisher in the 1920s. Fisher hypothesized that deleterious mutations become recessive or nearly recessive through a mutation at a second locus that modifies the dominance coefficient of the original mutation. These so-called "dominance modifier" mutations would be under positive selection as they dampen the effects of deleterious mutations. Essentially, this model is a case of epistasis (Fig. 2a; Fisher 1928). In the absence of the modifier allele, deleterious mutations would be additive when they first occurred (Fig. 2a). Because mutant homozygotes are rare for new mutations at very low frequency, selection would primarily occur on heterozygotes. Fitness would be maximized by reducing the heterozygous effect of the deleterious mutations, thus providing the opportunity for selection to act on the modifier mutation (Fig. 2a).

However, some observations and arguments do not support Fisher's theory. Sewall Wright pointed out that the strength of selection on the "dominance modifier" would be too small to be the main factor driving the fixation of the modifier mutation (Fisher 1928; Wright 1929, 1934). Further, according to Fisher's theory, the selective strength of the dominance modifier will not be influenced by the fitness effect of the original deleterious mutation. To see this, assume the selection coefficient of the derived deleterious allele at a locus is s and mutations to the derived allele occur at rate μ. The increase in fitness after reducing the dominance effect can be written as $s\Delta h$. The frequency of heterozygotes for the deleterious mutation at mutation–selection equilibrium is $2\mu/hs$. Thus, the net intensity of selection on the "dominance modifier" with the effect of Δh is $2\mu\Delta h/h$, the product of $s\Delta h$ and $2\mu/hs$. As the selection coefficient cancels out and will not influence the selection of the "dominance modifier," this model cannot explain observations that strongly deleterious or lethal mutations are more likely to be recessive (see below for more detail on this topic; Ohnishi 1977; Simmons and Crow 1977; Charlesworth 1979; Phadnis and Fry 2005; Agrawal and Whitlock 2011; Huber et al. 2018). Moreover, Fisher's theory predicts that selection for dominance modifiers would be drastically reduced in self-fertilizing populations because there are few heterozygotes to modify, as nearly all deleterious mutations would be homozygous (Haldane 1939). However, this prediction is inconsistent with empirical examples of dominance in mostly self-fertilizing populations (Haldane 1939). In addition, Fisher's theory is unable to explain the observation that mutations in haploid organisms also tend to be recessive (Orr 1991; Manna et al. 2011). However, the evolution of modified dominance in balanced polymorphisms can explain the evolution of dominance of genes controlling Batesian mimetic patterns (Charlesworth and Charlesworth 1975). Nevertheless, additional explanations for the recessiveness of deleterious mutations are required.

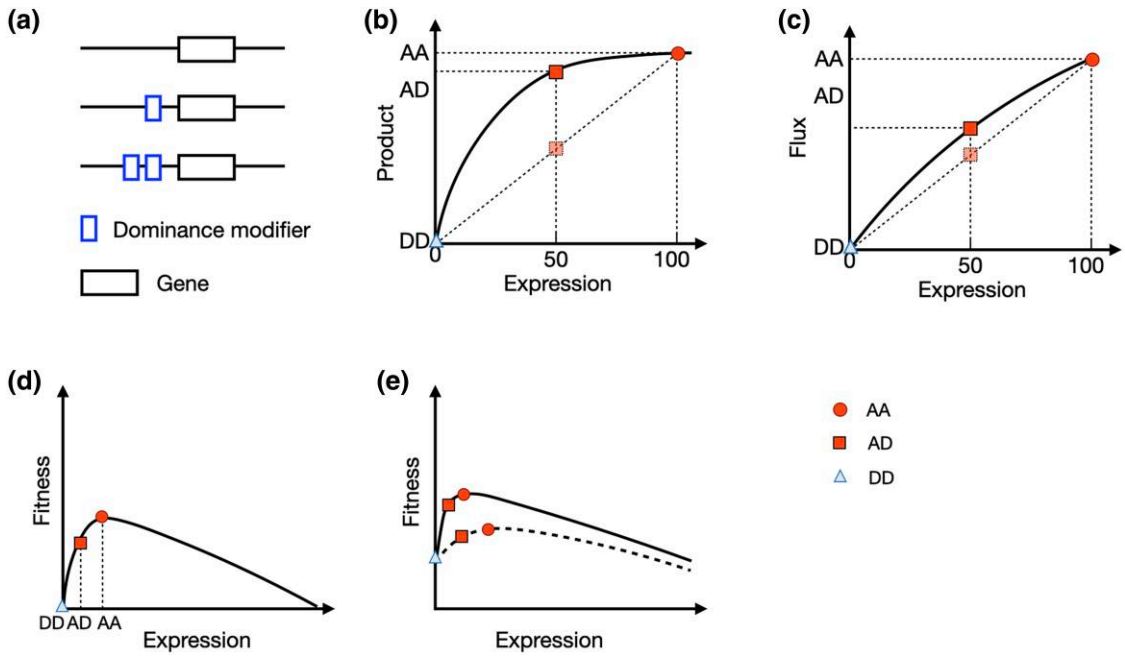

Fig. 2. Models explaining dominance. a) Fisher's dominance modifier model. Each line represents a small genomic region including a gene (black box). If there is no "dominance modifier" (the top line), a mutation in the gene is additive. Accumulation of the "dominance modifiers" increases the recessiveness of mutations. b) Wright's physiological model. Enzyme activity is shown on the *x*-axis. The *y*-axis indicates the enzyme-catalyzed reaction of different genotypes. The enzyme-catalyzed reaction can also be considered as an approximation to fitness in later models. The activity of the heterozygote (solid square) is very close to the fitness of the ancestral homozygote (circle) and much higher than the activity of the derived homozygote (triangle) because of the diminishing return. Adapted from fig. 6 of Kacser and Burns. c) Kacser and Burns' metabolic model. Similar to b) but in this example, there are fewer enzymes in the multiple-enzymes system. Consequently, here the flux of the heterozygote is close to the average of the ancestral and derived homozygotes. Adapted from fig. 4 of Kacser and Burns. d) Hurst and Randerson's optimal expression model. The *y*-axis indicates fitness and the *x*-axis indicates expression level, which is determined by the genotype. The ancestral homozygote has an optimal expression level (circle). Mutations decrease the expression level and fitness in the heterozygotes (square), and the expression and fitness drop to zero in the derived homozygote (triangle). Increasing gene expression beyond the optimal level decreases fitness due to the costs of increasing gene expression. e) Huber et al. (2018)'s modified optimal expression model is similar to d) but the fitness is above zero when the expression level of a non-essential gene is zero. The dashed line and solid line have different scale parameters reflecting different relationships between genotype and fitness. AA denotes genotypes homozygous for the ancestral allele, AD denotes heterozygous genotypes, and DD denotes genotypes homozygous for the derived allele.

Wright's "Physiological" Model

Another early attempt to explain dominance is Bateson and Punnett's presence and absence hypothesis (Bateson 1909). Here, the "presence" of a character is determined by a "positive entity" and most new mutations were simply a loss of this "positive entity" that brought loss of the corresponding character. These new mutations would usually be recessive because the presence of one dose of the "positive entity," which can be interpreted as an allele in a single-locus model, resulted in the character being closer to that seen with two doses of the positive entity compared with zero doses (Wright 1934). This is also a fundamental idea in Wright's "physiological model" which we explain in more detail.

Wright developed a physiological model to explain the existence of dominance that applies to genes encoding enzymes and mainly discussed the types of phenotypes associated with Mendelian mutations. Wright postulated that dominance is a byproduct of the physiology of enzyme kinetics (Wright 1934, 1956). In this model, the phenotype is related to the amount of product from a metabolic pathway, and the amount of the product increases as the amount of an enzyme increases. However, there is a diminishing return. After reaching a certain level, increasing the enzyme concentration further will no longer increase the amount of the product as rapidly as when enzyme levels were lower (Fig. 2b). The reason for this plateau in the amount of the product is that when the activity of the enzyme is high, it is the concentration of the substrate, not the enzyme, that limits the rate of the enzyme-catalyzed reaction. This type of product-to-enzyme concentration relationship can lead to recessive deleterious mutations. Here, the ancestral homozygous genotype is assumed to have high enzyme activity (circle in Fig. 2b). A derived mutation that reduces the enzyme activity by one-half that of the ancestral homozygote (squares in Fig. 2b) may not largely decrease the amount of the product, leading to the mutation having a recessive effect.

Kacser and Burns's "Metabolic Theory" Model

Wright's model was further developed by Kacser and Burns (Kacser and Burns 1981; Keightley and Kacser 1987; Keightley 1996b) into the metabolic theory of dominance. Here, an organism can be considered to be a system with many enzymes working in concert. The flux and enzyme activities are non-linearly correlated in a multiple-enzyme system according to the kinetic model of enzyme activities (Fig. 2b). Assuming the enzyme activity in mutant heterozygotes is intermediate to the two homozygous genotypes, because of the nonlinear relationship between the flux and enzyme activity in a multiple-enzyme system, the flux of the heterozygotes (up solid square in Fig. 2b) will be larger than the average fitness of the two homozygotes (circle and triangle in Fig. 2b). That is, the deleterious mutation will be recessive ($h < 0.5$). Kacser and Burns's metabolic theory also predicts that mutations in proteins in a pathway with more enzymes should be more recessive. In a system that involves more enzymes (Fig. 2b), the nonlinearity between flux and enzyme activity is higher than for a system that involves fewer enzymes (Fig. 2c). Thus, the dominance coefficient is negatively related (not linearly) to the number of enzymes participating in the system. Kacser and Burns also predicted that mutations with large effects are more recessive than mutations with small effects. In their model, for mutations resulting in a complete loss of enzyme activity in mutant homozygotes, h is approximately zero, but for mutations that mildly influence the enzyme activity, h is close to 0.5. Although Kacser and Burns's metabolic theory can explain the molecular basis of recessiveness and predicts a negative correlation between fitness effects and dominance coefficients, it has some limitations. This model is limited to partially recessive deleterious mutations and cannot explain dominant or overdominant mutations. The reason for this is that the fitness-enzyme function is shaped in such a way that the fitness of the heterozygote will always be within the range of the two homozygous genotypes.

In addition, as with Wright's physiology model, Kacser and Burns' metabolic model cannot explain recessive deleterious mutations in non-enzyme proteins and processes operating beyond cellular levels. Additionally, Kacser and Burn's metabolic theory model predicts that a system with many enzymes is less sensitive to changes in enzyme concentration than a system with fewer enzymes. However, experimental work does not always support this prediction, as some theoretical systems are quite sensitive to minor perturbations (Savageau and Sorribas 1989; Savageau 1992; Hurst and Randerson 2000). Also, Grossniklaus et al. (1996) showed that the flux through pathways with nonlinear kinetics may be more sensitive to enzyme kinetics than previously believed (Grossniklaus et al. 1996). Thus, while the metabolic theory of dominance has some theoretical and conceptual justification and can fit some empirical observations, it remains an incomplete explanation for the presence of dominance.

Gene Expression Models

Haldane (1930) proposed another model for the existence of dominance based on the increasing robustness of a system to deleterious mutations. In this model, increasing an enzyme concentration would be beneficial to the organism because it would make the organism more resistant to loss-of-function deleterious mutations. If the organism produces more enzymes than needed, reducing the concentration via deleterious mutations would not have an effect on fitness. Under this model, Hurst and Randerson found that an allele that would increase expression would increase in frequency in the population via positive selection due to its propensity to buffer the effects of other deleterious mutations (Hurst and Randerson 2000). However, Hurst and Randerson pointed out that this model is unrealistic. There are costs to the organism producing enzymes when they are not needed. When including such costs in this model, they found that the net fitness gain would probably be too small, being the same order of magnitude as the mutation rate. Thus, Haldane's model has the same limitations as Fisher's earlier model.

Hurst and Randerson then developed and analyzed a model that included costs to gene expression (Fig. 2d). Such a model includes stabilizing selection on an optimal amount of gene product. In other words, it features a concave relationship between fitness and the amount of gene product produced. Such a model relates the selection coefficients of mutations to the optimal expression levels of genes. This model predicts that deleterious mutations are likely to be recessive.

The Hurst and Randerson model of selection on the expression level has been further developed by Huber et al. (2018). They added two parameters to describe genes with different functionalities. First, Huber et al. (2018) included an intercept parameter that determines the fitness when the gene is not expressed (Fig. 2e) to reflect the fact that non-essential genes have been found in bacteria, yeast, and other organisms (Gao et al. 2015). In addition, fitness may increase with expression level at different rates in genes with different functions. To model this difference, Huber et al. (2018) introduced a scale parameter, which is the expression level at which fitness is exactly in the middle between the fitness at zero expression and infinite expression (assuming no expression costs; Fig. 2e, solid line and dashed line).

Fitness Landscape Models

Manna et al. (2011) further developed stabilizing selection models of dominance by using Fisher's geometric model of

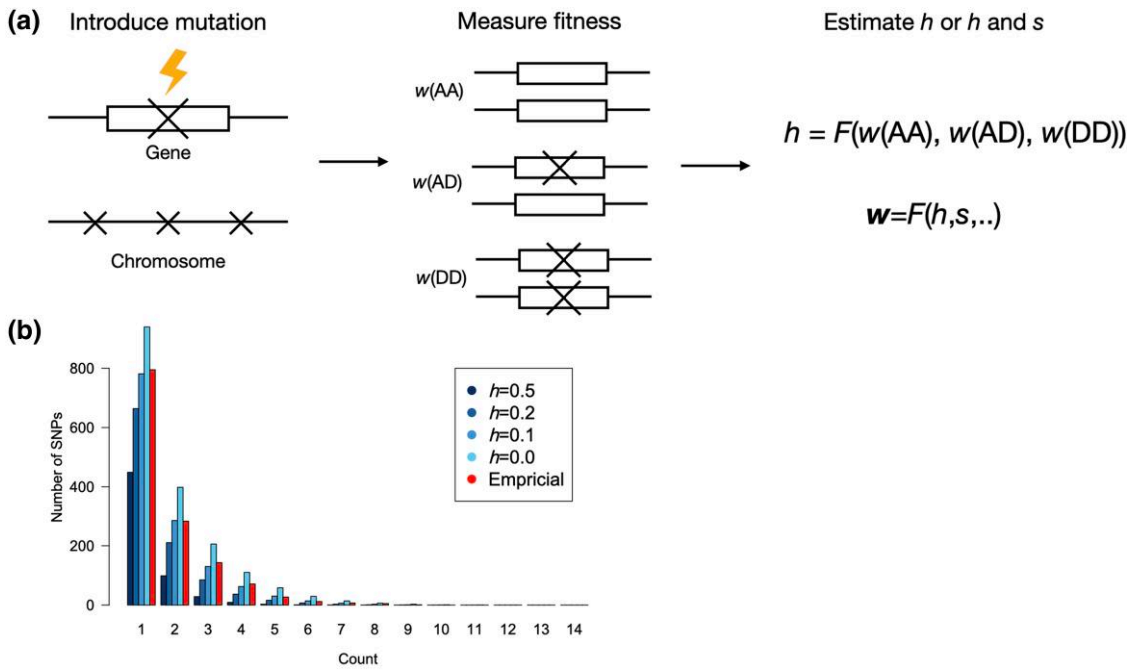

Fig. 3. Approaches to estimating dominance. a) Workflow of laboratory-based methods. The first step of experimental methods is to introduce mutations by deleting a gene, introducing mutations by chemical treatment, or allowing them to accumulate over time. Then, the fitness of heterozygotes, derived homozygotes, and ancestral homozygotes are either directly measured or approximated by some other measurements. Lastly, *h* and *s* are estimated from the fitnesses of the different genotypes using statistical models. b) The site frequency spectrum (SFS) is sensitive to the dominance of deleterious mutations. The SFS contains the number of variants (*y*-axis) at different frequencies in the sample (*x*-axis) from a population. Note, here we assume $\theta = 1{,}000$ and $2Ns = -5$. If deleterious mutations are more recessive, then they are more likely to segregate in the sample. To estimate *h*, the SFS from a hypothetical empirical sequencing dataset can be compared with that predicted by a model of dominance. Here the empirical SFS is closest to that from a model where $h = 0.1$. Note that the SFS was made using the theory presented in Williamson et al. (2004).

mutation and selection with a multivariate Gaussian fitness function relating phenotype to fitness. From this model, they made a number of observations regarding the dominance of mutations. First, the nonlinear relationship between phenotype and fitness when close to the optimum can naturally generate dominance for fitness even when mutations have additive effects on the underlying phenotype. Specifically, as the fitness function is likely to be concave, homozygous deleterious mutations are more likely to have an effect greater than twice that of the heterozygous genotype, making them recessive. Further, they found a mean $h < 0.5$ for deleterious mutations while the mean *h* is >0.5 for beneficial mutations, consistent with empirical estimates (see below). In contrast to other models, the model from Manna et al. also explains overdominance (Manna et al. 2011).

Approaches to Estimating Dominance Coefficients

Testing whether the conceptual models described above can explain the presence of dominance requires inferring dominance coefficients (*h*) from empirical data. These estimates of *h* also are important for answering various evolutionary questions which we will review later. Broadly, there are two types of methods for estimating *h* (Fig. 3). The first type uses measurements from organisms in the laboratory (Fig. 3a) while the second type uses polymorphism data from natural populations along with population genetic models (Fig. 3b). We will discuss specific estimates from the different methods after reviewing the methods.

Laboratory-Based Approaches

Here, individuals are bred in the lab and phenotypes or frequencies of heterozygous and homozygous individuals are measured to estimate dominance coefficients. This type of study can be used to study deleterious variation segregating in natural populations (Mukai et al. 1972) or de novo mutations that occur in the lab (Ohnishi 1977).

Different strategies can be used to generate mutant individuals in the lab. In one strategy, called mutation accumulation (MA), spontaneous mutations are allowed to accumulate over time. For example, in the experiments using *Drosophila* by Ohnishi (1977), spontaneous mutations accumulated up to 40 generations. MA experiments have also been used in multiple other studies of *Drosophila* (Mukai et al. 1964; Mukai and

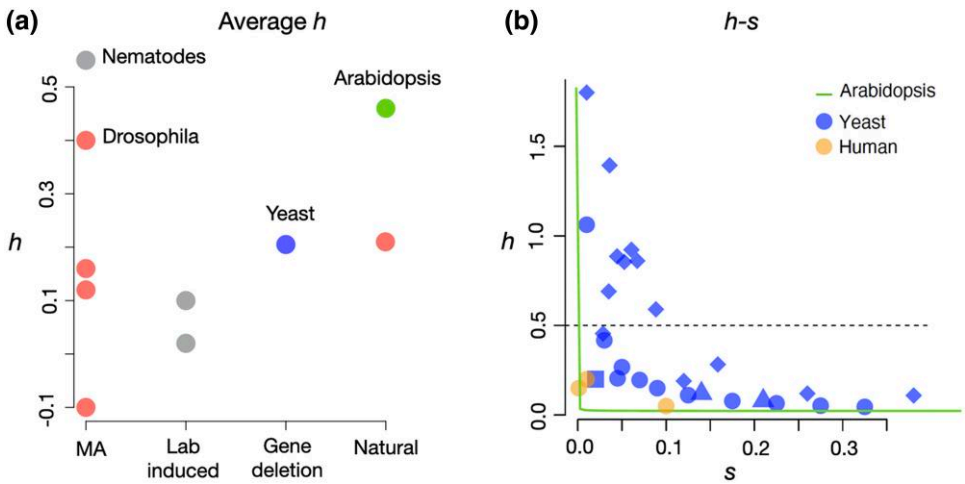

Fig. 4. Estimates of dominance coefficients. a) Average *h* estimated by different methods. Dominance coefficients vary between species and across methods used for estimation. From left to right are estimates from mutations in mutation accumulation experiments (MA), mutations induced by EMS treatment, mutations from gene deletions, and mutations segregating in natural populations. Together, there are five estimates from *Drosophila* (red), three estimates from nematodes (gray), one from yeast (blue), and one from *Arabidopsis* (green). Estimates of the same group of mutations may differ because of the different statistical methods employed (see main text). b) Negative relationship between *h* and *s*. Each dot indicates an estimate of *s* and *h*. Mutations that are more deleterious (larger *s*) tend to be more recessive (smaller *h*). This trend is observed in yeast and humans, with different shapes indicating different studies of yeast. Note that the *h*–*s* relationship inferred by Huber et al. (2018) for *Arabidopsis* (solid line) is highly recessive. See supplementary table S1, Supplementary Material online for the raw data in this figure.

Yamazaki 1968; Mukai 1969; Ohnishi 1977; Simmons and Crow 1977; Houle et al. 1997; García-Dorado and Caballero 2000; Chavarrías et al. 2001; Fry and Nuzhdin 2003) and nematodes (Vassilieva et al. 2000). In MA experiments, spontaneous mutations can accumulate in lab conditions where natural selection is minimal (Ohnishi 1977). Alternatively, new mutations are introduced by chemical mutagens (Ohnishi 1977; Simmons et al. 1978; Peters et al. 2003; Szafraniec et al. 2003) or X-ray treatment (Wallace 1958). For example, in the experiments by Ohnishi (1977), ethyl methanesulfonate (EMS) was fed to *Drosophila* to introduce mutations. In more recent studies of yeast, mutations were also directly introduced by disrupting individual genes (so-called gene-knockouts [Phadnis and Fry 2005; Agrawal and Whitlock 2011]). For example, data from the yeast protein-coding region deletions (Steinmetz et al. 2002) were used by Agrawal and Whitlock (2011) to infer dominance coefficients and study how the relationship between *h* and *s* varies across several functional categories of genes.

When studying either segregating variants in natural populations or de novo mutations that occurred in the lab, trait values, such as viability and growth rate, are measured for mutant homozygotes and heterozygotes (Fig. 3a) for inferring the dominance coefficients of mutations. This process is aided by using balancer chromosomes to mark the chromosome of interest and prevent crossing over (Mitchell 1977; Simmons and Crow 1977). Several statistical methods have been applied to calculate the average dominance coefficient. The first statistical method uses the decline in mean viability of mutant lines in both heterozygous and homozygous genotypes (Ohnishi 1977; Fry and Nuzhdin 2003; Peters et al. 2003). This method has been used differently across studies (Fry and Nuzhdin 2003; Peters et al. 2003) but essentially, *h* is estimated by the ratio of the difference between heterozygotes and controls versus homozygotes and controls. For example, in the study by Peters et al. (2003), the average dominance coefficient $\bar{h} = (Z_{\text{het}} - Z_{\text{an}})/(Z_{\text{hom}} - Z_{\text{an}})$, where Z_{het}, Z_{hom}, and Z_{an} are trait values among heterozygotes, homozygotes, and the ancestral genotypes, respectively. The second approach involves estimating the average *h* using a regression method (Mukai et al. 1972). Mukai et al. (1972) showed that the average *h* over the set of mutations can be measured by the regression coefficient of heterozygous viability on the sum of the two corresponding homozygotes (Mukai et al. 1972; Caballero et al. 1997; Fry and Nuzhdin 2003). In a two-allele model, assuming the genotypes at locus *i* are AA, AD, and DD, the relative fitness values are 1, $1-s_ih_i$, and $1-s_i$ and the genotype frequencies are p_i^2, $2p_iq_i$, and q_i^2 , respectively. If q_i is small, the genetic covariance between heterozygotes is $\sim 2p_iq_is_i^2h_i$ and the variance of the sum of homozygotes is $\sim 2p_iq_is_i^2$. Thus, the average of the dominance coefficient over loci, $\bar{h}$, can be approximated by the ratio of the genetic covariance between heterozygotes to the variance of the sum of homozygotes

$$h = \frac{\sum p_iq_is_i^2h_i}{\sum p_iq_is_i^2} = \sigma_{\text{het,hom}}/\sigma_{\text{hom}}^2$$

(eq. 8, Mukai et al. 1972; Peters et al. 2003). This average dominance coefficient is a weighted average over h of each mutation, being proportional to the genetic variance of the homozygotes. Several other variance component approaches have been developed to infer h. For example, Charlesworth and Hughes (2000) used the ratio of homozygous variances to the additive variances to infer h. Deng extended the covariance/variance approach described above to selfing organisms, eliminating the need to generate homozygous lines (Deng 1998). Other approaches relate h to the mean and variances of fitnesses in the parents and selfing as well as outcrossing progeny (Li and Deng 2000).

The different methods to estimate h have their advantages and disadvantages. The mean decline method is likely to be robust to noise but can be strongly biased by the omission of extreme points (Manna et al. 2011). The regression method is more robust to missing data (Manna et al. 2011). In addition, the regression method corrects for sampling error of the homozygous line means and even does not require a control line (Fry and Nuzhdin 2003). However, the estimates from the regression method can be biased by measurement error (Caballero et al. 1997; Peters et al. 2003; Manna et al. 2011). Compared with the regression method, which is weighted by s^2, the mean decline method is weighted by s (Mukai et al. 1972; Ohnishi 1977; Fry and Nuzhdin 2003; Peters et al. 2003). Indeed, the different statistical approaches have often led to dramatically different estimates of h from the same dataset (García-Dorado and Caballero 2000). For example, Fry and Nuzhdin (2003) estimated $h = -0.1$ using the mean decline method and $h = 0.16$ using the regression method (Fig. 4a; Fry and Nuzhdin 2003). A further challenge with MA experiments is the decline in viability during the experiment due to factors other than mutations. Such declines have been shown to bias inferences of h (García-Dorado and Caballero 2000).

More recent studies estimated dominance coefficients of gene-knockout mutations (Steinmetz et al 2002) by likelihood approaches (Peters et al. 2003; Phadnis and Fry 2005; Agrawal and Whitlock 2011). Steinmetz et al. (2002) created mutant yeast strains by deleting each gene one at a time and measured the growth rates of both heterozygous and homozygous mutant strains. Compared with the mean decline and the regression methods, recent studies fit more realistic models to the data and quantitatively co-estimated dominance and selection coefficients (Peters et al. 2003; Phadnis and Fry 2005; Agrawal and Whitlock 2011). For example, using the yeast growth rate data, Agrawal and Whitlock (2011) compared a series of models in a likelihood framework that described the distribution of s and h and the measurement error of growth rates. For each gene-knockout mutation, they calculated the likelihood of observing the growth rate of the pair of homozygotes and the heterozygote from the fitness data and simulated parameters. Summing the log-likelihoods over all the mutations, they found the parameters that maximize the likelihood.

However, there are drawbacks to the yeast knockout dataset. The fitness of the ancestral type is not included in the experiment by Steinmetz et al. (2002). Different studies used different approximations to circumvent this problem. In Phadnis and Fry (2005), the knockout strains with the highest fitness were used as putative ancestral types. However, Agrawal and Whitlock (2011) argued that this may result in an overestimation of the fitness of wild type because these strains with the highest fitness may have beneficial mutations (and may not be directly ancestral [Agrawal and Whitlock 2011]). Instead, Agrawal and Whitlock (2011) considered strains where the knockout genes do not have known functions as ancestral genotypes because these deletions are likely to be neutral.

Population Genetic Approaches

In addition to laboratory-based approaches to estimate dominance coefficients, patterns of genetic variation segregating in natural populations combined with population genetic models also have the potential to estimate dominance coefficients for newly arising mutations. Early theoretical work from Wright and Kimura showed the expected frequency of an allele in a finite population experiencing selection with arbitrary dominance coefficients (Wright 1938; Kimura 1964). This theory was later extended by Williamson et al. (2004) to be incorporated into the Poisson Random Field (PRF) framework that could be applied to data. In particular, PRF models use diffusion theory to model the change in allele frequency in the population as a function of mutation, drift, and selection. Then, binomial sampling is applied to the population allele frequencies to obtain the expected counts of variants at different frequencies in the sample (Sawyer and Hartl 1992; Hartl et al. 1994). This summary statistic is called the site frequency spectrum (SFS) and is a workhorse of population genetic inference (Fig. 3b; Williamson et al. 2005; Eyre-Walker et al. 2006; Gutenkunst et al. 2009). The key advance of Williamson et al. (2004) was that they derived expressions and a maximum likelihood inference scheme to estimate s and h for new mutations from genetic variation segregating in natural populations. However, this initial method assumed all mutations have the same values of s and h, which is probably not very biologically plausible. While extending the method to consider a distribution of fitness effects (DFE) and many values of h was possible, separately inferring both s and h is challenging as multiple combinations of s and h values can give rise to the same SFS. Further, the SFS-based approach only provides direct information about h for weakly deleterious mutations, as strongly deleterious mutations are less likely to be segregating in genetic variation datasets.

Perhaps due to the limited power to co-infer s and h, the approach of Williamson et al. (2004) was not applied to

empirical data for quite some time. Indeed, most of the population genetic inference of selection using the PRF approach assumed that $h = 0.5$ when inferring the DFE (Eyre-Walker et al. 2006; Keightley and Eyre-Walker 2007; Boyko et al. 2008; Huber et al. 2017; Kim et al. 2017; Tataru et al. 2017; Booker and Keightley 2018; Castellano et al. 2019; Tataru and Bataillon 2019; Huang et al. 2021). One of the first applications of the theory by Williamson et al. (2004) was part of a study of adaptive evolution on the X chromosome versus autosomes in 2014 (Veeramah et al. 2014). Here, using the PRF approach, Veeramah et al. inferred that nonsynonymous mutations on the human autosomes have $h = 0.3$ to 1, a large range encompassing many dominance coefficients. Essentially, they could only eliminate models where all nonsynonymous mutations were highly recessive ($h < 0.3$). More recent work found that the SFS has little power to detect strong recessive selection for individual genes in the human genome (Balick et al. 2022). These results illustrate just how hard it is to co-infer s and h.

A number of approaches have been developed to attempt to overcome the lack of identifiability between s and h. In an elegant study, Balick et al. used the observation that additive and recessive mutations will have different dynamics in populations that experience bottlenecks (Balick et al. 2015). Specifically, prolonged bottlenecks are expected to increase the number of derived deleterious additive alleles while decreasing the number of strongly recessive deleterious alleles, due to increased purging from the increase in homozygosity from the bottleneck effect. They developed a statistic called B_R that is the ratio of the number of derived deleterious alleles in a population that did not experience a bottleneck to the number of derived deleterious alleles in a population that experienced a bottleneck. $B_R > 1$ suggests that deleterious mutations are recessive. $B_R = 1$ suggests mutations are additive. Through simulations, Balick et al. showed that the power of this test is highly dependent on the specific biological parameters involved, including the timing and strength of the bottleneck, the selection coefficients, and the dominance coefficients. Unfortunately, researchers cannot control these parameters and this approach may not have sufficient power to infer the dominance coefficients for most mutations. Nevertheless, Balick et al. found that $B_R > 1$ for genes implicated in recessive diseases as identified in the Human Gene Mutation Database and the Laboratory for Molecular Medicine. In their more recent work, Balick et al. (2022) found that genes implicated in autosomal recessive disease are enriched for genes experiencing recessive deleterious mutations. These results suggest a link between disease phenotypes and evolutionary fitness.

In a different study, to co-infer the DFE and h, Huber et al. combined polymorphism data from selfing and outcrossing populations (Huber et al. 2018). The key insight that they had was that in a selfing population, all individuals are homozygous. Thus, patterns of genetic variation only reflect the homozygous effect of a deleterious mutation (i.e. there is information about s and no information about h). Conversely, in an outcrossing randomly mating population, most deleterious mutations are in the heterozygous state, thus providing information about s and h. By jointly leveraging both types of populations, it is possible to separately infer both s and h. Huber et al. implemented this idea in the PRF model, extending the work of Williamson et al. (2004). They then applied their inference method to *Arabidopsis thaliana* (largely selfing) and *Arabidopsis lyrata* (outcrossing).

Other studies have suggested that richer patterns of genetic variation beyond the SFS may provide additional power to infer s and h. Garcia and Lohmueller (2021) showed the decay of linkage disequilibrium (LD) between deleterious variants depended both on s and h, even when matching pairs of variants for allele frequency. Thus, LD contains additional information about s and h beyond that found in the single-variant SFS. Future work could use this idea in an inferential framework. Ragsdale (2022) developed a numerical approach for finding the 2-locus sampling distribution with arbitrary selection. Here, LD between pairs of variants depends on the specific parameters of both s and h. Beyond genetic variation data from a single generation, changes in allele frequency over time provide valuable information about s and h for beneficial and deleterious mutations. Methods have been developed to use time series data and transmission patterns in parent-offspring trios for inference of h (Mathieson and McVean 2013; Foll et al. 2015; Barroso and Lohmueller 2023), which may gain increased applicability as more ancient genomes and large samples of trios are generated.

Estimates of Dominance Coefficients

Studies have applied the methods described above to estimate dominance coefficients in a variety of organisms, sets of genes, and types of mutations. We discuss some general trends from these studies, focusing on deleterious mutations.

Estimates From Different Organisms

We compiled estimates of dominance coefficients of deleterious mutations from 15 studies of 5 organisms conducted over the past 55 years (supplementary table S1, Supplementary Material online). Estimates of the dominance coefficient vary dramatically across studies (Fig. 4a). For example, in *Drosophila*, estimates of the mean dominance coefficient ranged from −0.1 (Fry and Nuzhdin 2003) to 0.4 (Mukai and Yamazaki 1968). However, serious concerns have been raised about early estimates of *Drosophila* in Mukai and Ohnishi's experiments (García-Dorado and Caballero 2000). Estimates of the average dominance coefficient in nematodes ranged from 0.02 (Peters et al. 2003) to

0.55 (Vassilieva et al. 2000). More recent estimates from genetic variation in *Arabidopsis* found a dominance coefficient of nonsynonymous mutations of 0.46 (Huber et al. 2018).

In some studies, estimates of h are close to 0.5 (Vassilieva et al. 2000; Huber et al. 2018). However, this result should be interpreted cautiously because the mean h may be heavily influenced by the many weakly deleterious mutations segregating in experimental or natural populations. If the weakly deleterious mutations tend to be additive while strongly deleterious mutations are more recessive (Simmons and Crow 1977), then the average estimates of h may be biased toward higher values (i.e. more additive). Thus, the mean of dominance coefficients may not be informative or appropriate for comparison between studies (Houle et al. 1997; Agrawal and Whitlock 2011). To address this challenge, Agrawal and Whitlock (2011) proposed an s-weighted dominance coefficient, $E_s(h_{\mathrm{del}}) = \sum hs / \sum s$. However, the s-weighted average also may not be so informative because of the large variance of h.

A better way to compare the degree of dominance across studies may be to compare h within a small interval of selection coefficients (Phadnis and Fry 2005; Agrawal and Whitlock 2011). This comparison requires co-estimating h and s. Several studies that co-estimated h and s found that strongly deleterious mutations are more likely to be recessive (Simmons and Crow 1977; López and López-Fanjul 1993; Szafraniec et al. 2003; Phadnis and Fry 2005; Agrawal and Whitlock 2011; Huber et al. 2018). This pattern is especially evident in Fig. 4b where we show the estimates of dominance coefficients (h, y-axis) of mutations as a function of the selection coefficient (s, x-axis). For example, in the yeast gene-knockout experiments (Steinmetz et al. 2002), Agrawal and Whitlock (2011; Fig. 4b, blue circle) and Phadnis and Fry (2005; Fig. 4b, blue diamond) estimated the dominance and selection coefficients of coding region deletions. Within each study, h was small when s was large (Fig. 4b). This pattern held for studies that used different statistical approaches for estimation, suggesting that it is not entirely driven by the methodology used for inference (supplementary table S1, Supplementary Material online). In addition, Huber et al. co-estimated the DFE and h for nonsynonymous mutations in *Arabidopsis* (Huber et al. 2018). They found that a model where h was a function of s fit the data better than a model where all mutations were assumed to have the same value of h. In this best-fitting model, as s becomes more deleterious, mutations become more recessive (Fig. 4b). Nonsynonymous mutations across the human autosome were estimated to not be very recessive ($h > 0.3$; Veeramah et al. 2014). However, a smaller h was found to be paired with a larger s for mutations in autosomal recessive disease genes (Fig. 4b, Balick et al. 2015). In older studies of *Drosophila*, mutations were qualitatively categorized into a few groups, such as lethal and nonlethal. Lethal mutations or mutations of large effect were found to be more recessive than mildly deleterious mutations (Wright et al. 1942; Muller 1950; Ohnishi 1977; Simmons and Crow 1977; Crow and Simmons 1983; Keightley 1996a). In addition, the estimates of h for more deleterious mutations varied less between studies compared with estimates of h for less deleterious mutations (Fig. 4b).

Estimates of Dominance for Different Mutations & Genes

Estimated dominance coefficients also may vary between studies when different types of mutations are considered. Deleterious variants segregating in individuals sampled from natural populations may be more recessive compared with experimentally introduced mutations. The reason for this is that more dominant deleterious mutations are more easily removed by natural selection (see further discussion of this topic below). Empirical estimates from *Drosophila* found that h for segregating mutations in natural populations ranged from 0.20 to 0.35 while new mutations occurring in the lab have greater h, ranging from 0.35 to 0.5 (Simmons and Crow 1977).

Gene function also influences the degree of dominance of mutations. Mutations in enzymes might be more recessive than in structural proteins. A single-copy of an ancestral allele may be sufficient to maintain the whole or partial function of enzymes (Kondrashov and Koonin 2004), while structural proteins contributing to the structural integrity of a complex (GO:0005198, structural molecule activity [Ashburner et al. 2000; The Gene Ontology Consortium et al. 2023]) may need more molecules to build a functional part (Huber et al. 2018). Several empirical studies in different species suggest that mutations in catalytic genes are more recessive than those in structural proteins (Phadnis and Fry 2005; Agrawal and Whitlock 2011; Huber et al. 2018). For example, Huber et al. (2018) found that mutations in structural proteins are partially recessive, but for a given s, mutations in catalytic genes are more recessive. In addition, Agrawal and Whitlock (2011) showed that gene-knockout mutations in structural genes ($h = 0.127$, $s = 0.14$) may be more deleterious, but less recessive compared with knockout mutations in catalytic genes ($h = 0.046$, $s = 0.04$).

Expression level and gene connectivity to other genes are also features that may influence the dominance of mutations. Specifically, the model of Huber et al. (2018) predicts that mutations in genes with high optimal gene expression are more additive than genes with low optimal gene expression. Consistent with this prediction, Huber et al.'s (2018) empirical analyses of dominance coefficients in *Arabidopsis* found that mutations in genes with higher expression levels tended to be more additive than mutations in genes with lower expression levels. Empirical analyses also found that mutations in genes with higher connectivity to other genes are more additive than mutations in genes that are less connected (Phadnis and Fry 2005; Huber

et al. 2018). Thus, it appears that dominance patterns differ across functional categories in ways beyond just contrasting catalytic versus structural proteins.

Gene duplication or polyploidization has been proposed to be a buffer against deleterious mutations (Haldane 1933; Fisher 1935). However, although mutations in yeast duplicate genes have been found to have weaker homozygous effects than those in single-copy genes (Gu et al. 2003), for a given selection strength, the dominance of mutations in these genes is similar to that of duplicate genes (Phadnis and Fry 2005).

Estimates of Dominance Coefficients Provide a Test of Models of Dominance

As discussed earlier, the different conceptual models for the existence of dominance make different predictions that can be tested using empirical estimates of dominance. Fisher's "modifier model" does not predict a relationship between h and s. Thus, the ample evidence from different studies (Simmons and Crow 1977; López and López-Fanjul 1993; Szafraniec et al. 2003; Phadnis and Fry 2005; Agrawal and Whitlock 2011; Huber et al. 2018) that more deleterious mutations are more recessive cannot be explained by Fisher's model. The "physiological model" for dominance predicts that mutations in enzymes will be more recessive than mutations in structural proteins. At first glance, this prediction seems to be supported by Agrawal and Whitlock (2011) and Huber et al. (2018), who found that mutations in catalytic genes are more recessive than those in structural proteins. However, Huber et al. (2018) also found that while mutations in structural proteins were more additive than those in catalytic proteins, they were still partially recessive ($h < 0.1$ for mutations with s more deleterious than 5×10^{-4}). As the physiological model cannot fully explain the presence of recessive mutations in non-catalytic genes, it cannot be the entire explanation for the existence of dominance. Huber et al. (2018) also noted that there were differences in the h-s relationship across other functional categories of genes. Specifically, genes with higher expression tended to be more additive than genes with lower expression. The physiological model does not predict this pattern. Searching for a more general model for the existence of dominance, Huber et al. (2018) extended the model from Hurst and Randerson (2000). The Huber et al. (2018) model predicts that dominance is the inevitable consequence of stabilizing selection and maintaining genes at their optimal expression levels. This premise is supported by the stabilizing selection model used by Manna et al. (2011), which also predicted dominance as a consequence of stabilizing selection acting on multiple traits. Through simulations, Huber et al. (2018) showed that their model predicts that more deleterious mutations will be more recessive and that genes with higher optimal expression will be more additive than genes with lower optimal expression. Both of these predictions were supported by the empirical estimates of dominance. As such, the Huber et al. (2018) model for dominance explains key features of the data and is applicable to more genes than other existing models.

The Impact of Dominance on Evolution

Thus far, we discussed why dominance exists and how we can quantify it. Now we turn our attention to how dominance influences evolutionary processes. Once a selected mutation appears in a population, its fate is influenced by

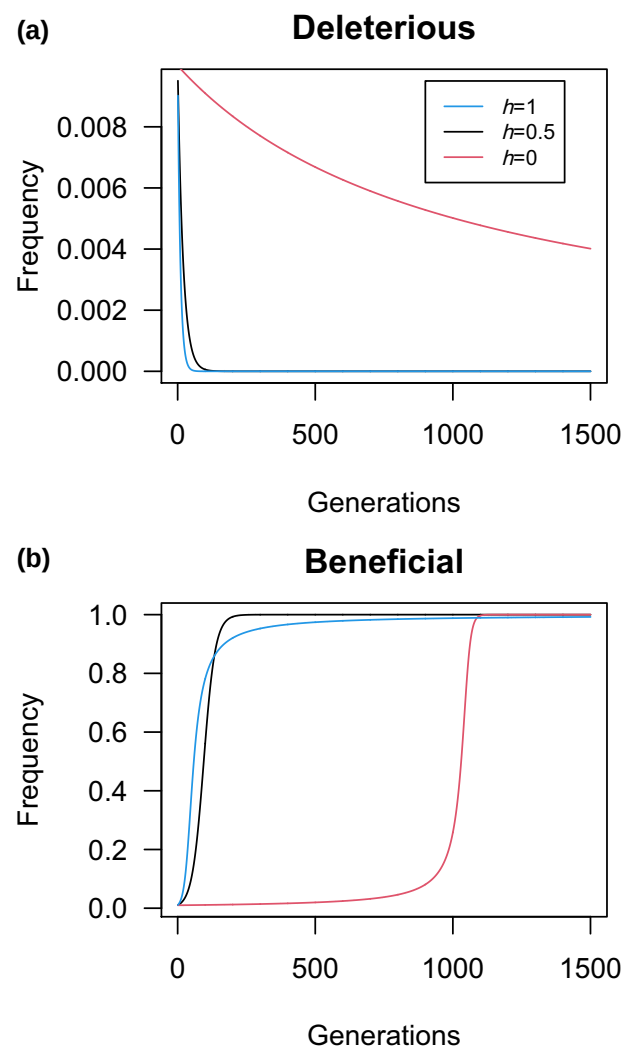

Fig. 5. Dominance affects the dynamics of selected mutations in populations. a) Recessive deleterious mutations (h=0) decrease in frequency much more slowly due to selection than additive or dominant deleterious mutations. Here, $s = 0.1$. b) Beneficial mutations that are dominant (h=1) initially increase in frequency due to selection faster than additive or recessive beneficial mutations. Here, $s = 0.1$.

the dominance and selection coefficients as well as other nonselective forces like drift. We outline a number of scenarios where the dominance coefficient of a selected mutation can have a fundamental impact on the dynamics of evolution.

Deleterious Mutations

The per-generation decrease in frequency of a deleterious allele by selection is heavily influenced by h (Fig. 5a). Note that recessive deleterious mutations take substantially longer to decrease in frequency by selection than do additive or dominant mutations. As most deleterious mutations are at low frequency in the population, they are typically carried in heterozygous genotypes. Thus, as long as $h > 0.2$, selection will primarily operate on heterozygotes and can reduce the frequency of the allele. For fully recessive mutations where $h = 0$, selection will only occur on homozygotes, and recessive deleterious mutations can be shielded from selection in the heterozygous state, resulting in a slower decrease in frequency. However, if strongly deleterious mutations have h even slightly > 0, as suggested by some early studies (Wright et al. 1942; Muller 1950; Ohnishi 1977; Crow and Simmons 1983; Keightley 1996a), then they should be eliminated by selection on the heterozygous carriers.

Thus, models of mutation–selection balance predict that h influences the equilibrium frequency of deleterious alleles. Specifically for non-recessive mutations, the equilibrium frequency is $q = \mu/hs$. For the recessive case ($h = 0$), $q = \sqrt{\mu/s}$. Thus, for a given s, the equilibrium frequency of the deleterious allele will be higher for more recessive alleles, sometimes dramatically so. For example, for a strongly deleterious mutations with $s = 0.1$, if $\mu = 1 \times 10^{-4}$, then q would be 0.2% if the mutation were additive, but would be 3.2% if recessive.

Because recessive deleterious mutations can be hidden in the heterozygous state, they may be significant to the evolutionary fate of populations. As inbreeding (i.e. mating among close relatives) increases homozygosity, it will unmask the heterozygous deleterious mutations, reducing fitness. This effect is known as inbreeding depression and has been documented in numerous laboratory and field populations (Keller and Waller 2002; Charlesworth and Willis 2009). Theoretical work has shown that the relationship between population size and inbreeding depression greatly depends on the dominance coefficients for strongly deleterious mutations ($s > 0.01$). Specifically, strongly deleterious mutations with $h < 0.05$ can accumulate in heterozygous genotypes in large populations compared with smaller populations. However, the population size has less of an effect on the frequency of partially recessive ($h > 0.05$) or additive strongly deleterious mutations (Hedrick 2002; Hedrick and Garcia-Dorado 2016; Kyriazis et al. 2021). These dynamics have also been shown to have important implications for how historical demography influences risk of extinction due to reduced fitness from inbreeding depression (Kyriazis et al. 2021; Robinson et al. 2022, 2023).

Dominance also plays a major role in determining the impact of a population bottleneck for deleterious variation (Balick et al. 2015). Specifically, a bottleneck is expected to increase the number of weakly deleterious additive alleles, since selection cannot overcome the increased drift due to the small population size in the bottleneck. However, the dynamics for partially recessive deleterious mutations are more complex. The bottleneck will lead to an increase in homozygosity for recessive deleterious mutations. This in turn will initially result in a transient reduction in fitness (due to the homozygous individuals), but later purging of deleterious alleles, thus increasing fitness even beyond that seen in the ancestral population. The specific quantitative effects of the bottleneck on patterns of deleterious variation are highly parameter-dependent, and we suggest using simulations to explore scenarios relative to cases of interest (Kirkpatrick and Jarne 2000; Balick et al. 2015; Kyriazis et al. 2023).

Partially recessive deleterious variation can produce complex dynamics on linked variants. Specifically, in a condition known as associative overdominance, recurrent partially recessive deleterious mutations can cause an apparent heterozygote advantage at linked neutral loci (Ohta 1971). Heterozygous genotypes mask the effects of the recessive deleterious alleles, resulting in higher fitness. In linked neutral regions, associative overdominance may increase genetic diversity above neutral levels, due to the greater sojourn time of weakly deleterious recessive mutations (Mafessoni and Lachmann 2015). This effect has been studied theoretically and empirically and is found most likely to occur in scenarios of weak selection (where $2Ns < 10$), $h < 0.25$, and when recombination rates are low (Zhao and Charlesworth 2016; Gilbert et al. 2020; Charlesworth and Jensen 2021; Charlesworth 2022). Empirical studies have documented this effect in low recombination rate regions of the human and Drosophila genomes (Becher et al. 2020; Gilbert et al. 2020).

This associative overdominance effect can be especially pronounced after population admixture. If deleterious mutations are recessive, then introgression from a different population would mask the recessive deleterious mutations in the original population, increasing fitness. Thus, introgressed haplotypes without any beneficial mutations per se can increase in frequency in a population, potentially mimicking the genomic patterns of adaptive introgression (Crow 1948; Whitlock et al. 2000; Harris and Nielsen 2016; Kim et al. 2018; Abu-Awad and Waller 2023). This effect is likely to be most pronounced when the populations that mix diverged a long time ago such that each population accumulated its own private set of deleterious variants. Further, this effect is likely to occur in regions of the genome with low recombination rates and high density of functional elements, like exons.

If deleterious mutations are additive, on the other hand, then the population sizes of the two parental populations will have a greater influence on the dynamics of introgression. This associative overdominance effect was recently shown potentially to have led to several false signatures of adaptive introgression in human populations at *HYAL2* and *HLA*. The 24 other candidate regions for adaptive introgression are likely not confounded by this effect (Zhang et al. 2020).

Beneficial Mutations

Dominance also affects positive selection and adaptation. For instance, the dynamics of a beneficial allele in a population is greatly influenced by the degree of dominance (Fig. 5b). The frequency of a recessive beneficial mutation that is rare in the population only slowly increases in frequency by positive selection. The reason for this slow change in frequency is that recessive beneficial mutations only experience the effects of selection when they become homozygous. Most rare alleles are carried as heterozygotes, and are thus shielded from selection. Once the beneficial allele becomes more common, however, then selection more quickly increases its frequency. New beneficial mutations that are dominant, on the other hand, quickly increase in frequency in the initial stages.

These differing dynamics of beneficial mutations with distinct dominance coefficients has several implications for understanding adaptation. Haldane argued that beneficial mutations that are dominant are more likely to become established and reach fixation compared with recessive beneficial mutations (Haldane 1924, 1927). Thus, if one were to examine cases where adaptation was complete, they would be enriched for dominant mutations. This concept was termed, "Haldane's sieve" (Turner 1981). Indeed, a number of examples of adaptation from dominant mutations have been noted for insecticide resistance (Bourguet and Raymond 1998; Charlesworth 1998) and in experimental evolution (Marad et al. 2018), though the degree of dominance for many beneficial mutations has not been quantified.

The conditions under which Haldane's sieve is expected to apply were further extended to consider a model where standing deleterious variation, after an environmental shift, becomes beneficial (Orr and Betancourt 2001). Importantly, Orr and Betancourt assumed that the degree of dominance for the deleterious effect was the same as that for the beneficial effect. They found that in a situation where deleterious alleles are at mutation–selection balance, the dominance coefficient does not influence the fixation probability of the beneficial allele, suggesting Haldane's sieve does not apply. The intuition behind this result is that while dominant beneficial mutations have a higher fixation probability than recessive ones, a dominant deleterious allele will be at a lower frequency at mutation–selection balance than the recessive allele. Put another way, recessive deleterious alleles can reach higher frequencies, and thus, when they become beneficial, they are already starting from higher frequencies. These effects cancel, leading to the breakdown in Haldane's sieve.

More recent work has explored models of selection on standing variation where the initial deleterious mutations are recessive, but become dominant when the mutations become beneficial (Muralidhar and Veller 2022). Evidence for strong-effect deleterious mutations being recessive has been discussed above. Gain-of-function mutations are thought to be primarily dominant. A change in dominance coefficient may be expected if the precise phenotype under selection changes with the environmental shift. As an example, mutations in the Ace locus in Drosophila alter the shape of the binding site and may be recessive deleterious, because they are less effective metabolizing acetylcholine when homozygous (Labbé et al. 2014; Zhang et al. 2015). However, if such mutations also decrease the binding of organophosphate pesticides, then such mutations may be beneficial by providing pesticide resistance. In this case, the specific phenotype affecting fitness would have shifted from enzyme kinetics to pesticide resistance. Such mutations when beneficial have been shown to be dominant (Bourguet and Raymond 1998; Charlesworth 1998). In their simulation study, Muralidhar and Veller (2022) found that the shift in dominance coefficient from recessive to dominant resulted in a higher probability of the beneficial mutation becoming established in the population. Further, this shift in dominance coefficient also resulted in a higher probability of the completed sweep occurring on standing genetic variation thus being soft (that is, the beneficial mutation being present on multiple haplotypes at the onset of selection) rather than hard.

The dominance coefficient also influences the ability to detect a selective sweep from genetic variation data. Due to the longer time of the recessive beneficial allele being at low frequency, it has more opportunity to recombine onto additional haplotypes at the start of selection as compared with an additive or dominant beneficial mutation. In doing so, recessive beneficial mutations are expected to leave less pronounced footprints in the surrounding neutral genetic variation compared with additive or dominant beneficial mutations (Teshima and Przeworski 2006).

Looking Ahead

While dominance has received considerable attention in evolutionary genetics over the past 100 years, additional challenges and unanswered questions remain. First, new methods to better estimate the DFE jointly with the dominance coefficients would be indispensable. Methods using genetic variation from natural populations could leverage multiple aspects of the data beyond allele frequencies. Machine learning

approaches offer a promising way forward for integrating multiple summaries of variation. For laboratory studies, technological improvements enabling higher-throughput experiments could result in more precise and robust estimates. Improving estimates of dominance for strongly deleterious mutations would be especially fruitful as existing estimates for these mutations are uncertain.

Second, while it appears the degree of dominance of mutations may differ between species (Fig. 4), interpreting these differences is challenging. Biological factors, such as the DFE, genome architecture, population size, etc., may differ across species and could potentially explain apparent differences in dominance between them. Further work co-estimating selection coefficients and dominance coefficients across species is necessary for comparative studies.

Third, all of the existing conceptual models of dominance struggle to explain some empirical observations or apply to all circumstances. So, continued development of conceptual models for the existence of dominance will be fruitful. Additionally, continuing to empirically test predictions from such models will be valuable as seeing where models do not fit the data may reveal new biological insights, beyond just falsifying the models.

Fourth, many population genetic methods and models for studying natural selection assume that selected mutations have additive effects on fitness. However, as the dynamics of selected mutations are influenced by the dominance coefficient, more realistic models of dominance may yield qualitatively different results. Consequently, using population genetic models with arbitrary dominance will enable more accurate inferences and predictions for understanding future evolutionary trajectories of populations and complex traits.

Overall, we are optimistic that the advances in high-throughput genomic technologies, cheaper genome sequencing, increases in computational tools, and creative thinking will enable progress toward resolving these questions.

Supplementary Material

Supplementary material is available at *Genome Biology and Evolution* online.

Acknowledgments

We thank Chris Kyriazis and Stella Yuan for helpful comments on the manuscript as well as other members of the Lohmueller lab for helpful discussions on these topics.

Funding

This work was supported by the National Institutes of Health grant R35GM119856 to K.E.L.

Data Availability

All data used in this paper can be found in the supplementary table.

Literature Cited

Abu-Awad D, Waller D. Conditions for maintaining and eroding pseudo-overdominance and its contribution to inbreeding depression. Peer Community J. 2023:3:e8. https://doi.org/10.24072/pcjournal.224.

Agrawal AF, Whitlock MC. Inferences about the distribution of dominance drawn from yeast gene knockout data. Genetics. 2011:187(2):553–566. https://doi.org/10.1534/genetics.110.124560.

The Gene Ontology Consortium, Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, Feuermann M, Gaudet P, Harris NL, et al. The gene ontology knowledgebase in 2023. Genetics. 2023:224(1):iyad031. https://doi.org/10.1093/genetics/iyad031.

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. Nat Genet. 2000:25(1):25–29. https://doi.org/10.1038/75556.

Balick DJ, Do R, Cassa CA, Reich D, Sunyaev SR. Dominance of deleterious alleles controls the response to a population bottleneck. PLoS Genet. 2015:11(8):e1005436. https://doi.org/10.1371/journal.pgen.1005436.

Balick DJ, Jordan DM, Sunyaev S, Do R. Overcoming constraints on the detection of recessive selection in human genes from population frequency data. Am J Hum Genet. 2022:109(1):33–49. https://doi.org/10.1016/j.ajhg.2021.12.001.

Barroso GV, Lohmueller KE. Inferring the mode and strength of ongoing selection. Genome Res. 2023:33(4):632–643. https://doi.org/10.1101/gr.276386.121.

Bateson W. Mendel's principles of heredity. Cambridge: Cambridge University Press; 1909. https://doi.org/10.1017/CBO9780511694462.

Becher H, Jackson BC, Charlesworth B. Patterns of genetic variability in genomic regions with low rates of recombination. Curr Biol. 2020:30(1):94–100.e3. https://doi.org/10.1016/j.cub.2019.10.047.

Booker TR, Keightley PD. Understanding the factors that shape patterns of nucleotide diversity in the house mouse genome. Mol Biol Evol. 2018:35(12):2971–2988. https://doi.org/10.1093/molbev/msy188.

Bourguet D, Raymond M. The molecular basis of dominance relationships: the case of some recent adaptive genes. J Evol Biol. 1998:11(1):103–122. https://doi.org/10.1046/j.1420-9101.1998.11010103.x.

Boyko AR, Williamson SH, Indap AR, Degenhardt JD, Hernandez RD, Lohmueller KE, Adams MD, Schmidt S, Sninsky JJ, Sunyaev SR, et al. Assessing the evolutionary impact of amino acid mutations in the human genome. PLoS Genet. 2008:4(5):e1000083. https://doi.org/10.1371/journal.pgen.1000083.

Caballero A, Keightley PD, Turelli M. Average dominance for polygenes: drawbacks of regression estimates. Genetics. 1997:147(3):1487–1490. https://doi.org/10.1093/genetics/147.3.1487.

Castellano D, Macià MC, Tataru P, Bataillon T, Munch K. Comparison of the full distribution of fitness effects of new amino acid mutations across great apes. Genetics. 2019:213(3):953–966. https://doi.org/10.1534/genetics.119.302494.

Charlesworth B. Evidence against Fisher's theory of dominance. Nature. 1979:278(5707):848–849. https://doi.org/10.1038/278848a0.

Charlesworth B. Adaptive evolution: the struggle for dominance. Curr Biol. 1998:8(14):R502–R504. https://doi.org/10.1016/s0960-9822(98)70318-5.

Charlesworth B. The effects of weak selection on neutral diversity at linked sites. Genetics. 2022:221(1):iyac027. https://doi.org/10.1093/genetics/iyac027.

Charlesworth B, Campos JL, Jackson BC. Faster-X evolution: theory and evidence from *Drosophila*. Mol Ecol. 2018:27(19):3753–3771. https://doi.org/10.1111/mec.14534.

Charlesworth D, Charlesworth B. Theoretical genetics of batesian mimicry III. Evolution of dominance. J Theor Biol. 1975:55(2):325–337. https://doi.org/10.1016/S0022-5193(75)80083-X.

Charlesworth B, Hughes KA. The maintenance of genetic variation in life-history traits. In: Singh RS, Krimbas CB, editors. Evolutionary genetics—from molecules to morphology. Cambridge: Cambridge University Press; 2000. p. 369–392.

Charlesworth B, Jensen JD. Effects of selection at linked sites on patterns of genetic variability. Annu Rev Ecol Evol Syst. 2021:52(1):177–197. https://doi.org/10.1146/annurev-ecolsys-010621-044528.

Charlesworth D, Willis JH. The genetics of inbreeding depression. Nat Rev Genet. 2009:10(11):783–796. https://doi.org/10.1038/nrg2664.

Chavarrías D, López-Fanjul C, García-Dorado A. The rate of mutation and the homozygous and heterozygous mutational effects for competitive viability: a long-term experiment with *Drosophila melanogaster*. Genetics. 2001:158(2):681–693. https://doi.org/10.1093/genetics/158.2.681.

Crow JF. Alternative hypotheses of hybrid vigor. Genetics. 1948:33(5):477–487. https://doi.org/10.1093/genetics/33.5.477.

Crow JF, Simmons MJ. The mutation load in *Drosophila*. In: Ashburner M, Carson HL, Thompson JN, editors. The genetics and biology of Drosophila. Vol. 3c. London: Academic Press; 1983. p. 1–35.

Deng H-W. Estimating within-locus nonadditive coefficient and discriminating dominance versus overdominance as the genetic cause of heterosis. Genetics. 1998:148(4):2003–2014. https://doi.org/10.1093/genetics/148.4.2003.

Eyre-Walker A, Woolfit M, Phelps T. The distribution of fitness effects of new deleterious amino acid mutations in humans. Genetics. 2006:173(2):891–900. https://doi.org/10.1534/genetics.106.057570.

Fisher RA. The possible modification of the response of the wild type to recurrent mutations. Am Nat. 1928:62(679):115–126. https://doi.org/10.1086/280193.

Fisher RA. The sheltering of lethals. Am Nat. 1935:69(724):446–455. https://doi.org/10.1086/280618.

Foll M, Shim H, Jensen JD. WFABC: a wright-fisher ABC-based approach for inferring effective population sizes and selection coefficients from time-sampled data. Mol Ecol Resour. 2015:15(1):87–98. https://doi.org/10.1111/1755-0998.12280.

Fry JD, Nuzhdin SV. Dominance of mutations affecting viability in *Drosophila melanogaster*. Genetics. 2003:163(4):1357–1364. https://doi.org/10.1093/genetics/163.4.1357.

Gao F, Luo H, Zhang C-T, Zhang R. Gene essentiality analysis based on DEG 10, an updated database of essential genes. In: Lu LJ, editors. Gene essentiality: methods and protocols. New York: Springer; 2015. p. 219–233. https://doi.org/10.1007/978-1-4939-2398-4_14.

García-Dorado A, Caballero A. On the average coefficient of dominance of deleterious spontaneous mutations. Genetics. 2000:155(4):1991–2001. https://doi.org/10.1093/genetics/155.4.1991.

Garcia JA, Lohmueller KE. Negative linkage disequilibrium between amino acid changing variants reveals interference among deleterious mutations in the human genome. PLoS Genet. 2021:17(7):e1009676. https://doi.org/10.1371/journal.pgen.1009676.

Gilbert KJ, Pouyet F, Excoffier L, Peischl S. Transition from background selection to associative overdominance promotes diversity in regions of low recombination. Curr Biol. 2020:30(1):101–107.e3. https://doi.org/10.1016/j.cub.2019.11.063.

Grossniklaus U, Madhusudhan MS, Nanjundiah V. Nonlinear enzyme kinetics can lead to high metabolic flux control coefficients: implications for the evolution of dominance. J Theor Biol. 1996:182(3):299–302. https://doi.org/10.1006/jtbi.1996.0167.

Gu Z, Steinmetz LM, Gu X, Scharfe C, Davis RW, Li W-H. Role of duplicate genes in genetic robustness against null mutations. Nature. 2003:421(6918):63–66. https://doi.org/10.1038/nature01198.

Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD. Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. PLoS Genet. 2009:5(10):e1000695. https://doi.org/10.1371/journal.pgen.1000695.

Haldane JBS. A mathematical theory of natural and artificial selection. Part II the influence of partial self-fertilisation, inbreeding, assortative mating, and selective fertilisation on the composition of Mendelian populations, and on natural selection. Biol Rev. 1924:1(3):158–163. https://doi.org/10.1111/j.1469-185X.1924.tb00546.x.

Haldane JBS. Mathematical theory of natural and artificial selection. Part V: selection and mutation. Proc Camb Philos Soc. 1927:28:838–844.

Haldane JBS. A note on Fisher's theory of the origin of dominance, and on a correlation between dominance and linkage. Am Nat. 1930:64(690):87–90. https://doi.org/10.1086/280299.

Haldane JBS. The part played by recurrent mutation in evolution. Am Nat. 1933:67(708):5–19. https://doi.org/10.1086/280465.

Haldane JBS. The theory of the evolution of dominance. J Genet. 1939:37(2):365–374. https://doi.org/10.1007/BF02982734.

Harris K, Nielsen R. The genetic cost of Neanderthal introgression. Genetics. 2016:203(2):881–891. https://doi.org/10.1534/genetics.116.186890.

Hartl DL, Moriyama EN, Sawyer SA. Selection intensity for codon bias. Genetics. 1994:138(1):227–234. https://doi.org/10.1093/genetics/138.1.227.

Hedrick PW. Lethals in finite populations. Evolution. 2002:56(3):654–657. https://doi.org/10.1111/j.0014-3820.2002.tb01374.x.

Hedrick PW, Garcia-Dorado A. Understanding inbreeding depression, purging, and genetic rescue. Trends Ecol Evol. 2016:31(12):940–952. https://doi.org/10.1016/j.tree.2016.09.005.

Houle D, Hughes KA, Assimacopoulos S, Charlesworth B. The effects of spontaneous mutation on quantitative traits. II. Dominance of mutations with effects on life-history traits. Genet Res. 1997:70(1):27–34. https://doi.org/10.1017/s001667239700284x.

Huang X, Fortier AL, Coffman AJ, Struck TJ, Irby MN, James JE, León-Burguete JE, Ragsdale AP, Gutenkunst RN. Inferring genome-wide correlations of mutation fitness effects between populations. Mol Biol Evol. 2021:38(10):4588–4602. https://doi.org/10.1093/molbev/msab162.

Huber CD, Durvasula A, Hancock AM, Lohmueller KE. Gene expression drives the evolution of dominance. Nat Commun. 2018:9(1):2750. https://doi.org/10.1038/s41467-018-05281-7.

Huber CD, Kim BY, Marsden CD, Lohmueller KE. Determining the factors driving selective effects of new nonsynonymous mutations. Proc Natl Acad Sci U S A. 2017:114(17):4465–4470. https://doi.org/10.1073/pnas.1619508114.

Hurst LD, Randerson JP. Dosage, deletions, and dominance: simple models of the evolution of gene expression. J Theor Biol. 2000:205(4):641–647. https://doi.org/10.1006/jtbi.2000.2095.

Kacser H, Burns JA. The molecular basis of dominance. Genetics. 1981:97(3-4):639–666. https://doi.org/10.1093/genetics/97.3-4.639.

Keightley PD. A metabolic basis for dominance and recessivity. Genetics. 1996a:143(2):621–625. https://doi.org/10.1093/genetics/143.2.621.

Keightley PD. Metabolic models of selection response. J Theor Biol. 1996b:182(3):311–316. https://doi.org/10.1006/jtbi.1996.0169.

Keightley PD, Eyre-Walker A. Joint inference of the distribution of fitness effects of deleterious mutations and population demography based on nucleotide polymorphism frequencies. Genetics. 2007:177(4):2251–2261. https://doi.org/10.1534/genetics.107.080663.

Keightley PD, Kacser H. Dominance, pleiotropy and metabolic structure. Genetics. 1987:117(2):319–329. https://doi.org/10.1093/genetics/117.2.319.

Keller LF, Waller DM. Inbreeding effects in wild populations. Trends Ecol Evol. 2002:17(5):230–241. https://doi.org/10.1016/S0169-5347(02)02489-8.

Kim BY, Huber CD, Lohmueller KE. Inference of the distribution of selection coefficients for new nonsynonymous mutations using large samples. Genetics. 2017:206(1):345–361. https://doi.org/10.1534/genetics.116.197145.

Kim BY, Huber CD, Lohmueller KE. Deleterious variation shapes the genomic landscape of introgression. PLoS Genet. 2018:14(10):e1007741. https://doi.org/10.1371/journal.pgen.1007741.

Kimura M. Diffusion models in population genetics. J Appl Probab. 1964:1(2):177–232. https://doi.org/10.2307/3211856.

Kirkpatrick M, Jarne P. The effects of a bottleneck on inbreeding depression and the genetic load. Am Nat. 2000:155(2):154–167. https://doi.org/10.1086/303312.

Kondrashov FA, Koonin EV. A common framework for understanding the origin of genetic dominance and evolutionary fates of gene duplications. Trends Genet. 2004:20(7):287–290. https://doi.org/10.1016/j.tig.2004.05.001.

Kyriazis CC, Robinson JA, Lohmueller KE. Using computational simulations to model deleterious variation and genetic load in natural populations. Am Nat. 2023:202(6):737–752. https://doi.org/10.1086/726736.

Kyriazis CC, Wayne RK, Lohmueller KE. Strongly deleterious mutations are a primary determinant of extinction risk due to inbreeding depression. Evol Lett. 2021:5(1):33–47. https://doi.org/10.1002/evl3.209.

Labbé P, Milesi P, Yébakima A, Pasteur N, Weill M, Lenormand T. Gene-dosage effects on fitness in recent adaptive duplications: ace-1 in the mosquito *Culex pipiens*. Evolution. 2014:68(7):2092–2101. https://doi.org/10.1111/evo.12372.

Lande R, Schemske DW. The evolution of self-fertilization and inbreeding depression in plants. I. Genetic models. Evolution. 1985:39(1):24–40. https://doi.org/10.2307/2408514.

Li J, Deng H-W. Estimation of parameters of deleterious mutations in partial selfing or partial outcrossing populations and in nonequilibrium populations. Genetics. 2000:154(4):1893–1906. https://doi.org/10.1093/genetics/154.4.1893.

López MA, López-Fanjul C. Spontaneous mutation for a quantitative trait in *Drosophila melanogaster*. II. Distribution of mutant effects on the trait and fitness. Genet Res. 1993:61(2):117–126. https://doi.org/10.1017/S0016672300031220.

Mafessoni F, Lachmann M. Selective strolls: fixation and extinction in diploids are slower for weakly selected mutations than for neutral ones. Genetics. 2015:201(4):1581–1589. https://doi.org/10.1534/genetics.115.178160.

Manna F, Martin G, Lenormand T. Fitness landscapes: an alternative theory for the dominance of mutation. Genetics. 2011:189(3):923–937. https://doi.org/10.1534/genetics.111.132944.

Marad DA, Buskirk SW, Lang GI. Altered access to beneficial mutations slows adaptation and biases fixed mutations in diploids. Nat Ecol Evol. 2018:2(5):882–889. https://doi.org/10.1038/s41559-018-0503-9.

Mathieson I, McVean G. Estimating selection coefficients in spatially structured populations from time series data of allele frequencies. Genetics. 2013:193(3):973–984. https://doi.org/10.1534/genetics.112.147611.

Mendel J. Verhandlungen des naturforschenden vereines in Brünn. Abhandlungen. 1865:4:3–47.

Mitchell JA. Fitness effects of EMS-induced mutations on the X chromosome of *Drosophila melanogaster*. I. Viability effects and heterozygous fitness effects. Genetics. 1977:87(4):763–774. https://doi.org/10.1093/genetics/87.4.763.

Mukai T. The genetic structure of natural populations of *Drosophila melanogaster*. VI. Further studies on the optimum heterozygosity hypothesis. Genetics. 1969:61(2):479–495. https://doi.org/10.1093/genetics/61.2.479.

Mukai T, Chigusa SI, Mettler LE, Crow JF. Mutation rate and dominance of genes affecting viability in *Drosophila melanogaster*. Genetics. 1972:72(2):335–355. https://doi.org/10.1093/genetics/72.2.335.

Mukai T, Chigusa S, Yoshikawa I. The genetic structure of natural populations of *Drosophila melanogaster*. II. Overdominance of spontaneous mutant polygenes controlling viability in homozygous genetic background. Genetics. 1964:50(4):711–715. https://doi.org/10.1093/genetics/50.4.711.

Mukai T, Yamazaki T. The genetic structure of natural populations of *Drosophila melanogaster*. V. Coupling-repulsion effect of spontaneous mutant polygenes controlling viability. Genetics. 1968:59(4):513–535. https://doi.org/10.1093/genetics/59.4.513.

Muller HJ. Our load of mutations. Am J Hum Genet. 1950:2:111–176.

Muralidhar P, Veller C. Dominance shifts increase the likelihood of soft selective sweeps. Evolution. 2022:76(5):966–984. https://doi.org/10.1111/evo.14459.

Ohnishi O. Spontaneous and ethyl methanesulfonate-induced mutations controlling viability in *Drosophila melanogaster*. III. Heterozygous effect of polygenic mutations. Genetics. 1977:87(3):547–556. https://doi.org/10.1093/genetics/87.3.547.

Ohta T. Associative overdominance caused by linked detrimental mutations. Genet Res. 1971:18(3):277–286. https://doi.org/10.1017/S0016672300012684.

Orr HA. A test of Fisher's theory of dominance. Proc Natl Acad Sci USA. 1991:88:11413–11415. https://doi.org/10.1073/pnas.88.24.11413.

Orr HA, Betancourt AJ. Haldane's sieve and adaptation from the standing genetic variation. Genetics. 2001:157(2):875–884. https://doi.org/10.1093/genetics/157.2.875.

Peters AD, Halligan DL, Whitlock MC, Keightley PD. Dominance and overdominance of mildly deleterious induced mutations for fitness traits in *Caenorhabditis elegans*. Genetics. 2003:165(2):589–599. https://doi.org/10.1093/genetics/165.2.589.

Phadnis N, Fry JD. Widespread correlations between dominance and homozygous effects of mutations: implications for theories of dominance. Genetics. 2005:171(1):385–392. https://doi.org/10.1534/genetics.104.039016.

Ragsdale AP. Local fitness and epistatic effects lead to distinct patterns of linkage disequilibrium in protein-coding genes. Genetics. 2022:221(4):iyac097. https://doi.org/10.1093/genetics/iyac097.

Robinson JA, Kyriazis CC, Nigenda-Morales SF, Beichman AC, Rojas-Bracho L, Robertson KM, Fontaine MC, Wayne RK, Lohmueller KE, Taylor BL, et al. The critically endangered vaquita is not doomed to extinction by inbreeding depression. Science. 2022:376(6593):635–639. https://doi.org/10.1126/science.abm1742.

Robinson J, Kyriazis CC, Yuan SC, Lohmueller KE. Deleterious variation in natural populations and implications for conservation genetics. Annu Rev Anim Biosci. 2023:11(1):93–114. https://doi.org/10.1146/annurev-animal-080522-093311.

Ronfort J, Glemin S. Mating system, Haldane's sieve, and the domestication process. Evolution. 2013:67(5):1518–1526. https://doi.org/10.1111/evo.12025.

Savageau MA. Dominance according to metabolic control analysis: major achievement or house of cards? J Theor Biol. 1992:154(1):131–136. https://doi.org/10.1016/S0022-5193(05)80194-8.

Savageau MA, Sorribas A. Constraints among molecular and systemic properties: implications for physiological genetics. J Theor Biol. 1989:141(1):93–115. https://doi.org/10.1016/S0022-5193(89)80011-6.

Sawyer SA, Hartl DL. Population genetics of polymorphism and divergence. Genetics. 1992:132(4):1161–1176. https://doi.org/10.1093/genetics/132.4.1161.

Simmons MJ, Crow JF. Mutations affecting fitness in *Drosophila* populations. Annu Rev Genet. 1977:11(1):49–78. https://doi.org/10.1146/annurev.ge.11.120177.000405.

Simmons MJ, Sheldon EW, Crow JF. Heterozygous effects on fitness of EMS-treated chromosomes in *Drosophila melanogaster*. Genetics. 1978:88(3):575–590. https://doi.org/10.1093/genetics/88.3.575.

Steinmetz LM, Scharfe C, Deutschbauer AM, Mokranjac D, Herman ZS, Jones T, Chu AM, Giaever G, Prokisch H, Oefner PJ, et al. Systematic screen for human disease genes in yeast. Nat Genet. 2002:31(4):400–404. https://doi.org/10.1038/ng929.

Szafraniec K, Wloch DM, Sliwa P, Borts RH, Korona R. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast *Saccharomyces cerevisiae*. Genet Res. 2003:82(1):19–31. https://doi.org/10.1017/s001667230300630x.

Tataru P, Bataillon T. polyDFEv2.0: testing for invariance of the distribution of fitness effects within and across species. Bioinformatics. 2019:35(16):2868–2869. https://doi.org/10.1093/bioinformatics/bty1060.

Tataru P, Mollion M, Glémin S, Bataillon T. Inference of distribution of fitness effects and proportion of adaptive substitutions from polymorphism data. Genetics. 2017:207(3):1103–1119. https://doi.org/10.1534/genetics.117.300323.

Teshima KM, Przeworski M. Directional positive selection on an allele of arbitrary dominance. Genetics. 2006:172(1):713–718. https://doi.org/10.1534/genetics.105.044065.

Turner JRG. Adaptation and evolution in heliconius: a defense of neoDarwinism. Annu Rev Ecol Syst. 1981:12(1):99–121. https://doi.org/10.1146/annurev.es.12.110181.000531.

Vassilieva LL, Hook AM, Lynch M. The fitness effects of spontaneous mutations in *Caenorhabditis elegans*. Evolution. 2000:54(4):1234–1246. https://doi.org/10.1111/j.0014-3820.2000.tb00557.x.

Veeramah KR, Gutenkunst RN, Woerner AE, Watkins JC, Hammer MF. Evidence for increased levels of positive and negative selection on the X chromosome versus autosomes in humans. Mol Biol Evol. 2014:31(9):2267–2282. https://doi.org/10.1093/molbev/msu166.

Wallace B. The average effect of radiation-induced mutations on viability in *Drosophila melanogaster*. Evolution. 1958:12(4):532. https://doi.org/10.2307/2405964.

Whitlock MC, Ingvarsson PK, Hatfield T. Local drift load and the heterosis of interconnected populations. Heredity (Edinb). 2000:84(Pt 4):452–457. https://doi.org/10.1046/j.1365-2540.2000.00693.x.

Williamson S, Fledel-Alon A, Bustamante CD. Population genetics of polymorphism and divergence for diploid selection models with arbitrary dominance. Genetics. 2004:168(1):463–475. https://doi.org/10.1534/genetics.103.024745.

Williamson SH, Hernandez R, Fledel-Alon A, Zhu L, Nielsen R, Bustamante CD. Simultaneous inference of selection and population growth from patterns of variation in the human genome. Proc Natl Acad Sci. 2005:102(22):7882–7887. https://doi.org/10.1073/pnas.0502300102.

Wright S. Fisher's theory of dominance. Am Nat. 1929:63(686):274–279. https://doi.org/10.1086/280260.

Wright S. Physiological and evolutionary theories of dominance. Am Nat. 1934:68(714):24–53. https://doi.org/10.1086/280521.

Wright S. The distribution of gene frequencies under irreversible mutation. Proc Natl Acad Sci U S A. 1938:24(7):253–259. https://doi.org/10.1073/pnas.24.7.253.

Wright S. Modes of selection. Am Nat. 1956:90(850):5–24. https://doi.org/10.1086/281903.

Wright S, Dobzhansky T, Hovanitz W. Genetics of natural populations. VII. The allelism of lethals in the third chromosome of *Drosophila pseudoobscura*. Genetics. 1942:27(4):363–394. https://doi.org/10.1093/genetics/27.4.363.

Zhang LJ, Jing YP, Li XH, Li CW, Bourguet D, Wu G. Temperature-sensitive fitness cost of insecticide resistance in Chinese populations of the diamondback moth *Plutella xylostella*. Mol Ecol. 2015:24(7):1611–1627. https://doi.org/10.1111/mec.13133.

Zhang X, Kim B, Lohmueller KE, Huerta-Sánchez E. The impact of recessive deleterious variation on signals of adaptive introgression in human populations. Genetics. 2020:215(3):799–812. https://doi.org/10.1534/genetics.120.303081.

Zhao L, Charlesworth B. Resolving the conflict between associative overdominance and background selection. Genetics. 2016:203(3):1315–1334. https://doi.org/10.1534/genetics.116.188912.

Associate editor: Adam Eyre-Walker