Title:

The effect of allelic molecular interactions on phenotypic dominance

Authors:

Tushar Pal[1,2,3,4,5], Angel F. Cisneros[1,2,3,4], Christian R. Landry[1,2,3,4]

Affiliations:

[1]Département de biochimie, de microbiologie et de bio-informatique, Faculté des sciences et de génie, Université Laval, G1V 0A6, Québec, Canada
[2]Institut de biologie intégrative et des systèmes, Université Laval, G1V 0A6, Québec, Canada
[3]PROTEO, Le regroupement québécois de recherche sur la fonction, l'ingénierie et les applications des protéines, Université Laval, G1V 0A6, Québec, Canada
[4]Centre de recherche sur les données massives, Université Laval, G1V 0A6, Québec, Canada
[5]Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata, Mohanpur, Nadia, West Bengal, 741246, India

Correspondance:
Tel: +1 418 656 3954; E-mail: christian.landry@bio.ulaval.ca

Abstract

Mendelian genetics provides us with a framework for studying allelic dominance relationships at a locus at the phenotypic level. These dominance relationships result from different molecular factors, including the mapping of fitness onto molecular activity. For homomeric proteins, physical interactions between alleles could affect protein activity, providing a mechanism for dominance. Here, we examine the molecular underpinnings of dominance using a model in which fitness is determined by the activities of monomeric and dimeric proteins. Our model recapitulates previous observations that monomeric proteins can show phenotypic dominance due to the underlying fitness function in the absence of molecular dominance. In turn, the landscape of phenotypic dominance for the dimer system results from the interplay of the fitness function, molecular dominance, and the selective homo- or heteromerization of the two alleles. Our results show the complex relationships between molecular and phenotypic dominance and highlight the fundamental difference in dominance landscapes for monomeric and homomeric proteins.

Keywords

Phenotypic dominance / molecular dominance / mutational effects / fitness / gene duplication / protein interactions

Introduction

A major question in genetics is how alleles combine to produce phenotypes. Mendel's pioneering experiments on inheritance patterns showed that some phenotypes are dominant over others (Mendel 1865). While these early results provided a framework to estimate the relative frequencies of phenotypes, the underlying molecular bases of dominance remained elusive.

Heterozygous individuals harbor different alleles for the same gene. Alleles are phenotypically dominant if heterozygotes carrying that allele have the same phenotype as homozygotes of that allele. Since phenotypes determine the reproductive success (fitness) of individuals, phenotypic dominance is typically inferred from experimental (Greenberg and Crow 1960) or population data (Le Poul *et al.* 2014; Huber *et al.* 2018). In both cases, the fitness of mutant homozygous individuals is compared relative to a reference homozygote to estimate its selection coefficient (s), which is positive for deleterious genotypes and negative for beneficial genotypes (see equation 13 in Methods). Then, a dominance coefficient (h, also called heterozygous effect) is inferred based on the fitness of heterozygous individuals and the previously calculated selection coefficient (see equation 14 in Methods) (Greenberg and Crow 1960). Based on this definition, alleles are typically considered recessive if $0 \leq h < 0.5$ or dominant if $0.5 < h \leq 1$, with $h=0$ and $h=1$ indicating cases of complete recessiveness and complete dominance, respectively (Greenberg and Crow 1960; Billiard *et al.* 2021; Di and Lohmueller 2024).

The dominance relationships among alleles are of major importance for human diseases and evolution. In terms of human health, loss-of-function alleles for most genes tend to be recessive (Lek *et al.* 2016; Karczewski *et al.* 2020), but more than half of known disease-associated genes annotated in the Online Mendelian Inheritance in Man (OMIM) database (Hamosh *et al.* 2005) are linked with dominant conditions (Zschocke *et al.* 2023). Examples of the role of dominance in evolution include how strongly dominant alleles in neotropical butterflies were favored by selection because they enhanced mimetic wing patterns that prevent predation (Le Poul *et al.* 2014). Similarly, experimental evolution studies with yeast populations have shown that dominance relationships are allele-dependent, with some providing a higher fitness for heterozygotes (overdominance) and others for the mutant homozygote (partial dominance) (Fisher *et al.* 2021; Aggeli *et al.* 2022; Bautista *et al.* 2024). In most cases, the biochemical underlying mechanisms of dominance relationship remain unknown.

Different models have aimed to explain the phenotypic dominance of certain alleles. Since the contribution of a gene to phenotype is often linked to the function and abundance of its encoded protein, these models generally assume that fitness depends on the total activity provided by the two alleles (Wright 1934; Kacser and Burns 1981; Hurst and Randerson 2000; Huber *et al.* 2018; Billiard *et al.* 2021; Di and Lohmueller 2024). For example, if the encoded protein is an enzyme, its activity would refer to the concentration of the product of its catalyzed reaction. In these models, higher activity generally leads to higher fitness, although with diminishing returns,

that is, further increases in activity provide ever smaller gains in fitness. In fact, beyond a given point, very high protein activity and the energetic cost associated with excessive protein synthesis could become toxic for the cell (Dekel and Alon 2005; Cherry 2010; Bhattacharyya *et al.* 2016; Eguchi *et al.* 2018). Because of the various nature of protein activity, the fitness functions of different genes vary widely with respect to the abundance of their encoded proteins due to the trade-off between the benefit of higher activity and the cost of protein synthesis (Qi *et al.* 2013; Keren *et al.* 2016; Otto *et al.* 2024). Since the shape of these fitness functions is typically nonlinear, a loss-of-function mutation in one of the two alleles might halve the total protein activity but lead to a stronger or weaker decrease in fitness, depending on the shape of that function. Whether a mutant allele is phenotypically dominant or recessive with respect to the WT can simply be a result of these nonlinearities. Thus, models of dominance have to consider the underlying fitness function, which is likely specific to each protein in given conditions (Qi *et al.* 2013; Keren *et al.* 2016; Wu *et al.* 2022; Després *et al.* 2022; Otto *et al.* 2024).

Physical interactions between encoded proteins are another factor that might shape dominance effects of alleles. For example, roughly around 20-45% of proteins self-assemble into homomers (Schweke *et al.* 2024). In diploids, the two corresponding alleles of homomeric proteins usually assemble into a mix of homo- and heteromers (Walker *et al.* 1997; Pereira-Leal *et al.* 2007)unless specific mechanisms mediate allele-specific assembly (Badonyi and Marsh 2023). These heteromers then provide a mechanism by which a mutant allele can interfere with the function of a wild-type (WT) copy, providing an opportunity for dominance to occur at the molecular level as well. Well-known but extreme examples of molecular dominance include dominant negative alleles, which capture WT subunits in catalytically inactive complexes (Herskowitz 1987). Since these mutations often result in drastic losses of protein function, they have been associated with a wide variety of human diseases (Bergendahl *et al.* 2019; Backwell and Marsh 2022; Meyer-Schuman *et al.* 2023; Zschocke *et al.* 2023). As a result, molecular interactions contribute to the emergence of dominant phenotypes.

Current models (Wright 1934; Kacser and Burns 1981; Hurst and Randerson 2000; Huber *et al.* 2018) show how dominance could simply emerge from nonlinear fitness functions. While these frameworks allow modeling dominance in independent monomeric proteins, they do not account for potential interactions between copies. Here, we ask how molecular dominance between alleles could translate into phenotypic dominance. We build a model that introduces molecular terms representing the reduction in the specific activity of a mutant homodimer relative to the WT (r) and the dominance of such effects on the activity of the heterodimer (g). These parameters are analogous to the parameters that describe the effect of mutations on phenotype: the selection (s) and dominance (h) coefficients. By comparing the cases for monomeric and homodimeric proteins, we validate that phenotypic dominance can be observed even when there is no molecular dominance depending on the shape of the underlying fitness function. We show that molecular dominance introduces a fundamental difference in the fitness landscapes of homodimeric versus monomeric proteins. Since the effect of molecular dominance depends on the assembly of heterodimers, we explore how mutations that affect the equilibrium between homo- and heterodimers contribute to phenotypic dominance. Finally, we show how this

framework could be applied to infer the parameters associated with molecular dominance from experimental data on fitness functions and the fitness of homo- and heterozygotes.

Results

A model of the effect of molecular dominance on phenotypic dominance

We built a model assuming a diploid organism, with a gene of interest, “a”, with three genotypes: 1) WT homozygous (with both alleles being the wild type), 2) mutant homozygous (with both alleles having the same mutation), and 3) heterozygous (with one allele being WT and the other being a mutant allele). From here onwards, we will refer to WT alleles as “a” and mutant alleles as “b”. After gene transcription and translation, proteins “A” and “B” are synthesized from alleles “a” and “b”, respectively. Here, we will focus on modelling molecular and phenotypic dominance for the specific cases of these proteins exerting their activity as either monomers or dimers. In the case of monomers, the two proteins are independent from one another, while in the case of dimers, the two proteins form a mix of homo- and heterodimers. Phenotypic dominance emerges from differences in fitness between the two homozygotes and the heterozygote, while molecular dominance refers to how physical interactions with mutant protein copies affect the function of WT proteins.

We define fitness as a function of protein activity at equilibrium. In the case of the monomeric proteins (**Fig 1a**), the equilibrium features only protein “A” or “B” in the case of homozygotes, or a combination of both in the case of heterozygotes. The concentrations of each protein at equilibrium are calculated as a ratio of synthesis and decay rates (see Methods). The total activity at equilibrium (φ) is then calculated based on the concentration of each protein and its specific activity. For the WT homozygous genotype, we assume a single WT monomer “A” has a reference specific activity of 1 (a_A=1). In the case of the mutant homozygous genotype, the mutation can have either a neutral, positive, or negative effect on the specific activity of the protein. We define a term “r” as the reduction in specific activity due to the mutation. Hence, a single mutant monomer “B” has an activity (a_B=1-r). If r is positive, the specific activity decreases, and vice-versa. The fitness (ω) of each genotype is then calculated as a function of its total activity ($f(\varphi)$), where “f” is the fitness function for that protein. We calculate a selection coefficient (s_m) as the difference in fitness between the WT homozygote and the mutant homozygote, that is, s_m is positive for deleterious mutations. Finally, the fitness of the heterozygote is defined as ω=1-$s_m h_m$, where h_m is the phenotypic dominance coefficient. The value of h_m can thus be inferred from heterozygote fitness and approaches 0 if the mutation is recessive or 1 if it is dominant.

Now, let us consider the case where the proteins form dimers (**Fig 1b**). In the WT homozygous and mutant homozygous systems, proteins A and B would form the homodimers AA and BB, respectively. Assuming that each dimer has two active sites with additive activity, we set the specific activity of the homodimers as twice that of their corresponding monomers. Thus, a_{AA}=2, a_{BB}=2(1-r). In the case of a heterozygous individual, the two proteins A and B can form homodimers AA and BB, but also the heterodimer AB. Assuming that the binding affinities of homo- and heterodimers are identical and all synthesized protein copies assemble into dimers, the resulting mix of dimers would have the following relative occupancies: 25% of homodimer

AA, 25% of homodimer BB, and 50% of heterodimer AB (Hochberg *et al.* 2018). Given the interaction between the WT and mutant proteins, we introduce a term for molecular dominance (g), which determines the specific activity of the heterodimer based on the effect the mutant proteins have on each other upon interacting (a_{AB}=2(1-gr)). If g is close to zero, the heterodimer has a specific activity similar to that of the WT homodimer AA, making the mutation recessive at the molecular level. Similarly, if g is close to 1, the mutation is dominant at the molecular level. While we assume that dimerization is required for full activity, we allow any free homomers at equilibrium to have residual activities, referred to as a_A and a_B. Next, we use the framework from Cisneros et al. (2024) to solve for the equilibrium concentrations c_i, which now depend not only on the synthesis and decay rates, but also on the binding affinities of the different dimers (see Methods). Finally, we calculate the total protein activity of a given genotype and the corresponding fitness using a fitness function, as described above. Based on the differences in fitness, we estimate the phenotypic coefficients s_d and h_d for dimers.

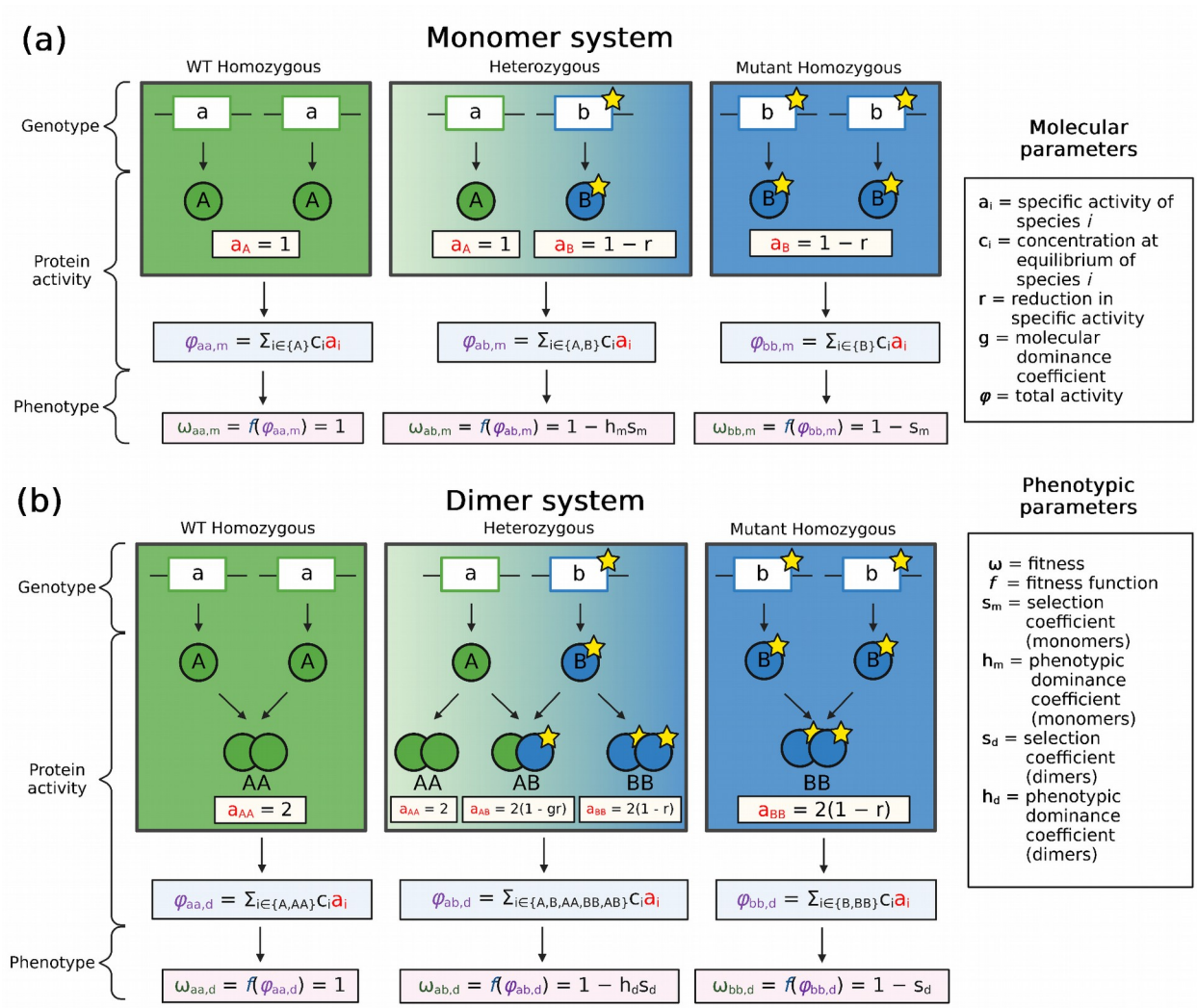


Figure 1. Effect of molecular dominance on phenotypic dominance. a-b) Dominance model at the molecular and phenotypic levels for genes encoding monomers (**a**) or dimers (**b**). In both cases, the model considers three possible genotypes with two copies of a focal gene: WT homozygote, heterozygote (WT + Mutant), and mutant homozygote. The specific activity of the WT allele is set to 1, whereas the mutant is set to 1-r. Molecular dominance results from

the physical interaction between the two proteins produced from the two alleles, thus modifying the specific activity of the heterodimer using the molecular dominance coefficient (*g*). The total protein activity is calculated as the weighted sum of the specific activity values of each functional unit (monomers in panel **a**, dimers in panel **b**), with their respective concentrations as weights. Fitness is then inferred based on a fitness function for the estimated values of total activity. Phenotypic dominance is observed in the two heterozygotes, with the value of the phenotypic dominance coefficient (*h*) being inferred based on the fitness values of the two respective homozygotes and the heterozygote. [Figure created with BioRender.com]

The fitness landscape of phenotypic dominance with and without molecular dominance

We asked how changes in molecular parameters influence phenotypic dominance within this framework. Namely, we can examine how the reduction in specific activity (*r*), molecular dominance coefficient (*g*), equilibrium concentrations of the different monomers and dimers (c_i), and the form of the fitness function (*f*) interact to impact phenotypic dominance. In nature, these parameters can be modified by mutations. We study these effects one at a time, by keeping some of them fixed and varying others to see how the total activity and fitness change.

We focus first on how different homozygous mutations can modify fitness. To do so, we assume fitness is given by a lognormal function (**Fig 2a**) where the maximum fitness (normalized to 1) peaks at a given optimal activity value (also normalized to 1). The optimal activity value is calculated for a WT homozygous genotype, which is the total concentration of the protein complex present at equilibrium (see **Methods**). Fitness falls off on either side of this optimum, but at a slower rate for increases in activity beyond the optimum value, following other models (Dekel and Alon 2005; Cherry 2010) (see Methods). We isolate the effect of different values of *r* and *g* on the resulting phenotypes by keeping synthesis rates, folding energies, binding affinities, and decay rates constant. In our framework, the mutant homozygous models for both monomers and dimers are identical, since in both cases the total protein activity depends solely on *r*. Accordingly, since all concentrations stay constant, the total activity in the mutant homozygous genotype varies linearly with the change in *r* (**Fig 2b**). In turn, these total activity values can be used to estimate the fitness of genotype (**Fig 2c**) using the fitness function.

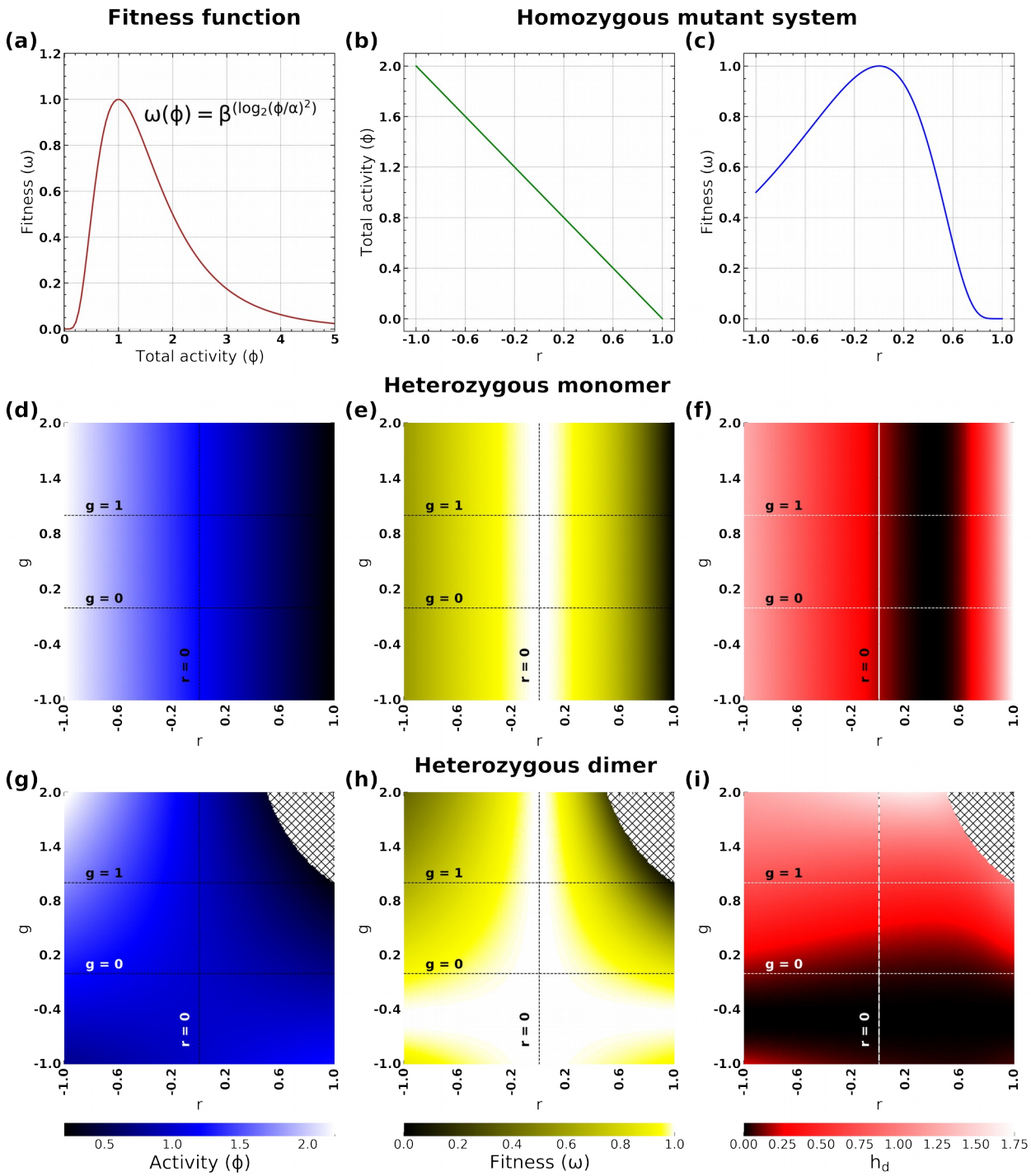


Figure 2. Molecular dominance introduces non-linearities in fitness and phenotypic dominance. a) Fitness function used to determine fitness from the protein activity of each genotype. Fitness peaks at a normalized activity value of 1 and decays as activity goes away from the optimum. **b-c)** Effect of different values of *r* on total activity (b) and fitness (c) for the mutant homozygous genotypes of both monomers and dimers. **d-f)** Effect of the interaction between *g* and *r* on the total activity (d), fitness (e), and phenotypic dominance (f) in a heterozygous system of monomers. **g-i)** Effect of the interaction between *g* and *r* on the total activity (g), fitness (h), and phenotypic dominance (i) in a heterozygous system of dimers. The hatched region indicates disallowed values that would result in negative values of total activity.

We then focused on how the molecular parameters combine to affect total activity, fitness, and phenotypic dominance for monomeric and homodimeric proteins. To do so, we considered the heterozygous genotypes and the associated molecular systems, and modified the molecular dominance coefficient and effects on activity while keeping all other parameters constant. Since there are no physical interactions between the two proteins in the monomer system (**Fig 2d-f**), the total activity, fitness, and phenotypic dominance are completely dependent on activity (see equations 15 and 16 in Methods). Importantly, our model allows for phenotypic dominance (**Fig 2f**) even in the absence of molecular dominance because the presence of the mutant allele affects fitness, in line with other models (Kacser and Burns 1981; Hurst and Randerson 2000). Conversely, the physical interactions between alleles in the dimer system (**Fig 2g-i**) produce a nonlinear landscape in terms of *r* and *g*. These landscapes reflect the asymmetry of the fitness function (**Fig 2a**) and how high fitness can result from different combinations of *r* and *g* in the monomer and dimer systems (**Fig 2e,h**). The magnitude of the phenotypic dominance coefficient for the dimer system is a function of both *r* and *g* (see equation 17 in Methods) and is maximized at high values of molecular dominance but moderate changes in specific activity (**Fig 2i**). Overall, phenotypic dominance can emerge either as a result of the sensitivity of the fitness function around the WT or high values of molecular dominance.

We asked if there are specific values of molecular dominance in the dimer system that could lead to phenotypically recessive mutations (h_d=0), regardless of the value of *r*. An intuitive expectation could be that mutations are phenotypically recessive when they are recessive at the molecular level (*g*=0). However, we showed previously that phenotypic dominance can also be produced by the shape of the fitness function in a system of monomers (**Fig 2f**). Indeed, when *g*=0 there is some variation in fitness (**Supp. Video 1**) because the heterodimer has the same activity as the WT, but the total protein activity of the genotype is still given by a mixture of homo- and heterodimers. Thus, the total activity of the genotype is not the same as that for a WT homozygote unless the activity of the mutant homodimer is the same as that of the WT homodimer (*r*=0). Instead, we identified that the effect of mutations is masked when *g*=-0.5 (**Supp. Video 1**). Due to the unequal concentrations of homo- and heterodimers, in this specific case the total contributions of the heterodimer and the mutant homodimer cancel out for any value of *r*, and the mutant allele becomes recessive (h_d=0). On the other hand, when *g*=1.5, mutations become phenotypically dominant regardless of the value of *r*, since the fitness of the heterozygote becomes identical to that of the mutant homozygote (h_d=1) (**Supp. Video 1**).

Homodimerization changes the fitness landscape as a result of molecular dominance

The presence of molecular dominance allowed by the dimerization between the product of two alleles (**Fig 2g-i**) fundamentally changes the landscapes of fitness and phenotypic dominance with respect to the reference system where the two alleles encode independent monomers (**Fig 2d-f**). We quantify the extent of this homodimerization effect as the differences in total activity, fitness, and phenotypic dominance coefficients between the monomer and dimer systems (**Fig 3**).

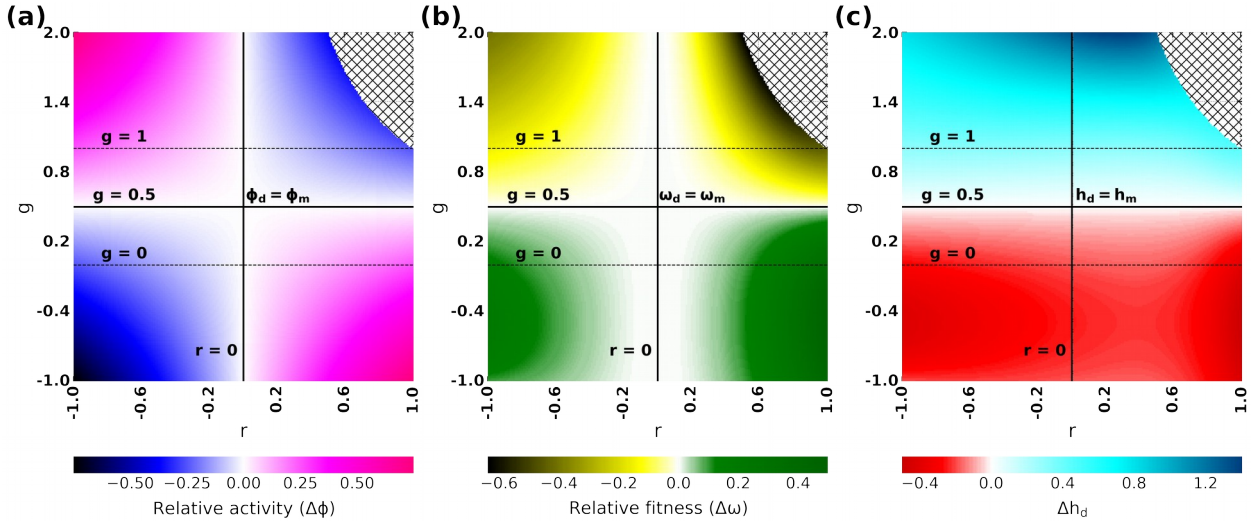


Figure 3. Homodimerization allows for molecular dominance and reshapes the landscape phenotypic dominance. a-c) Differences in the relative activity (a), relative fitness (b), and phenotypic dominance (c) observed between the monomer and dimer heterozygous systems as a function of *g* and *r*. Differences in each parameter are calculated by subtracting the values in the monomer system from those in the dimer system.

The magnitude of the homodimerization effect depends on both *r* and *g*. We derived an equation to capture these effects (see equation 18 in Methods). Using this expression, we asked if the homodimerization effect could disappear under specific combinations of *r* and *g*. A trivial example would be when *r*=0, since the mutant allele behaves exactly the same as the WT. Interestingly, when *g*=0.5, the homodimerization effect disappears, regardless of the value of *r* (**Supp. Video 1, Fig 3**). The effect disappears because the specific activity of the heterodimer becomes the average of those of the WT and mutant homodimers. The resulting total activity then becomes a weighted average identical to that of the monomer system. Note that when *g*=0, the heterodimer has the same specific activity as the WT homodimer, so the total activity is closer to that of the WT homozygote.

Shifts in equilibrium concentrations of homo- and heterodimers modify dominance effects

Amino acid substitutions can affect both the specific activities of proteins as well as their binding affinities (Smiley *et al.* 2002; Jankauskaitė *et al.* 2018; Starr *et al.* 2020; Després *et al.* 2022, 2024; Cisneros *et al.* 2023). As the binding affinities of a mixture of homo- and heterodimers are modified by mutations, so do their relative occupancies (Cisneros *et al.* 2024; Cortez-Romero *et al.* 2025). Thus, we examined how the landscapes of activity, fitness, and phenotypic dominance change as we relax the assumption of identical binding affinities of the dimers (**Fig 4**). We focus on the case of the heterozygous dimer because the monomer system is insensitive to molecular dominance. We vary the relative binding affinity of the heterodimer while keeping the binding affinities of the two homodimers constant to yield 0%, 10%, 25%, 50%, 75%, and 90% heteromers at binding equilibrium (see **Methods**). Since these landscapes also depend on *r* and *g*, we show the landscapes for *r*=-0.5 (**Fig 4a-c**) and *r*=0.5 (**Fig 4d-f**) with varying *g* as examples. We further study these landscapes for a wider range of relative binding affinities of

the dimer (or percentage of heteromers) with varying *g* and fixing more values of *r* (**Fig S1**), as well as the landscapes for a series of fixed *g* values with varying *r* (**Fig S2**).

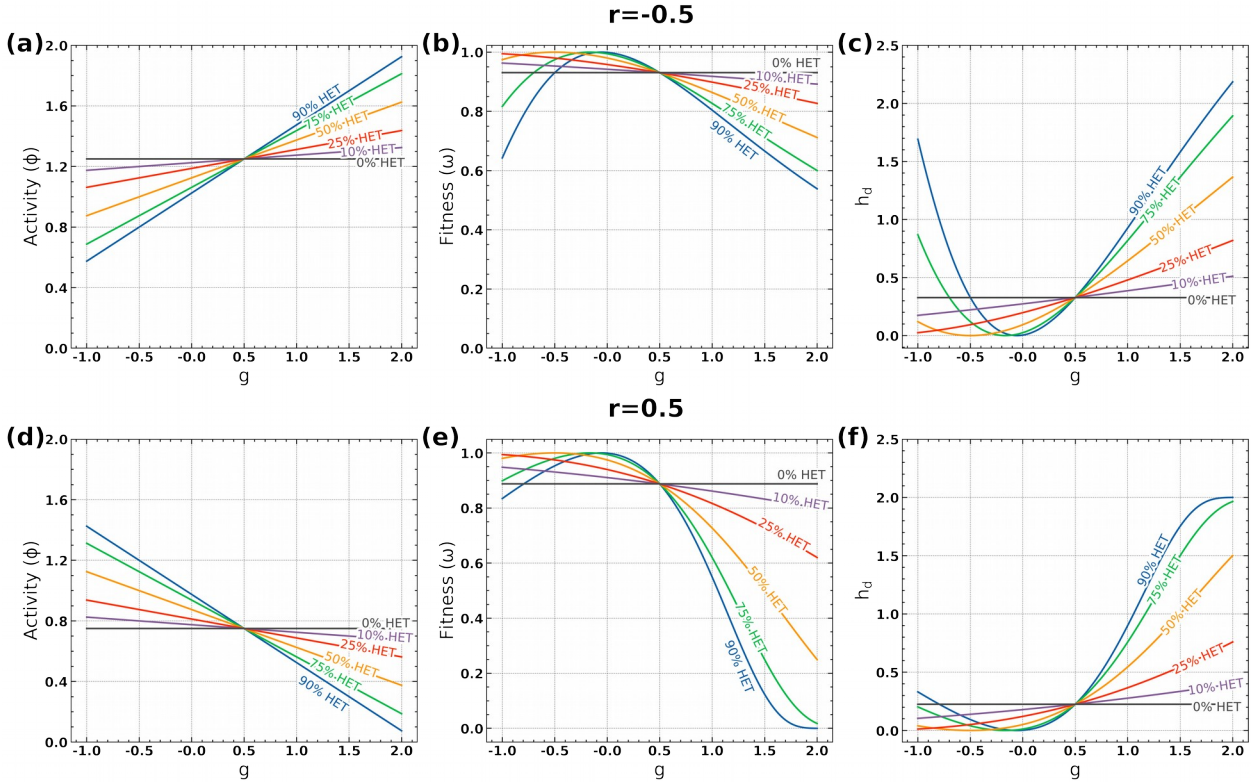


Figure 4. Shifts toward higher homo- or heterospecificity modify the landscape of dominance effects. The effect of mutations on activity has been fixed at *r*=-0.5 for **a-c**, and at *r*=0.5 for **d-f**. Landscapes for activity (**a,d**), fitness (**b,e**), and phenotypic dominance coefficient (**c,f**) are obtained for varying *g* and the percentage of heteromers at equilibrium. In all panels, heteromerization occupancy varied by keeping the binding affinity of both homodimers constant and varying the binding affinity of the heterodimer. Line labels were added using the matplotlib-label-lines Python package (doi.org/10.5281/zenodo.7428070).

The resulting landscapes illustrate how selective homo- or heteromerization modulates the effect of molecular dominance and phenotypic dominance. The effect of molecular dominance on activity, fitness, and phenotypic dominance is potentiated with higher heteromer occupancies and gradually disappears as the percentage of heterodimers approaches zero. For example, when the mutational effect increases specific activity of both the mutant homodimer and the heterodimer (*r*=-0.5, **Fig 4a, Fig S1a**), activity is maximized with higher values of molecular dominance. However, since both decreases and increases in activity have a fitness cost, the values of *r* and *g* required for optimum fitness depend on the fraction of heteromers (**Fig 4b**). Similarly, phenotypic dominance increases with higher magnitudes of molecular dominance, although the increase is more pronounced for higher concentrations of heteromers (**Fig 4c**). Conversely, decreases in specific activity (*r*=0.5) invert the landscape of activity (**Fig. 4d, Fig. S1**), but only rescale the landscapes of fitness and phenotypic dominance (**Fig 4e-f, Fig. S1**). Thus, shifts in heteromer occupancy lead to significant changes in fitness and phenotypic

dominance for the same underlying molecular parameters, often potentiating the effects of small values of *r* and *g*.

The interaction between heteromer occupancy and molecular dominance drives phenotypic dominance. As noted above, the homodimerization effect disappears when g=0.5 (**Supplementary Video 1**). Indeed, in all panels the curves for different percentages of heteromers at equilibrium intersect when g=0.5 (**Fig 4**), showing that all phenotypic parameters become the same regardless of heteromer occupancy. On the other hand, our previous analyses suggested that when g=-0.5, molecular dominance masks the effect of the mutation (h_d=0). This value is affected by the relative occupancy of heterodimers, as shown graphically in **Fig 4c,f** and analytically in equation 19. Overall, our results highlight the effects of interactions between molecular dominance and heteromer occupancy, as well as showing that specific values of molecular dominance can mask the effect of shifts in heteromer occupancy.

Applying the model to infer molecular parameters from phenotypic data and fitness functions

A major challenge pertaining to genetics and biochemistry is identifying how dominance emerges from biochemical interactions between alleles. Such connections are linked to human diseases (Bergendahl *et al.* 2019; Backwell and Marsh 2022; Zschocke *et al.* 2023) and influence the evolution of natural (Le Poul *et al.* 2014) and experimental populations (Fisher *et al.* 2021; Aggeli *et al.* 2022; Bautista *et al.* 2024). It is therefore key to be able to tell whether dominance at the molecular level is needed to explain the dominance seen at the phenotypic level.

With high-throughput experiments, it is now possible to measure the phenotypic impact of thousands of mutations in different genotypes at the phenotypic level (Fowler and Fields 2014), from which we would like to infer molecular mechanisms of action of mutations. While we can get direct measurements of the fitness of homo- and heterozygous individuals, it is much harder to deconvolute the relative activities of a mix of homo- and heteromers. Thus, we reasoned that our model could be used with fitness measurements and a fitness function to infer molecular parameters.

We tested the application of our model to a simulated distribution of mutational effects. These distributions are typically measured experimentally by generating a library of mutants and measuring their fitness in bulk competition experiments (Fowler and Fields 2014). Since distributions of mutational effects are typically bimodal (Després *et al.* 2022; Cisneros *et al.* 2023), we sampled selection coefficients from two separate normal distributions (see Methods), for a total of 2000 values. These values were paired to dominance coefficients sampled independently from a normal distribution to complete our simulated dataset (**Fig 5a**). Assuming that the WT homozygote has a fitness of 1 for a relative activity of 1, these parameters are enough to calculate the total fitness of the heterozygote and the mutant homozygote. We can then see how these values of phenotypes could be generated by *g* and *r* given a fitness function using its corresponding inverse function. We tested the effects of three fitness functions with

different shapes (**Fig. 5b**) since the relationship between activity and fitness varies widely (Keren *et al.* 2016) and can lead to phenotypic dominance in the absence of molecular dominance (**Fig. 2f**). We restricted our analysis to one-to-one fitness functions so that for each value of total activity there is only one unambiguous corresponding value of fitness and vice versa.

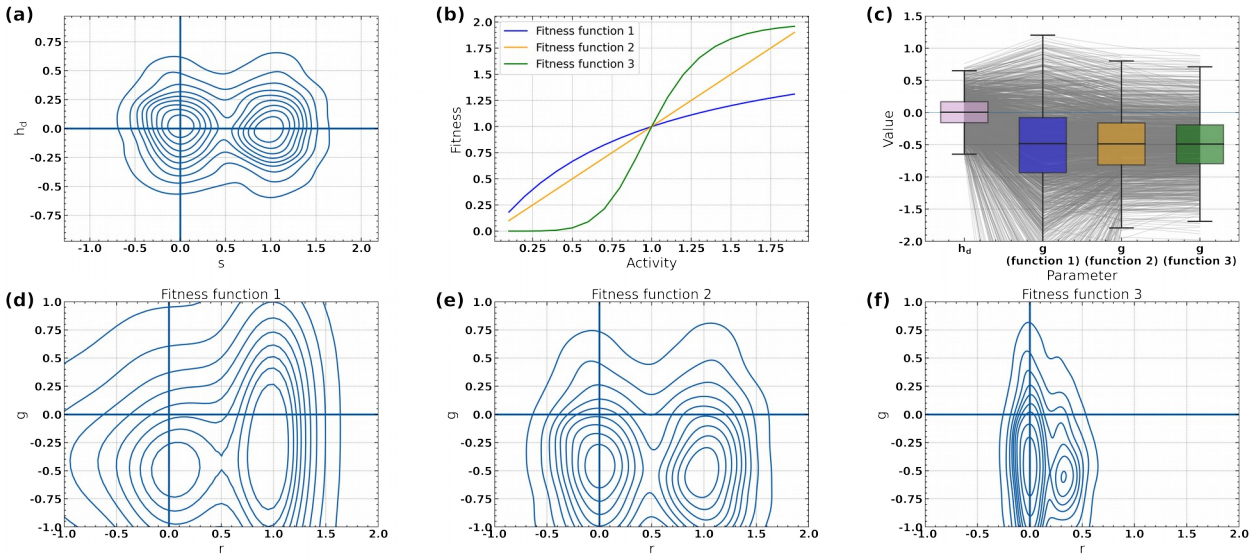


Figure 5. Relationship between molecular and phenotypic dominance under different fitness functions. a) Distribution of simulated dataset of phenotypic parameters (s, h_d). **b)** Different fitness functions tested. Functions 1 and 3 follow Hill equations with different parameters, while function 2 is linear. **c)** Distributions of molecular and phenotypic dominance coefficients inferred for each fitness function. **d-f)** Distributions of molecular parameters inferred for the pairs of phenotypic parameters in panel A and the inverse of fitness functions 1 (d), 2 (e), and 3 (f).

Our analysis showed that the same sample of pairs of phenotypic parameters implies different molecular parameters depending on the underlying fitness function (**Fig 5c-f**). The distributions of molecular dominance generally have higher magnitudes than the distribution of phenotypic dominance (**Fig 5c**), consistent with **Fig 2i**. Interestingly, the sensitivity of the fitness function to changes in activity around the WT influenced the range of values estimated for the molecular parameters. When fitness decays slowly around the WT (fitness function 1), the most deleterious mutations (s~1) are driven by mutations that strongly disrupt activity (r~1) with a wide range of variation of molecular dominance (**Fig 5d**). In turn, since fitness functions 2 and 3 are more sensitive around the WT (**Fig 5b**), the same changes in fitness are a result of smaller disruptions on specific activity and smaller magnitudes of molecular dominance (**Fig 5e-f**). Overall, our results show that the relationship between molecular and phenotypic parameters depends on the underlying fitness function and highlight a method by which molecular parameters can be inferred from experimental data.

Discussion

In humans and other diploid species, alleles combine to produce phenotypes. Multiple mechanisms have been proposed separately for the dominance relationships between alleles, including the underlying fitness function (Wright 1934; Kacser and Burns 1981) and molecular interactions between encoded proteins (Herskowitz 1987; Bergendahl *et al.* 2019; Backwell and Marsh 2022; Zschocke *et al.* 2023). Here, we developed a comprehensive model that disentangles the contributions of each of those effects to phenotypic dominance by joining them into a single, quantitative framework. Using this model, we show that the landscape of phenotypic dominance is different for monomeric and homomeric proteins because of the possibility of molecular dominance. While we show that phenotypic dominance can result from molecular dominance, it can also be observed in the absence of molecular dominance given a non-linear fitness function. Finally, we illustrate how the relationship between molecular and phenotypic dominance is modified by different fitness functions.

The equilibrium between homo- and heteromers is a key factor for the influence of molecular dominance on phenotypic dominance. Different studies have shown that homo- or heterospecificity can be achieved with very few mutations (Ashenberg *et al.* 2011; Hochberg *et al.* 2018; Emlaw *et al.* 2021; Cortez-Romero *et al.* 2025) or in the presence of barriers preventing heteromerization (Mallik *et al.* 2025). Indeed, proteins that assemble co-translationally into homomers tend to not form heteromers (Mallik *et al.* 2025), and they are thus less likely to harbor dominant negative mutations (Badonyi and Marsh 2023). Similarly, increasing the abundance of one of two proteins would increase the relative concentration of the corresponding homodimers (Cisneros *et al.* 2024). As a result, the relative contribution of heterodimers to fitness would decrease along with the observed dominance of the mutant allele, analogous to our observations for high homospecificity (**Fig 4**). However, regulatory mechanisms have been shown to modulate insufficient or excessive protein abundance for 10-30% of yeast genes (DeLuna *et al.* 2010; Springer *et al.* 2010; Hose *et al.* 2015; Ascencio *et al.* 2021), which could mask or potentiate dominant alleles. Further research could focus on deconvoluting these factors and evaluating the extent to which they contribute to dominance.

Our model focuses on dominance relationships in a diploid system, but it provides insights on other similar systems. For example, it could be applied to paralog interference since duplications of genes encoding homomeric proteins introduce heteromers of paralogs (Pereira-Leal *et al.* 2007; Kaltenegger and Ober 2015; Marchant *et al.* 2019; Mallik and Tawfik 2020). Analogous to the molecular dominance explained here, heteromers of paralogs represent a mechanism by which mutations in one gene can affect the function of its sister copy (Bridgham *et al.* 2008; Kaltenegger and Ober 2015). The distribution of the magnitudes of such effects might vary from one protein to another, but many of these effects could potentially be highly deleterious. The application of this model to similar datasets could illuminate the molecular parameters driving the observed phenotypes.

Molecular interactions introduce non-linearities in the landscape of phenotypic dominance.

While in general higher magnitudes of molecular dominance result in stronger phenotypic dominance, we also found that specific negative values of molecular dominance can mask mutational effects. While it is necessary to establish if the shape of the underlying fitness function for a given protein is sufficient to explain phenotypic dominance, the recessivity of alleles could be derived from molecular interactions between encoded proteins. Negative values of molecular dominance ($g<0$) imply that the mutant allele has opposing effects on the specific activities of the mutant homodimer and the heterodimer, which could lead to negative phenotypic dominance ($h<0$). These cases could correspond to hybrid heterosis (Boeven *et al.* 2020) or mutually complementary deleterious mutations (Walker *et al.* 1997; Després *et al.* 2024), although the extent to which negative molecular dominance coefficients occur in nature and the underlying molecular determinants are unknown.

We illustrate an application of our model to infer molecular parameters for simulated fitness data. Current methodologies provide simpler avenues to obtain high-throughput measurements of cellular growth rates than for specific protein activities. Since our model allows establishing a relationship between fitness and molecular activity, it could help produce high-throughput maps of effects of mutations on molecular activity between alleles of a locus based on fitness data and a small subset of validation experiments. Ultimately, our work provides a framework to study the determinants of the magnitude of molecular dominance and how their effects vary across proteins, allowing us to better understand how Mendelian genetics emerge from cellular biochemistry.

Methods

Calculation of concentrations of different species at equilibrium

We used our previously derived equations (Cisneros *et al.* 2024) to estimate the concentrations of different molecular species at equilibrium (**Fig 1**). The molecular species observed at equilibrium in the monomer system are only monomers A and B, whereas the dimer system contains a mixture of monomers (A and B) and dimers (AA, AB, BB). Briefly, these equilibrium concentrations (c_i) depend on the synthesis (s) and decay rates (d), folding energies (ΔG_{fold}), and dimerization binding affinities (ΔG_{bind}). Folding energies are used to estimate the fraction of proteins that attain a folded state (Sailer and Harms 2017), as shown in equation 1:

$$s_{folded,i} = s_i \left(\frac{1}{1+e^{\Delta G_{fold,i}}} \right) \qquad (1)$$

where $s_{folded,i}$ is the synthesis rate of properly folded monomers of protein i, s_i is the overall synthesis rate of monomers of protein i, and $\Delta G_{fold,i}$ is the folding energy of protein i.

In the homozygous monomer system, this synthesis rate is multiplied by two to represent the two alleles producing the same protein. For the monomer systems (both homozygous and heterozygous), since the proteins do not interact with each other, the concentration at equilibrium for a species is then simply given by the ratio of the modified synthesis rate of the corresponding monomer and its decay rate, as shown in equation 2:

$$c_i = \frac{s_{folded,i}}{d_i} \qquad (2)$$

where $s_{folded,i}$ is the synthesis rate of properly folded monomers of protein i and d_i is the decay rate of properly folded monomers of protein i.

In turn, for the dimer system, folded subunits assemble into homo- and heterodimers based on their binding energies, which can be converted into association constants using equation 3:

$$k_i = e^{\frac{-\Delta G}{RT}} \qquad (3)$$

where k_i is the association constant for dimer i, $\Delta G_{bind,i}$ is the binding energy of dimer i, R is the gas constant (1.987×10^{-3} cal K^{-1} mol^{-1}) and T is the temperature at 298 K.

These association constants then determine the concentrations at equilibrium based on equations 4-8 (Cisneros *et al.* 2024):

$$
\begin{aligned}
0 &= \\
&2 d_{AA} k_{AA} \left(4 \frac{d_{AA} k_{AA} d_{BB} k_{BB}}{d_{AB} k_{AB}} - d_{AB} k_{AB} \right) c_A^4 \\
&+ \\
&\left(d_A \left(8 \frac{d_{AA} k_{AA} d_{BB} k_{BB}}{d_{AB} k_{AB}} - d_{AB} k_{AB} \right) - 2 d_B d_{AA} k_{AA} \right) c_A^3 \\
&+ \\
&\left(\frac{2 d_{BB} k_{BB} d_A^2}{d_{AB} k_{AB}} + s_A \left(d_{AB} k_{AB} - 8 \frac{d_{AA} k_{AA} d_{BB} k_{BB}}{d_{AB} k_{AB}} \right) - d_A d_B - d_{AB} k_{AB} s_B \right) c_A^2 \\
&+ \\
&s_A \left(d_B - 4 d_A \frac{d_{BB} k_{BB}}{d_{AB} k_{AB}} \right) c_A \\
&+ \\
&2 s_A^2 \frac{d_{BB} k_{BB}}{d_{AB} k_{AB}}
\end{aligned}
\tag{4}
$$

$$
c_B = \frac{\left(s_{folded,i} / c_A \right) - 2 d_{AA} k_{AA} c_A - d_A}{d_{AB} k_{AB}} \tag{5}
$$

$$
c_{AA} = k_{AA} c_A^2 \tag{6}
$$

$$
c_{BB} = k_{BB} c_B^2 \tag{7}
$$

$$
c_{AB} = k_{AB} c_A c_B \tag{8}
$$

where c_i is the concentration of molecular species i, $s_{folded,i}$ is the synthesis rate of properly folded monomers of protein i, d_i is the decay rate of properly folded monomers of protein i, and k_i is the association constant for dimer i.

For all calculations in **Fig 2-4**, we set synthesis rates of A and B to 100 h^{-1} and decay rates to 1.3 h^{-1} based on estimates for the half-life of proteins and cell division rates in yeast (Hausser *et al.* 2019). The resulting total activity, 76.41 h^{-1}, was then taken as the normalization factor to use relative activity with the fitness function shown in **Fig 2a**.

Phenotypic dominance as a function of molecular parameters

Definitions of molecular and phenotypic parameters

Our model considers molecular parameters for the reduction of specific activity due to a

mutation and molecular dominance influencing the activity of a heterodimer. Thus, specific activities are given by equations 9-10:

$$a_{BB}=1-r \qquad (9)$$

$$a_{AB}=1-gr \qquad (10)$$

where a_{BB} is the specific activity of a mutant homodimer and a_{AB} is the specific activity of a heterodimer. In the monomer system, the activity of a mutant monomer (a_B) is defined using the above expression for a_{BB}.

With the calculated equilibrium concentrations (c_i) and the specific activities (a_i) of the different species, we calculate the total activity of each system (ϕ) using equation 11:

$$\phi=\sum_i c_i a_i \qquad (11)$$

The fitness for each genotype (ω_i) is then calculated based on the activity produced using the fitness function (f), as shown in equation 12 (see Fitness function section for more details):

$$\omega_i=f(\phi) \qquad (12)$$

Selection (s) and phenotypic dominance coefficients (h) can then be estimated by comparing the fitness of the WT homozygote (ω_{aa}=1) to that of the mutant homozygote (ω_{bb}) and the heterozygote (ω_{ab}), as shown in equations 13-14 (Greenberg and Crow 1960):

$$\omega_{bb}=1-s \qquad (13)$$

$$\omega_{ab}=1-hs \qquad (14)$$

These calculations of selection and phenotypic dominance coefficients apply to both the monomer (s_m, h_m) and dimer (s_d, h_d) systems.

Analytical expressions for the relationship between phenotypic and molecular dominance

If we assume identical synthesis, decay rates and folding energies (**Fig2, Fig3**), we can derive analytical expressions for phenotypic dominance as a function of molecular dominance, reduction in specific activity due to the mutation, and equilibrium concentration.

WT homozygote, monomer system: Let c be the equilibrium concentration of the monomer A given by equations 1 and 2. The total activity ($\varphi_{aa,m}$) would then correspond to c, following equation 11 and the definition that the specific activity of the WT is 1. This represents the optimal activity of the WT homozygote, which we use as the normalization factor for all other systems. This is then used as input for the fitness function, which in return yields a unit fitness

for the WT genotype. Thus, we have *f(1)* = 1.

Mutant homozygote, monomer system: Following the above assumptions, this system reaches the same concentration of monomers as the WT homozygote. Following equation 9, the specific activity of mutant monomers is 1-*r*, which yields a total activity of 1-*r*. Applying equation 13 and the fitness function yields equation 15, which directly expresses the selection coefficient as a function of the reduction in activity:

$$\omega_{bb}=1-s=f(1-r)$$

$$s=1-f(1-r) \qquad (15)$$

Heterozygous monomer system: Since now the two gene copies are different alleles, the concentration of each monomer will be *c*/2. Applying equation 11, the total protein activity becomes:

$$\phi=\sum_i c_i a_i=\frac{c}{2}+\frac{c(1-r)}{2}=c\left(1-\frac{r}{2}\right)$$

After normalizing by *c*, fitness becomes:

$$\omega_{ab}=f(1-\frac{r}{2})=1-h_m s_m$$

Since s_m is given by equation 15, the phenotypic dominance coefficient can be expressed as shown in equation 16:

$$h_m=\frac{1-f\left(1-\frac{r}{2}\right)}{s_m}$$

$$h_m=\frac{1-f(1-\frac{r}{2})}{1-f(1-r)} \qquad (16)$$

Homozygous dimer systems: Applying our above assumptions and assuming all protein copies assemble into dimers, the concentration of homodimers equals *c*/2. Since we allow dimers to have twice the specific activity of monomers to use the same normalization constant, we obtain that the total activity equals *c*. Thus, the homozygous dimer systems behave the same way as the monomeric systems: the fitness of the WT homozygote equals 1, and the selection coefficient of the mutant homozygote (s_d) is given by equation 15.

Heterozygous dimer system: After applying our assumptions, the system is composed of *c*/2 total dimers at equilibrium. Since this is a 1:2:1 mixture of homo- and heterodimers, their concentrations are:

$$c_{AA}=\frac{c}{8}, c_{AB}=\frac{c}{4}, c_{BB}=\frac{c}{8}$$

Substituting the specific activities from equations 9 and 10 into equation 11, the total protein activity is given by:

$$\phi=\sum_i c_i a_i=\frac{2c}{8}+\frac{c(2-2gr)}{4}+\frac{c(2-2r)}{8}$$

$$\phi=c\left(1-\frac{gr}{2}-\frac{r}{4}\right)$$

After normalizing by *c*, fitness becomes:

$$\omega_{ab}=f\left(1-\frac{gr}{2}-\frac{r}{4}\right)=1-h_d s_d$$

Since s_d is given by equation 15, the phenotypic dominance coefficient can be expressed as shown in equation 17:

$$h_d=\frac{1-f\left(1-\frac{gr}{2}-\frac{r}{4}\right)}{s_d}$$

$$h_d=\frac{1-f\left(1-\frac{gr}{2}-\frac{r}{4}\right)}{1-f(1-r)} \qquad (17)$$

Homodimerization effect

In **Fig 3**, we calculate the difference in total activity, fitness, and phenotypic dominance coefficient between the heterozygous dimer and monomer systems with identical synthesis and decay rates, given identical reductions in activity and molecular dominance values. Equation 18 shows an analytical expression for the magnitude of the homodimerization effect ($\Delta\phi$):

$$\Delta\phi=\phi_{ab,dimer}-\phi_{ab,monomer}=\left(1-\frac{gr}{2}-\frac{r}{4}\right)-\left(1-\frac{r}{2}\right)$$

$$\Delta\phi=\frac{r}{4}-\frac{gr}{2}=\frac{r}{2}\left(\frac{1}{2}-g\right)$$

$$\Delta\phi = \frac{r}{2}\left(\frac{1}{2} - g\right) \qquad (18)$$

We set $\Delta\phi$ = 0 in the above equation to identify values of *r* and *g* that eliminate the homodimerization effect. We found two solutions: *r*=0 and *g*=0.5 (**Fig 4**), For *r*=0, the mutation has no effect on the activity of the protein, and thus we always obtain the same total activity regardless of what system we choose. For *g*=0.5, the heterodimer activity falls right in the middle of the WT and mutant homodimers, which results in a system equivalent to the heterozygous monomer, with half WT and half mutant species (**Supplementary Video 1**).

Relationship between heterospecificity and masking of mutant alleles

A value of *g* that results in the masking of a mutant allele must provide the same total activity and fitness as the WT, which gives:

$$\phi_{ab} = \phi_{aa} = 1$$

$$\omega_{ab} = \omega_{aa} = 1$$

As shown in equation 11, the total activity of the heterozygous system is given by the sum of the contributions of all homo- and heterodimers. Assuming that all subunits assemble into dimers and that the two concentrations of homodimers are identical, the total activity of the heterozygous system can be written as shown in equation 19:

$$\phi_{ab} = 1 = \left(\frac{1-c_{AB}}{2}\right) + c_{AB}(1-gr) + \left(\frac{1-c_{AB}}{2}\right)(1-r)$$

$$1 = \frac{1}{2} - \frac{c_{AB}}{2} + c_{AB} - grc_{AB} + \frac{1}{2} - \frac{c_{AB}}{2} - \frac{r}{2} + \frac{rc_{AB}}{2}$$

$$1 = 1 - \frac{r}{2} + \frac{rc_{AB}}{2} - grc_{AB}$$

$$0 = r\left(\frac{c_{AB}}{2} - \frac{1}{2} - gc_{AB}\right)$$

$$\frac{1}{2} - \frac{c_{AB}}{2} = -gc_{AB}$$

$$g = \frac{c_{AB} - 1}{2c_{AB}} \qquad (19)$$

which indicates the value of *g* required to mask a mutant allele given the relative occupancy of heterodimers.

Fitness functions

We use a lognormal fitness function (equation 20) for all calculations in **Fig 2-4**, with the form:

$$\omega(\phi)=(\beta)^{(\log_2(\phi/\alpha)^2)} \qquad (20)$$

where α represents the optimal activity, and *β* is a scaling factor (0<*β*<1) that determines how rapidly fitness decays with increasing or decreasing activity. When total activity is half or twice the value of α, fitness equals *β* times the optimal fitness value. For all calculations in **Fig 2-4**, we set α=1 and *β*=0.5 (**Fig 2a**). This fitness function has a peak at *φ*=α, where it reaches an optimal fitness value of 1. The fitness function decays faster with decreasing activity, analogous to the fitness functions observed due to cost of excessive expression (Dekel and Alon 2005; Cherry 2010).

In **Fig 5**, we sought to illustrate how three different fitness functions affect the relationship between the molecular and phenotypic parameters of our model. Fitness function 2 was defined as a linear function with a slope of 1 and a y-intercept of 0. Fitness functions 1 and 3 were defined using Hill equations of the form shown in equation 21:

$$f(\phi)=\frac{a\phi^b}{c^b+\phi^b} \qquad (21)$$

where *ϕ* indicates total activity, *a* indicates the horizontal asymptote for maximum fitness as *ϕ*→∞, *b* indicates the Hill coefficient, and *c* indicates the activity value at which fitness is half of *a*. For fitness function 1, parameters were set as follows: *a*=2, *b*=1, *c*=1. For fitness function 3, parameters were defined as: *a*=2, *b*=6, *c*=1. The parameters of the three fitness functions were selected so that a WT activity value of 1 would have a fitness of 1.

Application of the model to simulated data

We simulated a distribution of mutational effects. Since such distributions are typically bimodal (Després *et al.* 2022; Cisneros *et al.* 2023), we sampled 1000 values of selection coefficients from a normal distribution of WT-like effects (mean=0, standard deviation=0.25) and 1000 from a normal distribution of loss-of-function effects (mean=1, standard deviation=0.25). These selection coefficients were then paired with 2000 phenotypic dominance coefficients sampled from a normal distribution (mean=0, standard deviation=0.25) (**Fig 5a**). These phenotypic dominance parameters were used to estimate the fitness of mutant homozygotes and heterozygotes using equations 13 and 14. In turn, fitness values were used with inverse of fitness functions 1-3 from the previous section to determine the corresponding total activities, which could then be related to the underlying molecular parameters following equations 9-11.

Attributions

TP, AFC, and CRL designed research; TP, AFC, and CRL developed the theoretical model; TP & AFC analyzed data. TP, AFC, and CRL wrote the paper. CRL supervised the project.

Acknowledgments

We thank the Landry lab members for their discussions during the project. We thank Simon Aubé, Pavithra Venkataraman, Philippe Després, and Soham Dibyachintan for their comments on the manuscript.

Funding

This work was funded by a Natural Sciences and Engineering Research Council of Canada (NSERC) grant to CRL RGPIN-2020-04844 and by a seed grant from PROTEO. CRL holds the Canada Research Chair in Cellular Systems and Synthetic Biology. TP was supported by a Mitacs Globalink Research Internship fellowship. AFC was supported by a PROTEO graduate fellowship, the Ministère de l'enseignement supérieur du Québec, the Agencia mexicana para la cooperación y el desarrollo internacional.

Conflict of interest

The authors have no conflict of interest to declare.

Data availability

All scripts used in this work are available at:
https://github.com/Landrylab/Dominance_model_2025.git

Supplementary figures

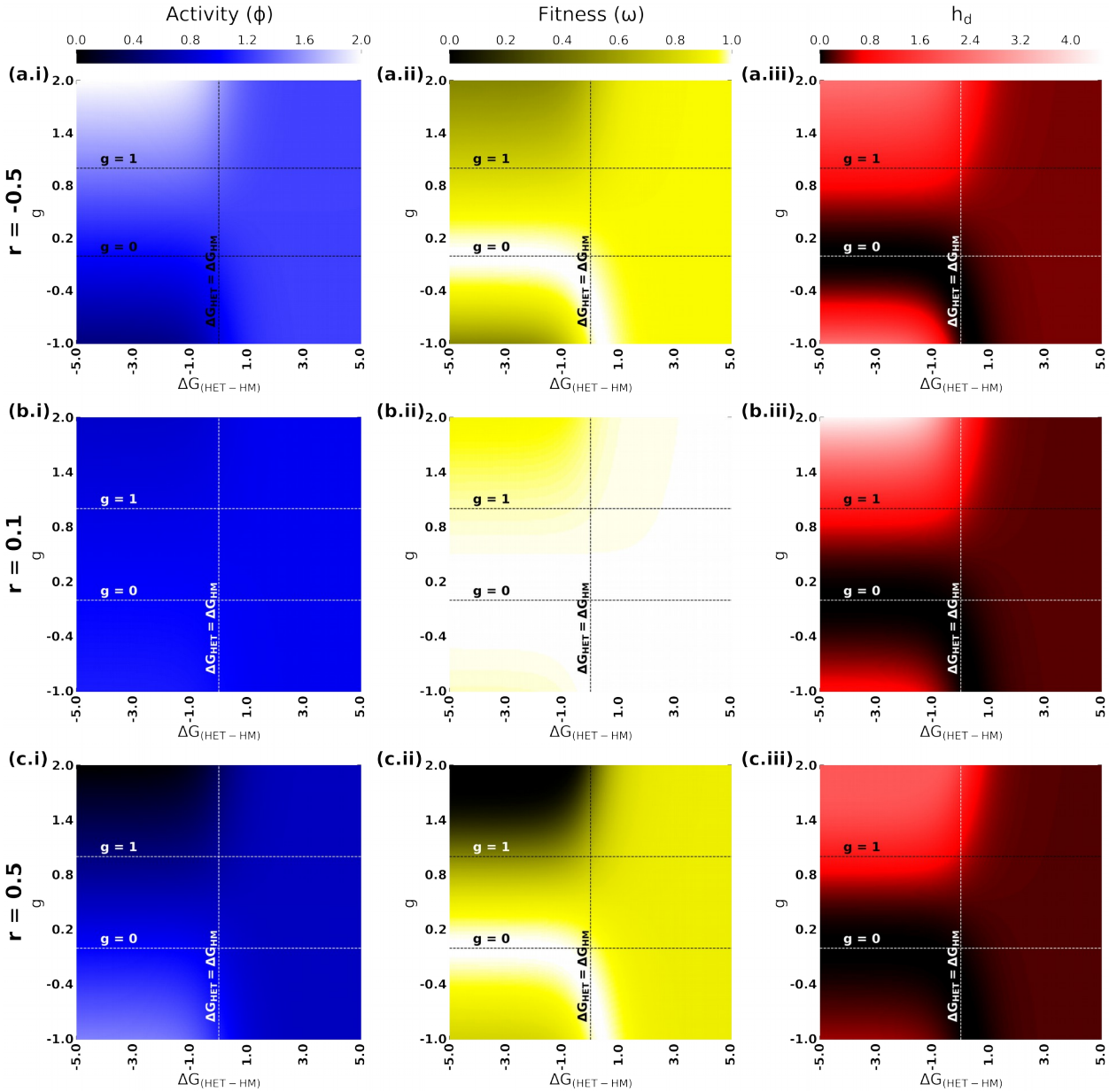


Figure S1. Changes in binding affinity modify the landscape of dominance effects (constant *r*). a-c) Landscapes of total activity (a.i, b.i, c.i), fitness (a.ii, b.ii, c.ii), and the phenotypic dominance coefficient (a.iii, b.iii, c.iii) for varying molecular dominance and heteromerization occupancies in the model with a heterozygous dimer system. The effect of mutations on activity (*r*) is kept constant for all panels in the same column: *r*=-0.5 for panels a.i-a.iii, *r*=0.1 for panels b.i-b.iii, and *r*=0.5 for panels c.i-c.iii. In all panels, heteromerization occupancy varied by keeping the binding affinity of both homodimers constant and changing the binding affinity of the heterodimer.

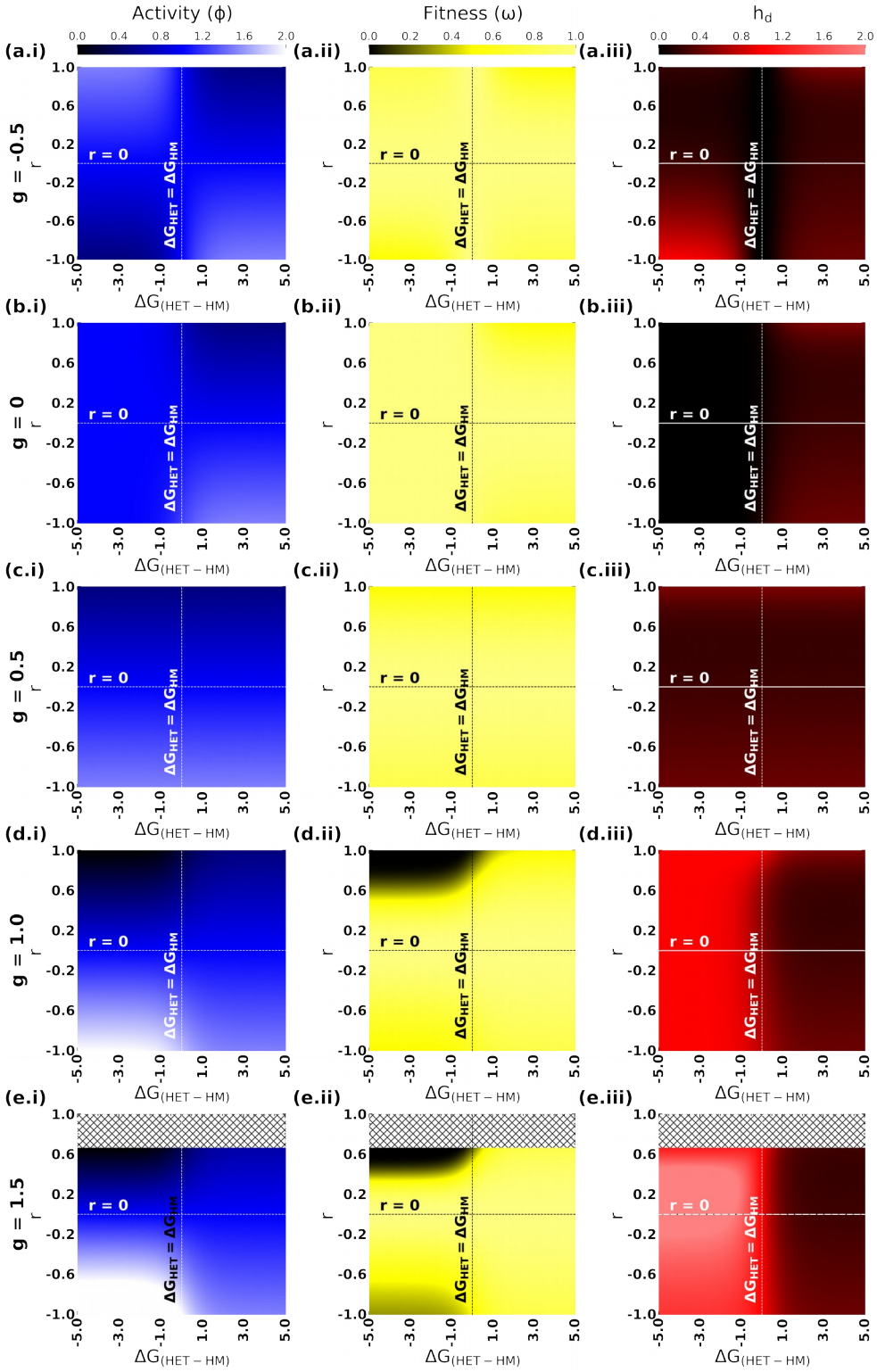
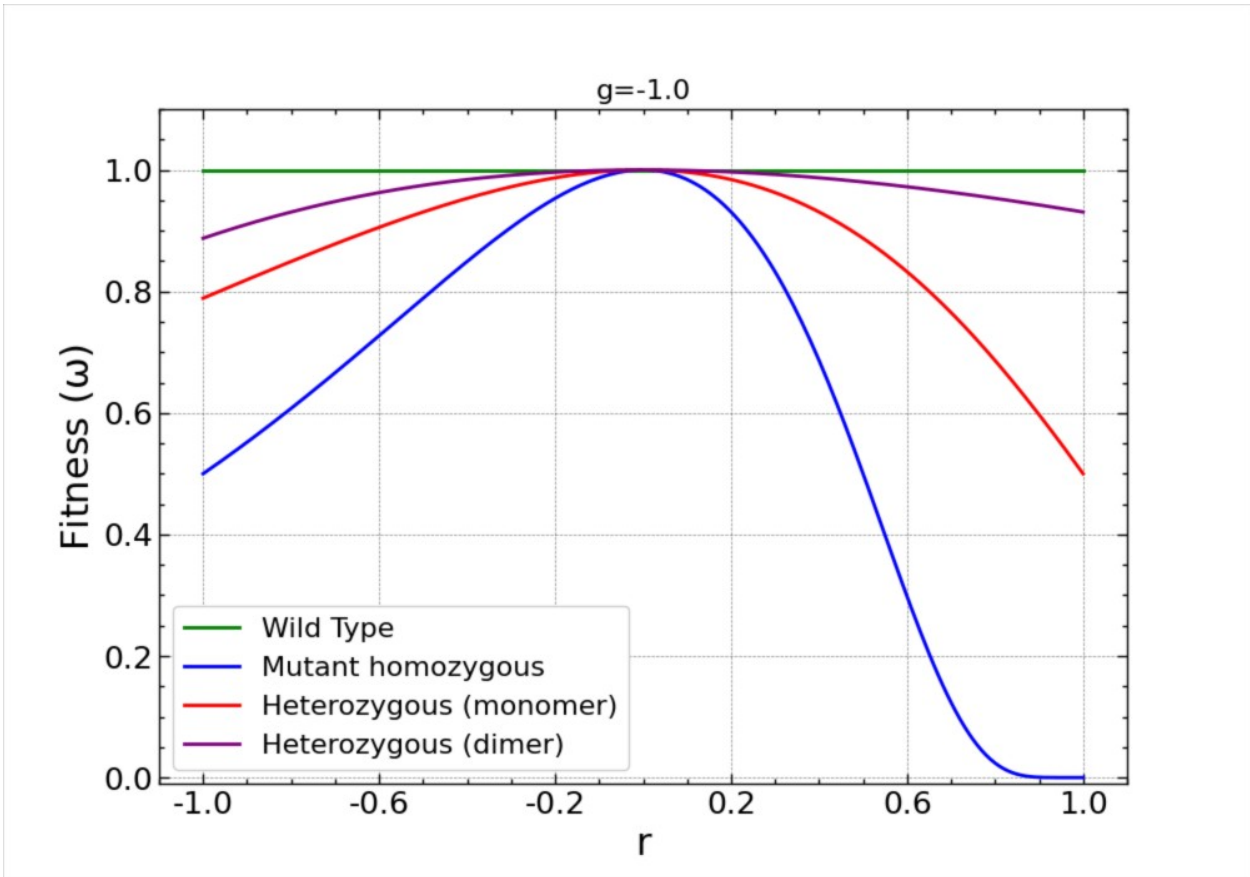


Figure S2. Changes in binding affinity modify the landscape of dominance effects (constant *g*). a-e) Landscapes of total activity (a.i, b.i, c.i, d.i, e.i), fitness (a.ii, b.ii, c.ii, d.ii, e.ii), and the phenotypic dominance coefficient (a.iii, b.iii, c.iii, d.iii, e.iii) for varying reductions in specific activity and heteromerization occupancies in the

model with a heterozygous dimer system. Molecular dominance (*g*) is kept constant for all panels in the same column: *g*=-0.5 for panels a.i-a.iii, *g*=0 for panels b.i-b.iii, *g*=0.5 for panels c.i-c.iii, *g*=1 for panels d.i-d.iii, and *g*=1.5 for panels e.i-e.iii. In all panels, heteromerization occupancy varied by keeping the binding affinity of both homodimers constant and changing the binding affinity of the heterodimer. The hatched region indicates disallowed values that would result in negative values of total activity.



Supplementary Video 1. Comparison of the heterozygous dimer system under different values of *g* and *r* to the other systems. While the fitness of the homozygous mutant dimer (blue) and the heterozygous monomer (red) systems are all determined solely by the reduction in specific activity (*r*), the heterozygous dimer system (purple) also depends on molecular dominance (*g*), which is varied from -1 to 2. The animation is paused briefly at particular values of *g* for which the phenotypic behaviour of the heterozygous dimer system is of interest: at ***g*=0** (complete molecular recessiveness); at ***g*=1** (complete molecular dominance); at ***g*=-0.5**, where it coincides with the WT phenotype, thus completely masking the effect of the mutant allele; at ***g*=0.5**, it coincides with the heterozygous monomer phenotype, indicating the elimination of the homodimerization effect; at ***g*=1.5**, it coincides with the mutant homozygous system, as the heterodimer's effect on φ overcompensates and nullifies that of the WT homodimer.

References

Aggeli D., D. A. Marad, X. Liu, S. W. Buskirk, S. F. Levy, *et al.*, 2022 Overdominant and partially dominant mutations drive clonal adaptation in diploid Saccharomyces cerevisiae. Genetics 221: iyac061.

Ascencio D., G. Diss, I. Gagnon-Arsenault, A. K. Dubé, A. DeLuna, *et al.*, 2021 Expression attenuation as a mechanism of robustness against gene duplication. Proc. Natl. Acad. Sci. U. S. A. 118: e2014345118.

Ashenberg O., K. Rozen-Gagnon, M. T. Laub, and A. E. Keating, 2011 Determinants of Homodimerization Specificity in Histidine Kinases. J. Mol. Biol. 413: 222–235.

Backwell L., and J. A. Marsh, 2022 Diverse molecular mechanisms underlying pathogenic protein mutations: Beyond the loss-of-function paradigm. Annu. Rev. Genomics Hum. Genet. 23: 475–498.

Badonyi M., and J. A. Marsh, 2023 Buffering of genetic dominance by allele-specific protein complex assembly. Sci Adv 9: eadf9845.

Bautista C., I. Gagnon-Arsenault, M. Utrobina, A. Fijarczyk, D. P. Bendixsen, *et al.*, 2024 Hybrid adaptation is hampered by Haldane's sieve. Nat. Commun. 15: 10319.

Bergendahl L. T., L. Gerasimavicius, J. Miles, L. Macdonald, J. N. Wells, *et al.*, 2019 The role of protein complexes in human genetic disease. Protein Sci. 28: 1400–1411.

Bhattacharyya S., S. Bershtein, J. Yan, T. Argun, A. I. Gilson, *et al.*, 2016 Transient protein-protein interactions perturb E. coli metabolome and cause gene dosage toxicity. Elife 5: e20309.

Billiard S., V. Castric, and V. Llaurens, 2021 The integrative biology of genetic dominance. Biol. Rev. Camb. Philos. Soc. 96: 2925–2942.

Boeven P. H. G., Y. Zhao, P. Thorwarth, F. Liu, H. P. Maurer, *et al.*, 2020 Negative dominance and dominance-by-dominance epistatic effects reduce grain-yield heterosis in wide crosses in wheat. Sci. Adv. 6: eaay4897.

Bridgham J. T., J. E. Brown, A. Rodríguez-Marí, J. M. Catchen, and J. W. Thornton, 2008 Evolution of a new function by degenerative mutation in cephalochordate steroid receptors. PLoS Genet. 4: e1000191.

Cherry J. L., 2010 Expression level, evolutionary rate, and the cost of expression. Genome Biol. Evol. 2: 757–769.

Cisneros A. F., I. Gagnon-Arsenault, A. K. Dubé, P. C. Després, P. Kumar, *et al.*, 2023 Epistasis between promoter activity and coding mutations shapes gene evolvability. Sci. Adv. 9: eadd9109.

Cisneros A. F., L. Nielly-Thibault, S. Mallik, E. D. Levy, and C. R. Landry, 2024 Mutational biases favor complexity increases in protein interaction networks after gene duplication. Mol. Syst. Biol. 20: 549–572.

Cortez-Romero C. R., J. Lyu, A. S. Pillai, A. Langanowsky, and J. W. Thornton, 2025 Symmetry facilitated the evolution of heterospecificity and high-order stoichiometry in vertebrate hemoglobin. Proc. Natl. Acad. Sci. U. S. A. 122: e2414756122.

Dekel E., and U. Alon, 2005 Optimality and evolutionary tuning of the expression level of a protein. Nature 436: 588–592.

DeLuna A., M. Springer, M. W. Kirschner, and R. Kishony, 2010 Need-based up-regulation of

protein levels in response to deletion of their duplicate genes. PLoS Biol. 8: e1000347.

Després P. C., A. F. Cisneros, E. M. M. Alexander, R. Sonigara, C. Gagné-Thivierge, *et al.*, 2022 Asymmetrical dose responses shape the evolutionary trade-off between antifungal resistance and nutrient use. Nat Ecol Evol 6: 1501–1515.

Després P. C., A. K. Dubé, M.-È. Picard, J. Grenier, R. Shi, *et al.*, 2024 Compensatory mutations potentiate constructive neutral evolution by gene duplication. Science 385: 770–775.

Di C., and K. E. Lohmueller, 2024 Revisiting dominance in population genetics. Genome Biol. Evol. 16: evae147.

Eguchi Y., K. Makanae, T. Hasunuma, Y. Ishibashi, K. Kito, *et al.*, 2018 Estimating the protein burden limit of yeast cells by measuring the expression limits of glycolytic proteins. Elife 7: e34595.

Emlaw J. R., C. J. G. Tessier, G. D. McCluskey, M. S. McNulty, Y. Sheikh, *et al.*, 2021 A single historical substitution drives an increase in acetylcholine receptor complexity. Proc. Natl. Acad. Sci. U. S. A. 118: e2018731118.

Fisher K. J., R. C. Vignogna, and G. I. Lang, 2021 Overdominant mutations restrict adaptive loss of heterozygosity at linked loci. Genome Biol. Evol. 13: evab181.

Fowler D. M., and S. Fields, 2014 Deep mutational scanning: a new style of protein science. Nat. Methods 11: 801–807.

Greenberg R., and J. F. Crow, 1960 A comparison of the effect of lethal and detrimental chromosomes from Drosophila populations. Genetics 45: 1153–1168.

Hamosh A., A. F. Scott, J. S. Amberger, C. A. Bocchini, and V. A. McKusick, 2005 Online

Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res. 33: D514–7.

Hausser J., A. Mayo, L. Keren, and U. Alon, 2019 Central dogma rates and the trade-off between precision and economy in gene expression. Nat. Commun. 10: 68.

Herskowitz I., 1987 Functional inactivation of genes by dominant negative mutations. Nature 329: 219–222.

Hochberg G. K. A., D. A. Shepherd, E. G. Marklund, I. Santhanagoplan, M. T. Degiacomi, *et al.*, 2018 Structural principles that enable oligomeric small heat-shock protein paralogs to evolve distinct functions. Science 359: 930–935.

Hose J., C. M. Yong, M. Sardi, Z. Wang, M. A. Newton, *et al.*, 2015 Dosage compensation can buffer copy-number variation in wild yeast. Elife 4: e05462.

Huber C. D., A. Durvasula, A. M. Hancock, and K. E. Lohmueller, 2018 Gene expression drives the evolution of dominance. Nat. Commun. 9: 2750.

Hurst L. D., and J. P. Randerson, 2000 Dosage, deletions and dominance: simple models of the evolution of gene expression. J. Theor. Biol. 205: 641–647.

Jankauskaitė J., B. Jiménez-García, J. Dapkūnas, J. Fernández-Recio, and I. H. Moal, 2018 SKEMPI 2.0: an updated benchmark of changes in protein–protein binding energy, kinetics and thermodynamics upon mutation. Bioinformatics 35: 462–469.

Kacser H., and J. A. Burns, 1981 The molecular basis of dominance. Genetics 97: 639–666.

Kaltenegger E., and D. Ober, 2015 Paralogue Interference Affects the Dynamics after Gene Duplication. Trends Plant Sci. 20: 814–821.

Karczewski K. J., L. C. Francioli, G. Tiao, B. B. Cummings, J. Alföldi, *et al.*, 2020 The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581: 434–443.

Keren L., J. Hausser, M. Lotan-Pompan, I. V. Slutskin, H. Alisar, *et al.*, 2016 Massively parallel interrogation of the effects of gene expression levels on fitness. Cell 166: 1282–1294.

Lek M., Exome Aggregation Consortium, K. J. Karczewski, E. V. Minikel, K. E. Samocha, *et al.*, 2016 Analysis of protein-coding genetic variation in 60,706 humans. Nature 536: 285–291.

Le Poul Y., A. Whibley, M. Chouteau, F. Prunier, V. Llaurens, *et al.*, 2014 Evolution of dominance mechanisms at a butterfly mimicry supergene. Nat. Commun. 5: 5644.

Mallik S., and D. S. Tawfik, 2020 Determining the interaction status and evolutionary fate of duplicated homomeric proteins. PLoS Comput. Biol. 16: e1008145.

Mallik S., A. F. Cisneros, C. R. Landry, and E. D. Levy, 2025 Co-translational assembly promotes functional diversification of paralogous proteins. bioRxiv 2025.01.22.634331.

Marchant A., A. F. Cisneros, A. K. Dubé, I. Gagnon-Arsenault, D. Ascencio, *et al.*, 2019 The role of structural pleiotropy and regulatory evolution in the retention of heteromers of paralogs. eLife 8: e46754.

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

Meyer-Schuman R., S. Marte, T. J. Smith, S. M. E. Feely, M. Kennerson, *et al.*, 2023 A humanized yeast model reveals dominant-negative properties of neuropathy-associated alanyl-tRNA synthetase mutations. Hum. Mol. Genet. 32: 2177–2191.

Otto R. M., A. Turska-Nowak, P. M. Brown, and K. A. Reynolds, 2024 A continuous epistasis model for predicting growth rate given combinatorial variation in gene expression and

environment. Cell Syst. 15: 134–148.e7.

Pereira-Leal J. B., E. D. Levy, C. Kamp, and S. A. Teichmann, 2007 Evolution of protein complexes by duplication of homomeric interactions. Genome Biol. 8: R51.

Qi L. S., M. H. Larson, L. A. Gilbert, J. A. Doudna, J. S. Weissman, *et al.*, 2013 Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152: 1173–1183.

Sailer Z. R., and M. J. Harms, 2017 Molecular ensembles make evolution unpredictable. Proc. Natl. Acad. Sci. U. S. A. 114: 11938–11943.

Schweke H., M. Pacesa, T. Levin, C. A. Goverde, P. Kumar, *et al.*, 2024 An atlas of protein homo-oligomerization across domains of life. Cell 187: 999–1010.

Smiley R. D., L. G. Stinnett, A. M. Saxton, and E. E. Howell, 2002 Breaking symmetry: mutations engineered into R67 dihydrofolate reductase, a D2 symmetric homotetramer possessing a single active site pore. Biochemistry 41: 15664–15675.

Springer M., J. S. Weissman, and M. W. Kirschner, 2010 A general lack of compensation for gene dosage in yeast. Mol. Syst. Biol. 6: 368.

Starr T. N., A. J. Greaney, S. K. Hilton, D. Ellis, K. H. D. Crawford, *et al.*, 2020 Deep mutational scanning of SARS-CoV-2 receptor binding domain reveals constraints on folding and ACE2 binding. Cell 182: 1295–1310.e20.

Walker D. C., J. Christodoulou, H. J. Craig, L. R. Simard, L. Ploder, *et al.*, 1997 Intragenic complementation at the human argininosuccinate lyase locus. Identification of the major complementing alleles. J. Biol. Chem. 272: 6777–6783.

Wright S., 1934 Physiological and evolutionary theories of dominance. Am. Nat. 68: 24–53.

Wu Z., X. Cai, X. Zhang, Y. Liu, G.-B. Tian, *et al.*, 2022 Expression level is a major modifier of the fitness landscape of a protein coding gene. Nat Ecol Evol 6: 103–115.

Zschocke J., P. H. Byers, and A. O. M. Wilkie, 2023 Mendelian inheritance revisited: dominance and recessiveness in medical genetics. Nat. Rev. Genet. 24: 442–463.