A dominance-assimilated liability model for complex fitness traits

Shobbir Hussain

Department of Life Sciences, University of Bath,

Claverton Down, Bath, BA2 7AY, United Kingdom

s.hussain@bath.ac.uk

https://researchportal.bath.ac.uk/en/persons/shobbir-hussain

Abstract

Opposing explanations for the evolution of dominance effects observed in genetic traits were first proposed by Fisher and Wright around a century ago. Over the last few decades, while Wright's theory and extensions of it have reached the status of accepted paradigm, Fisher's views have become widely disregarded. Here, a number of counterarguments are presented, including a modified version of his theory, which suggest that the core views of Fisher may instead be correct. Generalised implications for our current understanding of underlying genetic architectures of complex traits are also briefly discussed.

Introduction

Originally described by Mendel, the concept of dominance may be considered as old as the field of genetics itself (Mendel, 1865). Our modern conception considers that most genetic sequence variants hold a dominance coefficient, h, ranging in value anywhere between 0 and 1; thus for example a h-value of 0 reflects complete recessiveness, a h of 0.5 reflects perfect additivity, and a value of 1 reflects complete dominance (Di & Lohmueller, 2024). The relatively rarer, or minor/alternate, alleles at any given locus are much more rarely observed to display h > 0.5, and thus more generally their effects likely range from being recessive to additive. In the human complex genetic traits field, for such alleles (for example characterised from GWAS studies), the default assumption is that h = 0.5, or sufficiently close to it; thus, the possibility that the variants may display some significant 'dominance effect' where they might in reality be considerably, or even fully, recessive, considered generally negligible (Hill, Goddard, & Visscher, 2008; Polderman et al., 2015). Here, the possibility is considered further via theoretical reformulation of previous relevant work relating to the century-old debate on the evolution of dominance.

Previous explanations for the evolution of dominance

During the early periods of the 20th century, an important observation relating to suspected deleterious mutations that impacted on the fitness (reproductive potential) of an organism, was that they very often were recessive (Morgan, Bridges, & Sturtevant, 1925). Fisher was the first to suggest that such dominance effects may have arisen via evolutionary selection forces materialising as a result of recurrent deleterious mutations (Fisher, 1928). He proposed that recurrent deleterious mutations when they first appear should display additivity, and thus heterozygotes would have a phenotype midway between that observed for the wild-type homozygotes and mutant homozygotes. Since the heterozygotes would be much more frequent than the mutant homozygotes in the population, evolutionary selection would much more efficiently act on the heterozygotes, such that their phenotype eventually comes to resemble that of the wild-type homozygote phenotype. According to Fisher's proposed mechanism, such transformation of the heterozygote phenotype would occur via positive selection of a relevant beneficial genotype at some other 'modifier' locus distinct to the 'primary' locus where the deleterious mutation was occurring; thus, the positively selected modifier genotype would enhance/supplement the normal activity of the primary locus gene. Positive selection of the favourable modifier genotype would continue until it reaches fixation in the population and ideally the phenotype of the primary locus heterozygote comes to fully resemble that of the wild-type homozygote i.e. the deleterious mutant allele becomes completely recessive. Fisher qualified from the outset that this evolutionary process was likely to be 'extremely slow' since the selective forces acting on the modifier genotype would be very weak, for example being in the order of the rate of the deleterious mutation at the primary locus; nonetheless he maintained that, akin to Darwinian thought, given enough time, such evolutionary process was within the realms of possibility (Fisher, 1928).

However, Wright (who was akin to what Bohr was to Einstein i.e. a capable intellectual adversary) disagreed with Fisher (Wright, 1929a). Although Wright highlighted and agreed that the selective forces acting on Fisher's described modifier genotype would be in the order of the rate of the deleterious mutation, and thus very weak and slow-acting, what ultimately led him to completely disregard Fisher's theory was the fact that such weak selective force was extremely

unlikely to overcome other independent selection pressures likely to be acting on the modifier genotype i.e. owing to its independent/pleiotropic functions – and this was indeed a point raised that Fisher appeared not to have considered/made reference to in his original work (Additional Note 1). What Wright proposed instead as an explanation of dominance was quite drastically different: according to him it did not evolve via selection forces produced as a result of deleterious mutations, but instead dominance was more of an inherent by-product of physiological (biochemical) mechanisms (Wright, 1934). During that part of the century, enzymatic function was considered by far the most important component of cellular physiology and thus explanations of genetic mechanisms also tended to leverage relevant biochemical observations. Wright noted that enzymatic reaction output usually displayed a non-linear concave relationship with enzyme concentration, thus after a certain point, further increasing enzyme concentration yields diminishing returns in terms of catalytic output since the concentration of the substrate becomes the limiting factor. He suggested that enzymes were ordinarily set to be produced at high concentrations corresponding to high catalytic output levels, thus far along the plateau of the enzyme output-concentration curve, one advantage of which for example would be to offer some adaptability in response to adverse environmental conditions (Forsdyke, 1994). A consequence of the high gene activity/expression would be that should the concentration of an enzyme be halved, say due to a fully deleterious mutation on one copy of the specifying gene, then the reduction in catalytic output would in fact only be very minimal since this would occur within the vicinity of the plateau of the enzyme output-concentration curve; such mutations would thus be observed as being recessive. Under Wright's model, although gene activity/expression levels could conceivably have evolved to be high i.e. via stabilising selection mechanisms, the relevant selection pressures imagined are driven independently of deleterious mutations as described by Fisher; dominance is thus instead an inherent by-product of physiological mechanisms following evolution of optimally high gene activity/expression levels.

Not too long after, Haldane offered his views on the topic. Although he also strongly criticized and disregarded the modifier aspect of Fisher's theory (Haldane, 1930), similar to Fisher he also suggested a 'selectionist model' where selection forces materialising as a result of deleterious mutations at a considered gene could possibly cause the evolution of dominance, but in such cases the expression levels of the gene were proposed to evolve to provide a 'factor of safety' against the deleterious mutations (Haldane, 1939). Thus, a gene product would for example be produced in surplus and at around twice the concentration at which it is actually needed (irrespective of optimal gene activity/expression levels), and if one copy of the gene is inactivated by mutation then the remaining copy can still fulfil normal functional activity. Thus, by around the first third of the 20th century, the three founders of the quantitative genetics field had weighed in with their views on an explanation for dominance; the topic would perhaps be the most contentiously debated between them, but all three would remain steadfast to their opinions throughout the remainder of their careers.

During around the final third of the 20th century, further developments considered to be significant were made for example by Kacser and Burns, who presented their metabolic theory – essentially an extension of Wright's physiological model but relating to multi-step enzymatic pathways (Kacser & Burns, 1981). But the period was perhaps most notable owing to two prominent studies published claiming to provide empirical evidence against Fisher's and Haldane's models. Charlesworth (B. Charlesworth, 1979) presented a mathematical argument showing that Fisher's theory predicted that for a given allele, there shouldn't exist any type of relationship between the dominance coefficient, h, and its selection coefficient, s. s can be

taken as reflecting a measure of fitness for an allele and can range in value anywhere between 0 and 1; for example it holds a value of 1 if it results in complete and consistent loss of reproductive potential when homozygous, and a value of 0 if there is no impact at all on fitness when homozygous. It was noted by Charlesworth that studies that had investigated the effect of deleterious mutations on the egg-to-adult viability of *Drosophila* very clearly showed an inverse relationship between h and s; thus, relatively more deleterious, for example more frequently lethal, mutations were quite consistently observed to be more recessive (Mukai, Chigusa, Mettler, & Crow, 1972; Mukai & Yamazaki, 1968; Simmons & Crow, 1977). Since such observations were clearly contradictory to Charlesworth's mathematical demonstration that there should exist no type of h-s relationship according to Fisher's theory, this was taken as strong evidence against the model. Charlesworth also noted that an inverse h-s relationship was instead consistent with Wright's physiological model of dominance (discussed more fully later in the article). The second study which claimed to falsify Fisher's model was conducted by Orr (Orr, 1991) via studies using the usually haploid green algae, *Chlamydomonas reinhardtii*. During its reproductive process, this species of algae transforms very briefly to a non-vegetative diploid state before undergoing cell division and reverting back to its vegetative haploid state; the diploid state may thus be considered a very short-lived and functionally inert intermediate during the reproductive cycle. However, upon laboratory manipulation, the algae can be induced into and maintained in an artificial diploid vegetative state (Dutcher, 1988; Ebersold, 1967). Orr described that the majority of mutations observed in the artificial diploids were completely recessive (Orr, 1991). This was claimed as being starkly contradictory to Fisher's model since the theory suggests that dominance slowly evolves via a selectionist mechanism acting on heterozygotes. Orr also suggested that his findings were much more consistent with Wright's theory, since under that model, recessiveness of deleterious mutations could conceivably be an immediate by-product of gene activity/expression levels being independently set to optimally high levels. Orr thus claimed that his findings falsified Fisher's and Haldane's selectionist theories while providing strong support for Wright's physiological model. Along with conclusions drawn from Charlesworth's study (B. Charlesworth, 1979), this led to the eventual widely held view in the field, that for the most part at least, Fisher and Haldane were wrong and Wright was correct in their explanations of dominance (Billiard & Castric, 2011; Bourguet, 1999; Mayo & Burger, 1997).

Furthermore, Wright's model has also been extended and reformulated relatively more recently. Hurst and Randerson confirmed previous criticisms of both Fisher's and Haldane's theories mainly on the basis that the selection forces that the two models propose would be in the order of the considered deleterious mutation rate and thus very weak (Hurst & Randerson, 2000). They instead reasonably argued that high levels of gene expression of enzymes for example could more readily evolve simply via stabilising selection as a means of optimising their physiological activity, thus driving the existence of dominance. Furthermore, taking into consideration costs of transcription, which for example Haldane ignored, they suggest that an optimal level of gene expression with regard to the overall fitness effect on the organism is selected, rather than a significant surplus of gene expression as Haldane conversely imagined. However, a major drawback of Wright's model, and extensions of it, is that they relatively much more poorly yield an explanation of dominance for non-enzymatic proteins. Although there has been at least some attempt more recently by Huber et al. to extend the model to include other types of gene categories (Huber, Durvasula, Hancock, & Lohmueller, 2018), there is nonetheless recognition that this more generally is an unsolved problem that at the very least requires further examination (Di & Lohmueller, 2024).

'Fitness landscape' models have also been proposed (Manna, Martin, & Lenormand, 2011; McDonough, Ruzicka, & Connallon, 2024) which incorporate aspects of an independent adaptive evolution model proposed by Fisher (Fisher, 1930) known as his 'geometric model'. Also sometimes referred to as the 'Fisher-Orr model', the geometric model can be used model fitness effects of mutations that may contribute to adaptive evolution via stabilising selection (Orr, 1999, 2000). Although Fisher himself never used the geometric model to help provide an explanation of dominance, it can still be useful as it shares principles of stabilising selection consistent with formulations of Wright's dominance model (Hurst & Randerson, 2000; Manna et al., 2011; McDonough et al., 2024). Since such fitness-landscape models do not invoke the underlying selectionist principles of Fisher's dominance evolution theory, they are not subject to its major criticisms/purported counter-evidences, and have been used to model evolutionary explanations of dominance-effects for small-effect deleterious mutations (Manna et al., 2011) and much more recently for beneficial mutations (McDonough et al., 2024).

A final point perhaps worth of brief mention here is that while Wright's model and various reformulations of it have effectively reached the status of paradigm, there at least exists some suggestion that Fisher's original theory might apply to some rarer instances of dominance observations, for example in Batesian mimicry (D. Charlesworth & Charlesworth, 1975). A relatively more recent example was described as potentially occurring in flowering plants where small RNA genes may be able to modify dominance effects observed at nearby located protein coding genes (Tarutani et al., 2010); here the expressed small RNA modifiers can regulate methylation of the nearby protein coding gene promoter thus modifying it's expression levels while modulating dominance effects at the locus. Although identification of such modifiers of dominance would be consistent with what Fisher described in his model, such instances nonetheless are generally thought to only represent special cases (Di & Lohmueller, 2024).

A reformulated model for the evolution of dominance

Here it will be contested that the original Fisher (Fisher, 1928) and Haldane (Haldane, 1939) selectionist views on the evolution of dominance should be reconsidered as being potentially much more generally applicable explanations; a model essentially representing modification and unification of the two theories will be formulated. The objections raised to Fisher's theory by the studies of Charlesworth (B. Charlesworth, 1979) and Orr (Orr, 1991) have consistently since been the claimed 'damning' evidences against it, and to this day, findings of those two studies are also further used to benchmark the feasibility of dominance theories more generally (Di & Lohmueller, 2024; Huber et al., 2018). As such, a minimal requirement is illustration that the two studies do not actually provide the purported evidence against the selectionist view originally proposed in Fisher's model; via plain intuitive arguments, the bold claims made by Orr's study will be addressed at the outset here, while Charlesworth's study will be discussed following presentation of the modified model.

Orr's study collated data that showed that the majority of *Chlamydomonas* deleterious gene mutations analysed in artificial diploids were recessive (Orr, 1991); indeed, even mutations that would be lethal in haploids were recessive in the artificial diploids. Therefore, the mutations quite clearly have a propensity to display dominance effects without the opportunity for this to have evolved via selection acting on heterozygotes – particularly bearing in mind that the diploid state in nature is only an extremely brief non-vegetative state. Without any deeper examination, the reasoning provided by Orr is sound. However, a critical assumption not stated by Orr is that

he has imagined that the artificial diploids require more of the relevant gene products than that required by haploids in order to maintain physiological function within the observable normal range – this requires at least some empirical support that is never provided. If artificial diploids can function roughly as well with the same amount of gene activity that is expressed in haploids, then recessiveness of gene mutations in the artificial diploids is an *a priori* logical expectation, and not an evidence against Fisher's model. The argument here would therefore hinge on whether artificial diploids require more gene activity than in haploids in order to maintain physiological function within an observable normal range. Artificial diploid *Chlamydomonas* have a cell size around twice that of haploids (Ebersold, 1967) – although an observation on the face of it in Orr's favour, it is still a big leap to extrapolate this into meaning that more gene activity compared to haploids for all genes will be required to maintain normal functioning of diploids. It may be worth noting that around 10% of the *Chlamydomonas* gene mutations considered by Orr displayed semi-dominance (i.e. additive) effects – the possibility exists that it is the activity of these genes (but not those displaying recessiveness) that is required to be higher in artificial diploids in order to maintain levels of cellular function within a normal range, and thus heterozygous mutation more specifically of these genes leads to a measurable phenotypic effect. Orr specifically highlights the fact that flagellar gene mutations were recessive in artificial diploids, whereas flagellum are not expressed/required in the brief natural diploid phase of *Chlamydomonas*: therefore, there should in particular unequivocally exist no selection pressure for dominance effects to evolve via Fisher's mechanism for flagellar genes. However, there is an un-noted relevant fundamental consideration here that would be improper to ignore: do artificial diploids require more of the flagellar proteins compared to that expressed in haploids to maintain physiological function within the observable normal range? If not, recessiveness of flagellar gene mutations in the artificial diploids would be an *a priori* expectation, and falsification or proof of in fact very little. Considering that such aforementioned questions have not been addressed, Orr's observations are not necessarily an evidence against Fisher's theory, nor even necessarily a support of Wright's model.

Here it is suggested that selection forces materialising as a result of recurrent deleterious mutations may indeed be responsible for dominance evolution as suggested by both Fisher (Fisher, 1928) and Haldane (Haldane, 1939). In his theory, Fisher imagined that positively selected modifiers of dominance were located at loci alternate to the primary locus where the recurrent deleterious mutation occurred. As mentioned earlier, Haldane effectively suggested that an increase of gene expression, and thus a 'factor of safety', at the locus where the deleterious mutation is occurring is involved in the evolution of dominance. Here, Fisher's model is adopted but is adjusted such that the modifiers of dominance that Fisher spoke of are instead positively selected mutations actually within the primary locus gene promoter that act to increase the gene's expression. Such an adjusted model draws inspiration from previous examinations of the evolution of gene expression levels owing to host-parasite interactions using frameworks structurally similar to Fisher's dominance evolution model (Nuismer & Otto, 2005). However, the adjusted model, offers a couple of critical advantages over Fisher's original theory. First, Wright's most critical objection to the theory, regarding the modifiers being subject to other likely overwhelming independent selection pressures due to pleiotropic effects (Wright, 1929a, 1929b) (Additional Note 1), becomes virtually irrelevant since here it is envisaged that the modifier sequence variant occurs within the primary locus gene promoter - and it's function thus limited to regulating the expression of said gene. Second, a modifier variant specifically within the gene promoter has the potential to have a very significant effect in increasing the fitness of the deleterious mutation heterozygote, since sequence variants within promoters can

potentially substantially enhance expression of genes; Fisher and Wright for example agreed that the magnitude of the modifier effect would also be a key factor influencing the feasibility of Fishers' model (Fisher, 1928, 1929; Wright, 1929a, 1929b). Note that this adjusted model specifies modifier sequence variants within the gene promoter as opposed to within cis regulatory enhancer elements - an important distinction since the latter can often regulate the expression of multiple (sometimes distant) genes and thus would not escape Wright's pleiotropy argument; on similar reasoning, this model does not invoke potential modifier variants within other connected gene regulatory network loci. Thus, deviating from Fisher's original theory by specifically invoking modifier variants not subject to independent pleiotropic effects should significantly increase the likelihood that, given enough time, they will reach fixation in the population and could thus account for the evolution of dominance. There are a couple of other important considerations too. So far, we have considered relationships between single recurrent deleterious mutations and single modifier sequence variants; however, the reality is instead likely to be that in most cases, multiple distinct deleterious mutations will exist for a given gene across a population and also that the potential for multiple distinct modifiers of gene expression within the gene's promoter will exist. Thus in reality, under this model, evolution of dominance effects at a locus at a given point in time is unlikely to rely on a small selection force generated from just one specific deleterious mutation in the gene, but much rather, multiple deleterious mutations in the gene can potentially combine to conceivably yield a much greater selection force in order to drive the fixation of the dominance modifier variants. This is since the role of the gene promoter dominance modifiers is to increase the expression of the gene which would conceivably have a beneficial effect on all of the distinct deleterious heterozygotes for that gene present in the population. A related consideration is that the mutation rates for distinct genomic sites is known to significantly vary, with some genes for example displaying 'hypermutability' (Michaelson et al., 2012; Nesta, Tafur, & Beck, 2021); since the rate of deleterious mutation is the central factor determining the rate of dominance evolution under the selectionist model, this point also should be taken into account. Thus, for example a basic prediction regarding the evolution of dominance that can be made from this model is that genes with higher mutation rates, while taking into consideration the total combined rate of deleterious mutation across the whole gene, are more likely to evolve dominance effects; genes with lower total deleterious mutation rates are less likely to achieve the selection forces required and thus the relevant alleles more likely to retain additive effects.

The earlier mentioned study of Charlesworth described an inverse h-s relationship for mutations affecting egg-to-adult viability in *Drosophila*, thus purporting to support Wright's model while providing evidence against Fisher's (B. Charlesworth, 1979). Charlesworth notes that Wright's theory assumes that the function of most gene loci is to specify the structure of enzymes. Based on the concentration-catalytic output relationship of enzymes, mildly deleterious mutations (thus low s) are predicted to have additive effects (high h), while strongly deleterious mutations (high s) are predicted to display strong recessiveness (low h) (Figure 1A); Charlesworth thus also notes that an inverse h-s relationship is indeed predicted from Wright's model. However, a critical point not noted is that this reasoning only considers an enzyme or enzymes that for example help determine the manifestation of a particular trait (we'll call this trait y). To see why this is important, consider a second example here, where we have an enzyme encoding a distinct trait (trait z) (Figure 1B), but imagine that this trait has a much more severe effect on fitness. According to Wright's model, the h-value for the loss-of-function mutations from the two examples will be identical for both (Figures 1A and 1B). However, considering that trait z has a much more severe effect on fitness than trait y, then in reality the loss-of-function

mutation for trait z will in fact have a significantly higher s-value than that for the loss-of-function mutation for trait y; indeed in general, there would be a much less discernible inverse h-s relationship when analysing mutation effects across the two traits. Thus, when considering relevant situations within a given fitness trait, an inverse h-s relationship may indeed be predicted from Wright's theory, but when extending the analysis across independent traits with varying effects on fitness, such a relationship is not in fact a general expectation of the model. Critically, the *Drosophila* egg-to-adult viability mutation data considered by Charlesworth were in fact collated from several independent studies, including a larger meta-analyses performed by Simmons and Crow (Simmons & Crow, 1977), and would have included mutation effects across a number of independent fitness traits: the relevant measurement considered was simply the rate at which randomly-occurring spontaneous and chemically-induced mutations led to loss of viability (Mukai et al., 1972; Mukai & Yamazaki, 1968; Simmons & Crow, 1977). Therefore, in those studies it is quite possible, even likely, that mutations associated with relatively mild fitness traits would have much less frequently led to loss of viability (thus low calculated s), while mutations associated much more severe fitness traits would much more frequently have led to loss of viability (thus high calculated s). Thus, the observation of an inverse h-s relationship from such analyses could in fact have had little to do with the relevant physiological/enzymatic effects predicted from Wright's model, and therefore not necessarily a support for his theory as concluded by Charlesworth. Similar reasoning can be used to reconsider whether Charlesworth study is necessarily an evidence contrary to Fisher's model. Charlesworth (B. Charlesworth, 1979), as well as others more recently (Di & Lohmueller, 2024), present the same mathematical argument showing that Fisher's model predicts that dominance should evolve independently of the selection coefficient, and therefore that there shouldn't exist any kind of h-s relationship according to the theory. When considering the context of mutations within a fitness trait, this also indeed seems a reasonable and intuitive expectation from the herein presented modified version of Fisher's model: the deleterious effects of a mild mutation (lower s) would be easier to recover as less modification of the heterozygote fitness is needed, while the deleterious effects of a severe mutation (higher s) is more difficult to recover as more modification of heterozygote fitness is necessary (Figure 1C); thus, where s-values significantly differ, this is counterbalanced by an equally significant difference in the 'modification work' necessary for dominance effects to evolve, and therefore the value of s should provide no overall advantage or disadvantage. Thus, contrary to the expectation from Wright's model, small-effect mutations are for example just as likely to display recessiveness compared to large-effect mutations. However, similar to earlier, when considering a more general h-s relationship analysis, it is critical to consider likely mutation effects across traits while considering the impact of the traits themselves on fitness (Figure 1D) (Additional Note 2). Therefore, for a severe fitness trait, s-values in reality will on average be much higher than the s-values for a much milder fitness trait; s-values for severe fitness traits are therefore much more likely to significantly outweigh the modification work necessary for dominance to evolve, even if such modification work were for example to be equal for both traits, and thus h should in fact diminish toward zero much more readily in comparison (Figure 1D). Thus, when analysing mutation effects across traits which themselves have varying fitness effects, a prediction that rather follows from the modified Fisher model is that for severe fitness traits (higher s on average) dominance is more likely to evolve (lower h on average), and for mild fitness traits (lower s on average) dominance is less likely to evolve (higher h on average); therefore yielding an inverse h-s relationship, and thus consistent with the empirical observations collated by Charlesworth (B. Charlesworth, 1979) and with relatively more recent

similar findings of inverse h-s relationships described in other organisms (Agrawal & Whitlock, 2011; Huber et al., 2018; Phadnis & Fry, 2005).

A final point well worth making is that while Wright's model relatively more poorly explains dominance for non-enzymatic proteins, Fisher's theory would explain dominance equally well for all types of gene: this is a quite critical but a rather understated advantage of Fisher's model.

Implications for the genetic architecture of complex traits

If we were to consider the potential implications of the preceding discussions on the current complex genetic traits field, a couple of relevant aspects are worth highlighting. It is generally thought that complex traits usually have a similar genetic architecture irrespective for example of whether they are non-fitness traits (for example height) or fitness traits (for example disease), such that the genetic variance is generally thought to be largely additive (Hill et al., 2008; Polderman et al., 2015). However, a first prediction of the model presented above is that dominance may likely evolve for fitness traits but not for non-fitness traits; furthermore, dominance is likely for example to be more frequently observed for severe fitness traits compared to mild fitness traits. A second prediction is that, within a given fitness trait, genes with higher total deleterious mutation rates are much more likely to display dominance effects. The first prediction may in fact help explain why twin study-based estimates of the broad sense heritability of morphological traits such as human height for example is virtually fully recovered by its additive genetic variance characterised using genomic methods, whereas large respective gaps remain for disease traits including for example body mass index (BMI) (Wainschtein et al., 2022) which can be regarded as a significant fitness trait (Nguyen, Wilcox, Skjaerven, & Baird, 2007; Zhu et al., 2022). Indeed, similar observations are made for complex traits when comparing broad sense heritability estimated from identity-by-descent (IBD)-based studies (Sidorenko et al., 2024) and the respective total narrow sense heritability estimated from relevant GREML-based methods (i.e. whole-genome-sequencing/WGS-heritability) (Wainschtein et al., 2022). Thus for example, IBD-based heritability of height (~0.75) is virtually fully recovered by its WGS-heritability estimate (~0.7), whereas the WGS-heritability of BMI (~0.3) falls significantly short of its IBD-based heritability (~0.55). Reliable IBD-based broad sense heritability and WGS-based narrow sense heritability estimates are not yet available for other traits, but the 'still-missing heritability' (the gap between available WGS- or SNP-based heritability and respective broad sense heritability estimates) for BMI and potentially other fitness traits may be explained at least in part by dominance effects; indeed, the dominance model proposed here could provide a tangible evolutionary explanation for significant disparities in the potential amount of still-missing heritability observed for fitness traits compared to non-fitness traits. It may be noted that previous studies that investigated genetic dominance in large numbers of complex traits from UK biobank datasets found averaged across traits evidence of only a small contribution from dominance effects (Hivert et al., 2021; Palmer et al., 2023; Pazokitoroudi, Chiu, Burch, Pasaniuc, & Sankararaman, 2021), however, in the context of our discussions here, two relevant observations are worth making. First, the vast majority of complex traits included in those studies were non-fitness traits ranging from morphological traits to, for example, tea intake. Second, those studies only analysed SNP common variants, thus all common variant-rare variant and rare variant-rare variant dominance interactions would have been completely undetectable: a particularly important omission if the variants involved in proposed dominance effects are also large-effect ones (see below).

As discussed, under the proposed model, genes with high total combined deleterious mutation rates are much more likely to display dominance effects since this will be the pivotal factor driving its evolution. Such 'hypermutable' genes should also be much more susceptible to future recurrent *de novo* mutations as well. Disease associated germline *de novo* mutations are found to occur in protein coding sequence and are of large-effect, but they are relatively rare and thus tend to explain a very small fraction of the phenotypic variance in the population (Satterstrom et al., 2020; Singh et al., 2022). However, although poorly investigated for non-cancerous complex disease traits given the difficulties involved in adequately detecting them, somatic *de novo* mutations potentially may represent a largely 'untapped' source of relevant contributing variance (Hussain, 2024); these may also be predicted (under the model proposed here) to occur relatively more frequently at loci where dominance effects have evolved. Thus, the possibility exists that some heritable genetic variants associated with complex disease traits, detected via GWAS and some sequencing studies for example, may not act additively as assumed (and as predicted from Wright's model for small-effect mutations) but might in reality exhibit recessiveness; therefore, they would contribute in conferring liability to a disease owing to potential interactions with other associated inherited variants as well as with potentially large-effect *de novo* variants within the same locus.

Summary

Over the last few decades, Fisher's dominance evolution model has become apparently discredited and quite widely disregarded. The core principles of the theory are revived here, while the described modifications made to the model should enhance its general feasibility. There are a couple of predictions relating to our current understanding of the genetic architecture of complex traits that also emerge from the model. Firstly, the fundamental characteristics of their genetic architecture may not be the same for all complex traits, such that while non-fitness traits are likely to more generally display additivity of associated genetic variants, fitness traits will instead have an increased propensity to display dominance effects for associated variants at some loci. Secondly, such loci where dominance effects are observed are more likely to be those characterised by higher total combined rates of deleterious mutation across the gene. Such outcomes are not expected from the current consensus explanations of dominance, such as Wright's model and subsequent reformulations of it. Thus, testing of the aforementioned predictions in future would offer means to provide empirical support for the model presented.

Additional Notes

Additional Note 1

Although the debate between Fisher and Wright on the explanation of dominance has sometimes been described as a fierce one, the two did respect each other and the exchanges were cordial and fair. It is a common misconception that a matter of contention between the two was Wright's assertion that the selection forces described by Fisher's theory would be in the order of the mutation rate and thus very weak and slow-acting (Wright, 1929a); this was in fact a point already alluded to by Fisher during the presentation of the theory in his original paper (Fisher, 1928), and was a point that he consistently readily embraced (Fisher, 1929). There were

also a couple of mathematical inconsistencies which were resolved following exchanges between the two authors (Fisher, 1934; Wright, 1934).

As both Wright and Fisher pointed out, the real significant point of remaining contention between the two was whether the very weak selection forces acting on the described modifier genotypes could adequately compete with the other independent pleiotropic selection pressures also very likely to be acting on them; Wright clearly felt that this was likely impossible, while Fisher had to concede that this was the one well-directed criticism of his theory and could offer no tangible counterargument other than to reiterate his original point that dominance evolution under his model would be 'extremely slow' (Fisher, 1929; Wright, 1929b).

Additional Note 2

An objection that might be levelled against the reasoning presented here may be a counterargument that Fisher viewed dominance as being equally prevalent across traits irrespective of their effect on fitness (Fisher, 1922): in his ill-fated 'dominance ratio' study, he proposed that one third of the genetic variance would be accounted for by dominance across traits including non-fitness traits such as height (Fisher, 1922). However, it is important to note that this study was published some six years before he originally presented his model on the evolution of dominance (Fisher, 1928). Thus, although he may never have explicitly reclarified, it would be sensible to assume that his views would have been revised to imagine that dominance should for example be prevalent in fitness traits but absent in non-fitness traits, considering after all that his dominance evolution model is built around the concept of selection forces materialising as a result of deleterious mutations which compromise fitness. In any case, Fisher's dominance ratio study would not have relevance to the validity of the herein presented iteration of his dominance evolution model.

Declarations

Availability of data

There are no data underlying this work

Competing Interests

The author declares no competing interests

Funding

There are no funders of this work

References

Agrawal, A. F., & Whitlock, M. C. (2011). Inferences about the distribution of dominance drawn from yeast gene knockout data. *Genetics, 187*(2), 553-566. doi:10.1534/genetics.110.124560

Billiard, S., & Castric, V. (2011). Evidence for Fisher's dominance theory: how many 'special cases'? *Trends Genet, 27*(11), 441-445. doi:10.1016/j.tig.2011.06.005

Bourguet, D. (1999). The evolution of dominance. *Heredity (Edinb), 83 (Pt 1)*, 1-4. doi:10.1038/sj.hdy.6885600

Charlesworth, B. (1979). Evidence against Fisher's theory of dominance. *Nature, 278*, 848-849.

Charlesworth, D., & Charlesworth, B. (1975). Theoretical genetics of Batesian mimicry III. Evolution of dominance. *J Theor Biol, 55*(2), 325-337. doi:10.1016/s0022-5193(75)80083-x

Di, C., & Lohmueller, K. E. (2024). Revisiting Dominance in Population Genetics. *Genome Biol Evol, 16*(8). doi:10.1093/gbe/evae147

Dutcher, S. K. (1988). Nuclear fusion-defective phenocopies in Chlamydomonas reinhardtii: mating-type functions for meiosis can act through the cytoplasm. *Proc Natl Acad Sci U S A, 85*(11), 3946-3950. doi:10.1073/pnas.85.11.3946

Ebersold, W. T. (1967). Chlamydomonas reinhardi: heterozygous diploid strains. *Science, 157*(3787), 447-449. doi:10.1126/science.157.3787.447

Fisher, R. A. (1922). On the dominance ratio. *Proc R Soc Edinb., 42*, 321-341.

Fisher, R. A. (1928). The Possible Modification of the Response of the Wild Type to Recurrent Mutations. *American Naturalist, 62*, 115-126.

Fisher, R. A. (1929). The evolution of dominance; reply to Professor Sewall Wright. *American Naturalist, 63*, 553-556.

Fisher, R. A. (1930). The genetical theory of natural selection. *Clarendon Press*.

Fisher, R. A. (1934). Professor Wright on the theory of dominance. *American Naturalist, 68*, 370-374.

Forsdyke, D. R. (1994). The heat-shock response and the molecular basis of genetic dominance. *J Theor Biol, 167*(1), 1-5. doi:10.1006/jtbi.1994.1044

Haldane, J. B. S. (1930). A note on Fisher’s theory of the origin of dominance and a correlation between dominance and linkage. *American Naturalist, 64*, 87-90.

Haldane, J. B. S. (1939). The theory of the evolution of dominance. *American Naturalist, 37*, 365-374.

Hill, W. G., Goddard, M. E., & Visscher, P. M. (2008). Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet, 4*(2), e1000008. doi:10.1371/journal.pgen.1000008

Hivert, V., Sidorenko, J., Rohart, F., Goddard, M. E., Yang, J., Wray, N. R., . . . Visscher, P. M. (2021). Estimation of non-additive genetic variance in human complex traits from a large sample of unrelated individuals. *Am J Hum Genet, 108*(5), 786-798. doi:10.1016/j.ajhg.2021.02.014

Huber, C. D., Durvasula, A., Hancock, A. M., & Lohmueller, K. E. (2018). Gene expression drives the evolution of dominance. *Nat Commun, 9*(1), 2750. doi:10.1038/s41467-018-05281-7

Hurst, L. D., & Randerson, J. P. (2000). Dosage, deletions and dominance: simple models of the evolution of gene expression. *J Theor Biol, 205*(4), 641-647. doi:10.1006/jtbi.2000.2095

Hussain, S. (2024). A modeling of complex trait phenotypic variance determinants. *PNAS Nexus, 3*(11), pgae472. doi:10.1093/pnasnexus/pgae472

Kacser, H., & Burns, J. A. (1981). The molecular basis of dominance. *Genetics, 97*(3-4), 639-666. doi:10.1093/genetics/97.3-4.639

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

Mayo, O., & Burger, R. (1997). The evolution of dominance: a theory whose time has passed? *Biol. Rev. Cambridge Philos. Soc., 72*, 97-110.
McDonough, Y., Ruzicka, F., & Connallon, T. (2024). Reconciling theories of dominance with the relative rates of adaptive substitution on sex chromosomes and autosomes. *Proc Natl Acad Sci U S A, 121*(44), e2406335121. doi:10.1073/pnas.2406335121
Mendel, J. G. (1865). Verhandlungen des naturforschenden vereines in Brünn. *Abhandlungen, 4*, 3-47.
Michaelson, J. J., Shi, Y., Gujral, M., Zheng, H., Malhotra, D., Jin, X., . . . Sebat, J. (2012). Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell, 151*(7), 1431-1442. doi:10.1016/j.cell.2012.11.019
Morgan, T. H., Bridges, C. B., & Sturtevant, A. H. (1925). The Genetics of Drosophila. *Bibliographica Genetica, 2*, 1-262.
Mukai, T., Chigusa, S. I., Mettler, L. E., & Crow, J. F. (1972). Mutation rate and dominance of genes affecting viability in Drosophila melanogaster. *Genetics, 72*(2), 335-355. doi:10.1093/genetics/72.2.335
Mukai, T., & Yamazaki, T. (1968). The genetic structure of natural populations of Drosophila melanogaster. V. Coupling-repulsion effect of spontaneous mutant polygenes controlling viability. *Genetics, 59*(4), 513-535. doi:10.1093/genetics/59.4.513
Nesta, A. V., Tafur, D., & Beck, C. R. (2021). Hotspots of Human Mutation. *Trends Genet, 37*(8), 717-729. doi:10.1016/j.tig.2020.10.003
Nguyen, R. H., Wilcox, A. J., Skjaerven, R., & Baird, D. D. (2007). Men's body mass index and infertility. *Hum Reprod, 22*(9), 2488-2493. doi:10.1093/humrep/dem139
Nuismer, S. L., & Otto, S. P. (2005). Host-parasite interactions and the evolution of gene expression. *PLoS Biol, 3*(7), e203. doi:10.1371/journal.pbio.0030203
Orr, H. A. (1991). A test of Fisher's theory of dominance. *Proc Natl Acad Sci U S A, 88*(24), 11413-11415. doi:10.1073/pnas.88.24.11413
Orr, H. A. (1999). The evolutionary genetics of adaptation: a simulation study. *Genet Res, 74*(3), 207-214. doi:10.1017/s0016672399004164
Orr, H. A. (2000). Adaptation and the cost of complexity. *Evolution, 54*(1), 13-20. doi:10.1111/j.0014-3820.2000.tb00002.x
Palmer, D. S., Zhou, W., Abbott, L., Wigdor, E. M., Baya, N., Churchhouse, C., . . . Neale, B. M. (2023). Analysis of genetic dominance in the UK Biobank. *Science, 379*(6639), 1341-1348. doi:10.1126/science.abn8455
Pazokitoroudi, A., Chiu, A. M., Burch, K. S., Pasaniuc, B., & Sankararaman, S. (2021). Quantifying the contribution of dominance deviation effects to complex trait variation in biobank-scale data. *Am J Hum Genet, 108*(5), 799-808. doi:10.1016/j.ajhg.2021.03.018
Phadnis, N., & Fry, J. D. (2005). Widespread correlations between dominance and homozygous effects of mutations: implications for theories of dominance. *Genetics, 171*(1), 385-392. doi:10.1534/genetics.104.039016
Polderman, T. J., Benyamin, B., de Leeuw, C. A., Sullivan, P. F., van Bochoven, A., Visscher, P. M., & Posthuma, D. (2015). Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet, 47*(7), 702-709. doi:10.1038/ng.3285
Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., . . . Buxbaum, J. D. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell, 180*(3), 568-584 e523. doi:10.1016/j.cell.2019.12.036
Sidorenko, J., Couvy-Duchesne, B., Kemper, K. E., Moen, G. H., Bhatta, L., Asvold, B. O., . . . Yengo, L. (2024). Genetic architecture reconciles linkage and association studies of complex traits. *Nat Genet, 56*(11), 2352-2360. doi:10.1038/s41588-024-01940-2
Simmons, M. J., & Crow, J. F. (1977). Mutations affecting fitness in Drosophila populations. *Annu Rev Genet, 11*, 49-78. doi:10.1146/annurev.ge.11.120177.000405

Singh, T., Poterba, T., Curtis, D., Akil, H., Al Eissa, M., Barchas, J. D., . . . Daly, M. J. (2022). Rare coding variants in ten genes confer substantial risk for schizophrenia. *Nature, 604*(7906), 509-516. doi:10.1038/s41586-022-04556-w

Tarutani, Y., Shiba, H., Iwano, M., Kakizaki, T., Suzuki, G., Watanabe, M., . . . Takayama, S. (2010). Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility. *Nature, 466*(7309), 983-986. doi:10.1038/nature09308

Wainschtein, P., Jain, D., Zheng, Z., Group, T. O. A. W., Consortium, N. T.-O. f. P. M., Cupples, L. A., . . . Visscher, P. M. (2022). Assessing the contribution of rare variants to complex trait heritability from whole-genome sequence data. *Nat Genet, 54*(3), 263-273. doi:10.1038/s41588-021-00997-7

Wright, S. (1929a). Fisher's theory of dominance. *American Naturalist, 63*, 274-279.

Wright, S. (1929b). The evolution of dominance: comment on Dr. Fisher’s reply. *American Naturalist, 63*, 556-561.

Wright, S. (1934). Physiological and evolutionary theories of dominance. *American Naturalist, 68*, 25-53.

Zhu, L., Zhou, B., Zhu, X., Cheng, F., Pan, Y., Zhou, Y., . . . Xu, Q. (2022). Association Between Body Mass Index and Female Infertility in the United States: Data from National Health and Nutrition Examination Survey 2013-2018. *Int J Gen Med, 15*, 1821-1831. doi:10.2147/IJGM.S349874

Figure 1

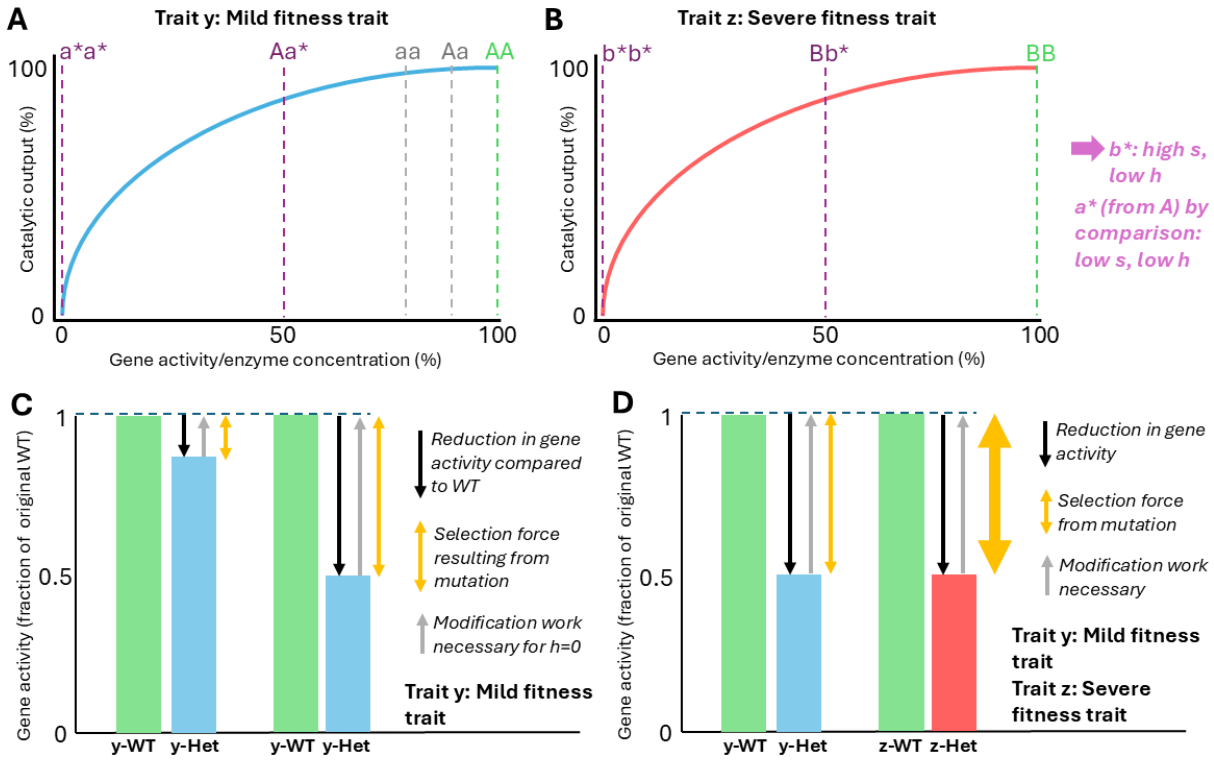

Figure 1. h-s relationship expectations from Wright's and Fisher's dominance models

(A) A typical profile of an enzyme concentration-catalytic output relationship is depicted. The enzyme is normally encoded by a wild-type homozygous genotype 'AA'. A mutation denoted by the allele 'a' that is mildly deleterious with regard to activity of the enzyme, and thus holds a low s-value, results in a roughly-equal decrease in the catalytic output when either heterozygous or homozygous; thus, an additive effect of the mutation and therefore a high h-value is expected. However, a complete loss-of-function mutation, denoted by 'a*', which would hold a relatively high s-value, results in only a relatively small decrease in catalytic output when heterozygous compared to complete loss of output when homozygous, thus yielding a low h-value. This is Wright's model of dominance and from which an inverse h-s relationship can be expected. Indeed, similar results can be expected even when comparing mutation effects across potential multiple genes (not shown here) that specify the trait being considered. Furthermore, if mutation effects across independent traits each with the same impact on fitness were to be considered (not shown here), then such an inverse h-s relationship can still similarly be expected from Wright's model.

(B) The profile of an enzyme concentration-catalytic output relationship for an enzyme encoding a second trait is depicted – this relationship is quantitatively identical to that which we observed for the enzyme in (A). However, this enzyme specifies a severe fitness trait, the wild-type genotype of which is denoted by 'BB'. Complete loss-of-function mutation, 'b*', results in identical reductions in catalytic output in both the heterozygote and mutant homozygote compared to those caused by the respective a* loss-of-function mutations encountered in (A); thus under Wright's model, both the b* and a* mutations will unequivocally hold the same value for h. If it were imagined that

the sole factor that determined the s-values for the a* and b* mutations was the impact of the mutations on the level of gene activity, then s would also be the same for both, and thus a h-s relationship is still maintained from Wright's model. However, this is clearly unrealistic: considering that s-values are also largely determined by the impact of the trait itself on fitness, in reality the s-value for the b* mutation will be much higher than the s-value of the a* mutation, thus meaning that a clear h-s relationship is unlikely to be observed when analysing mutation effects across the two independent traits. Thus, when considering mutation effects across traits which themselves have varying effects on fitness, a h-s relationship is not necessarily to be expected from Wright's model, especially if traits with widely differing effects on fitness are included in the analysis.

(C) The effects of a mildly deleterious mutation are shown on the left. Such a mutation only reduces gene activity to a limited extent in heterozygotes (black arrow) and thus the 'modification work' necessary to recover this (by elevating expression of the gene) is equally small (grey arrow). The effects of a complete loss-of-function mutation are depicted on the right, here the reduction in gene activity is much larger by comparison but this is associated with an equally larger amount of modification work necessary to recover the fitness of the heterozygote. Note that the relative strengths of the selection force resulting from the deleterious mutation which is proposed to drive the evolution of dominance under Fisher's model (depicted by the gold arrows) will be proportional to the magnitude of reduction in gene activity caused by each of the mutations (black arrows) – this is counterbalanced by equally proportional degrees of modification work necessary for dominance to evolve (grey arrows). Thus, in such cases, higher values of s for example would provide no overall advantage for dominance to evolve and a h-s relationship is unlikely to be observed, therefore consistent with Charlesworth's analysis of Fisher's theory. Similar would apply if we were to consider mutation effects across any of the potential multiple genes (not shown here) that specify the particular trait being considered, or even indeed across multiple traits each of which have same effect on fitness (not shown here). WT = wild-type; Het = heterozygote.

(D) The effects of a complete loss-of-function mutation for a mild fitness trait are shown on the left, while the same for a severe fitness trait are shown on the right. Since the relative magnitude of effects of the mutations on gene activity (black arrows) and the modification work necessary for recovery of heterozygote fitness (grey arrows) are all equal in both examples, we might expect dominance to evolve equally for both traits; however, this expectation is unrealistic as it ignores the effect of the trait itself on fitness. In reality, the selection force generated (depicted by gold arrows) for the severe fitness trait mutation will be much higher than that for the mild fitness trait mutation, meaning that h is driven toward zero much more readily in comparison given that the modification work necessary to recover heterozygote fitness (grey arrows) still remains equal for both traits. Thus, when analysing mutation effects across traits with differing effects on fitness, an inverse h-s relationship can much more readily be expected from the model. WT = wild-type; Het = heterozygote.