Unresolved Boundaries of Evolutionary Theory and the Question of how Inheritance Systems Evolve: 75 Years of Debate on the Evolution of Dominance

HOMAYOUN C. BAGHERI1,2,3*
1Max Planck Institute for Infection Biology, Berlin, Germany
2Santa Fe Institute, Santa Fe, New Mexico, USA
3Department of Ecology and Evolutionary Biology, Yale University, New Haven, Connecticut, USA

ABSTRACT One of the key issues in the evolution of life is the evolution of inheritance systems. In population genetics, the earliest attempt at addressing the latter problem revolved around Fisher’s theory on the evolution of dominance. Fisher’s hypothesis was that inheritance systems could be modified during the evolutionary process in such a way that wild-type phenotypes could become dominant with respect to mutant phenotypes. This would result in the buffering of a population against the deleterious effects of mutations. The debate that ensued on this topic has been one of the most longstanding in evolutionary theory. At present, the prevalent view is that dominance cannot evolve as a direct result of selection. Furthermore, it has been argued that due to inherent constraints in biochemical systems, the manifestation of dominance is a default expectation and hence evolutionary explanations are not necessary. This has led to the position that the subject is generally resolved and no further debate is necessary. However, there are also several studies indicating that dominance levels can be modified as a result of changes in the genetic background. Furthermore, other studies have indicated that dominance selection is possible in certain circumstances. To a large degree, conclusions from both of the latter types of studies have been ignored. In this article, the history of several intellectual and methodological traditions that have contributed to this debate are traced, including experimental genetics, theoretical population genetics and theoretical biochemistry. In the light of both old and contemporary works on this topic, it is argued that contrary to the prevalent view, the evolution of dominance is not a resolved issue. A re-examination of this issue is essential, given that dominance evolution is likely to be an important stepping stone towards understanding the evolution of inheritance systems.

In 1928, Ronald Fisher presented a population genetic model for the evolution of dominance (Fisher, ’28a,b). The novelty of this work was that it tried to address how the properties of a genetic system can change within a neo-Darwinian framework. In its intent, this was an attempt at understanding how an inheritance system can evolve.

In this article, an inheritance system is defined as the set of processes by which the phenotypic characteristics associated with a parent set are inherited, with some degree of reliability, by its offspring. As defined, such a process has to include development. Understanding the evolution of development is an important goal, because without it, the relation between genotype and phenotype is confined to the status of a fixed black-box. As such, we are not equipped to address macro-evolutionary problems such as morphological change and the origin of novel characters. If such a framework remains unaltered, we are hindered from further understanding the evolution of organismal organization (see for example Riedl, ’77; Buss, ’87; Müller and Wagner, ’91; Fontana

Grant sponsor: Yale University; Grant sponsor: Santa Fe Institute; Grant sponsor: The Max Planck Society
*Correspondence to: Homayoun C. Bagheri, Max Planck Institute for Infection Biology, Schumannstr. 21/22, Berlin 10117, Germany. E-mail: bagheri@molgen.mpg.de
Received 22 September 2004; Accepted 15 June 2005
Published online 13 September 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/jez.b.21069.
It is now well understood that Fisher’s model had several flaws, and that it is unlikely for dominance to evolve in accordance to the conditions of his model (Feldman and Karlin, ’71; Wagner and Bürger, ’85; Mayo and Bürger, ’97). But the rebuttal of Fisher has not simply ended in a rejection of his model. In some of the subsequent works on this topic, the rebuttal of Fisher has transformed into a rejection of the idea that dominance modification and its subsequent evolution can play a major role in population dynamics. This rejection has taken different forms and is used in different contexts. Some of these works will be addressed in this article. Current opinions on the extent and manner in which dominance can evolve are mixed (for reviews with varying perspectives, see Nanjundiah, ’93; Keightley, ’96; Porteous, ’96; Mayo and Bürger, ’97; Bourguet, ’99; Falk, 2001; Cornish-Bowden and Nanjundiah, in press; Veitia and Bost, in press). Given the long history and different disciplines involved in this topic, it is to be expected that the disagreements can also extend to perceptions of this history. Nonetheless, it is clear that since the 1980s, the debate on whether and how dominance can evolve took a new direction. Based on theoretical investigations of metabolic pathways, Kacser and Burns (’81) posed the argument that metabolic systems possess inherent constraints that cause dominance of the wild type to be the default expectation. The corresponding interpretation was that microevolutionary dynamics is irrelevant with respect to the origin and manifestation of dominance. This argument has taken a prominent position in many subsequent discussions on evolution. Periodic disagreements notwithstanding, the overall trend has been to explain the manifestation of dominance as a default—and invariant—property of biological systems. The details of this trend will be discussed later in this article. However, it is useful to note from the outset that objections to dominance evolution can be separated into two main classes: population genetic and biochemical. In principle, neither approach negates the process of evolution. If one were to broaden the scope of the argument, one could consider that any property of an organism is a result of evolution, since the organism is the result of evolution. However, the relevant scope of the debate is at a finer level, and there is a distinction between the population genetic and biochemical objections. The population genetic objection has been that selection for dominance modifiers is ineffective due to unfavorable population conditions. In this case, there is no objection to the possibility that dominance can be modified. On the other hand, the biochemical objection is that dominance modification itself is limited by innate biochemical constraints. Nonetheless, in either case, the simplified argument is that dominance evolution does not occur within a microevolutionary framework.

The consequence has been to arrest the further development of one of the earliest attempts at understanding the evolution of inheritance systems. This is not to say that other attempts have not been made. The evolution of sex (division of sexes, recombination and random chromosomal segregation) is a prime example of an analogous query with a longstanding tradition in evolutionary biology (see for example Maynard-Smith, ’78; Otto and Lenormand, 2002). Not surprisingly, many of the modelling frameworks used in the evolution of dominance also arise in the evolution of sex. A good example is the use of modifiers to model the effects of selection on recombination rates (Nei, ’69; Feldman, ’72; Karlin and McGregor, ’74; for an overview, see Christiansen, ’99, p 305–331). Common to both the sex and dominance problems is the concern about how
The variational properties of a genetic system can evolve within a population genetic framework. Central to both queries, is the additional question of whether variational properties can be selected due to their adaptive advantage.

In addition to its effects on the variational properties of a population, there are two more factors that make the evolution of sex an important query in evolutionary biology. One is that to many biologists, the deleterious “cost of sex” could be an equally deficient approach which takes us back to the conceptual outlook of the “developmental mechanics” of the 19th century. Such a deficiency can hold true even when a reductionist approach is used as the foundation for a systems-level approach.

In order to escape the restrictions of classical population genetics, a necessary step is to develop a framework for how inheritance systems can evolve. However, it is not clear whether the resulting body of theory will be anything akin to classical population genetics. In this respect, Fisher may have been aware of some of the limitations of the theoretical framework to which he had been a major contributor. At the 1932 conference of the Sixth International Congress of Genetics (International Congress of Genetics, ’32) in Ithaca, New York, Haldane speculated that the immediate task for evolutionary biologists during the coming years would be to elucidate the effects of heredity on evolution. With the benefit of hindsight, one is tempted to agree with him; a cursory review of evolutionary theory in the 20th century corroborates his presentment. Fisher was not in disagreement, but he had a complimentary vision. He declared that he also wanted to look at the flip side of the coin: the effects of evolution on heredity (Fisher, ’32, p 165). From Fisher’s point of view, the evolution of dominance was the first step towards expanding the scope of population genetics for addressing the evolution of inheritance systems. However, one could argue that he may have overestimated the range of issues that could be addressed without modifying the machinery of population genetics. Nonetheless, the motivation for Fisher’s question deserves further attention. In many respects, the prolonged focus on Fisher’s 1928 model has detracted attention from his original question. In this vein, it is likely that Fisher is not given due credit for his pursuit—in the subsequent years—of a line of inquiry that is likely to be crucial in the further development of evolutionary theory. Whether he was able to produce the right tools for achieving this goal is another matter. However, it may be instructive to note that more than 75 years later, biologists have still not produced a well-developed body of theory that can address the evolution of inheritance systems. As stated earlier, this is not for a lack of trying, given that a problem such as the evolution of sex has been a central concern in evolutionary biology. Despite many advances, the extent to which the latter problem has been resolved is still undetermined (see for example Rice, 2002; Otto and Lenormand, 2002). Furthermore, given that dominance is conceivably a simpler problem than sex, one might argue that

---

4There are many arguments put forth on the disadvantages of sex. The most frequently posed argument is the “two-fold cost” of sex. Consider two variants of a species—sexual and non-sexual—with the same female fertility rates. The males in the sexual variant do not bear offspring while the asexual variant only produces offspring-bearing females. In this simplified case, the rate of population growth—and hence fitness—for the non-sexual variant would be twice that of the sexual one.

5Nanjundiah (’93) and Mayo and Bürger (’97) make a similar argument with respect to the possible resolution of this debate.
if we cannot resolve the question of dominance evolution, it is unlikely that we can resolve the evolution of sex either. The humbling fact is that dominance and sex are merely “entry-level” problems in relation to our understanding of how inheritance systems evolve. We are still far from a theoretical framework that provides the tools for dissecting the origin and evolution of inheritance systems. In this regard, queries on the evolution of development are in part a direct attempt at approaching the latter problem.

In this article, I revisit the issue of dominance evolution by reviewing the history of this topic. I attempt to outline the roots and nature of the empirical and theoretical concepts put forth throughout this debate. The purpose is to focus on some of the main points of contradiction between different hypotheses, and the possible paths to resolution. In doing so, I shall argue that the theoretical argument against the evolution of dominance—as presented in the realm of biochemical systems—is not valid. Furthermore, I discuss the fact that many of the older empirical studies on dominance modification directly contradict the idea that dominance is an invariant system property. I shall also discuss some of the possible population genetic conditions under which dominance can evolve. The aim is to reopen a debate, which in my opinion was prematurely cast aside as concluded.

EVOLUTION OF DOMINANCE AS A SCIENTIFIC DEBATE

Prelude to a debate: questions leading to the origin

On the Origin of Species (Darwin, 1859) is an abstract based on 20 years of empirical and conceptual work. Its legacy stands on two concepts: heritable variation and selection. The apparent simplicity of these concepts can be deceptive. The integration of these concepts into a scientific hypothesis was far from simple. In fact the dispersed and contradictory nature of biological understanding prior to the Origin can warrant the argument that natural selection was one of the most difficult intellectual achievements at its time. Nonetheless, overemphasizing the importance of natural selection can leave hidden intellectual scars. Natural selection can serve as a general framework for addressing why organisms are set up the way they are. But in its simplistic interpretation, natural selection can appear as the only theoretical component required for explaining the evolution of adaptations. However, we still require a set of theories that can include the processes that transform genotypic variation into phenotypic variation. Population genetics was only the first step, whereby Mendelian inheritance could be placed within the framework of Darwinian selection. The latter union provided the framework for including mutation and heredity. But it did not provide a scheme for the mechanistic rules that govern the production of new organizational variants.

The need for a conceptual framework that can represent organizational variation and complement natural selection is a task that modern biologists repeatedly face (Riedl, ’77; Gould and Lewontin, ’79; Raff and Kaufman, ’83; Maynard Smith et al., ’85; Buss, ’87; Kaufman, ’93; Fontana and Buss, ’94; Rose and Lauder, ’96; Wagner and Altenberg, ’96; Gerhart and Kirschner, ’97; Maynard Smith and Szathmáry, ’97; Solé and Goodwin, 2000; Oyama et al., 2001; Lewontin, 2002). In brief, three interrelated problems are:

1. What are the developmental constraints on phenotypic variation?
2. What are the mechanisms for the production of the type of variation that can lead to complex adaptations?
3. Are there internally driven processes in biological systems that can lead to self-organization?

These concerns roughly fall under the rubric of organization and evolvability. Question 1 has its early roots in morphology and comparative anatomy (Riedl, ’77). To a large degree, 2 and 3 are modern analogs of questions that motivated Lamarck (Shaner, ’27, p 254). The fact that the solutions envisioned by Lamarck were erroneous does not mean that the queries themselves were misdirected. Darwin was privy to all three problems and became more sensitive to 1 and 2 in his later years. From a neo-Darwinian perspective, it is ironic that successive editions of the Origin tried to address problem 2 by leaning towards Lamarckian views involving inheritance of acquired characters. By the time Darwin wrote the Descent of Man (1877), he also acknowledged problem 1 by discussing the importance of the
logical relations between different parts of an organism.

The early participants in the neo-Darwinian synthesis were well aware of the earlier versions of the problems listed above and toyed with possible answers. During the 1920s and early 1930s, neither evolutionary biology nor genetics were institutionalized as a field and the respective scientists were groping for research directions. Population genetics was in the uneasy position of premiere bastard-child of the two fields. Here was an approach that combined the historico-teleonomic perspective of Darwinian evolution with the mechanistic reductionism of Mendelian particulate inheritance. The former asks why organisms are set up the way they are while the latter asks how they are set up the way they are. Accordingly, here were individuals prone to population thinking such as Fisher, Wright, Haldane and Dobzhansky talking to mechanistic geneticists like Muller and Morgan. The neo-Darwinian synthesis had been possible precisely because the historical perspective was wed with a mechanistic one (Dobzhansky, '37; Powell, '87).

This interposition of why and how questions did not have to stop at the synthesis. One may argue that in biology the two always go hand in hand. In 1932, the person who pressed the why question most was Ronald Fisher. This is ironic, because as the founding father of population genetics, he is regarded as the one who consolidated mechanistic reductionism into evolution. Fisher’s focus on treating selection within a deterministic framework, while disregarding some of Wright’s ideas on stochastic effects such as drift, may have contributed to this image.  

The question of dominance evolution: an abortive attempt at “evo-devo”

Fisher was struck by the observation that in diploid organisms a large proportion of mutant phenotypes are recessive with respect to the wild type (Fisher, ’28a,b, ’29, ’31, ’34, ’58). As examples, he referred to the work by Morgan et al. (’25) on Drosophila, and his own observations with Ford on melanic moths (Fisher and Ford, ’26; Fisher, ’27). Prior to Fisher, Tower had shown that dominance relations could change according to changes in environmental conditions (Tower, ’10). Bridges encountered a case of dominance modification due to changes in the genetic background (Bridges, ’13). Later, it would also become apparent that in some cases the phenotypic effects of a given allele could be dominant in relation to one trait and recessive with respect to another (Ford, ’30; Caspari, ’50). Hence, what was dominant was not necessarily the gene but rather the phenotypic effect (Wallace, ’68).

Over the years disagreements have emerged concerning the topic of dominance. The main issue has been the common observation that wild-type phenotypes are frequently dominant with respect to their less fit mutant counterparts (Wilkie, ’94). The main point of contention has been whether the dominance of the wild-type phenotype is an evolutionary adaptation that confers robustness against deleterious mutations, or whether it is simply an inherent property of genetic systems (Keightley, ’96; Porteous, ’96; Mayo and Bürger, ’97). A third possibility is that dominance is a side effect of selection for other properties.

At its core, the question of dominance evolution is a question about the evolution of mutational effects. As such it has direct relevance to the problem of the evolution of development: how can selective forces change the relationship between genotype and phenotype? Accordingly, the problem of robustness would also be taken up by developmental biologists and referred to as canalization (Waddington, ’42; Schmalhausen, ’49; Rendel, ’67; Scharloo, ’91; Gavrilets and Hastings, ’94; Stearns and Kawecki, ’94; Stearns et al., ’95; Wagner et al., ’97; Eshel and Matessi, ’98; Rice, ’98; Hartman et al., 2001; Nijhout, 2002; de Visser et al., 2003).

Fisher postulated that dominance could evolve via the selection of alleles at a modifier locus which could diminish the detrimental effects of mutant alleles at a primary locus. Up to this point, the logical structure of population genetics consisted of a feedback cycle, from genotypes to phenotypes, and back to genotypes again via selection. In other words, population genetics focused on the feedback effects of selection on gene frequencies and, as a consequence, genome composition. By considering the evolution of mutational effects, Fisher was bringing into play the evolution of gene interactions (i.e., epistatic interactions). Figure 1 summarizes this perspective. In its logical structure, such a framework is equivalent to part of what “developmental evolution” intends to address: the evolution of the genotype–phenotype map (Wagner and Altenberg, ’96). However, the seeds of controversy had already been sown prior

7For an authoritative history of the relevant issues, see Provine (’86).
to 1932. Wright and Haldane had objected to Fisher’s 1928 model on the evolution of dominance. Their calculations showed that selection pressure for modifier alleles would be very small and on the order of mutation rates (Wright, ’29a,b; Haldane, ’30, ’39). Hence, selection would not be much more effective than drift in the evolution of dominance. Given the conditions of Fisher’s model, their criticism is correct (see for example Feldman and Karlin, ’71; Wagner and Bürger, ’85). Nonetheless, as Provine (’86) notes, a central component of the disagreement was not necessarily the low selective coefficients for dominance modifiers. Rather, it was the difference of opinion on the potential for drift—or other forces—to overpower such weak selection for dominance modification. For Fisher (’29, p 555), the antidote to weak selection was to extend the time period for selection. Wright (’34a, p 51) acknowledged Fisher’s solution, but questioned its validity. The ensuing debate resulted in a public enmity between Fisher and Wright. By 1934, they appeared not to be on speaking terms any more (Provine, ’86).

### Early empirical evidence on dominance modification and the unspoken rift between physiological and population perspectives

During the early years of the discussions on dominance, Haldane and Wright had also argued that dominance in any form should be attributed to the underlying properties of the physiology. However, they fully accepted the possibility that differing degrees of dominance could evolve via modification of the physiology. In fact, in a series of experiments on guinea pig coat color, Wright had observed dominance modification effects in what he determined to be a seven-locus system (Wright, ’27). During that time, several other geneticists were beginning to grapple with experimental results that indicated that dominance effects could be modified by the genetic background (Bridges, ’13; Castle and Phillips, ’14, ’19; Jennings, ’17; Lancefield, ’18; Timofeeff-Ressovsky, ’27).

Haldane and Wright postulated that an increase in biochemical reaction rates could lead to a “factor of safety” against underlying perturbations (Haldane, ’30, ’39; Wright, ’34a,b, ’77). The idea of a safety factor was already floating around at the time. Haldane’s arguments were inspired by his exchanges with Goldschmidt. As a World War I exile working at Yale’s Osborn Zoological Laboratories, Goldschmidt (’16, ’17) had been trying to establish a link between what he termed as “quantitative combinations of factors [genes],” “quantity of enzyme reactions” and melanic wing coloration in moths. By the time of Haldane’s proposal, Muller (’32) and Plunkett (’33) were proposing similar ideas (see also Forsdyke ’94). In Muller’s version, the “factor of safety” was mainly arising as an evolutionary response to environmental perturbations rather than as a result of selection for dominance effects.

Once again, as in the case of the neo-Darwinian synthesis, there was an explicit desire to link a mechanistic perspective to an evolutionary one: the interplay of physiological constraints and evolution. Unfortunately, such a union did not materialize. In the coming years, a rift developed between some of the population and physiological perspectives on genetics and their respective view on dominance. Part of the reason was the logical independence of the questions posed by the two approaches. Population genetics was geared towards answering the following type of question: given a specific pattern of phenotypic variation and population conditions, what would be the outcome of evolution? The physiological viewpoint

---

8Figure 2 showcases an example in which Timofeeff-Ressovsky (’27) observed dominance modification for the manifestation of the “radius incompletus” mutation in *Drosophila* wings. His stock had been derived from a “wild population” that lived on “rotten potatoes in one of the buildings near the Hydrophysiological Station of the Institute of Experimental Biology, in the Zvenigorodsky Department of the Moscow Government.”

9Goldschmidt had developed an early interest in the problem of dominance modification. As a zoology professor at the University of Munich, in his *Einführung in die Vererbungswissenschaft* (Goldschmidt, ’11), which was an introduction to “inheritance science” for “students, doctors and breeders,” he had devoted a major part of a chapter (p 246–265) to reviewing evidence on “fluctuating and changing dominance.”
was focused more at the level of the individual: given a particular physiological organization, what would be the phenotype and what kind of variation could it exhibit? If in mutual isolation, it is possible for the two approaches to yield conclusions that are diametrically opposed. In one case, a physiological approach may indicate that a particular pattern of variation can include a given target phenotype. Subsequently, one may be tempted to conclude that the target phenotype can be selected. Meanwhile, in the same case, the population genetic approach may indicate that within a given set of population conditions, the target phenotype could not be effectively selected. To aggravate the situation the converse is also possible. A population genetic approach may determine that with a hypothetically postulated pattern of variation, a particular phenotype can be selected in a population. But then the physiological approach may indicate that the postulated pattern of variation is not possible for the given phenotype. The conclusion is almost inevitable: for an understanding of evolution, the physiological and population perspectives have to fit together.

The genetic views on the evolution of dominance problem would take two forms. One form had a physiological guise and another a population genetic one. Following Wright and Haldane’s criticism of Fisher’s model, the majority of work up to the 1970s corroborated and elaborated on the population genetics aspects of Wright and Haldane’s criticism. The consensus was that selection for the dominance modifier could not be much more effective than drift (Ewens, ’66; Sved and Mayo, ’70; Feldman and Karlin, ’71). All such models had one thing in common. They all assumed equilibrium starting conditions in which the frequency of heterozygotes at the primary locus was on the order of mutation rates and that the wild-type allele frequency at the primary locus had approached fixation. But these conditions did not capture the whole story. The situation could be different if non-equilibrium starting conditions were to be considered. Among the first to seriously consider this possibility was Haldane (’56). We shall return to the non-equilibrium arguments shortly.

In contrast to the population genetic perspective, the physiological perspective (in this case, combined with artificial selection experiments) indicated that dominance levels can be subjected to artificial selection. During the 1930s, Fisher (’38) conducted a series of breeding experiments using domesticated poultry that possessed well-known dominant phenotypes. He had been directly influenced by Castle and co-workers, who between 1907 and 1919 had been carrying out a series of artificial selection experiments on rats (Castle and Phillips, ’14; Castle, ’19). By 1914, Castle’s team had examined more than 25,000 rats in the laboratory. Castle had selected for modifications of a recessive pigmentation phenotype. He was able to establish distinct lineages with either decreased or increased manifestations of the mutant phenotype (in the mutant homozygotes). He could further break down this modification by backcrosses into a wild-type population. Consequently, Fisher was interested in morphological and pigmentation phenotypes which had been presumably selected by breeders.
during domestication. He found that when he outcrossed domesticated poultry to their ancestral wild jungle fowl *Gallus gallus*, the degree of dominance for some phenotypes was reduced to co-dominant levels (Fisher, '35, '38).

However, on this occasion, even though he could detect dominance modification, his results were inconclusive with regard to the effects of selection. One problem was that he was not able to develop a successful breeding program. From a historical point of view, these experiments are interesting because they mark Fisher’s shift to an experimental approach for addressing dominance evolution. They also indicate his interest in pinpointing developmental characteristics that could be quantitated and related to mutational effects. Figure 3 shows an example in which he was counting the number of hallux bones in order to address dominance for polydactyly. In this case, we see a direct link between dominance evolution and the evolution of development.

Fisher was not alone in observing dominance modification. The fact was that the decade of the 1930s was witness to an astounding barrage of experimental results that did not conform to additive notions of Mendelian inheritance. By that time, experimenters had some benefit of hindsight and a clearer idea of the pitfalls of genetics. Consequently, they were more prepared to detect modification effects. Dominance modification was being reported for morphological and pigmentation phenotypes of some of the staple genetic subject-animals of the time: moths, flies, mice and poultry (Snell, '31; Crew and Lamy, '32; Lebedeff, '32; Barrows, '34; Dubinin and Sidorov, '34; Dunn and Landauer, '34; Green, '36; Goldschmidt, '88, p 99–123; Helfer, '39). Gene interactions were rampant, mutational effects were not invariant and dominance modification was showing up everywhere (see Figs. 4 and 5 for examples).

A key evidence for modifiability and susceptibility of dominance to selection was to come from Ford’s ('40) artificial selection experiments (presented at the Seventh International Congress in 1939). Ford had started with a population of the moth *Abraxas grossulariata* whose heterozygotes exhibited a wing coloring that was intermediate between wild type and a mutant phenotype called *lutea*. Through artificial selection, Ford could produce separate lineages in which *lutea* could be either dominant or recessive with respect to the wild type. Dominance could also be broken down by backcrossing into the original wild-type population. Impressively, Ford could achieve the dom-

![Fig. 3. Dominance modification for polydactyly, a dominant mutation in domestic poultry. Fisher was able to reduce polydactyly to co-dominant levels by outcrossing to jungle fowl. (A) Frequency distribution of number of hallux bones and its relation to genotype in an outcrossed population. Distribution of heterozygotes (shaded in gray) reflects co-dominance. Given that polydactyly was dominant in domestic poultry, the distribution of heterozygotes in the domestic case would be presumably shifted to the right (Fisher does not show such data). (B) Examples from outcrossed individuals from the distribution in (A). Left column: examples of polydactyl homozygotes; right column: heterozygotes. Variation is exhibited in the number of hallux bones (lowest digit in diagrams, with a forked appearance due to polydactyly). Bones pertaining to digit IV are not shown (both (A) and (B) from Fisher, '35, p 221).](image)

11This paper is also a good counterexample to the notion that those who proposed that dominance could be modified and selected were somehow ignoring the underlying biochemical nature of gene action. In this paper, Ford ('40) goes to the extent of including a figure for the molecule responsible (apigenin) for the pigmentation phenotype he studied (Ford and Fisher were collaborators and maintained correspondence). In effect, the oft-hailed notion that
another artificial selection experiment, Fisher and Holt ('44) were successful in dominance modification for a mutation that reduced tail length in mice (Fig. 6). This was probably the most compelling experiment on dominance modification and selection with which Fisher was involved. 

Despite the accumulated evidence, the question still remaining from the perspective of theoretical population genetics was whether the conditions existed in natural populations for the selection of dominance modifications. Later in the 1950s and 1960s, Kettlewell's ('55, '61, '65) work on the evolution of industrial melanism in the moth *Biston betularia* showed some evidence that the observed dominance of the melanic form had changed during evolution. For one thing, the dominance has to do with biochemistry and not evolution can ring somewhat hollow. The issue was whether natural selection could mold the underlying biochemistry. The role of biochemistry in gene action was not doubted.

Note that there are some controversies regarding Kettlewell's hypothesis that moth coloration affects predation, and hence fitness (Majerus, '98). Nonetheless, although the topic deserves ...
heterozygotes of post-industrial specimens did not show an intermediate phenotype and were darker than those observed in collections of pre-industrial melanics (whether homozygote or heterozygote). In addition, work by Ford and Sheppard (’69) on the moth Panaxia dominula and by Kettlewell (’73 p 181–193) on the moth Amathes glareosa indicated that there could also be geographic variation in the degree of dominance for pigmentation phenotypes. However, the more telling evidence was in the experiments of Ford (’55) on Triphaen comes and of Kettlewell (’65) on Biston betularia. They showed that the dominance of the melanic forms of the moths could be broken down by outcrossing into populations where the melanic form was not prevalent (i.e., outcrossing into populations where the melanic form had presumably not been selected). Even more impressive was that dominance could be reinstated by backcrossing into the original genetic background (Kettlewell, ’65). The implication was that in the melanic population the genetic background contained dominance modifiers.13 These modifiers were presumably present as a result of selection during the evolution of melanism.14

Cases with a high frequency of heterozygotes: balanced polymorphisms and transient selective sweeps

Given the initial objections to Fisher’s model, one obvious solution would be to find cases where the frequency of heterozygotes would be relatively high. In such cases, selection for modifiers would not be weak anymore. One simple scenario where heterozygotes could be maintained in the population would be in populations with balanced polymorphisms. One of the simplest scenarios in which such polymorphisms could be maintained would be in the case of overdominance (i.e., heterozygote advantage) (Fisher, ’30, ’31; Ford, ’30; Sheppard, ’59, first published ’58, p 136–145). Under such conditions, dominance could evolve

more careful study, Kettlewell’s main conclusions seem to remain intact (Grant, ’99; Mani, ’99).

13For an interesting study in plants with similar implications, see Doebley et al. (’95).

14Interestingly, dominance breakdown and buildup could not be observed in some other species examined later (West, ’77; Mikkola, ’84). This could mean that in some populations the modifiers could be linked or that dominance does not always have to evolve through modifier selection. One possibility is that in some cases dominance can be modified due to mutations at the primary locus itself. See also Mayo and Bürger (’97).
via the modification of one of the homozygote phenotypes towards the heterozygote.\textsuperscript{15} Another type of balanced polymorphism are ones in which the heterozygote is at a disadvantage to the two homozygotes. In this case, the heterozygote would evolve to resemble one of the homozygotes. The latter types of polymorphisms are usually unstable, unless special conditions exist for their maintenance (e.g., frequency-dependent selection). A well-studied example is a polymorphism between cripsis and mimicry in the butterfly \textit{Papilio dardanus}, where evidence indicates that dominance effects had changed in natural populations (Clarke and Sheppard, ’60a, b, ’63; Clarke and O’Donald, ’64; Sheppard and Ford, ’66; O’Donald and Barrett, ’73; Ford, ’75; Mayo and Bürger, ’97). In general, further theoretical studies have reached the conclusion that dominance evolution is possible under balanced polymorphisms (Feldman and Karlin, ’71; Charlesworth and Charlesworth, ’75; Bürger, ’83c; Otto and Bourguet, ’99).

One recent set of studies is particularly indicative of the possibility that dominance modifiers can be selected when heterozygotes are maintained in a population (for a concise review, see Bourguet, ’99). For the case of insecticide resistance in the mosquito \textit{Culex pipiens}, Bourguet et al. (’96) found that dominance for insecticide resistance can vary with environmental conditions. They also found that dominance levels could vary depending on the resistance allele (Bourguet et al., ’97). Use of the Kacser and Burns (’81) framework led to conflicting results (Bourguet and Raymond, ’98). Subsequently, Otto and Bourguet (’99) proposed a population genetic framework for the evolution of dominance due to balanced polymorphisms. A novel aspect of the latter work is the treatment of spatial heterogeneity (with respect to selection) as a cause for the maintenance of balanced polymorphisms. Hence, spatial aspects of ecology are brought to bear on the problem of dominance evolution (see also Van Dooren, ’99). This is an avenue that deserves further attention.

We shall now turn to an argument that Haldane (’56) considered midway through the debate. It revolves around non-equilibrium conditions. In the 1950s, industrial melanism was considered to be one of the best candidate cases for the evolution of dominance in nature. Haldane’s argument was that after a change in environmental conditions, such as the advent of industrial pollution, a selectively advantageous allele such as that for melanism could be initially at a low frequency. Subsequently, during selection for the newly advantageous allele, there would be a transient period of high heterozygosity at the primary locus. This transient period could be propitious for modifier selection. Nonetheless, even though Haldane had proposed this possibility, in the end he was not certain of the efficacy of modifier selection during the selective sweep. He left the matter as an open question for the future. In the meantime, he opted for balanced polymorphisms as a more likely option. Parsons and Bodmer (’61) agreed that selection for the modifier would be more effective if the heterozygote were to be maintained under a balanced polymorphism. They also argued that linkage between the modifier and the primary locus would also increase the effectiveness of modifier selection (Bodmer and Parsons, ’62). Nonetheless, disagreements persisted on the degree to which non-equilibrium conditions could improve modifier selection (Ewens, ’66).

An understanding of the far-from-equilibrium dynamics was further extended during the early 1980s. Working at the University of Vienna, Wagner and Bürger re-examined this scenario. They approached the problem using a combination of nonlinear analysis and numerical simulations. They showed that during a selective sweep for a beneficial allele at the primary locus, it is concurrently possible to select for the modifying allele at the modifier locus (Wagner, ’81; Bürger, ’83a–c; Wagner and Bürger, ’85). They also confirmed that linkage could aid the process. In summary, the population genetics showed that in non-equilibrium conditions, dominance can evolve via selection for the dominance effects of a modifier allele. The caveat was that initial conditions had to be just right. Selection for the modifier would only occur when the frequency of the wild-type allele at the primary locus was still low. Furthermore, this implied that towards the end of a selective sweep for the primary allele, selection for the modifier allele would become slow again. This meant that after completion of the selective sweep at the primary locus, fixation at the modifier locus would once again depend on drift. Nonetheless, along with balanced polymorphisms, such high heterozygosity scenarios were the strongest hypotheses at the time for how
dominance could evolve. This view did not gain widespread attention in evolutionary circles. One reason may have been that consideration of modifiers in far-from-equilibrium starting conditions was not part of the mainstream repertoire of queries at the time. Secondly, arguments by theoreticians in metabolic physiology were starting to point in a very different direction (see next section).

**Metabolic control analysis and biochemical systems theory**

In the first half of the dominance debates, the population genetic view had been that dominance cannot be selected according to Fisher’s postulated conditions. By the early 1980s, some of the alternative conditions under which dominance could be selected had been elaborated. The trend in physiological conceptions was in a somewhat different direction. In the beginning, it was thought that dominance can be modified and selected. This view was mainly supported by artificial selection experiments. By the 1980s, a new generation argued that dominance was an “inevitable” property and that evolutionary explanations were unnecessary. This argument was first posed by Kacser and Burns (‘81), who while at the University of Edinburgh had pioneered a theoretical approach for investigating the properties of multi-enzyme pathways (Kacser and Burns, ’73; Kacser et al., ’95). Their particular approach, in unison with independent developments by Heinrich and Rapoport (’74), developed into a field referred to as metabolic control analysis (MCA) (for textbooks, see Heinrich and Schuster, ’96; Fell, ’97).

Kacser’s take on dominance is particularly interesting, given that earlier in his career, he had written an appendix on the biochemical underpinnings of biological organization in Waddington’s (’57) book on the *Strategy of the Genes*. Waddington (’42) had been previously exploring hypotheses on the evolution of canalization (of which dominance is a simple form (Rendel, ’67)). However, Kacser would eventually take a very different approach (in comparison to Waddington) on the causes of phenotypic robustness. Waddington’s view was that canalization could be selected, while Kacser would later conclude that dominance was inherently expected and not a result of microevolutionary dynamics or selection.

One of the physiological phenotypes being addressed using MCA was the steady-state flux (the rate at which metabolites are produced in a pathway). From their analysis, Kacser and Burns concluded that dominance is an invariant and *ipso facto* property of multi-enzyme pathways and that explanations based on dominance modification and selection were not required. The logic of Kacser and Burns’ argument and the associated claims will require some explanation. Readers who are interested in the mathematical aspects of their argument should refer to the Appendix of the present article. A verbal version of their argument is given in three parts below:

(a) “Changes in enzyme concentrations have small effects on flux.” At the core of the MCA approach was a mathematical proposition referred to as the “flux summation theorem.” One of its implications was that the higher the number of enzymes in a pathway, the smaller the average phenotypic effects (on flux) of mutations that changed underlying enzyme concentrations. This meant that most enzyme dosage effects would have a small phenotypic effect. Furthermore, the summation theorem implied an invariant relationship, whereupon if flux was to become sensitive to the concentration of a given enzyme, it would also become less sensitive to some other enzyme (see Fig. 7, with underlying model in the Appendix).

(b) “Small effects on the flux phenotype translate to dominance of the wild type.” Based on some observations at that time, it was argued that in a large number of mutant heterozygotes, the
mutated allelomorph leads to a less active or inactive copy of the enzyme. Such heterozygotes would have a lower active gene dosage and consequently lower enzyme dosage. Since the flux summation theorem postulated that the flux effects of the majority of enzyme dosage changes have to be small, it implied that wild-type flux phenotypes (with full dosage of underlying enzyme) would be similar to the heterozygote (intermediate dosage) and hence would be dominant with respect to the mutant homozygote (very low or zero dosage).

(c) “Dominance is an invariant property of metabolic systems.” Kacser and Burns’ models of metabolic pathways exhibited a low-gradient flux plateau that was consistent with their flux summation theorem. More importantly, experimental work seemed to confirm the plateau effect in the model. The fact was that the aesthetic of their argument was similar to Haldane and Wright’s “factor of safety.” The difference was that Haldane and Wright thought that the factor of safety was an evolved property while Kacser and Burns claimed that it was an invariant system property. The main reason for the difference was that the flux summation theorem represented an invariant relation, which implied that the general expectation of dominance would be invariant.

In sum, the predominant view that originated from MCA was that the predisposition of a specific mutation to be dominant or recessive is an invariant property of enzyme-catalyzed reactions in metabolic pathways (Kacser and Burns, ’81). Hence it was believed that for a given sequence of reactions, dominance relations for the whole system—as characterized by the summation theorem—cannot be altered (Keightley, ’96; Porteous, ’96). Two objections have been posed against the conclusions from MCA, but they have not been generally accepted in MCA or evolutionary biology. In the first place, Cornish-Bowden (’87) showed that in a sequential pathway if the maximal rate $V_{\text{max}}$ of consecutive enzymes is sequentially decreasing, then dominance is not a necessary property of the pathway. Hence, the possibility exists that dominance can evolve due to selection. This objection has been rejected in MCA on grounds that the specific arrangement of kinetic values suggested is a special case that is very unlikely to occur by chance (Kacser, ’87). A more general set of objections was then forwarded by Savageau and Sorribas. They argued that in pathways exhibiting nonlinear behavior that

17In the literature on this topic, alternative terms such as inevitable or inherent consequence of physiology have also been used. Although such terms are certainly acceptable when contextualized, they can sometimes be too vague and their usage has been generally avoided here.

18This citation is another example where mapping the belief system within scientific communities can enter gray areas. Despite the original objection by Cornish-Bowden (’87), in a subsequent work Cornish-Bowden (’89) defends what he deems to be the strengths of the MCA approach and does not pursue the dominance issue any further. In a more recent article, there is a conciliatory attempt at satisfying both viewpoints (Cornish-Bowden and Nanjundiah, in press).
arises from properties such as enzyme–enzyme interactions, feedback loops or non-sequential pathway structure, it can be shown that dominance is not inherent to the pathway (Savageau and Sorribas, ’89; Savageau, ’92). Savageau and Sorribas’ objections were rejected by most MCA proponents at the time. We shall return to this shortly.

MCA was an outgrowth of theoretical work by Higgins (’63) on the properties of sequential biochemical reactions. The origins of the MCA field itself were geographically divided between the United Kingdom (Kacser and Burns, ’73; Fell, ’97) and the former East Germany (Heinrich and Rapoport, ’74; Heinrich and Schuster, ’96). This division was due to the independent discovery of precepts, which were later referred to as MCA. However, the community that predominantly participated in the debate on dominance was centered around the UK group.19 In the meantime, MCA was not the only methodology that was being developed for addressing the behavior of multi-enzyme systems. During the late 1960s, a different methodology with overlapping aims had developed in the United States. Michael Savageau, based at the University of Michigan, pioneered a mathematical approach that is referred to as biochemical systems theory (Savageau, ’69a,b, ’76; Voit, ’91). The objective of biochemical systems theory, in the same vein as MCA, was to address the behavior of multi-enzyme systems. However, the biochemical systems theory methodology was more geared towards analyzing nonlinearities in metabolic systems. As a consequence, by using the latter methodology, it was more natural to find scenarios in which dominance was not an expected property of metabolic systems. In this regard, it is possible that the rejection of Savageau’s objections had as much to do with rivalries between communities as the logical content of the arguments put forth by Savageau and his colleagues (see for example Kacser, ’91; Savageau, ’92). Part of the disagreement may have also been caused by the mathematical formalism used in metabolic systems theory. To understand the latter approach, one has to first become familiar with a set of mathematical transformations that are commonly referred to as S-systems. This requirement may have acted as an unintentional barrier to entry for some. Furthermore, even though the results obtained with an S-system transformation can be valid, an understanding of the general characteristics of the S-system methodology is more complicated. Consequently, the alternative perspective afforded by biochemical systems theory may not have been given the due attention it deserves in some of the subsequent debate.20

Towards the end of his influential career, Kacser made his stance towards critics explicit:

There is no serious disagreement among those who have taken up and extended the subject (the so-called “controlniks”), while those who refuse, for a variety of reasons, to adopt the new insights (the so-called “refuseniks”) continue to linger in the outer darkness. (Kacser, ’95, p 388)

Such oratory may have seemed charming or entertaining at the time. As evidenced by the many acknowledgements and dedications that one can find in relevant papers, Kacser appears to have been well-liked and respected by many of his colleagues. Nonetheless, one cannot help but raise the possibility that the starkly confrontational dynamics that he was part of may have hindered a re-examination of the MCA perspective on dominance. In several of the writings that Kacser was associated with, there is an unwavering projection of confidence with regard to the general applicability of MCA precepts:

Discussion and experiment in the absence of an understanding of these summation and connectivity properties can only proceed in a kind of intellectual vacuum. [...] the modern metabolic control analysis outlined here has been applied experimentally to the quantitative elucidation of metabolic control in systems as diverse as isolated mitochondria, hepatocytes, yeast, erythrocytes, Neurospora crassa, mice and Drosophila. In each instance fruitless controversies arising from the purely qualitative, intuitive, teleological or metaphorical approaches of the past have been resolved. (Kacser and Porteous, ’87, p 14)

The MCA perspective on dominance was believed to be supported by a body of experimental work which indicated that control coefficients of

---

19Interestingly, there is no indication that Heinrich and Rapoport supported Kacser and Burns’ conclusions on dominance. In effect, Heinrich asserts that he did not subscribe to Kacser’s conclusions on dominance (personal communication).

20For a more recent study, with diligent attention given to physiological detail and parametrization, see Salvador and Savageau (2003).
enzyme concentration with respect to flux were generally low (for a review, see Fell, '92). This meant that most changes in enzyme concentrations had a small effect on flux. These studies essentially corroborated the fact that dominance could be very prevalent, but they did not specifically address whether or how such dominance could be modified throughout the system. Hence, what these studies did was to corroborate that in many organisms, metabolic flux was robust with respect to enzyme concentrations. However, they did not address whether such robustness was a result of selection, nor did they address the extent to which it could be modified.

At this juncture, it is relevant to consider an analysis by Orr ('91), which is frequently cited in support of Kacser and Burns’ perspective. The alga *Chlamydomonas rienhardtii* spends most of its life cycle as a haploid. However, there is a brief diploid zygotic state. Furthermore, zygotes occasionally divide mitotically, leading to vegetative diploids. Orr observed that most known mutations were recessive when observed in the diploid stages. Since the organism spends most of its time in a haploid stage, he argued that dominance could not be a result of selection for dominance modification effects and hence that “most mutations are, from the beginning, recessive.” The latter, is in essence, the “invariant expectation” argument posed by Kacser and Burns. Orr’s ('91) deduction is highly original, but strictly speaking, the results indicate that in the haploid yeast, dominance could not have evolved according to Fisher’s population genetic scenario. However, reliance on this work as a confirmation of the Kacser and Burns ('81) theory is unwarranted. From a logical standpoint, results from Orr’s work can be used to reject Fisher’s model in yeast, but they do not verify Kacser and Burns’ model.21 Furthermore, in the *C. rienhardtii* case, the inevitability of dominance argument neglects the fact that the underlying basis for the evolution of dominance is phenotypic robustness to mutation. Dominance is a specific manifestation of the latter phenomenon in diploids. However, there are no rules that presuppose that phenotypic robustness to mutations cannot be modified or evolve in haploids. If phenotypic robustness to mutations (canalization) is an evolved property in *C. rienhardtii*, we would still observe Orr’s ('91) results. Hence, one has to address whether the robustness that is leading to dominance is an inevitable (i.e., fixed) property or the consequence of an evolved property.22

Meanwhile, in evolutionary microbiology, the work of Dykhuizen, Hartl and Dean corroborated the plateau effect proposed by both the Wright/Haldane and Kacser/Burns conceptions of metabolic physiology (Dykhuizen and Hartl, '80; Hartl et al., '85; Dean et al., '86). Their experiments involved *Escherichia coli* populations whose fitness was dependent on nutrient flux. They showed that mutations that lead to changes in enzyme activity lead to fitness changes along a low-gradient plateau. The first impact of these experiments was to advance the idea of “evolved selective neutrality.” The original intent of these experiments was not to test the Kacser and Burns dominance theories (the microbes are haploid). Nonetheless, these experiments have been sometimes interpreted as lending support to the “invariant expectation” of dominance argument. However, these experiments can also be interpreted as lending support to Haldane’s “factor of safety,” which is in effect closer to the notion of “evolved selective neutrality.”

In the subsequent years, the position espoused by MCA proponents was that in the realm of biochemical physiology, the dominance issue had been largely resolved (Keightley, '96; Porteous, '96). Accordingly, a significant proportion of works in many branches of evolutionary biology have been influenced—to varying extents—by the assumption or interpretation that dominance effects are an invariant property of biochemical systems, or that they can be largely attributed to physiological causes (see for example Keightley, '89, '96b; Clark, '91; Orr, '91; Szathmary, '93; Turelli and Orr, '95; Hartl and Taubes, '96; Solé and Goodwin, 2000; Hartman et al., 2001; Nijhout, 2001; Meiklejohn and Hartl, 2002; Siegal and Bergman, 2002; Papp et al., 2003; True, 2003; Kondrashov and Koonin, 2004).23

21If the Kacser and Burns hypothesis were true, we would observe the Orr ('91) results. However, the veracity of Orr’s results does not necessarily imply the Kacser and Burns hypothesis. In other words, there is no if and only if relation between the two propositions.

22See Mayo and Bürgér ('97) for further arguments regarding Orr’s ('91) results.

23A convenient place to start is an entire issue of *J Theor Biol* (182(3), 1996), which is dedicated to Kacser and the MCA perspective (this was a dedicatory issue published after the death of Dr. Kacser).
The case of metabolism and the invariant expectation of dominance reconsidered

I believe that no one who is familiar, either with mathematical advances in other fields, or with the range of special biological conditions to be considered, would ever conceive that everything could be summed up in a single mathematical formula, however complex. (Fisher, ’32, p 166)

The dismissal of dominance evolution in the context of microevolutionary dynamics was not devoid of contradiction. For one thing, there seemed to be a disconnect between the biochemical views on dominance and the genetic views on dominance. Genetic and artificial selection experiments showed that dominance effects could be modified. Nevertheless, the view from MCA was that dominance was an invariant and expected property of a multi-locus system. There is a logical disconnect between the two positions, given that at some point the dominance modification seen in the genetic experiments had to have some biochemical underpinnings. If the genetics did not have a biochemical underpinning, why would the biochemical argument matter in the first place? The only way in which one could reconcile the two views is to conceptualize that as dominance is modified in the genetic experiments, the modification is being counteracted by a change in sensitivity to mutations at some other loci. However, the genetic and artificial selection experiments were not designed to test such a notion.

Despite some of the unresolved contradictions, the majority view was to combine the results from MCA and equilibrium population genetics to conclude that dominance could not evolve within the classical population genetic context. This fusion was so effective to the point that in the later literature on this topic, the viewpoints of Kacser and Burns on the one hand and Wright and Haldane on the other were viewed as the same. This was an erroneous interpretation. Wright and Haldane thought that dominance can have a physiological underpinning. But contrary to Kacser and Burns, they did maintain that it could be modified and evolve through mechanisms other than Fisher’s (Wright, ’27, ’29a,b, ’34a, ’77; Haldane, ’30, ’39, ’56). Hence, if anything, the Wright and Haldane views were in contraposition to the Kacser and Burns view. Simply focusing on the fact that both parties claimed that dominance is a “consequence of physiology” is not sufficient to put all parties in the same corner. One would even be hard pressed to argue that Fisher had a problem with the notion that physiology mediated the manifestation of dominance. In fact, he acknowledged Wright and Haldane’s physiological models as a possible underpinning for dominance modification (Fisher, ’31, p 358–359). What Wright and Haldane were asking was: how do population level processes interact with this physiological substrate? The perspective of Kacser and Burns (’81, p 640) on the other hand was that:

the recessivity of mutants is an inevitable consequence of the kinetic properties of enzyme-catalyzed pathways and that no other explanation is required.

Their view on Haldane and Wright’s perspective was:

there is one critical difference between Wright’s conclusions and ours. Although he correctly suggested a hyperbolic relationship between enzyme and flux, his treatment did not explain why the majority of enzymes should lie on the plateau of the relationship. Having rejected Fisher’s hypothesis of modifiers, he came down on Haldane’s (’30) and Plunkett’s (’33) selection for “safety factors,” that is, for increased activity of the wild type allele at the locus. The summation property eliminates the necessity of postulating selection to bring enzymes into such a position. (Kacser and Burns, ’81, p 664)

There are factors that may have subsequently contributed to the confusion of the different positions in the debate. One is that interpreting or prioritizing the different hypotheses espoused by different actors in the early stages of a debate is not a straightforward task. Another is that terms like “factor of safety,” “inevitable consequence” and “consequence of physiology” possess a certain degree of looseness and the meaning ascribed to them can vary depending on the author or reader. Furthermore, judging from successive publications, the actors in the debate change their emphasis as they try to gauge which lines of inquiry or hypotheses are the correct ones to pursue. Wright’s conception of the physiological

---

When Kacser and Burns refer to Wright’s work, they are almost certainly referring to his 1934 work entitled “Physiological and evolutionary theories of dominance,” published in Am Nat 68:24–53. They cite the latter work as “Molecular and evolutionary theories of dominance,” Am Nat 63:24–53.
basis of dominance provides a good example. In his first response to Fisher, he states that:

it is easy to show that increase in the activity of a gene should soon lead to a condition in which even doubling of its immediate effect brings about little or no increase in the ultimate effects. (Wright, '29a, p 278)

These are the beginnings of Wright’s biochemical conception, though he had yet to present a specific model. In a later work, he further states:

The suggestion that mutations most frequently represent inactivations of genes, and that, for simple physiological reasons, inactivation should generally behave as recessive, still seems adequate as a positive alternative hypothesis. (Wright, '29b, p 561)

It is possible for anyone who is focusing on Wright’s physiological views in his 1929 articles to think that his views are, in principle, the same as the MCA position. However, we should note that he also accepted the concept of dominance modification. In fact, Wright (’29a, p 277) conceded that Fisher’s model would work if the frequency of heterozygotes were high in a population. His physiological conception becomes clearer when he finally presents a model (Wright, ’34a) based on a simple set of differential equations that represent “the relation between amounts of catalyst and amounts of product.” His model supports his 1929 physiological arguments, and exhibits the plateau effect that also forms the basis of the MCA approach to dominance. However, he also notes that:

I have no objection to an evolutionary process by which the dominance of wild type over mutations may be increased provided the pressure toward fixation is sufficient to be effective. (Wright, ’34a, p 50)

His objection to Fisher’s model was the strength of selection, not constraints on dominance modification. As Kacser and Burns had noted themselves, there are no notions of an invariant expectation of dominance in Wright’s conception. Nonetheless, at this stage it was still not clear to what extent Wright thought that dominance of the wild type is a chance result of the “law of diminishing returns,” and how much of it is due to evolutionary history. This is an important issue, because it is precisely the evolutionary history involved in building up the “factor of safety” that holds the key to the difference between the physiological views of Wright and Kacser. How this factor of safety evolves is not clearly spelled out in the early literature. We shall return to this problem in the next section.

Why is the proposition that dominance can be modified or selected problematic, and thereby subject to resistance? For one thing, the evolution of dominance is a conceptual nuisance because it is at odds with some of the classical simplifications of both genetics and evolution. In genetics, it is easier to deal with context-independent gene effects. With context independence, one can easily conceptualize evolution as a gene-centric process focusing on the independent effects of alleles at different loci. On the other hand, with context dependence, genetics becomes much more difficult and we are forced to restrain any triumphalist claims regarding our understanding of the relation between genotype and phenotype. Furthermore, context dependence makes evolution a more difficult problem. Not only are we forced to deal with the classical paradigm, the effects of heredity on evolution, but also with its sister paradigm, the effects of evolution on heredity. In this context, it is the ultimate irony that Fisher, the founding father of population genetics, could point to solid evidence that physiology allows for dominance modification. Yet he could not come up with a successful population genetic scenario for the evolution of dominance. It did not help that from the beginning, the evolution of dominance was associated with Fisher’s idea that small selective coefficients were sufficient if given enough time (Fisher, ’29). This was part of Fisher’s deterministic conception of evolution. As a result, he resisted Wright’s ideas on the importance of other forces such as random drift, ideas that would later encounter support. Meanwhile, by the 1980s, the MCA perspective was in the process of being incorporated into the mainstream. By the time the far-from-equilibrium explanations matured in the 1980s, they may have been perceived to be dealing with a special case. Furthermore, from the logical perspective, relevance of the far-from-equilibrium scenarios would have to depend on the falsehood—or at least exceptions to—the invariant expectation argument. With regard to dominance, MCA provided the ideal compromise: a return to a non-historical view.25 It was no longer necessary to ask why dominance had arisen, it was sufficient to say

---

25This does not imply that from the MCA perspective, all aspects of metabolism had to be viewed in a non-historical manner. For example, in one interesting article, Kacser and Beeby (’84) use a
that it was there as a default property of the physiology. From this perspective, one did not explicitly have to deal with evolution. Meanwhile, from the evolutionary perspective, dominance evolution could be dispensed with. Somewhere in the shuffle, 50 years of genetic experiments on dominance modification were left in an inconspicuous attic.

As a contraposition to the general reach of Kacser and Burns' work, several theoretical examples indicate that if one examines molecular models involving nonlinear dynamics, then dominance is not an invariant property of the system. This has been shown in models involving enzymes that have high Hill coefficients or that are involved in oscillatory loops (Grossniklaus et al., '96). Models that encompass the effects of regulatory feedback also indicate that dominance is not an invariant property (Omholt et al., 2000). Models of development that do not involve metabolism lead to the same conclusions (Gilchrist and Nijhout, 2001). Models that consider the effects of protein–protein complexes also lead to similar conclusions (Veitia, 2003, 2004). These works lend support to earlier objections to the notion of the "inevitability" of dominance (Cornish-Bowden, '87; Savageau and Sorribas, '89; Savageau, '92). The common structural property of these models is that they allow for the possibility of nonlinearities. Nonlinearity can lead to epistasis.

Epistasis is defined as a situation in which mutations at one locus can modify the effects of mutations at a different locus (see for example Wagner et al., '98; Rice, 2000; Hansen and Wagner, 2001). If a model allows for epistasis, then dominance modification can occur. However, there are implicit assumptions in Kacser and Burns' theory of dominance that exclude epistasis.

---

<footnote continued>

theoretical approach for addressing the evolution of catalysts in the early stages of the evolution of life.

26See also Nijhout and Paulsen ('97) for a related model.

27Part of the analysis in Wagner et al. ('98) was concerned with epistasis in metabolism. A section of this paper made the argument that the Kacser and Burns ('73) approach did not address epistasis. However, the algebraic representation of what consists of a physiological phenotype \( \Psi \) was erroneous in the 1998 paper (a deficiency for which the present author is mainly responsible). The model in this paper exhibited no epistasis with respect to \( \Psi \) (as stated in the paper), but it did exhibit epistasis with respect to \( J \), which should have been the relevant phenotype. It took a more careful examination of the relevant mathematics to untangle and understand the consequences of the Kacser and Burns theory with respect to dominance and epistasis (Bagheri-Chaichian et al., 2003).

In effect, one can show by mathematical proof that in the case of finite changes in enzyme concentrations, Kacser and Burns' flux summation theorem is only valid in the absence of epistasis (Bagheri-Chaichian et al., 2003). This implies that their conclusions regarding the evolution of dominance can only hold in the absence of epistasis. Such a situation is very unlikely in a metabolic system (for more on this, see the Appendix). For example, it can be shown that the presence of enzyme saturation can produce epistasis. The latter possibility provides opportunities for dominance to evolve via mutations that change saturation levels (Bagheri-Chaichian et al., 2003; Bagheri and Wagner, 2004). The implication is that dominance is unlikely to be an invariant property of metabolic systems.

If we accept that dominance is not an invariant property of molecular systems in general, then evolutionary history becomes important. Hence, we come back to the original population genetic complications associated with the evolution of dominance. As discussed in an earlier section, previous theoretical works have shown that dominance can evolve in situations where the initial frequency of heterozygotes is relatively high (Wagner and Bürger, '85; Mayo and Bürger, '97). A good example is that of balanced polymorphisms (Bürger, '83c; Otto and Bourguet, '99). The main caveat to such assertions is that it is not clear how often such scenarios can occur. Can a high initial frequency of heterozygotes account for all the cases of dominance evolution?

One possibility is selection for alleles with multiple effects. Mechanistic models of enzyme kinetics indicate that not only can dominance be modified, but it can also evolve in a manner that is insensitive to the frequency of heterozygotes (Bagheri and Wagner, 2004). This essentially happens because in models that are based on Michaelis–Menten-type kinetics, mutations that have a dominance modification effect can also have an independent fitness effect. Due to such "dual effects," dominance evolution can occur as an incidental side effect of selection for the independent effect of a modifier allele (one could view this as "incidental selection"). The dual effects are not explicitly built into the metabolic models as an a priori assumption. They are the effect of Michaelis–Menten-type kinetics in the context of sequential reactions. Such conclusions may serve to resolve some of the controversies revolving around the issue of dominance modification and evolution. However, more work is needed to
establish their generality in other types of molecular models.

**Multiple-effect alleles and the evolution of dominance**

The possibility of modifier alleles with independent fitness effects was proposed by Wright ('29a, b, '34a, '64, '77) early in the debate. This possibility was one among many directions probed by him, and was not extensively pursued by others in later years. Nonetheless, compared to other scenarios, the population genetic framework for addressing selection on multiple effects in a two-locus two-allele setting can be more straightforward (Feldman and Karlin, '71). However, its physiological underpinnings may be more controversial.

Tracing the development of the idea of multiple effects can also serve to address the variety of ways by which different discussants approached the "factor of safety" issue. In his first response to Fisher, Wright ('29a, p 277) suggests that if modifier alleles have "multiple effects," then their modifying effects may not be a major factor in driving their fate in the population. In a later work, he reiterates this by suggesting the possibility that all genes have multiple effects and, through one or other of these, each in general is subject to direct selection which takes precedence in controlling its fate. (Wright, '29b, p 558)

The following year, Haldane ('30) took up this topic. In considering a case where AA is a wild-type genotype at the primary locus, and Mm is the genotype at the modifier locus, whereby the dominance modifier M itself is assumed to be dominant, he notes that:

It is clear that AAMm cannot have a viability greater than normal [i.e. AAmm], or M would spread through the population apart from its modifying effect. Hence its presence in the species would have nothing to do with the mutability of A. (Haldane, '30, p 87)

Although Haldane considered the possibility of multiple effects, he did not seem to believe that it was the likely explanation for dominance evolution. He then goes on to propose his factor of safety notion. First he suggests that "genes are catalysts acting at a definite rate," thus introducing a physiological component to his explanation, as Wright had also done earlier. He then considers a case where a mutant allele A2 represents a catalyst that is twice or more faster than a slower counterpart, A1. In such a situation, he posits that A1A2, A1A1, and A2A2 will be indistinguishable from each other. However, he proposes that A2a individuals—whereby a represents a dysfunctional allele—will be normal as in the case of A1A1 individuals. He then concludes that:

Hence A2a zygotes will have a better chance of survival than A1a, and A2 will be selected. (Haldane, '30, p 88)

Haldane's argument with regard to the factor of safety is somewhat paradoxical. He rejects Fisher's model, and introduces the factor of safety. In the process of proposing his alternative, Haldane's concern seems to have been whether the modification was occurring at the primary locus itself rather than at a separate modifier locus. But a close examination of his argument points to the fact that selection for A2 in lieu of the A1 allele is subject to the same kind of objections that he had posed against Fisher; the selective pressure for A2 depends on the frequency of the a allele. As far as one can tell, in Haldane's '30 scenario, one would have to depend on drift for the fixation of the A2 allele.28 The indeterminacy on how this factor of safety could evolve is likely to have added to the subsequent confusion surrounding this topic. Furthermore, it does not help that more than one valid scenario can account for the factor of safety. Muller ('32, p 240) thought that a "margin of stability and security" could also evolve as a result of selection for stability against environmental perturbations. Wright ('34a, p 50) also discusses such a possibility, by referring to a hypothesis that had been previously formulated by Plunkett ('33).

The physiological model that Wright ('34a) formulates, forms the central basis for much of the subsequent debate. The central idea that remains from this latter work is the platea-
shaped curvature of the relation between “gene activity” and the phenotype (Wright, ’34a, p 44). However, the 1934 article does not necessarily clarify the issue of how the factor of safety evolves. Wright did refer to the environmental theories of Muller and Plunkett as a possibility. He also reiterated his earlier hypothesis that:

if the combination AAM— has any advantage or disadvantage relative to AAm the pressure due to such selection is certain to take precedence over that due to its effect on the rare heterozygote. (Wright, ’34a, p 29)

The above statement is quite clear and is consistent with his previous position. However, in the 1934 article it is not clear how this statement maps to the physiological model he presents. His presentation of the physiological model corresponds more to an illustration of Haldane’s model, whereupon the modifier is the primary allele itself. This does not mean that the model could not be used to illustrate modifiers at a different locus; however, on that occasion, the model was not used to such an end. It was not until 43 years later that Wright (’77, p 503) presented such an illustration. Few articles on the evolution of dominance refer to Wright’s (’77, p 498–526) work on dominance, which is a chapter buried in a four-volume set. This chapter attempts to review many issues relevant to dominance evolution. Nonetheless, it may not have been successful in helping readers focus on any particular solution. More importantly, by the 1970s, models of biochemical kinetics had developed much further, and Wright’s model does not correspond to the specifics of what is known about enzyme kinetics. Nonetheless, even if Wright did not have the correct model for biochemical kinetics or gene regulation, his model seems to have captured a general property of nonlinear representations of multi-step chemical transformations. Not only does his model illustrate the plateau effect that Kacser and Burns noted, but his model also captures the possibility of modifier alleles with multiple effects: an eventuality that is for example present in Michaelis–Menten-type models (Bagheri and Wagner, 2004). It is likely that with systems with nonlinearity and epistatic interactions, it is actually quite difficult to find regimes where allele substitutions have single effects.\[31\] Such regimes do exist, but whether and why a system is in such a regime, whereby it exhibits additivity or modularity of effects, is itself an evolutionary question.

In order to bring the problem into sharper focus, there are important empirical questions that need to be addressed. How ubiquitous are modifier alleles that also have independent fitness effects that are correlated in the direction of higher dominance? Are such alleles common in molecular systems other than metabolism? More generally, given any relationship between a set of one or more perturbations (e.g., environmental change and mutations) and a set of one or more phenotypic traits, it would be pertinent to approach the problem of whether there exist regimes where there are correlations between the effects of different perturbations. As an example, in a theoretical study on RNA folding and stability, Ancel and Fontana (2000) use the term “plastogenetic congruence”\[32\] for an inverse correlation between phenotypic robustness to mutations and phenotypic plasticity with regard to microenvironmental variation. The issue of how we address such correlations may be central to our future understanding of how selection interacts with the mechanistic basis of development and heredity. Other than the well-publicized disagreements on the effects of drift, it is perhaps on the issue of such correlations that we find one of the most subtle—but deeply consequential—mistakes in Fisher’s assessment of the dominance problem. Upon considering the issue of multiple effects, in a personal letter to S.C. Harland he notes that:

On the question of modifying factors selected on their own account, there is a distinction worth making, […] it is exceedingly difficult for any factor to be mathematically neutral; […] Consequently, the factors actually used in dominance modification will necessarily be predominantly those which have some, perhaps slight, selective advantage on their

---

31Note that the term “multiple effects” is itself context dependent and may need to be further specified. It can be interpreted as “pleiotropic effects,” whereby a mutation can have effects on several phenotypic characters. But in cases where we are focusing on a single phenotypic character, it can refer to a modifier allele that exhibits a heterozygote rescue effect (i.e., dominance modification), in addition to an independent fitness effect in a wild-type homozygote. A third possibility is a modifier allele that confers robustness to both environmental and genetic perturbations. The distinction between the above interpretations of “multiple effects” is easy to blur, because all can serve as valid scenarios for the evolution of dominance.
own account. This, however, affords no explanation as to why dominance is modified in the right direction; the explanation lies in the additional selective advantage afforded by improvement in the heterozygotes, i.e. there is no need to postulate that those genes which make changes of dominance in the right direction do ipso facto enjoy any selective advantage other than that provided by the improved viability of the heterozygote. (Fisher to S.C. Harland, October 11, 1940, in Bennet, '83, p 216–217)

Here, Fisher acknowledges the likelihood of multiple effects. But he rejects it as an explanation; he bases his argument on the notion that there would be no general patterns of correlation between independent effects and modifying effects. He may have dismissed such a possibility too easily. For example, one can find the basis for such correlations in Wright’s physiological model. The occurrence of such correlations is precisely what needs to be addressed by future theoretical and empirical studies.

**DISCUSSION AND CONCLUSION**

The solution to the problem of dominance evolution (or evolution of development for that matter) is likely to lie in the direction of a further integration of population genetics and a mechanistic understanding of physiology and genetics. Population genetics arose as a result of combining a mechanistic perspective with an evolutionary one. We are now left at an impasse at which it seems that population genetics by itself is not sufficient to address the problem of evolution of mutational effects. Nor is an understanding of molecular mechanisms by themselves sufficient to address the problem. The solution will most likely have to come from a combination of the aims and methodologies used in both approaches. Hence, once again, the problem requires a combination of historical and mechanistic perspectives. In this respect, an MCA approach has the potential to contribute to the development of evolutionary approaches on dominance, provided that likely mistakes from the past are objectively redressed. Addressing the problem of dominance evolution requires theoretical frameworks that can accommodate the rich repertoire of behavior available to molecular systems. In pursuing this problem, it is likely that our previous conception of epistasis and dominance modification in molecular systems was too rigid.

In reviewing the debate on dominance evolution, one cannot help but ask whether due to excessive rivalries, objective reappraisal and flexibility in the formulation of scientific hypotheses have suffered. This includes rivalries between early figures such as Fisher and Wright, and subsequently between disciplines and philosophies. Stark opposition between differing views may be sometimes necessary. But such opposition can also result in a level of inflexibility with regard to the framing of subsequent questions and the interpretation of results. In the debate on dominance, this inflexibility may have been further compounded because of the passage of time. Given an extensive—and multi-disciplinary—literature, errors in transmission of ideas can accumulate, and the specifics of different positions can become blurred. Such a state of affairs is not an issue of blame on any single group, but rather a deficiency in the scientific discourse on this topic, whereby the relevant questions and possible answers have been excessively simplified. Given pressure for a post-debate scientific consensus, conclusions or interpretations have been cornered into restricted sound bytes that can be paraphrased as “dominance is a consequence of physiology” or that “Fisher was wrong.” When presented as the main conclusion, neither view is conducive to conceptual advances that can go beyond what was established by the 1970s. When used to find common ground between different disciplines, the term “consequence of physiology” lacks commitment, somewhat akin to saying that bird flight is due to physiology: a statement that is true regardless of whether one considers evolutionary history or not. In the meantime, the repeated refutation of Fisher, though once again valid, has ritual characteristics that are reminiscent of a civil war re-enactment. The added detriment is that with a steady accumulation of the accepted simplifications, articles that make a reference to dominance evolution are generally expected to conform to the conventional wisdom, otherwise they may find it more difficult to enter the literature.

Given the extensive time period that has accrued since Fisher’s (‘28a,b) hypothesis, it is worth re-examining Wright’s (‘29a) initial response. In the latter exchange, many of the open questions that need to be addressed have already taken form. We may now be in a better position to reformulate and answer some of these questions. Two factors may prove beneficial for such an endeavor. One is to

---

32 Instead of “physiology,” we can also substitute other insipid descriptors such as “biochemistry” or “network properties” and still have a valid statement.
reassess the issue of why the evolution of dominance is an important question. A reappraisal can be done regardless of whether the specifics of Fisher’s model were correct or not. In this process, the original context of the evolution of inheritance systems has to be considered. Secondly, it is now possible to move away from the high-profile rivalries, and allow for more flexibility in exploring the possible answers to the problem. Admittedly, the present framework within which the evolution of inheritance systems and dominance are discussed is restricted. However, in order to move on to a more general approach, the basic questions that were initially raised have to be clarified and resolved.

The conceptual history of the debate on dominance involves the interlacing of several intellectual and methodological traditions. With regard to the influence of ideas based on MCA, we have a case in which predictions derived from an abstract theoretical model, based on a highly specific and specialized case, have been pronounced as the basis for predictions with a general reach. Specifically, one possible end point of an evolutionary—physiological trajectory (i.e., the fixed assumption of no enzyme saturation, and hence no epistasis) has been used as the basis for making statements about the evolution of a system that allows for saturation and epistatic interactions. One can argue that such a theoretical model suffers from internal inconsistencies. In any event, the proposition that dominance is an invariant property of genetic systems should be subject to scientific scrutiny, and its validity will have to be addressed in a technical forum. Nevertheless, the historical aspects of this debate provide interesting questions in their own right. Mainly, that an extensive body of older empirical work indicated that dominance can be modified and artificially selected. However, this evidence was cast aside in deference to the union of restricted mathematical models in population genetics and restricted mathematical models of molecular systems. Such molecular models stated that dominance is a fixed property of metabolism and that there are inherent constraints on its modification. The results from the molecular models, when generalized as the principal model for the explanation of dominance, are in direct contradiction to the older empirical work. Given this contradiction, it is the older empirical work that has been ignored. However, it would be unfair to attribute the dismissal of dominance evolution as merely the consequence of the molecular models. For various reasons, the intellectual environment within evolutionary biology at the time was receptive towards the conclusions derived from the molecular models.

Admittedly, the integration of the different fields involved in this topic is a difficult task. One would expect that each of the different methodological traditions converge on the question of dominance evolution to exhibit non-overlapping weaknesses (and strengths). Nonetheless, given the importance of the topic, such an integration is necessary. The cross section of history presented here can serve as a worthwhile case study for considering the interdependent web of factors by which scientific communities decide which theories to adopt and which ones to disavow. In this regard, the aim of this article has been to put forth the proposition that, contrary to common claims, the evolution of dominance is not a resolved scientific problem, and that it needs to be addressed.

The possibility that selection for dominance-modification alleles can occur due to the independent fitness effects of such alleles is, in retrospect, a compelling alternative that deserves further attention. It is also interesting in the sense that it is a solution that does not lie at either extreme of the polar spectrum of competing hypotheses. In this scenario, dominance is not an inherent property of metabolism. Neither is it freely molded by selection. Rather, it is a result of some of the existing correlations between the independent fitness effects and the dominance-modification effects of alleles in a given type of physiological architecture (in this case, metabolic pathways). Here one is led to think of Gould and Lewontin’s (’79) notion of “spandrels”: unselected properties that can be a side effect of the structural requirements of an adaptation. Appropriately, there is no suggestion in the latter work that once spandrels have arisen, they cannot be subsequently modified or molded by selection. Similarly, in the metabolic models, there are no inherent metabolic constraints that prohibit modification or selection for higher dominance levels. The main impediments in the latter case are a result of the population conditions: ineffective selection due to low frequency of mutant heterozygotes at the primary locus. However, the same alleles can be selected due to their direct, rather than modifying, effects on fitness. In the latter case, dominance evolves as a result of selection for a property that is a physiological correlate of higher dominance. The issue underlined here is that organisms do
have constraints, but the effects of those constraints do not necessarily have to be divorced from their interaction with selection (see also Gould and Lewontin, '79, p 594).

The possibility that dominance or robustness levels can be changed within a microevolutionary framework has consequences that go beyond dominance. For example, the connection between the evolution of dominance and the evolution of sex may extend beyond the similarity between their respective conceptual frameworks. There can also be a mechanistic connection between the two. For example, consider the metabolic case where selection for higher flux can result in higher robustness and hence dominance of the wild type. Under conditions of evolved robustness, it is not difficult to imagine a situation whereupon single mutations have a small phenotypic effect, while an accumulation of several mutations has a larger effect than the sum of each mutational effect in isolation. The latter pattern is sometimes referred to as synergistic epistasis; it is one of the conditions that has been hypothesized as a cause for the evolution of sex (Kondrashov, '87). Hence, one can formulate a direct connection between the physiological properties that lead to the evolution of dominance and sex.

In terms of a methodological outlook for addressing evolution, the significance of the debate on dominance evolution is far reaching. However, it is not a simple task to sort out the competing theories on dominance evolution; this is partly due to some of the conflicting premises that can exist even within one position. The prime example is the view held in some quarters within evolutionary biology and biochemistry that dominance is a default expectation in metabolism. At first, one may be tempted to oversimplify the latter position as one in which dominance does not evolve. However, this is technically not a correct characterization. The difficulty is that the parties holding the “invariant expectation” position are not denying the basic tenets of population genetics or evolution. Rather, they are essentially saying that the basic population genetic framework does not apply to dominance evolution. Supposedly, dominance is already there by the time mutation and selection are of any relevance. Hence, one is left in a conceptual limbo, whereupon dominance cannot evolve in a microevolutionary (i.e., population genetic) framework, while somehow it has evolved as an expected outcome of the “macroevolutionary processes” that led to the existence of metabolism. That is, as soon as you evolve metabolism, you have dominance. The latter characterization is probably the best way in which one could reconcile the “invariant expectation” outlook with an evolutionary one. The problem with this outlook is that the “macroevolutionary processes” mentioned here are a rather vague black-box set in a deep evolutionary past—the origin of metabolism—to which dominance evolution is relegated. It is quite likely that for some evolutionary processes, we may indeed have to construct theoretical frameworks that are qualitatively different from the classical population genetic framework. Nonetheless, it seems that within the limited context of the established debate on dominance evolution, a qualitative restructuring that moves away from a microevolutionary framework is not necessary.

The argument for a reappraisal of dominance evolution, as presented in this article, is based on two premises. One is that the significance of a large body of evidence for dominance modification, in experimental and natural populations, has not been properly addressed. Secondly, it is suggested here that the theoretical argument for an inevitable expectation of dominance suffers from logical and empirical contradictions, and that alternative theoretical considerations have not been given due attention.

ACKNOWLEDGMENTS

Criticisms and suggestions from two reviewers and an editor were highly constructive. Consultations with Günter Wagner, Leo Buss and Rimas Vaisnys were invaluable. Discussions with Walter Fontana, Michelle Girvan, Erica Jen, David Krakauer and Philippe Roumagnac were much appreciated and highly instructive. I am thankful to R. Veitia and V. Nanjundiah for kindly providing preprints of respective book chapters that were in press. This work benefited from the readily accessible collection of old journals in the Yale University library system. The online collection of Fisher’s papers at the University of Adelaide was a valued resource.

APPENDIX

The flux summation theorem

In Kacser and Burns ('81), the theorem is stated as

$$\sum_{i=1}^{n} Z_i = 1,$$  \hspace{1cm} (1)
where
\[ Z_i = \frac{dF}{dE_i}F. \]  

\( E_i \) (denoted as \( E \) in the original) represents an underlying parameter for enzyme \( i \) and \( F \) is flux. As such, \( Z_i \) is a sensitivity coefficient for flux with respect to \( E_i \). The argument for the inevitable expectation of dominance stems from the idea that the average expectation for \( Z_i \), based on the summation theorem, is on the order of \( 1/n \). This is taken to mean that as \( n \) becomes large (i.e., a long pathway), most enzymes will have a small effect on flux \( F \).

The problem with the above argument is that when one considers finite—rather than continuous—changes in \( E \), then the finite version of the summation theorem does not hold unless the system is linear and devoid of epistasis (Bagheri-Chaichian et al., 2003). However, all changes in enzyme concentration are in the finite realm, including the null mutations for which Kacser and Burns develop their dominance arguments (based on 50% reductions in concentration). Furthermore, enzyme systems are nonlinear and exhibit epistasis. Hence, the summation theorem does not correspond to the facts of the phenomenon for which it was being used as the explanatory framework.

### Flux surfaces and the geometric version of the invariant expectation of dominance argument

The summation theorem was derived independently of any specific derivation of flux and constitutes the core of the invariant expectation of dominance argument. However, there is a frequently used flux derivation (Kacser and Burns, ’73, ’81) that is used for illustrating the application of the theorem. Much of the popular success of the invariant expectation argument is based on geometric illustrations using the latter flux derivation. There are certain characteristics associated with the use of the respective flux surfaces that have not been examined before. In this section, I will take the opportunity to illustrate some of the characteristics associated with the frequently used derivations. In the next two sections, I illustrate some of the problems associated with the use of such flux derivations.

By excluding the possibility of enzyme saturation, the flux \( F \) (denoted as \( J \) in later literature) for a sequential series of enzyme-catalyzed reactions is derived as
\[ F = \frac{(S_0 - S_n) \prod_{i=1}^{n} K_i}{(M_1/V_1) + \sum_{i=2}^{n} (M_i/V_i) \prod_{j=1}^{i-1} K_j}, \]

where \( n \) is the number of enzymes, \( S_0 \) and \( S_n \) are the source and sink substrates in the pathway, \( K_i \) is the equilibrium constant for each enzyme-catalyzed reaction from \( S_{i-1} \) to \( S_i \), and \( M_i \) is the Michaelis constant for the latter reaction.\(^{33}\)

In order to understand why variants of Eq. (3) have been used to illustrate that dominance is “inevitable,” it is preferable to first simplify the equation. Usually, the quantity \( (M_i/V_i) \prod_{j=1}^{i-1} K_j \) is simplified to one composite parameter and rewritten as \( 1/E_i \), the rationale being that the parameter \( E_i \) is proportional to total enzyme concentration \( E_i \) (given that \( V_i = k_{\text{cat}(i)} E_i \), where \( k_{\text{cat}(i)} \) is the catalytic turnover rate). Here, in a slight variation of this theme, we can define a different composite parameter \( p_i \) as
\[ \frac{1}{p_i} = \left( \frac{M_i}{k_{\text{cat}(i)} \prod_{j=1}^{i-1} K_j} \right) \times \left( \frac{1}{(S_0 - S_n) \prod_{i=1}^{n} K_i} \right). \]

Hence, for any given parameter vector \( P = (p_1, \ldots, p_n) \), one can define a function \( f(E) \) such that
\[ F = f(E) = \frac{1}{(\sum_{i=1}^{n} 1/(p_i E_i))}, \]

where \( E = (E_1, \ldots, E_n) \) is the vector of enzyme concentrations. We assume that mutations can change both \( E \) and \( P \).

We can now use Eq. (5) to illustrate the invariant expectation of dominance argument. Consider a normalized flux \( F^* \), such that for any given \( P \) and fixed operating point \( E^* \),
\[ F^* = \frac{f(E)}{f(E^*)}. \]

For the illustration in this Appendix, we will look at a case with \( n = 5 \) and \( E^* = \langle 10,10,10,10,10 \rangle \). Figure 7A shows the value of \( F^* \) with respect to cases where the concentration of a single enzyme has been changed in comparison to the operating point \( E^* \). The property to note here is the characteristic plateau effect of \( F^* \) with respect to changes in \( E_i \). The pathway in Figure 7A exhibits\(^{33}\)Note that in Kacser and Burns (’81), \( X_1 \) and \( X_n \) were used instead of \( S_0 \) and \( S_n \). The latter two substrates are generally assumed clamped and hence constant. The naming and indexing is changed here in order to avoid confusion with the original, which, due to what must have been a typing error, is missing a step \( (X_{i+1}) \) and \( K_{i+1} \).
the characteristic plateau effect indicative of low flux sensitivity with respect to $\delta E_i$. Let us assume that for any locus $i$, the enzyme concentration $E_i$ in a mutant heterozygote $A_Ai$ is half that of a wild-type homozygote $A_A$. Under such an assumption, Kacser and Burns argue for the default expectation of dominance. This is because the phenotype of the $A_AI$ individuals is closer to that of $A_AI$ individuals (i.e., dominance of wild type) rather than half that of $A_AI$ (co-dominance) or close to that of $aAi$ individuals (recessivity).

The argument that dominance modification cannot alter the general expectation of dominance is predicated on the constraints inherent to the summation theorem and flux functions of the type illustrated in Eq. (5). The latter equation obeys the continuous version of the summation theorem. Furthermore, when one changes the parameter vector $P$ in a way that flux becomes more sensitive to any given enzyme $E$, this increased sensitivity is compensated by decreased sensitivity to changes for other enzymes $j$ $\neq i$. Figure 7B is an example of the latter property. For illustration purposes, the $p_i$ values for each $P$ are offset from each other so that the different curves can be distinguished. In each case, what matters is the ratio between the $p_i$ values. For Figure 7A, $P = \langle 10,9,8,7,6 \rangle$. For Fig. 7B, $P = \langle 012,11,15,13,12 \rangle$.

Problems associated with the geometric argument

If one accepts the flux function in Eq. (5) as a general representation of metabolic pathways, then the robustness constraints as illustrated in Figure 7A and B cannot be avoided. However, Eq. (5) and equations related to it represent a very specific type of function derived under the simplifying assumption that no enzyme ever approaches saturation. Nonetheless, from a biochemical perspective, any enzyme must approach saturation at some point as $E_i$ is being decreased. Hence, we can be assured that the behavior of the function in Eq. (5) is not representative of enzyme-catalyzed systems as $E_i \rightarrow 0$. In fact, despite its nonlinear form, Eq. (5) behaves very much like a linear function when one compares heterozygotes with wild-type homozygotes. Consider a hypothetical case where we have a linear function such that

$$F = f_{\text{lin}}(E) = \sum_{i=1}^{n} p_i E_i.$$  \hspace{1cm} (7)

Eq. (7) can be normalized in the same fashion as shown in Eq. (6). Figure 8A and B show the normalized behavior of Eq. (7) in response to enzyme and parameter changes. On comparison with Figure 7A and B, it should become clear that Eq. (5) is subject to similar types of constraints in the domain of interest near the operating point (wild-type homozygotes and mutant heterozygotes) as the linear flux function in Eq. (7). The problem is that a representation that largely behaves like a linear function in the domain of interest is unlikely to capture many of the variational properties of biochemical systems, whose underlying processes are nonlinear. This becomes especially important when assessing the evolutionary history and variational possibilities of such systems.

As a consequence of the lack of saturation, Eq. (5) is an example of a flux function with restrictive constraints on the number of loci that can manifest sensitivity to enzyme changes. Despite the high restrictions, it still does not completely mesh with the summation theorem as the general explanation for an invariant expectation of dominance (or robustness). This has to do with the fact that even when the continuous version of the summation theorem holds, it is not a good predictor of robustness expectation with respect to finite changes. Let $C_i$ be the finite version of the sensitivity coefficient in Eq. (2), such that

$$C_i = \frac{\delta J}{\delta E_i}.$$  \hspace{1cm} (8)

In both Fig. 7A and B, as $\delta E_i \rightarrow 0$, then $\sum_{i=1}^{5} C_i = 1$. However, when $\delta E_i = -E_i/2$, then $\sum_{i=1}^{5} C_i = 1.66$ for Fig. 7A, while $\sum_{i=1}^{5} C_i = 1.05$ for Fig. 7B. Hence, the summation theorem does not map to a precise expectation of robustness. This is because even though Eq. (5) is a highly restricted case, it still exhibits sufficient epistasis for mild levels of dominance modification to occur. In fact, for any $|\delta E_i| > 0$, the equation $\sum_{i=1}^{n} C_i = \sum_{i=1}^{n} Z_i = 1$ holds if and only if flux is a linear function, and thereby devoid of epistasis (as is the case in Eq. (7) and Fig. 8). With regard to epistasis, the important concept to note here is that the term simply refers to the possibility of mutations at one locus modifying the effects of mutations at another locus.

$^{34}$Note that it is possible for a nonlinear system to evolve towards a parameter regime where it does behave like a linear system. However, barring restricted functions such as the one in Eq. (5), the linear state of affairs is unlikely to be a default expectation and cannot be separated from the evolutionary history or processes that led to such a state of affairs.
another locus. This is precisely what is required for dominance modification. On the other hand, a theory that indirectly implies a linear (i.e., additive) genotype–phenotype map is a theory that is based on an a priori exclusion of any kind of mutational modification via locus interaction.

It is important to note that if one includes saturation in the system, then the disparity between the continuous and finite versions of the sensitivity sum can become more pronounced. The latter eventuality means that the possibilities for dominance modification also become more pronounced. For simple Michaelis–Menten-type catalysis with saturation, if all $\delta E_i < 0$, then $\sum_{i=1}^{n} C_i \to n$ if all enzymes are approaching saturation (Bagheri-Chaichian et al., 2003). Accordingly, the modulation of saturation levels by mutations can allow for the modification and evolution of dominance in simple metabolic systems (Bagheri and Wagner, 2004).

With regard to the occurrence of saturation, a perennial objection (and hence defense of the Kacser and Burns model) has been that high saturation leads to the lack of a steady state, and therewith an interminable—and lethal—accumulation of intermediate substrates. Such arguments are predicated on simplifications of enzyme catalysis, whereby reactions are conceived as irreversible. For generic sequential pathways, if one includes reversible reactions (which is the physical expectation), substrates will generally not accumulate beyond ratios governed by the equilibrium constant—and hence thermodynamics—of the respective reaction. Nonetheless, when substrate concentrations are important to fitness, then one of the interesting issues that arise is that we have to consider the interaction of thermodynamic constraints with selection. In any case, once again, we cannot exclude fitness—and hence evolutionary history—from consideration (for a pertinent example, see Salvador and Savageau, 2003).

Fig. 8. A five-enzyme model based on the linear flux function in Eq. (7). In terms of the comparison of $A_iA_i$ to $A_ia_i$ individuals, the function behaves in a similar fashion to that shown for the nonlinear function in Eq. (5). (A) The wild-type phenotype is dominant with respect to mutations at all five loci. (B) If the wild-type phenotype is made more sensitive to mutations at any given locus (in this case, Enzyme 1), robustness to mutations increases at the other loci.

Richness of constraint variability associated with different types of biochemical circuitry

In its biological context, the debate about the inevitability of dominance is about whether the manifestation of robustness is due to immutable constraints within biochemical systems. Future works on this topic will have to investigate to what extent the variational properties of biochemical systems are subject to change. To this effect, the assumptions going into deriving theorems such as Eq. (1) and flux functions such as Eq. (5) are too restricted. It is likely that different functional requirements imposed on evolved biochemical circuitry can produce a large variety of network circuitry and kinetic schemes. Such circuitry can be selected to perform a variety of "logical"
functions. This is especially likely given the variety of situations in which biochemical systems can be involved (e.g., metabolism, signal transduction and gene regulation). The differences in the resulting structures are likely to lead to a wide range of variational properties. Hence, one could argue that from the very onset, biochemical models of the type exhibited in Eq. (5) are very limited in their scope and should not have been used as a general argument for an explanation of dominance.

Both Eqs. (5) and (7) are simplifications of chemical processes and exhibit very specific variational constraints. One can envision constructing other simplifications that are free from some of the constraints exhibited by these equations. However, the equations representing most kinetic models are too involved for consideration in this Appendix (for some examples, see Omholt et al., 2000; Salvador and Savageau, 2003; Bagheri and Wagner, 2004; Veitia, 2004). Here we shall restrain ourselves to a simple hypothetical case that can be placed in contraposition to the functions \( f \) and \( f_{\text{lin}} \) shown in Eqs. (5) and (7), respectively. The purpose is an illustration of principle rather than detailed modelling of kinetic processes.

We consider a case which is essentially an analog version of a logical AND function. Consider a set of \( n \) substrates and \( n \) independent saturable reactions such that the rate \( v_i \) of each reaction \( i \) is given by

\[
v_i = \frac{E_i}{E_i + p_i}.
\]

Note that \( E_i \) in Eq. (9) no longer plays the role of an enzyme as in metabolic reactions. Equations of the form as in Eq. (9) are sometimes used as approximations of the rates of mRNA transcription, whereby \( E_i \) represents the concentration of a transcription factor. We construct a function \( f_{\text{tr}} \), such that

\[
F = f_{\text{tr}}(E) = \prod_{i=1}^{n} \frac{E_i}{E_i + p_i}.
\]

In this model, assuming non-negative \( E_i \) and \( p_i \) values, \( F \) approaches its maximal value of 1 as all reactions \( v_i \) approach their maximal value of 1. Conversely, \( F \) approaches its minimal value of 0 if any reaction \( v_i \) approaches 0. Hence, the similarity to a logical AND function (note that when normalized, Eq. (5) can also approximate an AND function). This represents a simplification of a situation whereby a high phenotypic value \( F \) depends on high transcription rates for \( n \) independent reactions. Figure 9 shows the behavior of \( F^* \) when Eq. (10) is normalized in the same fashion as in Eq. (6), and where \( n = 5 \). The hypothetical function \( f_{\text{tr}} \) as illustrated in Fig. 9 is an example of a simplified case that is free of the robustness constraints exhibited in Figures 7 and 8.

![Figure 9](image-url)

**Fig. 9.** Simple model of an analog “AND” function with \( i = 1 \) to 5. Based on Eq. (10), this example is free of the robustness constrains in Eqs. (5) and (7). (A) The wild-type phenotype exhibits dominance with respect to mutations affecting all \( E_i \). (B) Sensitivity is increased with respect to mutations affecting all \( E_i \). Hence the wild type is co-dominant with respect to all mutations. (C) Robustness and hence dominance of the wild type is increased with respect to mutations affecting all \( E_i \).
In Figure 9A, as $\delta E_i \to 0$, then $\sum_{i=1}^{5} C_i \to 0.99$. Meanwhile if $\delta E_i = -E_i/2$, then $\sum_{i=1}^{5} C_i = 1.64$. Hence, the behavior of the function as shown in Figure 9A is not too dissimilar to that of Figures 7A and 8A. However, the function $f_{tr}$ can be made more robust, or sensitive, than its $f$ and $f_{lin}$ counterparts. In Figures 9B and C, as $\delta E_i \to 0$ then $\sum_{i=1}^{5} C_i \to 0.07$ and $\sum_{i=1}^{5} C_i \to 4.43$, respectively. Meanwhile, for the same figures, if $\delta E_i = -E_i/2$, then $\sum_{i=1}^{5} C_i = 0.14$ and $\sum_{i=1}^{5} C_i = 4.70$, respectively. The higher range of variability for $f_{tr}$ is due to the fact that, in this case, there is no constraint that requires high sensitivity to any given $E_i$ to be compensated by robustness to some $E_{j\neq i}$. Similarly, high robustness to any given $E_i$ is not compensated by sensitivity to some $E_{j\neq i}$.

From a kinetic perspective, the model in Eq. (10) is an artificial one: specifically, because it is predicated on the interdependence of rates rather than concentrations. The question to be addressed is the extent to which different classes of more realistic biochemical circuitry can be constructed that can exhibit behavior that is similar. Note that for models that can exhibit complete recessivity, consideration of sigmoidal regimes would be a natural choice.

**LITERATURE CITED**


Tower WL. 1910. The determination of dominance and the modification of behavior in alternative (Mendelian) inheritance, by conditions surrounding or incident upon the germ cells at fertilization. Biol Bull 18:285–352.


