Alleles versus mutations: Understanding the evolution of genetic architecture requires a molecular perspective on allelic origins

By: David L. Remington

Remington, D. L. (2015). Alleles versus mutations: understanding the evolution of genetic architecture requires a molecular perspective on allelic origins. *Evolution: international journal of organic evolution.* doi:10.1111/evo.12775

Abstract:

Perspectives on the role of large-effect quantitative trait loci (QTL) in the evolution of complex traits have shifted back and forth over the past few decades. Different sets of studies have produced contradictory insights on the evolution of genetic architecture. I argue that much of the confusion results from a failure to distinguish mutational and allelic effects, a limitation of using the Fisherian model of adaptive evolution as the lens through which the evolution of adaptive variation is examined. A molecular-based perspective reveals that allelic differences can involve the cumulative effects of many mutations plus intragenic recombination, a model that is supported by extensive empirical evidence. I discuss how different selection regimes could produce very different architectures of allelic effects under a molecular-based model, which may explain conflicting insights on genetic architecture from studies of variation within populations versus between divergently selected populations. I address shortcomings of genome-wide association study (GWAS) practices in light of more suitable models of allelic evolution, and suggest alternate GWAS strategies to generate more valid inferences about genetic architecture. Finally, I discuss how adopting more suitable models of allelic evolution could help redirect research on complex trait evolution toward addressing more meaningful questions in evolutionary biology.

Keywords: Adaptive evolution | allelic effects | complex traits | GWAS | QTL | selection

Article:

Solving the problem of how Mendelian inheritance of allelic differences can give rise to adaptive variation in complex traits was the cornerstone for the modern evolutionary synthesis. It has also provided the foundation for modern crop and livestock improvement, and for current efforts to understanding the genetic factors underlying human diseases. For the past century, it has been recognized that variation in quantitative traits could be explained by variation in multiple genes and environmental effects (Nilsson-Ehle 1909; East 1916). However, our understanding of the genetic architecture of complex trait evolution—how molecular variation at individual loci

contributes to trait variation and ultimately to evolutionary dynamics—remains in a state of flux. Hanging in the balance are key evolutionary questions, including the discoverability and evolutionary relevance of individual molecular variants (Barton and Turelli 1989; Rockman 2012; Travisano and Shaw 2013), the repeatability of molecular mechanisms in evolution (Stern and Orgogozo 2008; Martin and Orgogozo 2013; Lee et al. 2014), and the roles of new mutations versus preexisting "standing" genetic variation as the raw material for natural selection (Barton and Turelli 1989; Messer and Petrov 2013; Lee et al. 2014). My goal in this article is to argue that inadequate models relating molecular variation to complex phenotypes are severely hampering our understanding of these topics, and propose a more productive conceptual framework for gaining novel insights on the genetic basis of adaptive evolution.

Fisher (1918) and Wright (1921) first developed a statistical framework for genetic analysis of quantitative traits, giving rise to a model in which a very large number of genes, each with very small effects, shape quantitative trait variation (the infinitesimal model). Genetic models for quantitative traits have been subsequently extended to complex binary traits such as human diseases that lack a simple Mendelian basis, using threshold or relative-risk models (Wright 1934a, b; Lynch and Walsh 1998). Fisher (1930) extended the infinitesimal model into an evolutionary model of mutation and natural selection. His geometrical model described the evolution of phenotypes toward values that optimize fitness in a multidimensional trait landscape. The model predicted that only variants of small effect (or micromutations) had a reasonable probability of producing phenotypes with higher fitness, thereby providing an evolutionary basis for the infinitesimal model of genetic variation. The framework of Fisher's geometric model remains the primary lens through which the mutational basis of adaptive evolution is examined to this day.

Perspectives on the infinitesimal model began to shift in the 1980s when the widespread availability of molecular genetic tools made genome-wide mapping of quantitative trait loci (QTL) feasible (Paterson et al. 1988; Lander and Botstein 1989) (see Box 1, Terminology). Early QTL studies, generally in artificial populations, almost universally identified chromosomal regions with detectable and often large effects on traits, contrary to the infinitesimal model. Notable examples include floral traits involved in pollinator preference in Mimulus spp. (Bradshaw et al. 1995; Bradshaw et al. 1998; Schemske and Bradshaw 1999; Bradshaw and Schemske 2003), flowering and reproductive architecture in Arabidopsis (Alonso-Blanco et al. 1998; Johanson et al. 2000; Juenger et al. 2000; El-Assal et al. 2001; Ungerer et al. 2002, 2003), sensory bristle number in Drosophila (Mackay and Langley 1990; Lai et al. 1994; Long et al. 1998; Lyman and Mackay 1998), and morphological traits in sticklebacks (Colosimo et al. 2004; Shapiro et al. 2004). Large-effect polymorphisms were seen as explaining a substantial proportion, but not all, of the genetic variance for a typical trait (Orr and Coyne 1992; Mackay 2001a). The resulting quantitative genetic framework, what we might term the "QTL perspective," was still inherently polygenic but not strictly infinitesimal, consistent with Robertson's (1967) model of an exponential distribution of allelic effect sizes. The QTL perspective was bolstered theoretically by Orr's (1998) reevaluation of Fisher's geometric model. Orr showed that when the size of mutational effects and their probabilities of fixation are considered along with their probability of increasing fitness, some mutations fixed during an "adaptive walk" are likely to have relatively large phenotypic effects. The QTL perspective led to a great degree of optimism that the genes that matter in phenotypic evolution (Mackay 2001b;

Mackay 2001a; Feder and Mitchell-Olds 2003), crop breeding (Yano 2001) and complex human disease (Risch and Merikangas 1996) were discoverable. Accordingly, the QTL perspective was a key impetus behind the current revolution in genomic exploration.

Box 1: Terminology

I use the term allele (adj. allelic) in its classic textbook sense to represent a particular form of a gene; or from a molecular standpoint a DNA sequence haplotype at an individual gene, including its known or presumed cis-regulatory regions. I do so with the recognition that the term has been borrowed in many other contexts, representing genetic variants at levels ranging from individual single-nucleotide polymorphisms (SNPs) to entire genes to extended haplotypes over longer regions.

Depending on the context, I use gene to refer either to the specific coding sequences or to include also the cis-regulatory regions. This is consistent with other similar treatments (Stern and Orgogozo 2008; Martin and Orgogozo 2013), and I believe my meaning will be clear from the context. Similarly, I use locus generally to refer to a gene locus, or the chromosomal location of a single gene with its cis-regulatory regions, or in some cases a larger multigenic region as indicated by the context (e.g., the bab locus; Bickel et al. 2011).

A complex trait is a trait in which variation results from effects of variation in multiple genes and environmental factors; this includes both quantitative traits and categorical traits such as disease susceptibility that do not show simple Mendelian segregation. By contrast, complex alleles are alleles in which differential effects on phenotype result from the cumulative effects of multiple molecular polymorphisms, following the usage of Martin and Orgogozo (2013). Quantitative trait locus (QTL) can refer to a particular and possibly unknown locus with allelic variants that affect a quantitative or complex trait, to the statistical association of trait values with molecular polymorphisms, or in many cases to both at once. The former usage is consistent with the traditional definition that equated QTL with causal Mendelian factors (Paterson et al. 1988; Lander and Botstein 1989), so I retain it here rather than adopting a new term such as "quantitative trait gene" (e.g., Martin and Orgogozo 2013). I use the term QTL region when referring specifically to statistically associated regions, consistent with my practice elsewhere (Remington 2009; Remington et al. 2013). While the QTL terminology relates directly to quantitative traits, the same principles extend quite readily to complex binary traits, so the lessons here are directly relevant to research on complex human diseases.

More recent research, however, has challenged the QTL perspective on many fronts. These have included the discovery of inherent upward biases in QTL effects estimates (Beavis 1994), fractionation of large-effect QTL regions into multiple smaller QTL (Graham et al. 1997; Pasyukova et al. 2000; Huang et al. 2010; Steinmetz et al. 2012), cryptic effects of QTL that are

epistatic or linked to QTL with opposite effects (Weinig et al. 2003; Kroymann and Mitchell-Olds 2005), and traits that simply lack large-effect QTL (Buckler et al. 2009). Fine-scale genetic mapping studies to find the quantitative trait nucleotides (QTNs) underlying QTL have identified multiple polymorphisms with individually small effects on traits (Stam and Laurie 1996; Long et al. 2000; Palsson et al. 2005). Single-nucleotide polymorphisms (SNPs) showing significant trait associations in genome-wide association studies (GWAS) have typically accounted for only a small fraction of the estimated genetic variance for studied traits, with individual SNPs explaining only tiny fractions of the variance (Gudbjartsson et al. 2008; Lettre et al. 2008; Weedon et al. 2008; Manolio et al. 2009). Much of this "missing heritability" in association studies has been attributed to the prevalence of loci with undetectably small effects (Manolio et al. 2009; Yang et al. 2010; Gibson 2012). Perspectives on complex trait variation in recent years have shifted away from an emphasis on the relevance of large-effect QTL (Mackay 1995; Mackay 2001a) to one in which variation is seen as being governed mainly by many small-effect QTL (Flint and Mackay 2009; Mackay et al. 2009). Several simulation studies have even shown that approximations of the infinitesimal model can explain commonly observed distributions of QTL effects (Visscher and Haley 1996; Liu and Dekkers 1998; Noor et al. 2001; Cornforth and Long 2003).

The rise and fall of the QTL perspective came full circle with a paper by Rockman (2012), who pronounced the entire QTL perspective largely irrelevant to evolutionary biology. According to Rockman's analysis, very few reported large-effect QTL have survived subsequent scrutiny, and even fewer causal QTNs have been found. Moreover, he argued, the large-effect QTL that have been found are probably unrepresentative of the molecular basis for quantitative variation overall, and thus their characterization provides little useful insight on evolutionary processes. Rockman argued that, with few exceptions, complex-trait genetics would be far better served using statistical approaches consistent with the infinitesimal model than by ultimately futile attempts at molecular discovery. Travisano and Shaw (2013) argued further that studies of proximate molecular mechanisms have shed little light on ultimate evolutionary processes, reaffirming the dichotomy between functional and evolutionary biology articulated by Mayr (1961, 1982). While not all critics of the QTL perspective carry their arguments to these levels, it is clear that we have entered a new micromutationist era.

In the remainder of this article, I will argue that both the rise and fall of the QTL perspective are based in part on the limitations of premolecular models that conflate beneficial mutations with allelic variation. These conceptual limitations lead to contradictory insights on the adaptive process and to misleading interpretations of the data from QTL studies. My central premise is twofold. First, mutational effects and allelic effects (see Box 1) are not equivalent. The molecular processes that actually give rise to allelic variation in individual genes can include effects of multiple mutations and, importantly, intragenic recombination. Second, different selective regimes operating within and between populations can give rise to very different patterns of complex trait architecture. Contrary to the implicit assumptions behind much of the debate over the QTL perspective, there is no "one size fits all" expectation for what this architecture should look like. I will describe a large body of research that shows small-effect mutations and intragenic recombination indeed frequently give rise to adaptively important large-effect alleles under relevant evolutionary scenarios. I will explain why prevailing GWAS approaches consequently do not actually estimate effects of allelic variation, requiring

reexamination of GWAS-based conclusions about genetic architecture in both evolutionary biology and biomedical settings. Finally, I will suggest ways in which a more suitable framework for understanding the origin of alleles could refocus our attention on more productive paths forward in understanding the genetics of adaptation.

Many of my arguments parallel points about mutational versus allelic effects that were raised by Martin and Orgogozo (2013) in a recent examination of genetic "hotspots" of phenotypic variation. In proposing a more coherent framework for understanding complex trait evolution, my objectives are complementary to theirs. I acknowledge at the outset that none of what I have to say is terribly novel; the literature I discuss will make that clear. The issue is one of synthesis, in which the relevant concepts are largely in place but are not being brought together in a framework that adequately informs our search for answers to some of the big questions in evolutionary biology.

The Fisherian Perspective on Mutations and Allelic Variation

When Fisher's The Genetical Theory of Natural Selection (1930) was published, there was no understanding of how the genetic information encoded in genes and their allelic variants was organized. Discovery of the linear double-helix structure of DNA was still more than two decades in the future. It would be another 14 years before Avery et al. (1944) would publish their discovery that DNA was the "transforming agent" and thus the basic stuff of heredity. In 1930, the protein rather than the nucleic-acid component of chromosomes was widely assumed to be the hereditary material.

Given this state of knowledge, it was reasonable to treat mutational effects and the effects of allelic variation in individual genes as one and the same. Fisher (1930, p. 70) specifically considered genes to have no more than two alleles in almost all cases, differing in properties much like different-colored gumballs. By the time Orr (1998) revisited Fisher's conclusions about the sizes of fixed mutational effects, the state of knowledge about the molecular nature of genetic variation had been fundamentally transformed by discoveries in molecular biology that Fisher could not have anticipated. However, Orr's model of sequential fixation of mutations was essentially the same as Fisher's, in part because his primary objective was to revisit Fisher's micromutational conclusions. In essence, the twin poles around which the theoretical debate about the size of mutational effects revolved, represented by Orr's and Fisher's models, remained premolecular, a limitation Orr himself has been careful to point out (Orr 2005).

With the advent of molecular genetic techniques, it became possible to evaluate individual single-nucleotide polymorphisms (SNPs), first indirectly as allozymes and then as restriction fragment polymorphisms, and later directly by sequencing segments of DNA in samples from populations and across species. The notion of allelic differences now became reduced to the level of individual causative SNPs, and the search for QTL evolved into a search for QTNs. A "one QTL—one SNP" paradigm was the foundation for the ensuing development of association mapping strategies from the beginning, consistent with Fisher's view of alleles. In their groundbreaking paper that proposed the GWAS concept, Risch and Merikangas (1996) made head-to-head statistical power comparisons of linkage mapping and family-based association methods, explicitly assuming that a single SNP ("the actual polymorphism") was responsible for

the phenotypic effects associated with a single gene. Long and Langley (1999) tested the power of several association study designs to detect causative loci for disease risk, again specifically assuming a single causative SNP was responsible for allelic substitution effects. The one QTL—one SNP paradigm at least implicitly dominates the GWAS field to this day. In the extensive debates about the nature of the "missing heritability" in GWAS, the idea that alleles differ from one another at a single causative SNP is seldom specifically questioned.

A Molecular Perspective on Genes, Mutations, and Allelic Variation

The question we must answer, then, is whether the Fisherian allelic perspective and its genome-era counterpart, the one QTL—one SNP paradigm, are realistic. The understanding of molecular genetics and associated data that have accumulated over the past 60 years reveals a relationship between mutations and alleles that is not only more complicated than the Fisherian perspective but fundamentally different in nature. Several key components of a molecular-era perspective were discovered in the 1950s and 1960s with the groundbreaking studies of Benzer (1959, 1961) and Crick et al. (1961), both using mutations in the rII genes of the T4 bacteriophage. These studies revealed for the first time that mutations at multiple sites along the linear structure of an individual gene could give rise to phenotypic differences, and that intragenic recombination involving existing polymorphisms could further give rise to new alleles. The phenotypic effect of any single mutation was shown to be dependent on its broader allelic context, as a given mutation could either produce a mutant phenotype or restore wild-type function.

With the advent of DNA sequencing and molecular population genetic analysis, countless studies have shown that populations typically harbor a complex set of haplotypes in individual genes. Pervasive intragenic recombination has been well-understood to be a major contributor to allelic complexity (Hudson and Kaplan 1985; Templeton et al. 2000a; Templeton et al. 2000b). Adaptive evolution often appears to occur from positive selection on standing variation rather than classic selective sweeps on single variants (Messer and Petrov 2013). However, these population genetic insights have done little to inform our understanding of complex trait variation. Several quantitative genetic models have incorporated consecutive mutations at a single locus (Kimura 1965; Latter 1970; Lande 1976; Barton 1999), but are largely limited to exploring the overall maintenance of genetic variation rather than the details of genetic architecture. The modeling of genetic architecture by Yeaman and Whitlock (2011), which treats allelic effects as the cumulative effects of individual SNPs, is a noteworthy exception. Of particular importance, one cannot consider mutation alone in models for the origin of alleles, but intragenic recombination (including cis-regulatory and coding regions) needs to be explicitly incorporated as well. One implication is that typical mutation-accumulation experiments, whose designs severely limit effective recombination, can capture only a fraction of the potential sources of allelic variability.

These points highlight the limitations of the Fisher–Orr geometric models as a framework for understanding the nature of adaptive evolution. First, the notion of sequential fixation of mutations ignores the role of standing allelic variation in adaptation. Both the phenotypic and fitness effects of a novel mutation will depend on the mean phenotype of the allele in which a mutation occurs, not to mention allele frequencies at other loci. Second, steps in adaptation are likely to include not only advantageous mutations, but also intragenic recombination events that

generate novel higher fitness alleles. Intragenic recombination, in turn, can bring together portions of multiple alleles, probably with genotypic values at different points on the phenotypic surface and each having accumulated different sets of mutations. The incremental effect of a mutation at the time it first occurred might be substantially different than the current net contribution of the resulting polymorphism to the effects of the alleles in which it is found. Unlike the Fisher–Orr model with its single path of sequential fixation (Fig. 1A), differences between alleles will generally involve effects of multiple DNA sequence polymorphisms, arranged in various combinations due to intragenic recombination, which may act additively or nonadditively (Fig. 1B). At the very least, we should expect the sizes of mutational effects and allelic effects to follow different distributions, with allelic effects being much larger in many cases. Depending on the degree to which individual polymorphisms within a gene interact non-additively, the very concept of mutational effects may be ambiguous in some cases. Moreover, both novel mutations and intragenic recombination events occur at low frequencies, so new alleles, once formed by these processes, will be inherited intact from generation to generation.

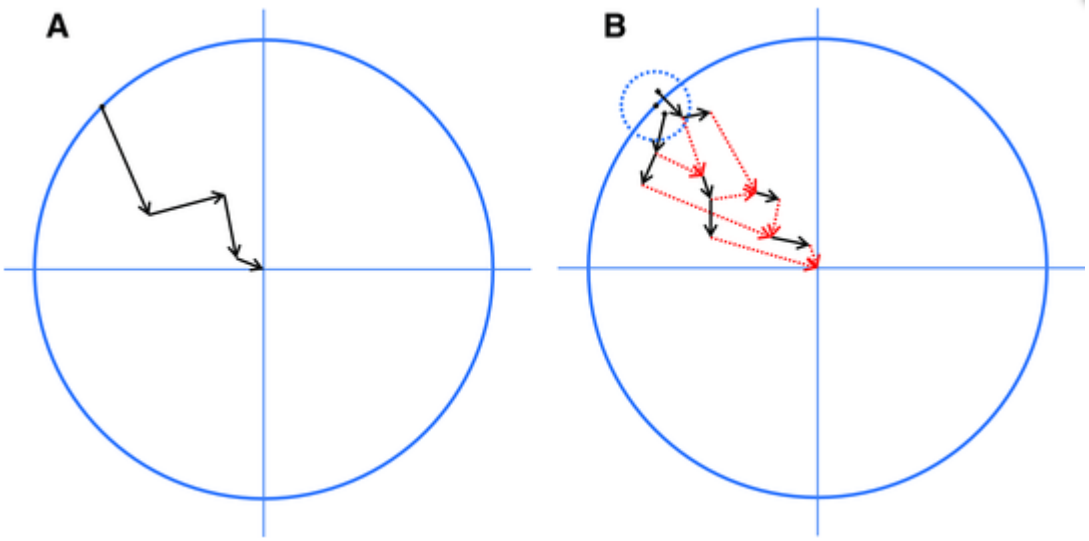

Figure 1. (A) Conventional Fisher–Orr model of an adaptive walk in a two-dimensional phenotypic space (modified from Orr 1998). The population starts with fixation for an invariant genotype conferring a suboptimal phenotype (solid circle), and evolves by sequential fixation of new mutations by natural selection toward the fitness optimum for both traits (intersection of the two axes). (B) A model of adaptation at an allelic level, incorporating a molecular perspective on gene structure. The initial population harbors standing allelic variation at a single locus affecting both traits (smaller dotted circle) centered on the same suboptimal phenotype as in (A). Adaptation involves selection on multiple alleles from the initial population (small diamonds), which gain additional standing variation by incurring new DNA sequence mutations (solid arrows) and undergo intragenic recombination between alleles (dotted arrows). Changes in allele frequencies over time at other loci affecting the same traits would also shift the mean phenotypes of alleles, and for simplicity are not shown here.

Mutational and Allelic Effects in Natural Populations

There is plenty of empirical evidence that the molecular insights into the nature of alleles described above apply to actual complex trait variation in natural populations. Some of the most

enlightening studies involve molecular dissection of alleles at genes already known or strongly inferred to be large-effect QTL.

The first fine-mapping studies of QTL candidate genes in Drosophila melanogaster showed evidence that multiple polymorphic sites contribute incrementally to the allele-substitution effects of single QTL. Stam and Laurie (1996) found that Drosophila Adh alleles with large effects on enzyme levels and activity consist of combinations of smaller-effect polymorphisms. Studies of the delta and achaete-scute complex loci involved in sensory bristle number variation showed similar results (Long et al. 1998; Long et al. 2000). Since then, numerous studies in Drosophila and other organisms have suggested that allelic variation at multiple functional sites may be the norm for QTL.

Reduced production of larval trichomes in Drosophila sechellia relative to related species has been found to result from expression differences in the shavenbaby (svb) gene (Sucena and Stern 2000). These expression differences are the cumulative result of sequence differences in multiple enhancer regions (McGregor et al. 2007). Recombinant alleles with different combinations of enhancers from D. sechellia and sibling species D. mauritiana produced various intermediate patterns, suggesting that trichome loss in D. sechellia evolved through a series of small mutations. Moreover, phenotypic effects of one of these enhancer regions result from the cumulative effects of multiple SNPs that act synergistically (Frankel et al. 2011). Parallel sets of mutations in svb account for similar loss of larval trichomes in D. ezona (Frankel et al. 2012).

Rebeiz et al. (2009) found that quantitative variation in abdominal pigmentation in Ugandan D. melanogaster populations from different altitudes was caused largely by variation at the ebony gene, and was associated with differences in gene expression. Using chimeric reporter constructs, Rebeiz et al. found that at least five substitutions in cis-regulatory sequences contributed to expression differences. Patterns of sequence variation at these sites in natural populations showed that some of the variants were specific to the high-elevation dark-pigmented lines in Uganda. Other contributing variants, however, were widely segregating in various combinations across Africa. These results suggested that a combination of selection on standing variation and novel mutations gave rise to the extant variation. Moreover, intragenic recombination events that brought together the segregating higher-pigmentation enhancer variants contributed to formation of the dark-pigment Ugandan alleles.

In an especially revealing study, Bickel et al. (2011) examined sequence variation at the bab locus, which is responsible for 60% of the variation in female abdominal pigmentation in a different D. melanogaster population. The bab locus consists of two tandemly duplicated paralogs (bab1 and bab2), which were sequenced along with the flanking regions in a set of 94 inbred lines. Variation in pigmentation in a sample of multiple females from these lines was associated with multiple sites in three noncoding regions, and individual polymorphisms were associated with only small effects (~1% of variation on average). Analyses of allele-specific expression in a cross to tester strains suggested that expression levels of bab2 but not bab1 were associated with SNPs in these regions.

Similar complex patterns of allelic variation at the Agouti locus were found to distinguish Nebraska deer mice with light versus dark coats inhabiting light sand dunes versus dark soils, an

apparent adaptation to hide from predators (Linnen et al. 2009; Linnen et al. 2013). In populations inhabiting contact zones between these two environments, separate sets of polymorphisms in Agouti coding and regulatory sequences were found to affect different aspects of pigment intensity and patterning in a partially modular fashion, but cumulatively resulted in alleles with major contrasting effects on coat pigmentation in mice living in the contrasting environments (Linnen et al. 2013). Extensive decay in LD across the Agouti locus in the contact zones suggests a role for recombination in generating the contrasting major effect alleles.

Several studies in maize have also identified evidence of allelic series, in which alleles at a QTL have a range of average phenotypic effects (Harjes et al. 2008; Buckler et al. 2009; Hung et al. 2012; Studer and Doebley 2012). In one example, polymorphisms in multiple regions of the maize lcyE locus incrementally affect the ratio of α- to β-carotenoids, with alleles containing particular combinations of variants in each region having large effects on levels of β-carotenoids (Harjes et al. 2008).

These studies underscore the importance of a gene-level perspective both for understanding both the genetic architecture and the functional basis for complex trait variation and evolution. In each case, finding the responsible loci in the first place involved linkage mapping and/or testing of a priori candidate genes, which allowed the specific or collective contributions of individual polymorphisms to be subsequently evaluated. One has to question whether a "bottom-up" approach starting with detection of individual SNPs would have shed as much light on either the allelic or mutational effects in these genes. In terms of function, these studies support the idea that such large-effect "complex alleles" consisting of many contributing polymorphisms might characterize evolutionary "hotspot" genes (McGregor et al. 2007; Martin and Orgogozo 2013). Contributing factors might include genes that are poised at key positions in developmental networks (McGregor et al. 2007), and modular regulatory organization that allows molecular fine-tuning of phenotypic effects of alleles (Linnen et al. 2013).

Some might argue that incremental and sometimes modular contributions of individual polymorphisms on phenotype and fitness in loci such as Agouti (Linnen et al. 2013) make the very notion of the gene as the functional unit of evolution irrelevant. The complex and intertwined regulatory complexity of the genome revealed by projects such as ENCODE (Gerstein et al. 2007; Djebali et al. 2012; Mudge et al. 2013; Keller 2014) might be seen to undermine further the applicability of a gene-level focus. Are the sheer number and diversity of mechanisms by which multiple SNPs in a single gene affect fitness simply too great for individual genes to matter much? I would argue just the opposite. The regulatory complexity in Agouti and the bab genes, for example, are adaptively important precisely because the genes encode proteins involved in mouse pigmentation and Drosophila abdominal patterning, respectively (Kopp et al. 2003; Linnen et al. 2009; Bickel et al. 2011). Studies showing that complex allelic variation leads to complex effects on phenotype affirm rather than undermine the relevance of gene function for understanding evolution. A molecular understanding of the nature of allelic variation suggests complex allelic effects on phenotype could be quite common. Testing this prediction will require genetic dissection of many more genes in a variety of organisms.

Allelic Effects and Evolutionary Forces

The studies described above make it clear that the accumulation of many small-effect mutations and recombination between them do indeed produce alleles with large contrasting effects on complex phenotypes in nature. However, much of the recent debate has to do with whether such large-effect complex alleles are a common feature of genetic variation, or simply an asterisk of sorts in a world where evolution of "typical" complex traits is dominated by alleles with small effects (Rockman 2012). If small-effect SNPs can contribute to large-effect complex QTL alleles in some circumstances, as described above, why have so many QTL mapping studies in mice, humans, Drosophila, yeast and maize (Flint and Mackay 2009; Mackay et al. 2009; Huang et al. 2010) revealed a lack of large-effect QTL? These questions bring me to my second premise: there is no reason to expect there to be such a thing as a "typical" complex trait in the first place. I would argue that different combinations of evolutionary forces are likely to generate fundamentally different genetic architectures under suitable models of allelic origin. These points have been argued conceptually by others (Fishman et al. 2002; Seehausen et al. 2014) and to some extent explored theoretically (Yeaman and Otto 2011; Yeaman and Whitlock 2011), but again their implications seem to be greatly underappreciated.

Most of the studies reviewed by Mackay and colleagues (Flint and Mackay 2009; Mackay et al. 2009), which suggest that large-effect loci are atypical, were designed to investigate genetic variation within more-or-less panmictic populations. Adaptively important traits in such populations are likely to be under some form of direct or indirect stabilizing selection, which would tend to reduce the frequency of large-effect alleles since individuals carrying them would be more likely to deviate from optimal phenotypes. Likewise, frequencies of alleles that increase disease risk (possibly as a pleiotropic effect of effects on functional phenotypes) should be reduced at rates proportionate to their effects due to purifying selection. Individual polymorphisms that affect phenotype, however, are likely to be found in multiple allelic backgrounds due to subsequent mutations and intragenic recombination, and would tend to be favored in some backgrounds and selected against in others due to differences in the mean phenotype of each allele. Under such circumstances, loci that are hotspots for trait variation (Martin and Orgogozo 2013) could still accumulate causative polymorphisms and account for substantial phenotypic variation. However, causative polymorphisms would tend to be distributed among many alleles in largely balanced configurations, reducing the frequency of alleles with extreme effects and thus reducing the variance among alleles.

By contrast, most QTL studies where large-effect QTL have been found or strongly inferred have been done to investigate the genetic basis for differences between divergently selected populations or closely related species. Examples include the studies described above in Drosophila (McGregor et al. 2007; Rebeiz et al. 2009; Frankel et al. 2011), and deer mice (Linnen et al. 2013), as well as marine versus freshwater sticklebacks (Shapiro et al. 2004; Colosimo et al. 2005), bee- versus hummingbird-pollinated Mimulus spp. (Bradshaw and Schemske 2003; Yuan et al. 2013), Müllerian mimicry in Heliconius butterflies (Baxter et al. 2010; Joron et al. 2011; Kronforst and Papa 2015), domesticated maize versus teosinte (Wang et al. 1999; Studer et al. 2011; Hung et al. 2012), and tomato domestication from wild relatives (Frary et al. 2000; Fridman et al. 2000). In such cases, divergent selection seems more likely to favor contrasting alleles with large effects on selected traits (Fishman et al. 2002; Seehausen et al. 2014) . One of the few studies in which development of complex alleles was specifically

modeled found that divergent selection was likely to generate large-effect alleles, especially in the presence of gene flow (Yeaman and Whitlock 2011). Thus, identification of large-effect QTL in the above studies should come as no surprise. While these large-effect alleles may still represent only the tip of the polygenic iceberg, they clearly have been important in the evolution of divergent phenotypes in these organisms.

The contrasting objectives and sampling designs in these two classes of studies reveal another limitation of the Fisher–Orr geometric model formulation for understanding quantitative trait variation. Orr's (1998) analysis predicting fixation of some large-effect mutations only applies to the distribution of QTL effects when divergently selected populations are being compared. It makes no prediction about variation segregating within individual populations under either stabilizing or directional selection. Even when comparing divergent populations, however, the Orr model only applies when the shift to a new habitat is a single discrete change, such as the adaptation of marine sticklebacks to freshwater environments or adaptation to a new pollinator in flowering plants—a point Orr himself makes clear. Where the shifts in habitat occur gradually or on a continuum, one might not expect to see large-effect substitutions (Rockman 2012). However, this does not preclude large-effect allelic differences from arising when cumulative small-effect substitutions occur and recombine within the same loci. This latter situation appears to be the case with the examples of adaptive evolution in the ebony (Rebeiz et al. 2009) and svb loci (McGregor et al. 2007; Frankel et al. 2011) in Drosophila spp.

Based on the above considerations, we would expect to find both large-effect alleles and individual large-effect polymorphisms most often in cases where large and discrete environmental shifts differentiate two populations. This is precisely what was seen in adaptive evolution of light pigmentation in coastal populations of beach mice (Peromyscus polionotus), where a single coding-region polymorphism in the Mc1r gene makes a large contribution to the light coat color of Gulf coast beach populations (Hoekstra et al. 2006). Similarly, adaptation of marine sticklebacks to freshwater lakes has involved pelvic reduction through independent large-effect deletions in a specific enhancer region in Pitx1 (Chan et al. 2010). Even in these systems, however, large-effect alleles have arisen through other means. Convergent evolution of light pigmentation in Atlantic coast beach mice does not involve the Mc1r mutation found in Gulf coast populations (Steiner et al. 2009), and similar evolution of light coat color in deer mice inhabiting Nebraska sand dunes involved multiple smaller-effect mutations in the Agouti gene, as described above (Linnen et al. 2013). Other adaptations in freshwater sticklebacks have involved independent episodes of selection on standing variation involving alleles segregating in marine populations, such as large-effect alleles in EDA affecting armor plating (Colosimo et al. 2005; Jones et al. 2012; Roesti et al. 2014). It is unclear how many separate functional polymorphisms distinguish alleles in this instance. Finally, Müllerian mimicry in wing patterning in Heliconius melpomene butterflies involves discrete adaptive phenotypes. Nevertheless, there is evidence that large-effect alleles at wing patterning genes result from variation at multiple functional sites (Baxter et al. 2010).

The Trouble with GWAS

Nowhere has the conflation of mutational versus allelic effects been more of a constraint to genetic and evolutionary understanding than the recent proliferation of genome-wide association

studies. This is a key issue due to the growing interest in using GWAS to address evolutionary questions in organisms other than humans (Atwell et al. 2010; Holliday et al. 2010; Fournier-Level et al. 2011; Hancock et al. 2011; Ingvarsson and Street 2011; Filiault and Maloof 2012; Sork et al. 2013; McKown et al. 2014). As explained above, the foundational papers for association mapping were explicitly based on a one QTL—one SNP model (Risch and Merikangas 1996; Long and Langley 1999). The prevailing approaches used in association mapping studies growing at least implicitly out of this foundation are particularly unsuited to detecting effects at the level of functional alleles.

Both the strengths and limitations of association mapping are due to the structure of DNA sequence variation in populations. Association mapping is unequivocally not the same as QTL linkage mapping at much finer resolution. In contrast with linkage-based QTL studies, association studies involve a large number of parents, each segregating at different sites in different loci, sampled across a population. The genetic architecture detected in these studies is inherently more complex than that found in linkage-based studies because they represent the entire spectrum of variation in the sampled populations, not just that of the parents of a particular cross. Moreover, while association studies provide much finer resolution of QTL due to the many meioses separating sampled chromosomes, the relationship between distance and linkage disequilibrium (LD) is much more capricious. At any given location in the genome, alleles have complex and unobserved coalescent relationships based on common ancestry and broken up by past recombination (Nordborg and Tavaré 2002). The strength of association between a sampled SNP and a causative site depends both on the history of recombination events separating them and on where each mutation occurred in the coalescent tree. Thus, unlike linkage-based QTL mapping, SNP-trait associations do not decay monotonically with distance.

Consequently, individual SNPs assayed in GWAS detect trait variation only in proportion to their LD with one or more causative polymorphisms. An assayed SNP will detect the entire phenotypic effect of variation in an associated gene only if it is in absolute LD with every coding or regulatory polymorphism at the locus affecting phenotype. In recent years, there has been a growing recognition that low-frequency causal variants, possibly with large phenotypic effects, are poorly associated with the intermediate-frequency SNPs that are generally assayed in GWAS SNP panels (Dickson et al. 2010). Incomplete LD between sampled SNPs and causal polymorphisms is recognized as one of the factors underlying the "missing heritability" in association studies, with the remainder attributed to loci with effects too small to detect (Yang et al. 2010; Gibson 2012) and possibly to inflated estimates of overall additive genetic variance due to epistasis (Zuk et al. 2012; Hemani et al. 2013). In general, association studies provide little power to separate the contributions of effect size and allele frequency of causative SNPs to the genetic variance the SNPs explain.

These arguments, however, still miss the point that alleles with contrasting effects on quantitative traits or disease risk are likely to differ at multiple causative sites. Unlike the models involving large-effect rare variants, the individual causative SNPs in complex alleles may have small effects, segregate at intermediate frequencies, and occur in various combinations. Standard GWAS techniques that test individual SNP-trait associations one at a time simply cannot capture the magnitude of variation associated with complex alleles. When all SNPs at the bab locus in Drosophila melanogaster were tested individually for associations with abdominal pigmentation,

the smallest P-values were about 10−5 (Bickel et al. 2011), which is weaker than typical genome-wide significance thresholds in GWAS. Thus, a GWAS study would have needed both a larger sample size and a fortuitous choice of sampled SNPs even to identify the bab locus as a contributor to variation in abdominal pigmentation, let alone discover the large size of its effects.

A recent analysis by Gusev et al. (2013) found that inclusion of all sampled SNPs at loci containing a significant SNP for complex disease risk substantially increased the detected heritability. Additional inclusion of all sampled SNPs at loci harboring significant SNPs for other related diseases increased the detected heritability for multiple sclerosis and rheumatoid arthritis five- to sevenfold over that detected using only the significant SNPs for these diseases. However, even these results only include examples for which significant SNP-trait associations were found in the first place. One recent study identified a major QTL region for von Willebrand factor levels using linkage analysis, which was completely missed by a GWAS involving the same cohorts (Desch et al. 2013). These authors proposed that unsampled large-effect rare variants could be responsible, but the possibility that a single gene harbors complex alleles differing at multiple causative sites seems equally plausible. Similarly, in a combined linkage-association study of cystic fibrosis lung-infection severity, a large-effect QTL region identified by linkage analysis showed no significant SNP-trait associations until weighting by the linkage results was applied (Wright et al. 2011).

As these examples show, the limitations of the one QTL—one SNP paradigm extend far beyond evolutionary biology, and have major implications for interpreting the prodigious output of GWAS studies of complex human diseases over the past decade. The extensive body of literature summarized above on the actual genetic architecture of known QTL is simply not being cited in the GWAS literature, and is doing little to inform GWAS design and interpretation. There is no denying that GWAS has succeeded in finding causative genes in many instances, which have been subsequently characterized by detailed follow-up studies. At best, however, SNP-by-SNP GWAS is an inefficient and underpowered tool for estimating the effects of genes underlying complex traits and diseases. One cannot help but wonder how many loci with large effects on complex human diseases have been missed or underestimated by GWAS, and about the extent to which complex alleles account for “missing heritability” (Gusev et al. 2013).

Another property of association studies limits their utility in studies of divergent populations, the very cases where large-effect alleles might be most likely to segregate. Effects of population structure and admixture will produce associations among unlinked SNPs due to shared ancestry of particular variants, and thus generate nonfunctional SNP-trait associations. Consequently, association studies are generally not useful for identifying variants that distinguish phenotypic differences between divergent populations. Including population structure and kinship in structured association studies can substantially factor out spurious genotype–phenotype associations (Yu et al. 2006; Yu et al. 2008). However, if molecular differentiation between populations is strongly correlated with the phenotypic differences between them, the effects of loci underlying between-population differences will also be factored out (Atwell et al. 2010; Ingvarsson and Street 2011; Larsson et al. 2013; McKown et al. 2014). A number of factors besides population structure can also generate misleading associations between assayed polymorphisms and traits (Platt et al. 2010). One of the first studies to use a structured association approach identified strong associations between sequence variants in Dwarf8 and

flowering time variation in maize (Thornsberry et al. 2001). Subsequent research has indicated that this association is largely spurious, and possibly results from selection on multiple traits that are affected by other genes near Dwarf8 (Larsson et al. 2013).

Implications and Recommendations

The current uncertain fate of the QTL perspective has led to ongoing debate about the path forward in understanding the genetic basis of adaptation. One view is that larger sample sizes and more sophisticated study approaches will allow us at last to capture the complexity of genotype-to-phenotype networks (Ayroles et al. 2009; Mackay et al. 2009). Another is that we need to adopt methods more consistent with the infinitesimal model such as genomic prediction, in which all scored polymorphisms are simply incorporated into a random model to predict breeding values, at the expense of gaining functional insights (Ober et al. 2012; Rockman 2012). Still others argue that proximate molecular mechanisms are inherently uninformative about ultimate evolutionary processes in the first place (Travisano and Shaw 2013). Somewhat lost in this debate have been the results of the many studies summarized above that have found large-effect QTL with important adaptive evolutionary consequences segregating between divergent populations or closely-related species in a variety of study systems.

A more biologically appropriate model of allelic variation, which correctly distinguishes mutational and allelic effects in the context of different selective scenarios, could lead to fresh insights in three ways. First, it could lead us to start asking better questions. Continuing to rehash questions about the generalized size distributions of QTL and SNP effects (i.e., QTNs), as if they were the same thing, might simply be leading us farther and farther down the wrong path. Adopting a suitable molecular framework leads to a more relevant set of questions that consider when we might expect to see different patterns of genetic architecture, taking into account both variation at the level of alleles and the steps that generate it. For example, how do new mutations, selection on standing variation, and recombination among existing polymorphisms contribute to allelic diversity? Do different selective regimes affect these respective processes, and thus the evolution of genetic architecture, in predictable ways? Do the same loci that contribute to adaptive divergence between populations in contrasting environments also tend to be polymorphic at the same causal sites within more-or-less stable populations, and thus provide the raw material for repeated evolution when environments change? New research focused on addressing these kinds of questions, using relevant sets of taxa, could provide much more fruitful insights about the connections between functional and evolutionary biology.

Second, adopting a better model of allelic variation could enhance cross-disciplinary awareness of relevant research. Human GWAS studies might lead to richer insights if they incorporated the potential implications of the research on the molecular architecture of complex trait variation highlighted here. Similarly, I think that evolutionary biology research would benefit greatly from a theoretical framework that embraces the full implications of the research on molecular architecture. Rockman's (2012) criticism that we have yet to find any QTNs for key “typical” complex traits is a case in point. Studies of loci such as svb, ebony, and bab1/2 in Drosophila spp. (Macgregor et al. 2005; Rebeiz et al. 2009; Bickel et al. 2011; Frankel et al. 2011) and Agouti in deer mice (Linnen et al. 2013) seem to indicate that the contributions of QTNs to genetic variation might be much more discoverable when evaluated in the context of allelic

variants at the gene level. The catalog of evolutionarily relevant loci by Stern and Orgogozo (2008; updated in Martin and Orgogozo 2013) provides a useful resource for identifying loci and taxa in which further dissection of allelic variation may be fruitful. It seems premature to dismiss such loci as atypical (Rockman 2012) until we have more information on just how frequently they contribute to instances of between-population divergence. Studies of smaller-effect QTL will also be essential in the long run in order to determine just how far an allele-based perspective can take us in understanding how individual causative SNPs contribute to allelic effects. For example, are different effect size distributions at different QTL due to differences in the size of SNP effects or in the number of causative variable sites per gene?

Third, an up-to-date model of allelic variation should lead to more appropriate use of the available tools. The need to replace the SNP-by-SNP approach to GWAS, which is implicitly based on the one QTL—one SNP model, is a prime example. Fortunately, approaches exist for addressing some of the limitations of association mapping. Analytical approaches that incorporate all SNPs within a physical region into a multivariate or random-effects model can be used to evaluate the entire effect of gene locus-sized regions on phenotypic variation (Nagamine et al. 2012; Gusev et al. 2013; Paré et al. 2015). Given the fact that cis-regulatory sequences can be distal to the genes they regulate and may overlap other genes, applying such approaches genome-wide in sliding windows will still be imprecise, but would much more closely approximate the population-level phenotypic variation associated with individual loci. The effectiveness of such an approach will be further enhanced as full-genome next-generation resequencing replaces array-based and bead-based SNP assays, thus assuring that causative polymorphisms are sampled. This can in turn set the stage for more detailed dissection of the role of individual polymorphisms and the process of allelic evolution. The analysis by Rebeiz et al. (2009) of evolution in the Drosophila ebony locus illustrates the potential of such approaches to reveal the molecular history of adaptive allelic evolution. The methods used in that study essentially reconstruct the "adaptive walk" depicted conceptually in Figure 1B based on a modern molecular understanding of allelic variation.

As discussed above, theoretical approaches that model allelic variation more appropriately than the Fisher–Orr framework are available. Finding theoretical distributions for allelic effects and those of the underlying polymorphisms under different selective regimes will be an important area for future theoretical studies. The modeling framework of Yeaman and Whitlock (2011) seems especially promising, as it incorporates alleles originating from multiple mutations, arranged on chromosomes that undergo recombination (though not intragenic), evolving under a flexible set of evolutionary forces.

In closing, I recognize that my arguments may not satisfy evolutionary biologists who see little relevance of genotype-to-phenotype studies for addressing evolutionary questions in the first place. Some proponents of this "back to basics" perspective argue in essence that the plethora of molecular and developmental processes underlying organismal phenotypes are so intertwined that they provide no useful predictive power for evolutionary trajectories, nor have they provided generalizable insights (Travisano and Shaw 2013). In this view, the summary statistical approaches to quantitative genetics developed by Fisher and others still hold the most promise for understanding adaptation, and the main need is for more direct studies of evolutionary processes (Rockman 2012; Travisano and Shaw 2013).

Others have responded in detail to those who take such a strong view of the dichotomy between functional and evolutionary biology; that is proximate versus ultimate causes (Watt 2000; Laland et al. 2011; Lee et al. 2014), and I will not repeat their arguments here. I agree with Travisano and Shaw's (2013) argument that genotype-to-phenotype studies cannot substitute for direct research on evolutionary processes. To my thinking, however, studies of molecular detail have added a great deal of richness to evolutionary research by providing insights on how genotypic variation affects evolutionary processes. By using more suitable models to understand the nature of genotypic variation, I believe we can do a better job of evaluating these connections. My hope is that in doing so we might indeed discover new generalizable principles of adaptation.

Acknowledgments

I thank John Willis and Malcolm Schug for providing encouragement and helpful comments, and Ruth Shaw, Walter Salzburger, and two anonymous reviewers for comments and suggestions that have greatly improved this manuscript.

Literature Cited

Alonso-Blanco, C., S. E. D. El-Assal, G. Coupland, and M. Koornneef. 1998. Analysis of natural allelic variation at flowering time loci in the Landsberg *erecta* and Cape Verde Islands ecotypes of *Arabidopsis thaliana*. *Genetics* **149**:749–764.

Atwell, S., Y. S. Huang, B. J. Vilhjalmsson, G. Willems, M. Horton, Y. Li, D. Meng, A. Platt, A. M. Tarone, T. T. Hu, et al. 2010. Genome-wide association study of 107 phenotypes in Arabidopsis thaliana inbred lines. *Nature* **465**:627–631.

Avery, O. T., C. M. MacLeod, and M. McCarty. 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J. Exp. Med.* **79**:137–158.

Ayroles, J. F., M. A. Carbone, E. A. Stone, K. W. Jordan, R. F. Lyman, M. M. Magwire, S. M. Rollman, L. H. Duncan, F. Lawrence, R. R. H. Anholt, et al. 2009. Systems genetics of complex traits in *Drosophila melanogaster*. *Nat. Rev. Genet.* **41**:299–307.

Barton, N. H. 1999. Clines in polygenic traits. *Genet. Res. Camb.* **74**:223–236.

Barton, N. H. and M. Turelli. 1989. Evolutionary quantitative genetics: how little do we know? *Ann. Rev. Genet.* **23**:337–370.

Baxter, S. W., N. J. Nadeau, L. S. Maroja, P. Wilkinson, B. A. Counterman, A. Dawson, M. Beltran, S. Perez-Espona, N. Chamberlain, L. Ferguson, et al. 2010. Genomic hotspots for adaptation: the population genetics of Müllerian mimicry in the *Heliconius melpomene*clade. *PLoS Genet.* **6**:e1000794.

Beavis, W. D. 1994. The power and deceit of QTL experiments: lessons from comparative QTL studies. Proceedings of the 49th annual corn and sorghum industry research conference, Chicago **49**:250–266.

Benzer, S. 1959. On the topology of the genetic fine structure. *Proc. Natl. Acad. Sci. USA* **45**:1607–1620.

Benzer, S. 1961. On the topography of the genetic fine structure. *Proc. Natl. Acad. Sci. USA* **47**:403–415.

Bickel, R. D., A. Kopp, and S. V. Nuzhdin. 2011. Composite effects of polymorphisms near multiple regulatory elements create a major-effect QTL. *PLoS Genet*. **7**:e1001275.

Bradshaw, H. D., Jr., K. G. Otto, B. E. Frewen, J. K. McKay, and D. W. Schemske. 1998. Quantitative trait loci affecting differences in floral morphology between two species of monkeyflower (*Mimulus*). *Genetics* **149**:367–382.

Bradshaw, H. D., Jr. and D. W. Schemske. 2003. Allele substitution at a flower color locus produces a pollinator shift in monkeyflowers.*Nature* **426**:176–178.

Bradshaw, H. D., Jr., S. M. Wilbert, K. G. Otto, and D. W. Schemske. 1995. Genetic mapping of floral traits associated with reproductive isolation in monkeyflowers (*Mimulus*). *Nature* **376**:762–765.

Buckler, E. S., J. B. Holland, P. J. Bradbury, C. B. Acharya, P. J. Brown, C. Browne, E. Ersoz, S. Flint-Garcia, A. Garcia, J. C. Glaubitz, et al. 2009. The genetic architecture of maize flowering time. *Science* **325**:714–718.

Chan, Y. F., M. E. Marks, F. C. Jones, G. Villarreal, M. D. Shapiro, S. D. Brady, A. M. Southwick, D. M. Absher, J. Grimwood, J. Schmutz, et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**:302–305.

Colosimo, P. F., K. E. Hosemann, S. Balabhadra, G. Villarreal Jr., M. Dickson, J. Grimwood, J. Schmutz, R. M. Myers, D. Schluter, andD. M. Kingsley. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science***307**:1928–1933.

Colosimo, P. F., C. L. Peichel, K. Nereng, B. K. Blackman, M. D. Shapiro, D. Schluter, and D. M. Kingsley. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biol*. **2**:635–641.

Cornforth, T. W. and A. D. Long. 2003. Inferences regarding the numbers and locations of QTLs under multiple-QTL models using interval mapping and composite interval mapping. *Genet. Res. Camb.* **82**:139–149.

Crick, F. H. C., L. Barnett, S. Brenner, and R. J. Watts-Tobin. 1961. General nature of the genetic code for proteins. *Nature***192**:1227–1232.

Desch, K. C., A. B. Ozel, D. Siemieniak, Y. Kalish, J. A. Shavit, C. D. Thornburg, A. A. Sharathkumar, C. P. McHugh, C. C. Laurie, A. Crenshaw, et al. 2013. Linkage analysis identifies a locus for plasma von Willebrand factor undetected by genome-wide association. *Proc. Natl. Acad. Sci. USA* **110**:588–593.

Dickson, S. P., K. Wang, I. Krantz, H. Hakonarson, and D. B. Goldstein. 2010. Rare variants create synthetic genome-wide associations. *PLoS Biol*. **8**:e1000294.

Djebali, S., C. A. Davis, A. Merkel, A. Dobin, T. Lassmann, A. Mortazavi, A. Tanzer, J. Lagarde, W. Lin, F. Schlesinger, et al. 2012.Landscape of transcription in human cells. *Nature* **489**:101–108.

East, E. M. 1916. Studies on size inheritance in *Nicotiana*. *Genetics* **1**:164–176.

El-Assal, S. E. D., C. Alonso-Blanco, A. J. M. Peeters, V. Raz, and M. Koornneef. 2001. A QTL for flowering time in Arabidopsis reveals a novel allele of CRY2. *Nat. Genet*. **29**:435–440.

Feder, M. E. and T. Mitchell-Olds. 2003. Evolutionary and ecological functional genomics. *Nat. Rev. Genet*. **4**:651–657.

Filiault, D. L. and J. N. Maloof. 2012. A genome-wide association study identifies variants underlying the *Arabidopsis thaliana* shade avoidance response. *PLoS Genet*. **8**:e1002589.

Fisher, R. A. 1918. The correlation between relatives on the supposition of Mendelian inheritance. *Trans. Roy. Soc. Edinburgh***52**:399–433.

Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford (Variorum ed. 1992.Oxford Univ. Press, New York).

Fishman, L., A. J. Kelly, and J. H. Willis. 2002. Minor quantitative trait loci underlie floral traits associated with mating system divergence in *Mimulus*. *Evolution* **56**:2138–2155.

Flint, J. and T. F. C. Mackay. 2009. Genetic architecture of quantitative traits in mice, flies, and humans. *Genome Res*. **19**:723–733.

Fournier-Level, A., A. Korte, M. D. Cooper, M. Nordborg, J. Schmitt, and A. M. Wilczek. 2011. A map of local adaptation in *Arabidopsis thaliana*. *Science* **334**:86–89.

Frankel, N., D. F. Erezyilmaz, A. P. McGregor, S. Wang, F. Payre, and D. L. Stern. 2011. Morphological evolution caused by many subtle-effect substitutions in regulatory DNA. *Nature* **474**:598–603.

Frankel, N., S. Wang, and D. L. Stern. 2012. Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc. Natl. Acad. Sci. USA* **109**:20975–20979.

Frary, A., T. C. Nesbitt, A. Frary, S. Grandillo, E. van der Knaap, B. Cong, J. Liu, J. Meller, R. Elber, K. B. Alpert, et al. 2000. *fw2.2*: a quantitative trait locus key to the evolution of tomato fruit size. *Science* **289**:85–88.

Fridman, E., T. Pleban, and D. Zamir. 2000. A recombination hotspot delimits a wild-species quantitative trait locus for tomato sugar content to 484 bp within an invertase gene. *Proc. Natl. Acad. Sci. USA* **97**:4718–4723.

Gerstein, M., C. Bruce, J. S. Rozowsky, D. Zheng, J. Du, J. O. Korbel, O. Emanuelsson, Z. D. Zhang, S. Weissman, and M. Snyder.2007. What is a gene, post-ENCODE? History and updated definition. *Genome Res*. **17**:669–681.

Gibson, G. 2012. Rare and common variants: twenty arguments. *Nat. Rev. Genet.* **13**:135–145.

Graham, J. I., D. W. Wolff, and C. W. Stuber. 1997. Characterization of a yield quantitative trait locus on chromosome five of maize by fine mapping. *Crop Sci*. **37**:1601–1610.

Gudbjartsson, D. F., G. B. Walters, G. Thorleifsson, H. Stefansson, B. V. Halldorsson, P. Zusmanovich, P. Sulem, S. Thorlacius, A. Gylfason, S. Steinberg, et al. 2008. Many sequence variants affecting diversity of adult human height. *Nat. Genet.* **40**:609–615.

Gusev, A., G. Bhatia, N. Zaitlen, B. J. Vilhjalmsson, D. Diogo, E. A. Stahl, P. K. Gregersen, J. Worthington, L. Klareskog, S. Raychaudhuri, et al. 2013. Quantifying missing heritability at known GWAS loci. *PLoS Genet.* **9**:e1003993.

Hancock, A. M., B. Brachi, N. Faure, M. W. Horton, L. B. Jarymowycz, F. G. Sperone, C. Toomajian, F. Roux, and J. Bergelson. 2011.Adaptation to climate across the *Arabidopsis thaliana* genome. *Science* **334**:83–86.

Harjes, C. E., T. R. Rocheford, L. Bai, T. P. Brutnell, C. B. Kandianis, S. G. Sowinski, A. E. Stapleton, R. Vallabhaneni, M. Williams, E. T. Wurtzel, et al. 2008. Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* **319**:330–333.

Hemani, G., S. Knott, and C. Haley. 2013. An evolutionary perspective on epistasis and the missing heritability. *PLoS Genet*.**9**:e1003295.

Hoekstra, H. E., R. J. Hirschmann, R. A. Bundey, P. A. Insel, and J. P. Crossland. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**:101–104.

Holliday, J. A., K. Ritland, and S. N. Aitken. 2010. Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytol*. **188**:501–514.

Huang, Y.-F., D. Madur, V. Combes, C. L. Ky, D. Coubriche, P. Jamin, S. Jouanne, F. Dumas, E. Bouty, P. Bertin, et al. 2010. The genetic architecture of grain yield and related traits in *Zea maize* L. revealed by comparing intermated and conventional populations. *Genetics* **186**:395–404.

Hudson, R. R. and N. L. Kaplan. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics* **111**:147–164.

Hung, H.-Y., L. M. Shannon, F. Tian, P. J. Bradbury, C. Chen, S. Flint-Garcia, M. D. McMullen, D. Ware, E. S. Buckler, J. F. Doebley, et al. 2012. *ZmCCT* and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc. Natl. Acad. Sci. USA* **109**:E1913–E1921.

Ingvarsson, P. K. and N. R. Street. 2011. Association genetics of complex traits in plants. *New Phytol*. **189**:909–922.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino, and C. Dean. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. *Science* **290**:344–347.

Jones, F. C., M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, et al. 2012.The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**:55–61.

Joron, M., L. Frezal, R. T. Jones, N. L. Chamberlain, S. F. Lee, C. R. Haag, A. Whibley, M. Becuwe, S. W. Baxter, L. Ferguson, et al.2011. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. *Nature* **477**:203–206.

Juenger, T. E., M. D. Purugganan, and T. F. C. Mackay. 2000. Quantitative trait loci for floral morphology in *Arabidopsis thaliana*.*Genetics* **156**:1379–1392.

Keller, E. F. 2014. From gene action to reactive genomes. *J. Physiol*. **592**:2423–2429.

Kimura, M. 1965. A stochastic model concerning the maintenance of genetic variability in quantitative characters. *Proc. Natl. Acad. Sci. USA* **54**:731–736.

Kopp, A., R. M. Graze, S. Xu, S. B. Carroll, and S. V. Nuzhdin. 2003. Quantitative trait loci responsible for variation in sexually dimorphic traits in *Drosophila melanogaster*. *Genetics* **163**:771–787.

Kronforst, M. R. and R. Papa. 2015. The functional basis of wing patterning in *Heliconius* butterflies: the molecules behind mimicry.*Genetics* **200**:1–19.

Kroymann, J. and T. Mitchell-Olds. 2005. Epistasis and balanced polymorphism influencing complex trait variation. *Nature* **435**:95–98.

Lai, C., R. F. Lyman, A. D. Long, C. H. Langley, and T. F. C. Mackay. 1994. Naturally occurring variation in bristle number and DNA polymorphisms at the *scabrous* locus of *Drosophila melanogaster*. *Science* **266**:1697–1702.

Laland, K. N., K. Sterelny, J. Odling-Smee, W. Hoppitt, and T. Uller. 2011. Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science* **334**:1512–1516.

Lande, R. 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res. Camb.***26**:221–235.

Lander, E. S. and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics***121**:185–199.

Larsson, S. J., A. E. Lipka, and E. S. Buckler. 2013. Lessons from *Dwarf8* on the strengths and weaknesses of structured association mapping. *PLoS Genet*. **92**:e1003246.

Latter, B. D. H. 1970. Selection in finite populations with multiple alleles. II Centripetal selection, mutation, and isoallelic variation.*Genetics* **66**:165–186.

Lee, Y. W., B. A. Gould, and J. R. Stinchcombe. 2014. Identifying the genes underlying quantitative traits: a rationale for the QTN programme. *AoB Plants* **6**:plu004.

Lettre, G., A. U. Jackson, C. Gieger, F. R. Schumacher, S. I. Berndt, S. Sanna, S. Eyheramendy, B. F. Voight, J. L. Butler, C. Guiducci, et al. 2008. Identification of ten loci associated with height highlights new biological pathways in human growth. *Nat. Genet*.**40**:584–591.

Linnen, C. R., E. P. Kingsley, J. D. Jensen, and H. E. Hoekstra. 2009. On the origin and spread of an adaptive allele in deer mice.*Science* **325**: 1095–1098.

Linnen, C. R., Y.-P. Poh, B. K. Peterson, R. D. H. Barrett, J. G. Larson, J. D. Jensen, and H. E. Hoekstra. 2013. Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**:1312–1316.

Liu, Z. and J. C. M. Dekkers. 1998. Least squares interval mapping of quantitative trait loci under the infinitesimal genetic model in outbred populations. *Genetics* **148**:495–505.

Long, A. D. and C. H. Langley. 1999. The power of association studies to detect the contribution of candidate genetic loci to variation in complex traits. *Genome Res*. **9**:720–731.

Long, A. D., R. F. Lyman, C. H. Langley, and T. F. C. Mackay. 1998. Two sites in the *Delta* gene region contribute to naturally occurring variation in bristle number in *Drosophila melanogaster*. *Genetics* **149**:999–1017.

Long, A. D., R. F. Lyman, A. H. Morgan, C. H. Langley, and T. F. C. Mackay. 2000. Both naturally occurring insertions of transposable elements and intermediate frequency polymorphisms at the *achaete-scute* complex are associated with variation in bristle number in *Drosophila melanogaster*. *Genetics* **154**:1255–1269.

Lyman, R. F. and T. F. C. Mackay. 1998. Candidate quantitative trait loci and naturally occurring phenotypic variation for bristle number in *Drosophila melanogaster*: the *Delta-Hairless* gene region. *Genetics* **149**:983–998.

Lynch, M. and B. Walsh. 1998. *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, MA.

Macgregor, S., S. A. Knott, I. White, and P. M. Visscher. 2005. Quantitative trait locus analysis of longitudinal quantitative trait data in complex pedigrees. *Genetics* **171**:1365.

Mackay, T. F. C. 1995. The genetic basis of quantitative variation: numbers of sensory bristles of *Drosophila melanogaster* as a model system. *Trends Genet*. **11**:464–470.

Mackay, T. F. C. 2001a. The genetic architecture of quantitative traits. *Ann. Rev. Genet.* **35**:303–339.

Mackay, T. F. C. 2001b. Quantitative trait loci in *Drosophila*. *Nat. Rev. Genet.* **2**:11–20.

Mackay, T. F. C. and C. H. Langley. 1990. Molecular and phenotypic variation in the *achaete-scute* region of *Drosophila melanogaster*. *Nature* **348**:64–66.

Mackay, T. F. C., E. A. Stone, and J. F. Ayroles. 2009. The genetics of quantitative traits: challenges and prospects. *Nat. Rev. Genet.***10**:565–577.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff, D. J. Hunter, M. I. McCarthy, E. M. Ramos, L. R. Cardon, A. Chakravarti, et al. 2009. Finding the missing heritability of complex diseases. *Nature* **461**:747–753.

Martin, A. and V. Orgogozo. 2013. The loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution***67**:1235–1250.
Mayr, E. 1961. Cause and effect in biology. *Science* **134**:1501–1506.
Mayr, E. 1982. *The growth of biological thought*. Harvard Univ. Press, Cambridge, MA.
McGregor, A. P., V. Orgogozo, I. Delon, J. Zanet, D. G. Srinivasan, F. Payre, and D. L. Stern. 2007. Morphological evolution through multiple *cis*-regulatory mutations at a single gene. *Nature* **448**:587–590.
McKown, A. D., J. Klápště, R. D. Guy, A. Geraldes, I. Porth, J. Hannemann, M. Friedmann, W. Muchero, G. A. Tuskan, J. Ehlting, et al.2014. Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytol*. **203**:535–553.
Messer, P. W. and D. A. Petrov. 2013. Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol. Evol.***28**:659–669.
Mudge, J. M., A. Frankish, and J. Harrow. 2013. Functional transcriptomics in the post-ENCODE era. *Genome Res*. **23**:1961–1973.
Nagamine, Y., R. Pong-Wong, P. Navarro, V. Vitart, C. Hayward, I. Rudan, H. Campbell, J. Wilson, S. Wild, A. A. Hicks, et al. 2012.Localising loci underlying complex trait variation using regional genomic relationship mapping. *PLoS One* **7**:e46501.
Nilsson-Ehle, H. 1909. Kreuzungsuntersuchungen and Hafer und Weizen. Lunds Univ. Arsskrift, n.s., Series 2 **5**:1–122.
Noor, M. A. F., A. L. Cunningham, and J. C. Larkin. 2001. Consequences of recombination rate variation on quantitative trait locus mapping studies: simulations based on the *Drosophila melanogaster* genome. *Genetics* **159**:581–588.
Nordborg, M. and S. Tavaré. 2002. Linkage disequilibrium: what history has to tell us. *Trends Genet.* **18**:83–90.
Ober, U., J. F. Ayroles, E. A. Stone, S. Richards, D. Zhu, R. A. Gibbs, C. Stricker, D. Gianola, M. Schlather, T. F. C. Mackay, et al. 2012.Using whole-genome sequence data to predict quantitative trait phenotypes in *Drosophila melanogaster*. *PLoS Genet*.**8**:e1002685.
Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution***52**:935–949.
Orr, H. A. 2005. The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* **6**:119–127.
Orr, H. A. and J. A. Coyne. 1992. The genetics of adaptation: a reassessment. *Am. Nat.* **140**:725–742.
Palsson, A., J. Dodgson, I. Dworkin, and G. Gibson. 2005. Tests for the replication of an association between *Efgr* and natural variation in *Drosophila melanogaster* wing morphology. *BMC Genet.* **6**:44.
Paré, G., S. Asama, and W. Q. Deng. 2015. Contribution of large region joint associations to complex traits genetics. *PLoS Genet*.**11**:e1005103.
Pasyukova, E. G., C. Vieira, and T. F. C. Mackay. 2000. Deficiency mapping of quantitative trait loci affecting longevity in *Drosophila melanogaster*. *Genetics* **156**:1129–1146.
Paterson, A. H., E. S. Lander, J. D. Hewitt, S. Peterson, S. E. Lincoln, and S. D. Tanksley. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. *Nature* **335**:721–726.

Platt, A., B. J. Vilhjálmsson, and M. Nordborg. 2010. Conditions under which genome-wide association studies will be positively misleading. *Genetics* **186**:1045–1052.

Rebeiz, M., J. E. Pool, V. A. Kassner, C. F. Aquadro, and S. B. Carroll. 2009. Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* **326**:1663–1667.

Remington, D. L. 2009. Effects of genetic and environmental factors on trait network predictions from quantitative trait locus data.*Genetics* **181**:1087–1099.

Remington, D. L., P. H. Leinonen, J. Leppälä, and O. Savolainen. 2013. Complex genetic effects on early vegetative development shape resource allocation differences between *Arabidopsis lyrata* populations. *Genetics* **195**:1087–1102.

Risch, N. and K. Merikangas. 1996. The future of genetic studies of complex human diseases. *Science* **273**:1516–1517.

Robertson, A. 1967. The nature of quantitative genetic variation. Pp. 265–280 *in* A. Brink, ed. *Heritage from Mendel*. Wisconsin Univ. Press, Fort Collins, CO.

Rockman, M. V. 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* **66**:1–17.

Roesti, M., S. Gavrilets, A. P. Hendry, W. Salzburger, and D. Berner. 2014. The genomic signature of parallel adaptation from shared genetic variation. *Mol. Ecol.* **23**:3944–3956.

Schemske, D. W. and H. D. Bradshaw, Jr. 1999. Pollinator preference and the evolution of floral traits in monkeyflowers (*Mimulus*).*Proc. Natl. Acad. Sci. USA* **96**:11910–11915.

Seehausen, O., R. K. Butlin, I. Keller, C. E. Wagner, J. W. Boughman, P. A. Hohenlohe, C. L. Peichel, G.-P. Saetre, C. Bank, Å. Brännström, et al. 2014. Genomics and the origin of species. *Nat. Rev. Genet.* **15**:176–192.

Shapiro, M. D., M. E. Marks, C. L. Peichel, B. K. Blackman, K. S. Nereng, B. Jonsson, D. Schluter, and D. M. Kingsley. 2004. Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**:717–723.

Sork, V. L., S. N. Aitken, R. J. Dyer, A. J. Eckert, P. Legendre, and D. B. Neale. 2013. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genet. Genomes* **9**:901–911.

Stam, L. F. and C. C. Laurie. 1996. Molecular dissection of a major gene effect on a quantitative trait: the level of alcohol dehydrogenase expression in *Drosophila melanogaster*. *Genetics* **144**:1559–1564.

Steiner, C. C., H. Römpler, L. M. Boettger, T. Schöneberg, and H. E. Hoekstra. 2009. The genetic basis of phenotypic convergence in beach mice: similar pigment patterns but different genes. *Mol. Biol. Evol.* **26**:35–45.

Steinmetz, L. M., H. Sinha, D. R. Richards, J. I. Spiegelman, P. J. Oefner, J. H. McCusker, and R. W. Davis. 2012. Dissecting the architecture of a quantitative trait locus in yeast. *Nature* **416**:326–330.

Stern, D. L. and V. Orgogozo. 2008. The loci of evolution: how predictable is genetic evolution? *Evolution* **62**:2155–2177.

Studer, A., Q. Zhao, J. Ross-Ibarra, and J. Doebley. 2011. Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat. Genet.* **43**:1160–1163.

Studer, A. J. and J. F. Doebley. 2012. Evidence for a natural allelic series at the maize domestication locus *teosinte branched1*.*Genetics* **191**:951–958.

Sucena, E. and D. L. Stern. 2000. Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven-baby*. *Proc. Natl. Acad. Sci. USA* **97**:4530–4534.

Templeton, A. R., A. G. Clark, K. M. Weiss, D. A. Nickerson, E. Boerwinkle, and C. F. Sing. 2000a. Recombinational and mutational hotspots within the human lipoprotein lipase gene. *Am. J. Hum. Genet.* **66**:69–83.

Templeton, A. R., K. M. Weiss, D. A. Nickerson, E. Boerwinkle, and C. F. Sing. 2000b. Cladistic structure within the human *lipoprotein lipase* gene and its implications for phenotypic association studies. *Genetics* **156**:1259–1275.

Thornsberry, J. M., M. M. Goodman, J. Doebley, S. Kresovich, D. Nielsen, and E. S. Buckler, IV. 2001. *Dwarf8* polymorphisms associate with variation in flowering time. *Nat. Genet.* **28**:286–289.

Travisano, M. and R. G. Shaw. 2013. Lost in the map. *Evolution* **67**:305–314.

Ungerer, M. C., S. S. Halldorsdottir, J. L. Modliszewski, T. F. C. Mackay, and M. D. Purugganan. 2002. Quantitative trait loci for inflorescence development in *Arabidopsis thaliana*. *Genetics* **160**:1133–1151.

Ungerer, M. C., S. S. Halldorsdottir, M. D. Purugganan, and T. F. C. Mackay. 2003. Genotype-environment interactions at quantitative trait loci affecting inflorescence development in *Arabidopsis thaliana*. *Genetics* **165**:353–365.

Visscher, P. M. and C. S. Haley. 1996. Detection of putative quantitative trait loci in line crosses under infinitesimal genetic models.*Theor. Appl. Genet.* **93**:691–702.

Wang, R.-L., A. Stec, J. Hey, L. Lukens, and J. Doebley. 1999. The limits of selection during maize domestication. *Nature***398**:236–239.

Watt, W. B. 2000. Avoiding paradigm-based limits to knowledge of evolution. *Evol. Biol.* **32**:73–96.

Weedon, M. N., H. Lango, C. M. Lindgren, C. Wallace, D. M. Evans, M. Mangino, R. M. Freathy, J. R. B. Perry, S. Stevens, A. S. Hall, et al. 2008. Genome-wide association analysis identifies 20 loci that influence adult height. *Nat. Genet.* **40**:575–583.

Weinig, C., L. A. Dorn, N. Kane, Z. M. German, S. S. Halldorsdottir, M. C. Ungerer, Y. Toyonaga, T. F. C. Mackay, M. D. Purugganan, and J. Schmitt. 2003. Heterogeneous selection at specific loci in natural environments in *Arabidopsis thaliana*. *Genetics***165**:321–329.

Wright, F. A., L. J. Strug, V. K. Doshi, C. W. Commander, S. M. Blackman, L. Sun, Y. Berthiaume, D. Cutler, A. Cojocaru, J. M. Collaco, et al. 2011. Genome-wide association and linkage identify modifier loci of lung disease severity in cystic fibrosis at 11p13 and 20q13.2. *Nat. Genet.* **43**:539–546.

Wright, S. 1921. Systems of mating. *Genetics* **6**:111–178.

Wright, S. 1934a. An analysis of variability in number of digits in an inbred strain of guinea pigs. *Genetics* **19**:506–536.

Wright, S. 1934b. The results of crosses between inbred strains of guinea pigs, differing in number of digits. *Genetics* **19**:537–551.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, et al. 2010. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**:565–569.

Yano, M. 2001. Genetic and molecular dissection of naturally occurring variation. *Curr. Opin. Plant Biol.* **4**:130–135.

Yeaman, S. and S. P. Otto. 2011. Establishment and maintenance of adaptive genetic divergence under migration, selection, and drift.*Evolution* **65**:2123–2129.

Yeaman, S. and M. C. Whitlock. 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution***65**:1897–1911.

Yu, J., G. Pressoir, W. H. Briggs, I. V. Bi, M. Yamasaki, J. Doebley, M. McMullen, B. S. Gaut, D. M. Nielsen, J. B. Holland, et al. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet*. **38**:203–208.

Yu, J. M., J. B. Holland, M. D. McMullen, and E. S. Buckler. 2008. Genetic design and statistical power of nested association mapping in maize. *Genetics* **178**:539–551.

Yuan, Y.-W., J. M. Sagawa, R. C. Young, B. J. Christensen, and H. D. Bradshaw, Jr. 2013. Genetic dissection of a major anthocyanin QTL contributing to pollinator-mediated reproductive isolation between sister species of *Mimulus*. *Genetics* **194**:255–263.

Zuk, O., E. Hechter, S. R. Sunyaev, and E. S. Lander. 2012. The mystery of missing heritabilty: genetic interactions create phantom heritability. *Proc. Natl. Acad. Sci. USA* **109**:1193–1198.