Trends in Genetics

CellPress
REVIEWS

Review

Analysis of Epistasis in Natural Traits Using Model Organisms

Richard F. Campbell,[1] Patrick T. McGrath ,[1,2,*] and Annalise B. Paaby[1]

The ability to detect and understand epistasis in natural populations is important for understanding how biological traits are influenced by genetic variation. However, identification and characterization of epistasis in natural populations remains difficult due to statistical issues that arise as a result of multiple comparisons, and the fact that most genetic variants segregate at low allele frequencies. In this review, we discuss how model organisms may be used to manipulate genotypic combinations to power the detection of epistasis as well as test interactions between specific genes. Findings from a number of species indicate that statistical epistasis is pervasive between natural genetic variants. However, the properties of experimental systems that enable analysis of epistasis also constrain extrapolation of these results back into natural populations.

Highlights

Identification of statistical epistasis in natural populations remains challenging due to the relationship between allele frequency and statistical power.

Artificial populations have been constructed in model organisms to detect statistical epistasis between two regions of the genome; however, it is difficult to use these results to understand how epistasis operates in natural populations.

Studies of focal perturbations in defined genetic backgrounds suggests that natural selection can influence the types of nonadditive relationships that exist.

The Genotype to Phenotype Map in Natural Populations

Complex phenotypes, including common diseases, are controlled by many genes and environmental factors; potentially in unexpected ways. The extent to which epistasis, or interaction between genes, complicates the relationship between genotype and phenotype is controversial. Does a genetic variant essentially play the same role in each human individual? Or, does its effect depend on epistatic interactions with other causal alleles, potentially obstructing the mapping of genotype to phenotype? Epistasis could be important for traits with a simple genetic basis, such as rare monogenic disorders, as well. For example, cystic fibrosis is caused by mutations in the *CFTR* gene, but alleles at other loci influence how the disease manifests in the lungs, intestines, pancreas, and metabolism [1]. Consequently, knowledge of an individual's **genetic background** (see Glossary) is necessary to fully understand the role of a genetic variant in influencing phenotype. These dynamics illustrate why, despite being difficult to study, epistasis remains a critical topic in genetics and evolutionary biology research, and carries important implications for human traits.

Understanding the connection between genotypic and phenotypic variation in natural populations motivates a large amount of research in the fields of human genetics, quantitative genetics, and evolutionary biology. Genetic variants are the cause of differences in genetic disease risk in humans, the basis for selectable traits in animal and plant breeding, and the substrate for evolutionary processes in all species. Identifying the causal genetic variants responsible for **natural traits** has the potential to improve: applied efforts in precision medicine for the prediction of individual disease risk and application of personalized therapies; efforts in agriculture for engineering crops and livestock with increased production; and efforts in evolutionary biology to understand the genetic mechanisms responsible for adaptation and speciation. However, connecting genotypic variants to their phenotypic effects is difficult for most natural traits. The difficulty is a consequence of extreme **polygenicity**, as most natural trait variation is caused by many alleles of small effect; sometimes acting in concert with rarer

[1]Department of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, 30332 USA
[2]Department of Physics, Georgia Institute of Technology, Atlanta, GA, 30332 USA

*Correspondence:
patrick.mcgrath@biology.gatech.edu
(P.T. McGrath).

alleles of small-to-medium effect [2,3]. Furthermore, interactions between alleles may be nonadditive, in that the effect of any one allele is conditional upon the genotypes of additional loci throughout the genome. This interaction is typically referred to as epistasis.

Determining the magnitude and nature of epistasis in the expression of natural trait variation is essential for resolving the relationship between genotype and phenotype. However, investigation into epistasis is constrained by experimental limitations, including: low incidence of rare alleles in the population; insufficient statistical power to interrogate all possible allelic combinations; and inability to separate putatively functional variants within haplotypes segregating in mapping populations due to **linkage disequilibrium**. Perhaps as a consequence of these limitations, epistasis has not been identified at any large scale in humans, and the importance of epistasis in the determination of common diseases and other traits of interest is still debated. Resolution of this question is important for predicting human traits from genotypic information, especially when comparing models across populations separated geographically (e.g., different countries) or temporally (e.g., Neanderthals), as epistasis will modify the apparent effect of a genetic variant when the genetic backgrounds of individuals within a population change. Identification of epistasis and inferences about its nature have primarily emerged from work in model organisms, whose experimental tractability permits manipulation of genotypes and allows for creation of **artificial populations** that overcome some of the obstacles listed above. In a variety of model organisms, epistasis has been shown to be pervasive in these populations and sometimes dramatic in the determination of biological traits. Bridging the gap between findings from experimental systems and observations made in human and other natural populations remains an important yet unresolved issue.

In this review, we highlight approaches for identifying epistasis in natural traits using model organisms, and how different approaches interrogate different aspects of gene interaction. We first provide a short primer to help understand how epistasis is defined and measured, including at both the population and organism level. Next, we describe how epistasis is identified in model organisms using mapping populations and genetic manipulations in experimental lines. Finally, we describe how various factors make generalization of results from experimental systems to natural populations difficult, and how idiosyncrasies between populations and species govern the nature and extent of epistasis in natural populations.

What Is Epistasis?

As has been lamented at the beginning of many reviews, the term epistasis may refer to multiple different phenomena (Box 1). Attempts to define these usages [4–6] and establish common terminology has failed to gain much traction, so typically a reader must rely on context to infer the type of epistasis under discussion [5]. Epistasis can sometimes refer to proteins that interact in some biologically meaningful way – either biophysically or in functional pathways – however, we avoid such usage as it creates additional confusion in an already confusing field. We instead restrict usage of epistasis to interactions between genic elements in the expression of phenotype. In studies of epistasis in natural traits, which is the topic of this review, gene interaction is nearly always measured statistically and without regard to biological mechanism; the interacting loci may be known or unknown. Because experimental methodology essentially defines the type of epistasis that is measured, we begin by describing historical and contemporary approaches to measuring epistasis.

Bateson versus Fisher Epistasis

The most important split in the meaning of epistasis stems from its two independent origins: the work of William Bateson in the application of the Mendelian approach to the study of inheritance

Glossary

Allele pair: designation for the specific alleles present within a diploid animal at a single variable site.
Artificial population: population generated by an experimenter to study the genetic basis of a trait, typically through controlled breeding and initiated from two or more individuals sampled from a natural population. A number of common breeding designs are traditionally used in quantitative genetics, which can be expressed as a pedigree.
Balancing selection: maintenance of two or more alleles at a locus by the action of natural selection.
Biometry: use of statistics to analyze biological data.
Breeding value: average value of an individual's phenotype, as judged by the phenotypes of the progeny.
Classical epistasis: approach used in classical genetics to analyze genetic pathways and order genes within them. This is traditionally recognized through analysis of ratios of phenotypes.
Compositional epistasis: epistasis arising from nonlinear interactions among allele pairs in an otherwise invariant genetic background. A typical test of compositional epistasis uses introgression or genome editing to generate all possible genotype combinations at two loci.
Genetic background: genomic context of an individual, that is, the full genotype across all loci.
Epistatic variance: one of the components of genetic variance (i.e., amount of phenotypic variance that can be attributed to genetic factors). Epistatic variance is sometimes called statistical epistasis.
Hardy–Weinberg equilibrium: if a population follows random mating, the genotype frequencies for a given allele pair can be estimated from the allele frequencies. Furthermore, if two loci are unlinked, the genotype frequencies of both allele pairs can be found by multiplying the individual probabilities.
Higher-order epistasis: nonadditive interactions between three or more different allele pairs. This is sometimes measured by comparing pairwise epistasis in different genetic backgrounds.

Box 1. Different Types of Epistasis

Epistasis might be the most defined and subcategorized terms in all of biology. Dozens of different types of epistasis can be found in the literature. While we have defined some of the major types of epistasis for the purposes of the paper discussion, it is useful for the naïve reader to also consider some of the motivating reasons that lead to the definitions of subtypes of epistasis. For example, consider classical epistasis, where phenotype ratios depart from the expected ratios of 9: 3: 3: 1. While Bateson was primarily concerned with color characters that followed a 9: 3: 4 ratio, for other traits, different ratios can occur. This motivated the definitions of new forms of epistasis for many of these ratios, such as recessive epistasis (9: 3: 4), duplicate recessive epistasis (9:7), dominant epistasis (12: 3: 1), duplicate recessive epistasis (13: 3), and duplicate gene epistasis (15: 1). A lack of consistency in naming conventions also means that these different ratios are sometimes referred to with different names.

For measuring statistical epistasis in haploids, additional definitions are used to summarize the exact relationship of phenotypes between four genotypes. For example, positive or synergistic epistasis occurs when a two-locus combination of alleles displays a phenotype beyond that expected from the individual effect of the alleles. Negative epistasis, sign epistasis, and reciprocal sign epistasis are also terms used to describe additional types of phenotypic regimes of the four genotypes.

As discussed in Box 2, measuring epistasis in diploids requires estimating a four-component vector. Each component can be named, such as additive–additive epistasis, additive–dominance epistasis, dominance–additive epistasis, or dominance–dominance epistasis.

Finally, epistasis is sometimes described in terms of physical or biological interactions between two proteins, irrespective of any genotype–phenotype relationship. This is sometimes called functional epistasis, mechanistic epistasis, or biological epistasis.

patterns of categorical traits; and the work of R.A. Fisher in applying **biometry** to studies of natural selection on continuously varying traits. The former is called Bateson's epistasis or **classical epistasis** [7] and is most commonly associated with classical geneticists' study of molecular and cellular pathways that control phenotypes of interest. The latter is called Fisher's epistasis or **statistical epistasis** [5,8], and typically assumes high polygenicity for the trait of interest and is measured through the use of regression or linear mixed models. While we will use statistical epistasis throughout this review (and give it a more precise definition below), we note that the term is one of the most inconsistently used in the literature.

Bateson first defined epistatic genes in studies of color inheritance: epistasis was defined as a masking phenomenon wherein alleles of one gene hide the effect of the alleles of a second gene. Masking of alleles can be measured using phenotypic ratios of categorical traits in an F_2 intercross, for example, when the loss of one phenotypic class produces 9: 3: 4 ratios instead of the expected 9: 3: 3: 1 [9]. Although Bateson worked with natural variation, contemporary molecular geneticists typically use laboratory-induced, loss-of-function mutations in the classical approach, which has proven powerful for analyzing and ordering genes into functional pathways [10,11].

Most analyses into the genetic basis of natural trait variation require a departure from the classical approach, for several reasons. First, most traits vary continuously rather than categorically, such as height or the expression of common diseases, which are modeled by an underlying liability scale. Calculating F_2 ratios is impossible for such traits. Second, classical epistasis experiments primarily test interactions between pairs of genes [11], but natural populations harbor large numbers of allelic variants that may influence gene function and mediate trait expression in myriad ways. In 1918, Fisher published a research paper on the application of biometry to bridge the divide between Mendelian genetics and natural populations [12], which accounts for both the quantitative action of Mendelian factors and interactions between two or more **allele pairs**. The Fisherian paradigm is useful as it allows a

Introgression: replacement of a section of the genome, of one strain by another, by a genetic cross between strains and then repeated backcrossing with the initial recipient strain.

Linkage disequilibrium: nonrandom association of alleles at different loci in a population. In an outbred, sexually reproducing population, linkage disequilibrium is short (~1 kb). Two loci must be close together to segregate nonindependently. For artificial populations, the physical distance between two linked loci is larger (of the order of 1 Mb), due to the limited number of generations of intercrossing between the parental lines. Linkage disequilibrium largely limits experimenters from identifying causative mutations without additional experiments.

Minor allele frequency (MAF): frequency of the second most common allele in a given population. The minor allele frequency is typically used to categorize alleles as rare (MAF <0.5%) or common (MAF >5%).

Natural traits: we refer to a trait as a distinct variant of a phenotypic characteristic of an organism (e.g., blue eyes). A trait is also sometimes used to refer to a phenotypic characteristic (e.g., eye color), and trait values are the variation in the trait. Natural traits indicate variation in a character that occurs within a natural population (e.g., spectrum of eye colors in humans).

Pairwise epistasis: common approach to the detection of epistasis, by limiting the measurement of nonadditivity to two different allele pairs at different loci.

Polygenicity: quality of having multiple genes. Polygenic inheritance occurs when one characteristic is controlled by two or more genes.

Quantitative trait locus (QTL): section of DNA that correlations with variation in a phenotype. The size of a QTL is related to the level of linkage disequilibrium in a population, along with the size of the effect of the causative allele. QTLs can contain two or more causative alleles in the same gene or in different nearby genes.

Statistical epistasis: we refer to statistical epistasis as the nonadditive

researcher to construct a model to quantitatively assess the effect of genetic variants (and their paired alleles) on phenotype. The model is flexible with respect to the number of allele pairs as well as the relationships between them. In the simplest case, the model evaluates the effect of an allele pair at a single locus on the observed phenotypic variance. A model with a second allele pair at a second locus can also include a term to represent the interaction between them, and in theory there is no limit to the number of terms representing direct contributions of allele pairs (additive effects) and all possible combinations of interdependence between them (interaction effects). A model with interaction terms estimates coefficients for the additive and nonadditive terms; the latter of which are estimates of what we call statistical epistasis (Box 2). For the rest of this paper, we discuss statistical epistasis as it is defined and measured in this sense.

terms of a linear regression or linear mixed model, estimated by averaging its effect across the genetic backgrounds within a population, or, in the case of compositional epistasis, from individuals that are otherwise genetically invariant. Some authors only consider statistical epistasis in the context of populations or refer to the epistatic component of variance as statistical epistasis.

Measurement of Statistical Epistasis in a Defined Genetic Background

Intuitively, it is easiest to understand epistasis between two allele pairs, with either four possible genotypes in a haploid individual, or nine possible genotypes in a diploid (the overall approach to modeling pairwise epistasis is described in Box 2 and visualized in Figure 1). While there is no reason to expect interactions to be limited to two loci – especially since most naturally varying traits are polygenic – most investigations only estimate **pairwise epistasis** for reasons of tractability. If the analysis is conducted in an invariant genetic background, it estimates what is called **compositional epistasis**, or the nonadditive effect of the two allele pairs in an individual that results from the combinatorial substitution of one allele for another [5]. Practically, combinatorial substitution could be accomplished using introgression or genome editing to create all possible genotypes (four genotypes for a haploid or nine genotypes for a diploid) within a single reference strain. A large number of genetically identical individuals can be phenotyped and statistical epistasis can be estimated using least squares or a similar procedure (Box 2). Tests of compositional epistasis are the purvue of model-system geneticists, who can manipulate genotypes and are often interested in directly assessing the effect of an allele substitution within an individual genotype. For most species, including humans, compositional epistasis serves more of a thought experiment that is useful for building intuition on what is being measured.

Compositional epistasis for a given set of alleles can be different in different genetic backgrounds. This is due to **higher-order epistasis** [13], where differences in genetic background across individuals can mask or modify the pairwise epistasis measured between two allele pairs. While higher-order epistasis is difficult to detect, it has been identified in multiple species [14–17], and theory and simulations suggest it can play an important role in evolution [18,19]. A recent systematic analysis of trigenic interactions in yeast suggests that higher-order genetic interactions play a key role in the genotype-to-phenotype relationship [20].

Measurement of Statistical Epistasis in Natural Populations

Statistical epistasis, as defined in Box 2, can be estimated from natural populations including extant human populations. In this case, an experimenter is limited to using individuals that naturally occur within a population; each with their own unique genetic background. In these natural populations, allele frequency (Box 3) plays an important role in the ability to detect statistical epistasis, because allele frequency determines the expected number of individuals to contain each of possible genotypic combinations and, thus, the statistical power of the test. Unlike measurements of compositional epistasis, where an experimenter ensures all genotype combinations are produced, measurements of statistical epistasis are dependent upon what exists in the natural population. Thus, even if epistatic gene action occurs between two variants, if those genotypes are not represented in the population then epistasis will not be detected. If a

Box 2. Measurement of Statistical Epistasis and Epistatic Variance in Haploids and Diploids

To understand statistical epistasis, it is useful to demonstrate how it is measured. It is most intuitive to start with pairwise epistasis measurements in a haploid species, considering allele pairs at two loci (A_1 or A_2 at locus A and B_1 or B_2 at locus B). We assume that these two loci are the only genetic factors that affect the phenotype. In the two-locus scenario, the question is whether the allelic variants contribute independently to phenotypic variation, or if the effect of an allele at one locus depends upon the identity of the allele at the other locus. For a haploid, each member of the population can take one of four possible genotype combinations: A_1B_1, A_1B_2, A_2B_1, and A_2B_2. To measure statistical epistasis between these allele pairs (see Figure 1 in main text), a population is genotyped and phenotyped and the resulting data matrix is fit to an equation containing four free parameters:

$$y = a_1x_1 + a_2x_2 + ix_1x_2 + \mu \quad \text{(I)}$$

where y is the phenotype and x_1 and x_2 are binary variables that represent the genotype of the individual for the two loci. The parameters represent the independent contribution of each locus to phenotype, that is, the additive effects (a_1, a_2); the mean phenotype of the population (μ); and variation in phenotype due to the (statistical) interaction between loci, or statistical epistasis (i). For haploids, the parameter i is what we call statistical epistasis throughout this paper. Through the above approach, an experimenter can estimate the value of i from a population.

This statistical framework also allows a biometrician to fit the data without accounting for nonadditive (interactive) effects between the two loci, using an alternative equation where i is excluded:

$$y = a_1x_1 + a_2x_2 + \mu \quad \text{(II)}$$

This second approach does not mean that epistasis does not exist between two allele pairs, simply that the experimenter chooses not to include it in their model. Assuming epistasis does exist in the system, the fit of the models given by equations (I) and (II) to the data will differ: the estimates of a_1 and a_2 will change.

Once a model is fitted, the amount of variance in the initial population that is explained by genetic factors is estimated. A standard approach for this is to decompose the total genetic variance (V_G) into two orthogonal components, additive variance (V_A) and epistatic variance (V_I, for interactive), with the following equation [30]:

$$V_G = V_A + V_I \quad \text{(III)}$$

V_G represents the total amount of phenotypic variation caused by genetic factors and is estimated from the amount of variation captured in equation (I). While the names of V_A and V_I suggest that they are simply the genetic effects decomposed into additive and interactive components, their calculation and meanings are actually unintuitive. Following Falconer and Mackay [30], Chapter 7, V_A is calculated using the **breeding values** of each of the four genotypes. The breeding value of a genotype represents the average value of the genotype as estimated by randomly mating individuals with members of the population. Assuming a population is in Hardy–Weinberg equilibrium, the breeding value of a genotype, and thus V_A, can be estimated using equation (II). V_I is then estimated using equation (III), that is, by subtracting V_A from V_G. Failure to observe epistatic variance is sometimes interpreted as an absence of interactive gene action, but this is not necessarily the case [79]. As noted by Falconer and Mackay, 'one cannot speak of an individual's breeding value without specifying the population in which it is to be mated' [30]. In other words, additive variance (as well as epistatic variance) is not only determined by the genotype–phenotype relationship, but also the allele frequencies of the causal alleles in the population from which it is measured. Statistical epistasis, in contrast, exists if there is nonadditive interaction between genotypes in the determination of phenotype, but its ability to be detected depends heavily on experimental conditions, including allele frequency (see Figure 1 in main text). Newer approaches to estimating variance, which deviate from this framework and are beyond the scope of this review, can deal with departures from Hardy–Weinberg equilibrium in the mapping population [80].

To extend this approach to diploid animals, one must account for the possibility of allelic interaction at heterozygous loci and extend the number of possible genotypes from four to nine. One standard approach is to modify the fit equation as such:

$$y = \mu + a_1x_1 + d_1z_1 + a_2x_2 + d_2z_2 + i_{aa}x_1x_2 + i_{ad}x_1z_2 + i_{da}z_1x_2 + i_{dd}z_1z_2 \quad \text{(IV)}$$

where x_1, x_2, z_1, and z_2 represent the full genotype of each individual, d_1 and d_2 represent the average dominance, and i_{aa}, i_{ad}, i_{da}, and i_{dd} represent the statistical epistasis [4]. Epistasis is necessarily more complicated to measure in diploids (i.e., a four-component vector) in order to account for all possible interactions. Likewise, the partitioning of the variance now includes a dominance term:

$$V_G = V_A + V_D + V_I \quad \text{(V)}$$

It should be noted that if epistasis is a deviation from the expectation that alleles act on phenotype independently, then additivity is not necessarily the only, or even the obvious, null model. Multiplicative models have also been used to represent independence among alleles, which may be more appropriate for traits like mortality where genetic effects might modify probabilities of death; in any event, it is often the case that phenotype data can be transformed to a different scale [5,7].

variant is at low frequency, then the effect at the interacting locus will appear additive since it will be predominantly measured in only one genotypic background for the second locus (for a thorough treatment, see [21]). The ability to detect epistasis is maximized at allele frequencies of 50% because the representation of genotypes is maximized (Figure 1, Box 3). In human populations, most genetic variants in a population are rare; >80% of genetic variants have a **minor allele frequency** <0.5% (Figure 2) [22]. For two variants in **Hardy–Weinberg equilibrium** with this frequency, only three of the nine possible diploid genotypes are sampled at an appreciable frequency (Figure 3). The constraints of allele frequencies mean that it is largely

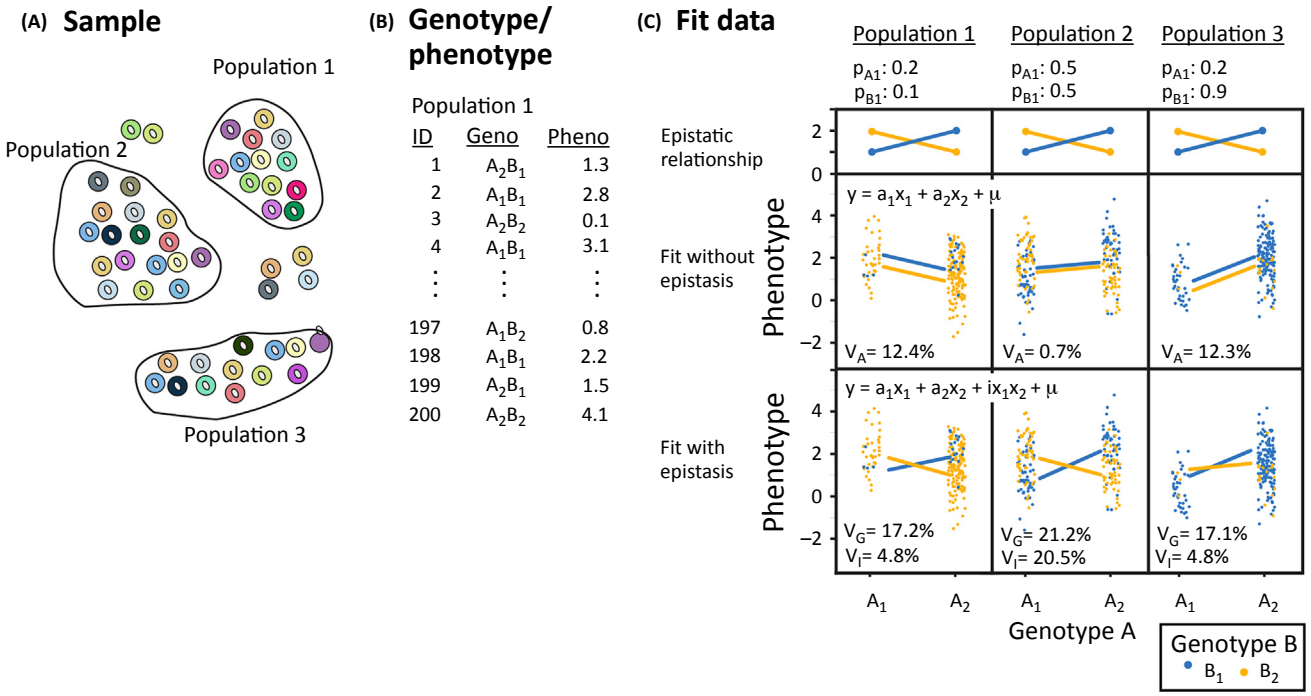

Figure 1. Primer on Measuring Pairwise Epistasis in a Natural Population. There are three primary steps to measuring statistical epistasis in a pairwise manner. Here, we illustrate its detection in a haploid population (e.g., *Saccharomyces cerevisiae*) between two loci (A and B) containing two alleles. We have chosen this example to demonstrate how allele frequencies can affect estimates of statistical epistasis and the additive terms of the model. Reading Box 2 in concert with this figure should be helpful. (A) To estimate epistasis, a number of individuals must be sampled from a larger population. Here, we show how three subpopulations can be sampled from a larger population of genotypically unique individuals (illustrated by different colors). These populations might be geographical in nature (i.e., yeast from the US vs China). (B) The sampled individuals are then genotyped and phenotyped to create a data matrix (n = 200 for each population). For a haploid, there are four possible genotypes between two allele pairs. Only one data matrix is displayed, but each population is assumed to have its own data matrix. The allele frequency at A and B can be calculated for each population using the Geno column. The phenotypic variance can be calculated for each population using the Pheno column. In this case, the allele frequencies for both loci are different in the three populations. (C) The data from (B) are fitted using linear regression. In these plots, the genotypes of both loci are represented by the x axis (A) or color of the points. Nonadditivity can be recognized by nonparallelism between the blue and yellow lines; the three populations differ in allele frequency across loci. To simulate how allele frequency affects the distribution of phenotypes in a population, first, we assume an epistatic relationship between the loci, as represented in the top row. This relationship does not vary across populations because we assume no higher-order epistasis exists. Next, random noise (from a Gaussian distribution) is added to the individual phenotypes, to represent variation contributed by the environment, stochasticity, or other loci (the average of which is constant across populations). Each individual's phenotype is represented by a single dot in the middle and bottom rows. The regression lines between the middle and bottom rows differ according to the regression model; in the middle, the additive model given by Equation (II) in Box 2 is used, in the bottom, the nonadditive model that includes statistical epistasis, given by Equation (I) in Box 2, is used. The amount of variance that is captured by the fit is also shown in each panel. V_G, V_A, and V_I are defined in Box 2. As can be seen, for two allele pairs with high levels of epistasis, allele frequency plays an important role in the slope of the fit (i.e., the direction of the effect size) and the amount of variance captured by the strictly additive model (middle row).

impossible to measure statistical epistasis between the majority of genetic variants that are present in human populations. These issues are compounded when studying interaction between three or more loci.

Just as genetic background can influence compositional epistasis, statistical epistasis between two allele pairs can be different in different populations (e.g., between European and African human populations) due to the actions of higher-order epistasis. The composition and allele frequencies of other variants participating in higher-order epistasis can vary across populations.

Due to the importance of epistasis for understanding complex diseases [23,24] or modifiers for mongenic disorders [25], there has been remarkable activity in the development of new

Box 3. Allele Frequency in Populations

One of the most important parameters in population genetics is allele frequency, the measure of how common or rare a specific allele is in a population. Allele frequency must be calculated with respect to a well-defined population at a specific timepoint, such as the set of individuals currently found on a given island. For haploid organisms, this measure is calculated by taking the total number of individuals carrying the allele of interest and dividing the total number of individual samples. For diploids, the frequency must include both chromosomes. This can be achieved by adding the frequency of homozygotes of an allele to the half the frequency of heterozygotes [30].

Changes in allele frequencies over time or between populations can be used to identify evolutionary forces (positive, balancing, or purifying selection), changes in demography (population size, migration, and substructure), or nonrandom mating. In quantitative genetics, these frequencies play an important role in the ability of statistical models to identify genetic variation that influence a phenotype of interest.

The distribution of allele frequencies for a large number of genetic variants is referred to as the allele frequency spectrum. For natural outbred populations, most variants have low allele frequencies, with a minor allele frequency <0.5%. When quantitative geneticists make artificial populations, via intercrossing and inbreeding designs initiated with a limited number of parental strains, the allele frequency spectrum is dramatically changed. Allele frequencies for a large number of alleles are inflated, which promotes the detection of rare variants and variants of small phenotypic effect.

methods to detect epistasis in humans [26]. Analyses performed in the context of genome-wide association studies in human populations usually benefit from increased statistical power due to large sample sizes. Still, since validation of putative interactions is difficult, the extent and nature of statistical epistasis in humans remains controversial [26,27]. A number of confounding factors can potentially create the false appearance of epistasis, including demographic structure in the data arising from sampling multiple populations, missing genotypes, and mismeasurement of phenotypes in the tails of the distribution [27].

Epistatic Variance – An Alternative Definition of Statistical Epistasis

In this review, we use the term statistical epistasis to explicitly refer to the interaction terms estimated by a linear model, as described above and explained in detail in Box 2 (for more thorough treatment, see [4,28,29]). However, much confusion in how epistasis is discussed and measured in naturally varying traits arises from conflation of two distinct uses of the term statistical epistasis. For example, a linear model with an interaction term, as described above, can be compared to an otherwise identical model without the interaction term [30]. The difference in the fit of these models to the empirical data is an estimate of **epistatic variance** (Box 2). Epistatic variance is sometimes called statistical epistasis [28,31], but these measures are distinct. Epistatic variance is the fraction of the total phenotypic variance for a given experimental design or population that requires the use of nonadditive interaction terms to capture the effect of a genotype on a trait. Confusion regarding what is captured by epistatic variance is exacerbated by the fact that the additive and interaction terms in a linear model both contribute to the additive variance captured by an additive-only model. This is because alleles that interact epistatically also show additivity, so long as their effects, when averaged over the genotypes at the other locus, are different. For the same reason, only a portion of the interaction term in a full linear model is responsible for the epistatic variance (Box 2). Working through an example problem, such as in Figure 1, can be helpful to a new student encountering epistatic variance for the first time. This example also was chosen to demonstrate how allele frequency can modify the estimated effect of an allele pair, despite the underlying compositional epistasis between the two genes remaining unchanged.

Epistatic variance can also be estimated without any assumption about the number of loci contributing to a trait using information about genetic relatedness between individuals within a mapping population [12]. For example, line crossing produces individuals of known genetic

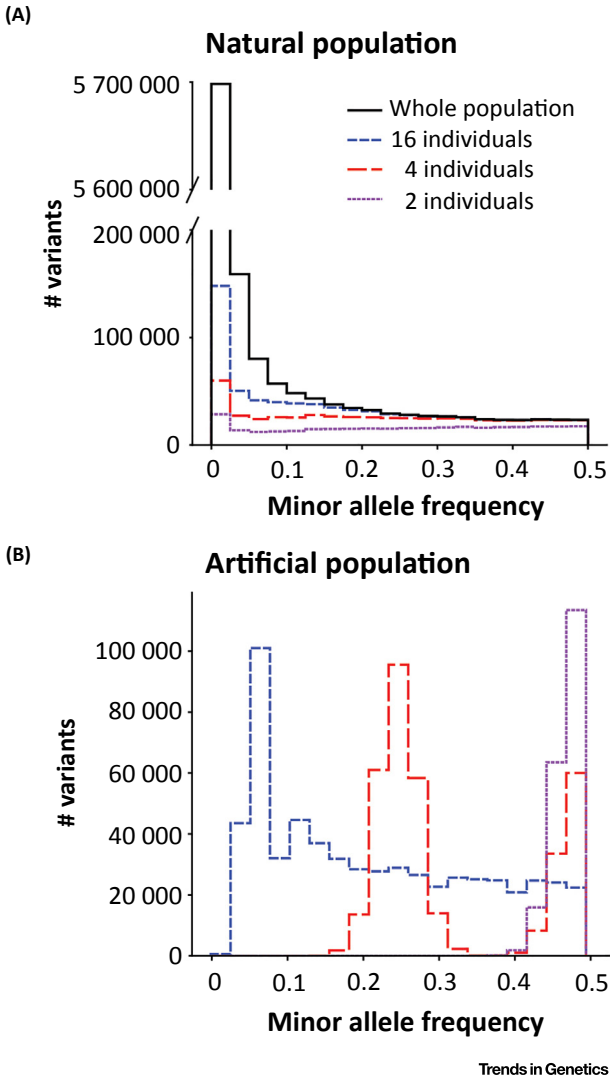

Figure 2. Allele Frequencies in Natural and Artificial Populations. (A) Histogram of allele frequencies of genetic variants in a human population (black line). These data were taken from 2504 sequenced individuals as part of the 1000 Genomes project, limiting data to chromosome I. The population allele frequencies of genetic variants that differ between two, four, or 16 individuals is also plotted, showing that while most genetic variation in a population is rare, most genetic variation between two individuals is common. (B) Histogram of allele frequencies of genetic variants in an artificial mapping population constructed from either two, four, or 16 parental lines (colors and line types follow A). The amount of genetic variation captured in the parents and initial allele frequencies are taken from data in (A). While the exact allele frequency histograms will vary between species due to idiosyncratic differences, these panels illustrate the inflation of allele frequencies that will occur due to construction of an artificial population.

relatedness and degree of heterozygosity, which can then be used to estimate epistatic variance and, furthermore, decompose it into multiple components that also include dominance effects [32]. These approaches are possible because the different components of variance are reduced at different rates as genetic relatedness decreases. Because these approaches do not require explicit genotyping information, they played a major role in first demonstrating the influence of epistasis in the expression of natural traits at a time when

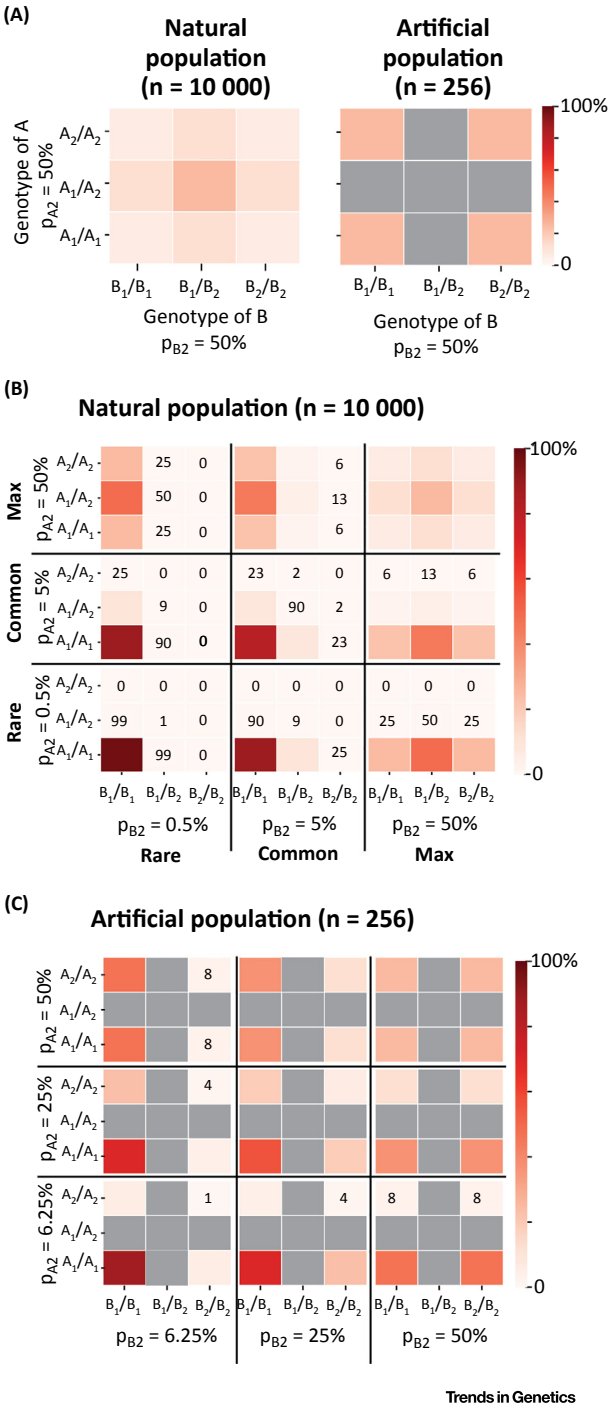

Figure 3. Genotype Frequencies in Natural and Artificial Populations. The genotype frequencies of a natural or artificial (inbred) diploid population (n = 10 000 or 256 respectively), shown as a heatmap. The artificial population was created from either a two-, four-, or 16-parent standard RIL cross design. The different population sizes (n) match typical sizes used with natural or artificial populations. There are nine possible genotypes for two allele pairs in a diploid, here represented as colored cells within each 3 × 3 matrix, which vary in the allele frequencies of either A or B (given by p_{A1} and p_{B2}). The genotype of locus A is shown on the y axis and the genotype of locus B is shown on the x axis. For rare genotypes

(Figure legend continued on the bottom of the next page.)

sequencing was expensive. Recent work in humans using pedigree information has also been used to estimate epistatic variance [33].

Use of Model Organisms to Study Epistasis

Because of the difficulties in studying epistasis in natural populations, much of the research into the nature and extent of epistasis has been performed in model organisms. The ability to manipulate genotypic combinations in model systems offers several powerful approaches that bypass the limitations imposed by natural populations [34]. The approaches are described below; all have revealed pervasive epistasis within experimental systems, including budding yeast, nematodes, fruit flies, mice, and a number of plant species. In the following section, we are specifically referring to the use of model organisms to study epistasis that occurs between natural genetic variants (i.e., those that are found in a natural population). The haploid and diploid models for estimating statistical epistasis (Box 2) can be used here as well.

Artificial Populations Can Identify Regions of the Genome That Interact with Each Other

Statistical epistasis can be estimated from **artificial populations** created by an experimenter through controlled breeding. These populations are created from a small number of individuals (typically 2–16) sampled from a larger natural population. These individuals capture a subset of the natural genetic variation segregating in a population, with a bias towards alleles that are common (Figure 2). Artificial populations can be well powered to map the genetic basis of a given phenotype. In part, they derive power from inflation of allele frequencies across the genome towards 50% (Figure 2) due to controlled mating and the use of a small number of parental lines to initiate the populations. A second advantage common to these approaches is the ability to replicate individual genotypes by creating true-breeding inbred lines, which improves the estimation of genetic contributions to the phenotype.

At its simplest, two inbred parental strains are crossed, and hundreds of independent F_2 individuals are inbred to create stable, recombinant inbred lines (RILs). Using multiple parental strains and adding several generations of intercrossing before inbreeding increases the amount of genetic diversity in the population and expand the genetic map. Once the RILs are genotyped and phenotyped, associations between them can identify large recombinant sections of the genome (often ~1 megabase) that contain one or more genetic variants that influence the phenotype, known as **quantitative trait loci** (QTLs). Overall, this approach has been successful; QTLs have been identified for many phenotypes across the range of eukaryotic life.

Identification of significant QTLs permits tests of statistical epistasis between them, and many empirical examples have been published (reviewed in [29,35]). If the genetic architecture of

(<100 in A and <10 in B), the exact number of expected individuals is also shown. If sufficient rounds of inbreeding occur, genotypes for artificial populations are homozygous. For clarity, the impossible heterozygous genotypes are shown in grey for artificial populations. (A) Comparison of genotype frequencies for natural and artificial populations. The allele frequency of each allele is 50%. For natural populations, individuals were assumed to follow Hardy–Weinberg equilibrium, and the most-likely genotype is the double heterozygote. For artificial populations, each of the four possible homozygote combinations are equally likely. (B) Comparison of genotype frequencies in natural populations for three different allele frequencies. The case of p = 0.5% represents rare variants, p = 5% represents common variants, and p = 50% represents allele frequencies where detection of epistasis is maximally powered (Max). Rare variants do not explore much of the genotype space (i.e., rare–rare only one genotype is >100, rare–common only two genotypes are >100, and rare–max only three genotypes are >100). (C) For artificial populations, individuals are assumed to be completely inbred, resulting in only four possible genotypes (the corners of the square). The case of p = 6.25% represents allele frequencies that are possible in a 16-parent RIL, p = 25% represents allele frequencies that are possible in a four- or 16-parent RIL, and p = 50% represents allele frequencies that are possible in a two-, four- or 16-parent RIL. Abbreviations: RIL, recombinant inbred line.

complex traits is generalizable, then the accumulation of examples of epistasis in these experimental systems would suggest that statistical epistasis is likely to be pervasive in other outbred populations, like humans [23,29]. While we do not dispute that epistasis is important for human traits, there are a number of challenges in generalizing results from experimental model organisms to the studies of natural populations that are worth explicitly discussing. These challenges make it nearly impossible to translate results from current approaches in model organisms into quantitative predictions of the pervasiveness of epistasis in any natural populations, including humans.

Four Issues with Translating Results on Epistasis from Artificial Populations to Natural Populations

Increased Linkage Disequilibrium

Due to the finite number of meiotic recombination events that occur in the construction of an artificial population, QTLs contain a large number of genetic variants, typically thousands, in linkage disequilibrium on common haplotypes (a more comprehensive description of breeding design and linkage disequilibrium can be found in [36]). Consequently, an outstanding issue in estimating epistasis from artificial populations is whether an epistatic QTL harbors more than one interacting allele pair. This possibility is reasonable, since fine-mapping QTLs, for example by **introgressing** fragments into an isogenic background, has been shown to fractionate the QTL into multiple loci [37]. If detection bias for primary QTLs is skewed towards genome regions with aggregates of causal alleles, this may contribute to the observation of pervasive epistasis in mapping populations.

One way to address this is through improved cross design to increase the number of generations of outcrossing. This approach was recently used in *Caenorhabditis elegans*, taking advantage of its short generation time (~3 days), to create a large mapping population using 140–190 generations of outcrossing [38]. This mapping population provides a substantial increase in mapping resolution, to approximately the single gene, with little population substructure. When used to map fertility and growth traits, the authors found that epistasis was extensive and explained up to 40% of the trait variance. While these results may be idiosyncratic to *C. elegans*, they confirm that statistical epistasis between QTLs is not strictly a consequence of interactions across large haplotypes and that identification of the interacting genes is possible.

Distortion of Allele Frequencies between Artificial and Natural Populations

For a two-parent mapping population, the creation of RILs balances the representation of genetic differences between individuals; regardless of its allele frequency in a natural population, each genetic variant is inflated to 50% in the artificial population. Similar inflation will occur for multiparent panels (Figure 3). While this change in allele frequency maximizes the potential for detection of epistasis [29], it also masks whether epistasis identified in mapping populations arises from allelic combinations that are rare or common in a population. This is a major issue for translating these observations to natural populations, as rare variants contribute less to the genetic variance of a trait. Furthermore, rare alleles are fundamentally different from common alleles. For example, deleterious variants are more likely to be rare [39] and rare alleles have been shown to have larger effect sizes for human height [2]. This may be explained by an assumption that rare alleles are newer and less likely to have acquired compensatory modifiers, with further implications for the role of epistasis in natural populations [29].

This issue can be resolved through the identification of the causative alleles, followed by calculating the allele frequency in a natural population. However, this typically requires

extensive fine mapping and additional genetic manipulation. CRISPR-enabled genome editing can potentially overcome some of this difficulty, and this approach has been used to identify causative mutations within a single QTL in multiple systems [40–43]. In the future, these approaches can be extended to identify allele pairs with nonadditive effects.

Homozygosity in Experimental Lines

Diploid mapping populations generated by forced inbreeding do not follow Hardy–Weinberg equilibrium and explore only four of the nine genotypes (Figure 3). Consequently, statistical epistasis in these artificial populations is estimated using the simpler linear model for haploids (Box 2). For a sexually reproducing species, the genotype space explored by an artificial population rarely explores genotypes representative of a randomly sampled allele pair from a natural population (Figure 3). In other words, there is no general way to translate the haploid estimate of statistical epistasis [see Equation (I) in Box 2] to the four-component diploid version of epistasis that is relevant to natural populations.

For mapping populations derived from naturally outbred wild strains, inbreeding also has significant effects on fitness and trait expression. The same problem arises in tests for compositional epistasis or genetic background effects, since a controlled genetic background must be homozygous. For example, *Drosophila melanogaster* and *Mus musculus* live as outbred populations in nature, and artificial mapping populations in these systems exhibit so much inbreeding depression that the majority of attempted lines typically go extinct in the process of constructing the population [44,45]. It is likely that rare deleterious mutations are unmasked by this process, and selection bias fixes alleles in strong epistasis in the recombinant lines that suppress the deleterious phenotype. Not only are inbred genetic backgrounds unlikely to represent the genetic backgrounds of outcrossed individuals found in the wild [21], but phenotypes of inbred lines may depart substantially from those in naturally outbred populations in the presence of dominance variance [46]. The experimental power provided by forced homozygosity comes at a price; genotypic effects and genetic interactions detected in these experimental systems are real, but not necessarily relevant outside the laboratory.

Idiosyncratic Differences between Species

Epistasis is likely idiosyncratic across species, given differences in recombination rates, population structure, and mating system. For example, can one generalize observations of epistasis from *C. elegans* to human populations? *C. elegans* hermaphrodites primarily reproduce by selfing, and experience rare outcrossing events through mating with males. A combination of high levels of linkage disequilibrium between different chromosomes and strong outbreeding depression is consistent with the existence of compensatory evolution in isolated lineages [47,48]. Therefore, the substantial epistasis observed in *C. elegans* artificial populations [38] could be caused by new allelic combinations that disrupt the stable haplotypes found in the wild. Limited numbers of sperm in hermaphrodites may encourage outcrossing in nature and consequently generate epistasis in the wild [49], but probably not to the degree generated by forced outcrossing in an artificial population. Overall, model organisms exhibit a wide range of mating systems and demographic patterns, which makes generalization of results between them difficult.

Some patterns of epistasis that appear divergent across species might participate in species-specific occurrences of **balancing selection**. *D. melanogaster* exhibits short-term cyclical changes in allele frequency across seasons that imply the presence of balancing selection, that is, selection for some alleles during cold months and others during warm months [50,51]. A quantitative model that can explain this maintenance of polymorphism is optimized under an

assumption of high polygenicity and diminishing returns epistasis [52]. *C. elegans* also exhibits evidence of pervasive balancing selection, but by stable signatures of long-term selection [53–56]. Genomic analysis suggests that hundreds of these loci segregate in the population [57,58]. It is currently unknown why balancing selection might persist in *C. elegans*, but its transition from sexual to androdioecious mating system might have been involved [56].

Recent Approaches for Using Model Organisms to Study Epistasis

While the caveats listed above must be kept in mind, a number of recent publications illustrate why model organisms will remain critical for understanding epistasis. We have chosen to highlight a few recent examples that we believe go beyond QTL mapping to analyze epistasis in greater detail.

Use of Compositional Epistasis to Estimate the Full Diploid Model of Statistical Epistasis

The identification of interacting QTLs via an artificial population, including putatively epistatic variants, may be followed by experiments that test for interactions in a more controlled setting. As discussed above, compositional epistasis can be measured by the placement of specific alleles in a single genetic background, for example, to evaluate sequential mutational steps in adaptive evolution [59] or dissect the genetic basis of a trait [15]. Interrogation of compositional epistasis in a diploid, two-locus scenario is rare in the literature, but one striking example demonstrated substantial epistasis between loci and, furthermore, provided an empirical estimate of the V_A contributed by the interacting effects [60]. In this study, alleles at seven previously identified QTLs for floral traits in *Mimulus* were distributed as pairwise combinations across experimental lines. The authors created all nine genotype combinations to fully measure all four components of statistical epistasis for 11 of the possible 21 pairs of loci for a variety of traits. With this information, the authors could predict the effect that allele frequency change would have on the effect of these QTLs. In some cases, epistasis amplified V_A, in others it reduced V_A, but overall the magnitude of V_A was largely determined by the frequency of the alleles. While this study and others like it cannot provide estimates of epistasis in natural populations, they can characterize how epistasis between wild-type alleles might influence population dynamics, including response to selection. This work is important, as selection will modify allele frequencies, and populations will explore different parts of the genotype space.

Epistatic Interactions with a Focal Allele Can Be Modified by Natural Selection

Another approach that requires the use of model organisms is to examine the effect of a focal allele in different genetic backgrounds. Magnitude or even direction of effects between alleles can change when they are tested in different wild lines or strains, demonstrating a nonadditive interaction between the focal locus and the genetic background [31]. This approach departs from our discussion of evaluating two-locus statistical epistasis by the linear models described above, but it is an especially tractable method in model systems, and interactions with genetic background are extensively reported in the literature [61]. Statistical interaction between a focal locus and the genomic background is typically estimated in the haploid or homozygous diploid state, which eliminates complexity contributed by dominance. A model with terms for the focal locus genotype, the genetic background, and the interaction between them can then be fitted to phenotype data, which permits analysis of more than two genetic backgrounds [62,63].

A special case of genetic background interaction is when perturbation of the focal gene, sometimes called a capacitor [64], reveals phenotypic variation arising from natural genetic variation that was previously silent, or cryptic. Much of the motivation behind research into this phenomenon lies in its potential to facilitate adaptive evolution, by increasing the additive genetic variance available to selection [65,66]. One of the most well-characterized examples of genetic background interaction involves the classic capacitor HSP90. The activity of this heat

shock protein is presumed to mask genetic variants by compensating for instability in protein conformation through its chaperone activity, and interruption of HSP90 function has been shown to unmask variation in many systems, from microbes to vertebrates. Cryptic variation has been revealed by laboratory-derived mutant alleles of HSP90 [67], natural allelic variants of HSP90 [68,69], and gene perturbations that are not technically genetic, including reduction of HSP90 function by RNAi [70] and, most commonly, drug inhibitors [62,67,71].

In addition to being pervasive, genetic background effects are also often dramatic [31], and unlikely to represent typical interactions among alleles that are common in natural populations. However, they are valuable in resolving genetic interaction networks and may shed light on the genetic architecture of modifiers of Mendelian genetic diseases [61,72]. Furthermore, they demonstrate how epistasis can seriously influence the consequence of a new mutation entering a population. For example, HSP90 activity masks the effects of standing genetic variation in many systems (e.g., [67,70,71]), but it has also been shown to potentiate the expression of *de novo* mutations in yeast [62]. Natural selection purges deleterious alleles, but only when they penetrate to phenotype in the background in which they occur; this work shows that standing genetic variation in natural populations represents both the mutations and the epistatic interactions that have escaped selection. Although examples outside of HSP90 [68,69] are rare, other naturally occurring capacitors probably do segregate within populations; if so, they should increase the additive genetic variance available to selection when combined with potentially cryptic alleles. One example comes from artificial selection in chickens, in which a QTL was determined to interact with other QTLs such that selection on the capacitating 'hub' released variation in the 'spokes' [73]. Recent work in yeast uncovered a similar hub and spoke architecture [74], further demonstrating the potential role of epistasis in directing evolutionary trajectories. In these scenarios, selection on hubs releases new phenotypic variation during an ongoing trajectory of selection. If this phenomenon is not rare, it may explain the ability of some populations to respond to selection over many generations without the assumption of new mutations [75].

Conclusion Remarks

In this paper, we have argued that the manipulations that enable detection and characterization of epistasis in experimental systems also limit extrapolation of these findings into the context of wild populations. While we have made this point strongly, we do not seek to deride the work that has been done in model organisms towards identifying epistasis. Rather, this work was essential to show that epistasis occurs in the expression of natural genetic variation and should not be ignored despite the difficulty in its study.

The challenge now in model organisms is to move away from detecting its existence and towards understanding its nature (see Outstanding Questions). Our review, like most treatments of epistasis in the context of natural genetic variation, mostly considers interactions between two loci, but there is reason to expect higher-order epistasis exists and matters for evolution [13,17]. Testing beyond the two- or three-locus case becomes a problem of combinatorics that will always exceed our capacity to increase sample size. Thus, experimental investigations into higher-order epistasis will have to rely on planned comparisons of genetic combinations, for examples of variants within described molecular pathways or between epistatic loci identified in natural or artificial populations from lower-order analyses. The discovery and development of genome editing technologies opens this door. Now, expedient replacement of specific alleles is possible, especially in microbial systems that are already capable of automated phenotyping in well-replicated experimental designs [76]. The ability to test a mutation, or combinations of mutations, in a large number of genetic backgrounds is now reasonable.

Outstanding Questions

With the advent of better methods to detect statistical epistasis, can epistasis be identified in natural populations including humans?

What is the contribution of epistasis to the additive variance of a complex trait?

What is the role of higher-order epistasis in natural traits?

Will hub alleles that interact with a large number of different loci be important for understanding the response to long-term selection?

Another key question that model organisms can address is to understand how natural selection influences the nature of epistatic relationships that persist among segregating variants. An example of this was provided by the recent work of Geiler-Samerotte [62]; however, the generalizability of this work remains unknown. Furthermore, since the effect of a mutation depends on the genetic background in which it lands, its fate in a population and the evolutionary trajectory of the population itself are shaped by epistasis [77,78]. Along these lines, model organisms will help to understand the role of epistasis in response to long-term directional selection. While the additive variance is important for predicting how a phenotype under selection will change in the immediate generation, little is known about the underlying genetic architecture that is responsible for the additive variance, and how this genetic architecture changes in response to selection as allele frequencies are changed by selection [68,73,74]. Cheap resequencing technologies have enabled the possibility to follow allele frequency over time in these experiments to identify regions of the genome under selection. Exploring the role of epistasis between these regions is now possible.

References

1. Cutting, G.R. (2010) Modifier genes in Mendelian disorders: the example of cystic fibrosis. *Ann. N. Y. Acad. Sci.* 1214, 57–69
2. Marouli, E. *et al.* (2017) Rare and low-frequency coding variants alter human adult height. *Nature* 542, 186–190
3. Visscher, P.M. *et al.* (2012) Five years of GWAS discovery. *Am. J. Hum. Genet.* 90, 7–24
4. Cordell, H.J. (2002) Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.* 11, 2463–2468
5. Phillips, P.C. (2008) Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* 9, 855–867
6. Wang, X. *et al.* (2010) The meaning of interaction. *Hum. Hered.* 70, 269–277
7. Aylor, D.L. and Zeng, Z.B. (2008) From classical genetics to quantitative genetics to systems biology: modeling epistasis. *PLoS Genet.* 4, e1000029
8. Moore, J.H. and Williams, S.M. (2005) Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *Bioessays* 27, 637–646
9. Bateson (1909) Discussion on the influence of heredity on disease, with special reference to tuberculosis, cancer, and diseases of the nervous system: introductory address. *Proc. R. Soc. Med.* 2 (Gen Rep), 22–30.
10. Avery, L. and Wasserman, S. (1992) Ordering gene function: the interpretation of epistasis in regulatory hierarchies. *Trends Genet.* 8, 312–316
11. Huang, L.S. and Sternberg, P.W. (2006) Genetic dissection of developmental pathways. *WormBook* 1–19
12. Fisher, R.A. (1918) The correlations between relatives on the supposition of Mendelian inheritance. *Trans. R. Soc. Edinb.* 52, 399–433
13. Taylor, M.B. and Ehrenreich, I.M. (2015) Higher-order genetic interactions and their contribution to complex traits. *Trends Genet.* 31, 34–40
14. Pettersson, M. *et al.* (2011) Replication and explorations of high-order epistasis using a large advanced intercross line pedigree. *PLoS Genet.* 7, e1002180
15. Taylor, M.B. and Ehrenreich, I.M. (2014) Genetic interactions involving five or more genes contribute to a complex trait in yeast. *PLoS Genet.* 10, e1004324
16. Taylor, M.B. *et al.* (2016) Diverse genetic architectures lead to the same cryptic phenotype in a yeast cross. *Nat. Commun.* 7, 11669
17. Weinreich, D.M. *et al.* (2013) Should evolutionary geneticists worry about higher-order epistasis? *Curr. Opin. Genet. Dev.* 23, 700–707
18. Sailer, Z.R. and Harms, M.J. (2017) High-order epistasis shapes evolutionary trajectories. *PLoS Comput. Biol.* 13, e1005541
19. Sailer, Z.R. and Harms, M.J. (2017) Detecting high-order epistasis in nonlinear genotype–phenotype maps. *Genetics* 205, 1079–1088
20. Kuzmin, E. *et al.* (2018) Systematic analysis of complex genetic interactions. *Science* 360. http://dx.doi.org/10.1126/science.aao1729 eaao1729
21. Hill, W.G. *et al.* (2008) Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genet.* 4, e1000008
22. 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526, 68–74
23. Mackay, T.F. and Moore, J.H. (2014) Why epistasis is important for tackling complex human disease genetics. *Genome Med.* 6, 124
24. Webber, C. (2017) Epistasis in neuropsychiatric disorders. *Trends Genet.* 33, 256–265
25. Gallati, S. (2014) Disease-modifying genes and monogenic disorders: experience in cystic fibrosis. *Appl. Clin. Genet.* 7, 133–146
26. Wei, W.H. *et al.* (2014) Detecting epistasis in human complex traits. *Nat. Rev. Genet.* 15, 722–733
27. Fish, A.E. *et al.* (2016) Are interactions between *cis*-regulatory variants evidence for biological epistasis or statistical artifacts? *Am. J. Hum. Genet.* 99, 817–830
28. Cheverud, J.M. and Routman, E.J. (1995) Epistasis and its contribution to genetic variance components. *Genetics* 139, 1455–1461
29. Mackay, T.F. (2014) Epistasis and quantitative traits: using model organisms to study gene-gene interactions. *Nat. Rev. Genet.* 15, 22–33
30. Falconer, D.S. and MacKay, T.F. (1996) *Introduction to Quantitative Genetics.* (4th edn), Pearson
31. Sackton, T.B. and Hartl, D.L. (2016) Genotypic context and epistasis in individuals and populations. *Cell* 166, 279–287
32. Wolf, J.B. *et al.* (2000) *Epistasis and the Evolutionary Process,* Oxford University Press
33. Kaplanis, J. *et al.* (2018) Quantitative analysis of population-scale family trees with millions of relatives. *Science* 360, 171–175
34. Ehrenreich, I.M. (2017) Epistasis: searching for interacting genetic variants using crosses. *G3 (Bethesda)* 7, 1619–1622
35. Mackay, T.F. (2015) Epistasis for quantitative traits in *Drosophila*. *Methods Mol. Biol.* 1253, 47–70
36. Rockman, M.V. and Kruglyak, L. (2008) Breeding designs for recombinant inbred advanced intercross lines. *Genetics* 179, 1069–1078

37. Stam, L.F. and Laurie, C.C. (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
38. Noble, L.M. *et al.* (2017) Polygenicity and epistasis underlie fitness-proximal traits in the *Caenorhabditis elegans* Multiparental Experimental Evolution (CeMEE) Panel. *Genetics* 207, 1663–1685
39. MacArthur, D.G. *et al.* (2012) A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335, 823–828
40. Ding, Y. *et al.* (2016) Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* 536, 329–332
41. Large, E.E. *et al.* (2016) Selection on a subunit of the NURF chromatin remodeler modifies life history traits in a domesticated strain of *Caenorhabditis elegans*. *PLoS Genet.* 12, e1006219
42. Yin, J.A. *et al.* (2017) Genetic variation in glia-neuron signalling modulates ageing rate. *Nature* 551, 198–203
43. Zdraljevic, S. *et al.* (2017) Natural variation in a single amino acid substitution underlies physiological responses to topoisomerase II poisons. *PLoS Genet.* 13, e1006891
44. Mackay, T.F.C. and Huang, W. (2018) Charting the genotype-phenotype map: lessons from the *Drosophila melanogaster* Genetic Reference Panel. *Wiley Interdiscip. Rev. Dev. Biol.* 7. http://dx.doi.org/10.1002/wdev.289
45. Shorter, J.R. *et al.* (2017) Male infertility is responsible for nearly half of the extinction observed in the mouse collaborative cross. *Genetics* 206, 557–572
46. Lynch, M.W.B. (1998) *Genetics and Analysis of Quantitative Characters,* Sinauer
47. Cutter, A.D. (2006) Nucleotide polymorphism and linkage disequilibrium in wild populations of the partial selfer *Caenorhabditis elegans*. *Genetics* 172, 171–184
48. Dolgin, E.S. *et al.* (2007) Inbreeding and outbreeding depression in *Caenorhabditis* nematodes. *Evolution* 61, 1339–1352
49. Large, E.E. *et al.* (2017) Modeling of a negative feedback mechanism explains antagonistic pleiotropy in reproduction in domesticated *Caenorhabditis elegans* strains. *PLoS Genet.* 13, e1006769
50. Bergland, A.O. *et al.* (2014) Genomic evidence of rapid and stable adaptive oscillations over seasonal time scales in *Drosophila*. *PLoS Genet.* 10, e1004775
51. Paaby, A.B. *et al.* (2014) A highly pleiotropic amino acid polymorphism in the *Drosophila* insulin receptor contributes to life-history adaptation. *Evolution* 68, 3395–3409
52. Wittmann, M.J. *et al.* (2017) Seasonally fluctuating selection can maintain polymorphism at many loci via segregation lift. *Proc. Natl. Acad. Sci. U. S. A.* 114, E9932–E9941
53. Ben-David, E. *et al.* (2017) A maternal-effect selfish genetic element in *Caenorhabditis elegans*. *Science* 356, 1051–1055
54. Ghosh, R. *et al.* (2012) Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. *Science* 335, 574–578
55. Greene, J.S. *et al.* (2016) Balancing selection shapes density-dependent foraging behaviour. *Nature* 539, 254–258
56. Seidel, H.S. *et al.* (2008) Widespread genetic incompatibility in *C. elegans* maintained by balancing selection. *Science* 319, 589–594
57. Greene, J.S. *et al.* (2016) Regulatory changes in two chemoreceptor genes contribute to a *Caenorhabditis elegans* QTL for foraging behavior. *Elife* 5. http://dx.doi.org/10.7554/eLife.21454 e21454
58. Thompson, O.A. *et al.* (2015) Remarkably divergent regions punctuate the genome assembly of the *Caenorhabditis elegans* Hawaiian strain CB4856. *Genetics* 200, 975–989
59. Khan, A.I. *et al.* (2011) Negative epistasis between beneficial mutations in an evolving bacterial population. *Science* 332, 1193–1196
60. Monnahan, P.J. and Kelly, J.K. (2015) Epistasis is a major determinant of the additive genetic variance in *Mimulus guttatus*. *PLoS Genet.* 11, e1005201
61. Chandler, C.H. *et al.* (2013) Does your gene need a background check? How genetic background impacts the analysis of mutations, genes, and evolution. *Trends Genet.* 29, 358–366
62. Geiler-Samerotte, K.A. *et al.* (2016) Selection transforms the landscape of genetic variation interacting with Hsp90. *PLoS Biol.* 14, e2000465
63. Richardson, J.B. *et al.* (2013) Histone variant HTZ1 shows extensive epistasis with, but does not increase robustness to, new mutations. *PLoS Genet.* 9, e1003733
64. Siegal, M.L. and Leu, J.Y. (2014) On the nature and evolutionary impact of phenotypic robustness mechanisms. *Annu. Rev. Ecol. Evol. Syst.* 45, 496–517
65. Gibson, G. and Dworkin, I. (2004) Uncovering cryptic genetic variation. *Nat. Rev. Genet.* 5, 681–690
66. Paaby, A.B. and Rockman, M.V. (2014) Cryptic genetic variation: evolution's hidden substrate. *Nat. Rev. Genet.* 15, 247–258
67. Rutherford, S.L. and Lindquist, S. (1998) Hsp90 as a capacitor for morphological evolution. *Nature* 396, 336–342
68. Chen, B. and Wagner, A. (2012) Hsp90 is important for fecundity, longevity, and buffering of cryptic deleterious variation in wild fly populations. *BMC Evol. Biol.* 12, 25
69. Sgro, C.M. *et al.* (2010) A naturally occurring variant of Hsp90 that is associated with decanalization. *Proc. Biol. Sci.* 277, 2049–2057
70. Sangster, T.A. *et al.* (2008) HSP90-buffered genetic variation is common in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2969–2974
71. Jarosz, D.F. and Lindquist, S. (2010) Hsp90 and environmental stress transform the adaptive value of natural genetic variation. *Science* 330, 1820–1824
72. Chandler, C.H. *et al.* (2017) How well do you know your mutation? Complex effects of genetic background on expressivity, complementation, and ordering of allelic effects. *PLoS Genet.* 13, e1007075
73. Carlborg, O. *et al.* (2006) Epistasis and the release of genetic variation during long-term selection. *Nat. Genet.* 38, 418–420
74. Forsberg, S.K. *et al.* (2017) Accounting for genetic interactions improves modeling of individual quantitative trait phenotypes in yeast. *Nat. Genet.* 49, 497–503
75. Crow, J.F. (2010) On epistasis: why it is unimportant in polygenic directional selection. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 365, 1241–1244
76. Sadhu, M.J. *et al.* (2018) Highly parallel genome variant engineering with CRISPR-Cas9. *Nat. Genet.* 50, 510–514
77. Draghi, J.A. *et al.* (2011) Epistasis increases the rate of conditionally neutral substitution in an adapting population. *Genetics* 187, 1139–1152
78. Kryazhimskiy, S. *et al.* (2014) Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* 344, 1519–1522
79. Huang, W. and Mackay, T.F. (2016) The genetic architecture of quantitative traits cannot be inferred from variance component analysis. *PLoS Genet.* 12, e1006421
80. Alvarez-Castro, J.M. *et al.* (2008) How to perform meaningful estimates of genetic effects. *PLoS Genet.* 4, e1000062