DOI: 10.1534/genetics.111.128900

Perspectives

Anecdotal, Historical and Critical Commentaries on Genetics

New Experiments for an Undivided Genetics

Michael R. Rose,[1] Laurence D. Mueller and Molly K. Burke

Department of Ecology and Evolutionary Biology, University of California, Irvine, California 92697-2525

ABSTRACT

There used to be a broad split within the experimental genetics research community between those who did mechanistic research using homozygous laboratory strains and those who studied patterns of genetic variation in wild populations. The former benefited from the advantage of reproducible experiments, but faced difficulties of interpretation given possible genomic and evolutionary complexities. The latter research approach featured readily interpreted evolutionary and genomic contexts, particularly phylogeny, but was poor at determining functional significance. Such burgeoning experimental strategies as genome-wide analysis of quantitative trait loci, genotype–phenotype associations, and the products of experimental evolution are now fostering a unification of experimental genetic research that strengthens its scientific power.

MOST empirical research in genetics during the 20th century can be crudely lumped into two main aggregations of researchers. On one hand, there were those who performed functionally or mechanistically oriented research using homozygous or clonal strains, as well as their crosses, recombinants, and segregants. Experiments in this branch of the field often achieved high levels of reproducibility and qualitatively clear results. Much of what we learned about the machinery of inheritance was based on experimentation of this type, notwithstanding the many contributions that biochemistry and molecular biology have made to our understanding of the foundations of genetics using non-genetic experimental methods. This type of research will be labeled "Mendelian genetics" here for convenience because it is relatively akin to the work that Mendel and the founders of 20th-century genetics such as T. H. Morgan performed. The term Mendelian genetics is in conformity with relatively common usage, although this specific field was sometimes called "transmission genetics" and is now often syncretically grouped with "molecular genetics." A common feature of such research was analysis of the functional effects of specific allelic differences in the laboratory.

On the other hand, there was 20th-century experimental research on the genetic variation of natural populations living in, or recently isolated from, the wild, research that is commonly referred to as "experimental population genetics." This usage also has its difficulties, with alternative terms such as "experimental evolutionary genetics" and "quantitative genetics" sometimes overlapping or even subsuming the term experimental population genetics in some cases. This type of empirical genetic research was less common than Mendelian genetics and was often afflicted with controversies, most notably the neutralist debate of the late 20th century. In particular, as we will discuss, this type of research was more concerned with overall patterns of genetic variation and the evolutionary mechanisms that produced such genetic variation.

It will not be argued here that one of these experimental approaches was better than the other, but it is becoming ever more apparent that they cannot persist through this century as "twin solitudes" within the scientific community. Instead, it will be argued that experimental genetics is now being unified by means of genome-wide experimental strategies that bring its sundered parts together. This point of view is not entirely novel or unheralded (*cf.* Stern 2000, 2010; Houle 2010) We should also emphasize at the outset that we are not proposing a new genetic *theory* of any kind, nor do we have any notably original views concerning the likely results of the presently emerging genome-wide experimentation. Instead, we want to call attention to a new class of experiments that presage a fruitful reunification of genetics and to encourage others to overcome

[1]*Corresponding author:* Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697-2525.
E-mail: mrrose@uci.edu

their inhibitions about practicing an undivided experimental genetics.

PAST DIVISION WITHIN EXPERIMENTAL GENETICS

It might surprise some Mendelian or molecular geneticists to learn that a number of population geneticists are critical of the extrapolation of Mendelian experimental results to natural populations in the wild. It might further surprise them that population geneticists are often highly critical of attempts to infer function in nature from experimental data. But those of us who were trained on the population-genetics side of biology have been dealing with this controversy for decades. Nielsen (2009) has supplied a recent example of this debate in population genetics. A key feature of Nielsen's review is skepticism specifically about the use of functional information derived from Mendelian genetics: "the combination of a functional effect and selection does not demonstrate that selection acted on the specific trait in question" (p. 2488), and, in referring to research on the *microcephalin* locus, "Mutations in many different genes might cause microcephaly, but changes in these genes may not have been the underlying molecular cause for the increased brain size occurring during the evolution of man" (p. 2489).

Within the field of aging research, in which Mendelian and evolutionary geneticists have had some notable conflicts, the difficulty of proceeding from the study of mutant strains to the evolution of natural populations has likewise been discussed (*e.g.*, Rose 1991; Van Voorhies *et al.* 2006). Specifically, the presence or absence of antagonistic pleiotropy in studies of specific mutant strains measured in a particular laboratory environment is demonstrably not a reliable guide to functional genetics in other environments, not even other laboratory environments. This is due, at least in part, to experimentally established propensities for genotype-by-environment interaction in functional characters (*e.g.*, Leroi *et al.* 1994; Khazaeli *et al.* 2005). Moreover, random mutations derived from genotoxic treatments are not notably representative of variants that arise or persist in natural populations.

A relatively direct test of the applicability of Mendelian genetics results to the genetics of wild populations has been supplied by experiments focusing on loci that affect bristle number in *Drosophila melanogaster*. Several studies have ascertained the effects of specific polymorphisms that have been shown to be important in isogenic laboratory strains, such as genetic polymorphisms at the *hairy* locus in wild-caught flies, but have failed to find a correspondence (*e.g.*, Macdonald and Long 2004). This is an important result because bristle-number phenotypes are easily scored, while *D. melanogaster* is a species that has been studied extensively in both Mendelian and population genetics. These features of fly bristle number favor our ability to identify commonalities between the findings of Mendelian and population genetics, but those commonalities are apparently not reliably found.

Thus there are good arguments to be made for skepticism concerning the empirical inference of population-genetics significance from typical experimental results in Mendelian genetics, however well-conducted and reproducible such experiments might be. If changes in genetic background make inferences about genetics strictly "local"—that is to say specific to the particular set of genotypes under study and the methods used to study them—then there are major problems making general inferences about the functional significance of particular allelic variants from the experiments of Mendelian genetics on their own. This is not to deny the possibility of such inferences in particular cases. But what is exciting at the present moment in the development of genetics is the emergence of genome-wide experimental techniques that offer general methods of connecting genetic variation to functional phenotypes.

On the other hand, there were significant problems with the interpretation of the results obtained from 20th-century studies of the experimental population genetics of wild populations. From its inception, there was a key difficulty facing such research, a difficulty that Sewall Wright may have realized sooner than anyone else as a result of his extensive collaboration with Theodosius Dobzhansky (Provine 1986). That difficulty is the problem of inferring which particular population genetics mechanism, or which combination of such mechanisms, is producing a particular pattern of genetic variation in a wild population. For example, genetic variation can be maintained by selectively balanced polymorphism or by the persistent immigration of individuals from genetically differentiated populations. Quite often, fluctuations in population structure, or "demography" as it is now called, can generate local patterns of genetic change similar to those of selection (Thornton and Andolfatto 2006), and if neither the extant population structure nor the mechanisms of natural selection are known with certitude, the mere characterization of changing genetic variation by population geneticists will only rarely yield useful conclusions. [For a recent example of this problem, see Hernandez *et al.* (2011).] These problems surfaced 70 years ago in the collaboration of Dobzhansky and Wright, and over the course of the development of their field in the 20th century, evolutionary geneticists became steadily more guarded in their inferences about the action of selection in wild populations. These difficulties did not escape the notice of other types of geneticists, who often had little use for population genetics research that they commonly found overly descriptive and mechanistically unfocused.

Mendelian genetics has yielded many promising findings concerning the basic features of gene transmission and expression, and experimental population

genetics has revealed the abundant genetic variation to be found in many wild populations with great clarity. But both experimental strategies have faced clear obstacles impinging on their ability to address many of the most important questions in biology, such as the forces maintaining genetic variability in natural populations or the genetic constraints on the limits of organismal function. Further, it is at least reasonable to suggest that the separation of these strategies sustained these obstacles.

Stern (2000, 2010), to give just one example of a view somewhat similar to our own, has argued that Mendelian experimentation, as we define it here, largely ignored the problem of variation, while experimental population genetics unduly focused on aggregate statistical descriptions of genetic variation. Thus the former approach has revealed elements of Lewontin's (1974) genotype-to-phenotype map, but little about the causes of population-level genetic variation. The latter approach has revealed a great deal about population-genetics variation, but little about the specifics of pleiotropy, epistasis, and genotype-by-environment interaction. Thus Mendelian experiments have provided characteristically reductionist insights that are deficient in context, while experimental population genetics has not achieved much more than an overview of some of the "holistic" genetic properties of wild populations, as far as authors like Stern are concerned.

Here, however, the case will be made that empirical strategies for building a more fruitful, undivided experimental genetics are now available. For those who have always rejected the split between experimental research in Mendelian and population genetics, they can regard the present review as written in the spirit of reconciling disparate strands of experimental genetics. In either verbal formulation, the present intent is the same.

GENOME-WIDE DATA AND THE OLDER MENDELIAN AND POPULATION GENETICS

With the advent of relatively inexpensive whole-genome sequencing, whole-transcriptome sequencing, and genome-wide gene-expression assays, modern biology has turned a major corner. Previously, we argued that these new genomic technologies have led to a "new biology" of great promise (Rose and Oakley 2007). Perhaps less noticed is the degree to which whole-genome research has imperiled purist Mendelian or population genetics.

Key to the problems now facing Mendelian genetics is the degree to which the newly abundant genomic information reveals both extremely complex networks of genes affecting phenotypes and widespread epistatic effects of single substitutions on genome-wide gene-expression patterns. While it may be too early to claim that this high degree of complexity is more common than otherwise, what is indubitable is that such functional genomic complexity must now be considered a possibility. Thus, elegant Mendelian experiments that work out the effects of particular substitutions or mutations at a single locus or at a few loci face fundamental uncertainties. It may be that the mechanistic pathways that these experiments reveal are indeed a complete and sufficient genetic analysis if the phenotypes of interest are not underlain by a complex genetic network, but at present complexity appears to be more the rule than the exception.

This in turn poses a kind of induction problem. How, practically speaking, will Mendelian geneticists know when they have done enough experimental work to establish that the network components that they have already uncovered are all the components of importance? The number of experiments to be performed to achieve such empirical closure can be astronomically large, thanks to the combinatoric properties of genetic experiments. If 10 loci, each with three significant alleles, are to be combined in every possible way, there are 6^{10} or >60 million genotypes to be assayed. This is not a reasonable experiment, even with automation.

In a sense, this problem goes right to the core of what genetics should be about. Should genetics be about all the genotypes that can be constructed in a genetics laboratory, including genotypes with artificially induced mutations? If so, then it seems to be an astronomically vast endeavor.

Twentieth-century experimental population genetics, particularly as practiced by many of the intellectual descendants of Dobzhansky, had a solution to this problem. It focused instead on the genotypes that are to be found in natural populations in the wild. This is, at least logically, a tenable solution to the aforementioned combinatoric problem of Mendelian genetics experiments on all possible combinations of all possible alleles, while it also focuses attention on those genetic variants that are most relevant to study from the standpoint of high-function alleles.

But population genetics research on wild populations sacrifices the experimental power that Mendelian experiments provide. At the core of this sacrifice is that, to take genetic samples into the laboratory to breed for experimental work, the organisms sampled will be subjected to different environmental conditions and different population structures in almost all cases. Generically, 20th-century population geneticists created inbred lines from wild-caught founders and housed them in laboratories or greenhouses. If inbreeding proceeds with sufficient rapidity, it can overwhelm selection in the laboratory and preserve, over a large ensemble of inbred lines, allele frequencies that are close to those of the natural population(s) from which these lines were derived. But the resultant genotypes of these laboratories will not be a genome-wide reflection of the genotypes found in natural populations, if for no other reason than that they will be an extremely small

subset of the genotypes to be found in a genetically polymorphic wild population that undergoes frequent recombination.

If the phenotypes associated with these inbred lines are assayed, then they can present problems of inbreeding depression when their ancestral wild population had high heterozygosity. Few experimental results are more common in genetics than hybrid vigor in crosses of such inbred laboratory strains, particularly when important functional characters closely related to fecundity, longevity, or viability are assayed. This kind of experimental result reveals the extent to which the genotypic structure of the ancestral wild population has been destroyed by the creation of inbred laboratory strains.

But a potentially more insidious problem for the study of laboratory strains derived from wild populations will be genotype-by-environment interaction. Unlike inbreeding depression, in which the sign of the effect is predictable, the effect of an evolutionarily novel laboratory environment on functional characters will be unpredictable as to both magnitude and sign. Sometimes the laboratory environment may artificially boost a functional character, and sometimes it may depress it. These effects in turn may depend on the particular genotype of an inbred laboratory strain. This problem afflicts the derived samples of population genetics as much as it does the arbitrarily assembled genotypes of Mendelian genetics.

Taken together, these problems hamper attempts to infer the causal factors that shape the population genetics of wild populations, in that the only straightforwardly interpretable experiments are those in wild conditions and the only entirely relevant genotypes are those that occur naturally in wild population(s). Only with plants and sessile invertebrates will such experiments in the wild normally be feasible if studies of functional characters are to be performed, and then the experimenter faces the palpable uncertainties of wild conditions, including weather, human encroachment, and sheer accident.

BREAKING FREE OF TRADITIONAL EXPERIMENTAL STRATEGIES IN GENETICS

A possible key to escaping from the dilemmas of experimental Mendelian and population genetics may lie in accepting principles similar to those of strong-inference experimentation in physical science. Instead of making doctrinal fetishes of the experimental concreteness of Mendelian genetics or the "naturalness" of the population genetics of wild populations, genetic experiments could instead focus on powerful tests of generalities, which should apply to broad classes of genetic systems without exception. Given this goal, experimental strategies could be chosen to maximize their power, rather than any other sense of appropriateness. Rather than continuing with established practices, geneticists might focus on new combinations of experimental tactics and technologies.

To be concrete, we will consider the now-burgeoning genome-wide experimental strategies that offer the prospect of an undivided experimental genetics: quantitative trait locus (QTL) analysis, genome-wide association (GWA) studies, and genomic studies of experimental evolution. We do not propose that these are the only strategies for an experimental genetics that seeks to overcome past dichotomies of empirical research. For example, Houle (2010) has offered an alternative strategy that seeks to match genome-wide data to more extensive characterization of what he calls the "phenome," a strategy that is in some respects based on the biometrical research tradition and its evolutionary quantitative genetics offshoots of the late 20th century. While such an approach is somewhat beyond the scope of this review, we consider it also of significant promise.

QTL ANALYSIS

Before genome-wide sequencing technologies became inexpensive, QTL mapping was the most practical method for identifying genes involved in complex traits. This approach involves crossing individuals from stocks with very-well-characterized phenotypes and genotypes and determining the recombined regions of chromosomes that can be statistically associated with phenotypes in the hybrid offspring. Currently, the genetic dissection of quantitative traits is most feasible in well-characterized model systems; Drosophila and *Caenorhabditis elegans* are model organisms that have all the tools necessary for identifying QTL and characterizing them at the molecular level (Ayyadevara *et al.* 2001, 2003; Mackay 2004) and will serve as key illustrations for the purpose of the present discussion. In particular, we endeavor to point out the difficulties that face QTL analysis.

The question of how many QTL affect variation in a quantitative trait is not easily answered. The number of QTL mapped in any one experiment is almost certainly an underestimate. First, the two parent strains used in an experiment represent only a limited sample of the species-wide genetic variation for the trait in question. It should therefore not be surprising if different studies point to different QTL, even in a single species. QTL mapping experiments that involve parental stocks derived from laboratory selection experiments, however, will contain a more representative fraction of segregating variation than will crosses of two inbred lines. Second, the number of QTL is expected to increase with sample size, where sample size is the number of recombinant individuals. Increasing the sample size allows mapping of more QTL and of QTL with smaller effects. The frequency of overlap among QTL discovered in distinct interstrain crosses

does allow pairwise estimates of the total number of loci of similar or greater significance to those observed in individual experiments. Three such comparisons agree in implicating 11–24 longevity QTL of comparable effect size (Shmookler Reis *et al.* 2006).

The substantial body of work that has aimed to identify QTL affecting longevity in Drosophila serves as an appropriate example of the strengths and weaknesses of the QTL mapping approach in general. These results come both from recombinant inbred (RI) lines constructed from two parental stocks that were not selected for longevity (Nuzhdin *et al.* 1997; Pasyukova *et al.* 2000; Vieira *et al.* 2000; Leips *et al.* 2006) and from RI lines constructed from parental stocks that had undergone long-term selection for postponed aging (Curtsinger and Khazaeli 2002; Luckinbill and Golenberg 2002; Forbes *et al.* 2004; Valenzuela *et al.* 2004).

The authors of these studies admit that mapping longevity QTL is an imprecise initial step toward identifying genes responsible for longevity, as they can be localized only to approximate chromosomal regions. In addition, the extent to which different QTL mapping results from one or more laboratories can be compared is ill-defined (Shmookler Reis *et al.* 2006). Drosophila longevity QTL studies tend to identify 10–20 longevity "genes" that have large enough phenotypic effects to be detected (but see Curtsinger 2002). These large-effect QTL have been localized to the centromeric region of chromosome 2, and the left arm of chromosome 3, in independent studies (Valenzuela *et al.* 2004). Certainly, many loci with smaller phenotypic effects exist and have yet to be detected.

A recurring theme from studies of Drosophila QTL is that genotype-by-sex, genotype-by-environment, and epistatic interactions are common and complex (*e.g.*, Dilda and Mackay 2002). Drosophila QTL are often sex and environment specific, and longevity QTL often show antagonistic pleiotropy (reviewed in Mackay 2001, 2004). If Drosophila QTL have variable effects depending on the sex, physical environment, and genetic environment in which the QTL are expressed, similar properties are to be expected for QTL in other organisms (*cf.* Shmookler Reis *et al.* 2006; Rockman *et al.* 2010).

Less than 2 decades ago, studies of the genetic architecture of Drosophila sensory bristle number were taken to imply that natural variation for this trait could be localized to polymorphisms in relatively few candidate genes (Mackay 1996). But even in these early studies, complications arose from sex-specific QTL effects and interactions between QTL. The challenge for the future will thus be to incorporate a systems-biology perspective into traditional QTL approaches to assess how the particular alleles of many genome-wide loci affect multiple quantitative traits and networks of transcriptional interactions.

MEDICAL GENOME-WIDE ASSOCIATION STUDIES

One of the few organisms that geneticists study in detail in very large numbers is the human. Furthermore, there has been abundant funding for genome-wide assays of human genetic variation, thanks to concern over the possible medical significance. Furthermore, there are abundant data concerning human medical phenotypes, most importantly, the diagnosis of chronic endogenous conditions such as diabetes, hypertension, obesity, and other ailments that are long-sustained and not direct outcomes of infection.

Genome-wide association (GWA) studies use high-throughput methods to genotype panels of individuals at hundreds of thousands of sites and relate those sites to traits of clinical importance. GWA studies represent an important advance in discovering genetic variants influencing disease, but also have important limitations, including their potential for false-positive and false-negative results and for biases related to selection of study participants and genotyping errors (McCarthy *et al.* 2008).

The GWA approach permits surveys of the entire human genome in thousands of unrelated individuals, unconstrained by specific *a priori* mechanistic hypotheses regarding genetic associations with disease (Hirschhorn and Daly 2005). The genome-wide nature of GWA studies represents an important step beyond candidate gene studies that attempt to probe patterns of inheritance by focusing on single loci at a time, using the methods of Mendelian or population genetics. For conditions that are not traditional genetic diseases, GWA studies also represent a valuable advance over family-based linkage studies in which inheritance patterns of affected families are analyzed and related to genetic markers throughout the genome. These family-based linkage studies can successfully identify genes of large effect for traditional genetic diseases such as cystic fibrosis, but have been far less successful for common, complex disorders (*e.g.*, Altmuller *et al.* 2001).

Some of the most important, and widely cited, human genetic studies at the present time are the medical case-control GWA studies, such as that of the Wellcome Trust Case-Control Consortium (2007) study. These studies have been able to uncover new single-nucleotide polymorphisms (SNPs) that are statistically associated with the onset of chronic diseases and often SNPs at loci that were not known to Mendelian human genetics prior to the advent of GWA studies. GWA studies are predicated on the "common disease, common variant" hypothesis, which generally assumes that common diseases can be attributed to genetic variants present in ~5% of the population (Collins *et al.* 1997). If rarer disease-causing variants exist, or the effects of individual loci are small, we are unlikely to detect them with this approach.

Unfortunately, there are major inferential problems associated with the results of human GWA studies. As for all genome-wide research, there is a considerable multiple-inference statistical test problem with the use of numerous statistical tests over the hundreds of thousands of SNPs, and often multiple clinical endpoints, tested in such experiments. Without correcting for multiple-hypothesis testing, there will be a high false-discovery rate. Furthermore, given the significant linkage disequilibrium that characterizes the human genome, it is only rarely possible to specifically identify the SNP that functionally affects a disease risk, when a significant GWA result has been obtained (Shmookler Reis *et al.* 2006). For chronic diseases, the SNPs found to be statistically significant in GWA studies account for only a small fraction of their heritability (Manolio 2009). But GWA analysis of human height, which is the characteristic for which the most data are available and is more easily measured than chronic disease status, suggests that larger bodies of data and a different strategy of data analysis can account for much of the heritability of this character (Yang *et al.* 2010).

But there are still other difficulties with GWA research. Like laboratory populations of model organisms, humans in industrialized nations live in relatively novel environments. It is certain that many of the selective forces and demographic features that have shaped human genetic variation over the past 100,000 years are no longer present under modern conditions, even though determining just what the ancestral conditions were is itself a formidable project. Thus, genotype-by-environment interaction is a problem that may afflict medical GWA studies. The now-common invocation of dietary change as an etiological factor in human disease (*e.g.*, Lindeberg 2010) is itself an indication of this problem, particularly since the medical conditions that are known to be affected by diet, such as obesity, type II diabetes, cardiovascular disease, and many cancers, are the concern of much GWA research.

Nonetheless, there is a great deal to commend medical GWA research. It systematically studies variation over entire genomes, with its level of resolution depending on the state of the technology for genome characterization as well as the degree of linkage disequilibrium among locations across the genome. This is a great improvement over studies of a few candidate loci, even though GWA still faces grave limitations with respect to its ability to detect effects due to rare alleles. Although there are significant problems with genotype-by-environment interaction effects, GWA brings genetic information together with function, making it content-laden, particularly compared to the lack of functional content in much of population genetics. While the inference of specific functional roles for identifiable sequence differences from GWA data is plagued with significant uncertainties, as just outlined and discussed further below, at least functional questions are being addressed genome-wide.

GENOME-WIDE STUDIES OF EXPERIMENTAL EVOLUTION

While experiments in which model populations are made to evolve in response to culture conditions are of significant vintage, since 1980 they have been carried out with enough attention to population size, controls, and replication to make experimental evolution a relatively reliable experimental strategy (see Garland and Rose 2009). Of particular importance for functional and evolutionary interpretation, experimental evolution seeks to control the circumstances of both the present state and the ancestry of the populations that it studies. In Mendelian genetics, the evolutionary histories of the homozygous strains that it employs are often little known, and certainly are haphazardly controlled. While experimental population genetics can sometimes infer the ecology and the ancestry of the wild populations that it studies, the lines that are derived from the populations are characteristically subject to either an abrupt course of intense inbreeding or an ill-defined process of evolutionary domestication (*cf.* Simões *et al.* 2009).

The genetics of experimental evolution are of great relevance for the conundrums adduced here. This point has been conceded, at least implicitly, even by some of the most determined skeptics of the prospects for an undivided genetics (*e.g.*, "We can repeat the experimental evolution of phages in the laboratory, and demonstrate that the same mutations go to fixation in repeated experiments conducted under the same conditions" (Nielsen 2009, p. 2488). In the same way, Dykhuizen and Dean (2009) show that experimental evolution in bacteria can be used to rigorously test the adequacy of functional hypotheses concerning specific genotypes when they are assembled in the laboratory. But perhaps more interesting, from the standpoint of integrating Mendelian and population genetics, are the open-ended experiments, in which populations undergo selection without having their genotypes "pre-assembled." These experiments use populations of two basic kinds: clonal populations that accumulate *de novo* mutations and outbreeding sexual populations that have abundant standing genetic variation to begin with. Up until recently, it was chiefly the former type of population that was most amenable to genetic analysis (*e.g.*, Riehle *et al.* 2001), but more recently inexpensive molecular-genetic technologies have allowed fairly good genetic characterization of experimental evolution in sexual populations (*e.g.*, Teotónio *et al.* 2009). When such genetic characterization is extensive, it allows the experimenter to infer both the genetic substratum of a laboratory-defined adaptation and, conversely, the functional significance of genetic variants. As such, the genetics of experimental evolution constitute a natural bridge between the questions of Mendelian genetics and those of population genetics.

Experimental evolution is not without its difficulties and limitations. The population sizes used in laboratory

studies of evolution are necessarily small and thus no doubt systematically smaller than those of most wild populations. With many populations handled in parallel, often by numerous experimenters, contamination between populations is always a risk, particularly in cases where the genetic variation that undergoes selection is generated by *de novo* mutation. The timescales of laboratory evolution experiments are quite limited relative to the timescales of evolution in the wild. Linkage disequilibrium will make the ascertainment of causally important sequences ambiguous, as it characteristically does. Again, there will be statistical problems of false discovery with genome-wide assays of the effects of evolution. Even though experimental evolutionists may think that they know which characters selection is targeting, they will not necessarily be right (see Leroi *et al.* 1994). Most importantly, the idea that an experimental evolution study necessarily serves as a particularly pertinent guide to the evolution of wild populations is at least dubious (*cf.* Huey and Rosenzweig 2009). But in pursuit of universal principles of genetics, this is a secondary point. If a law of genetics is alleged to be exceptionless, then we could falsify it using the genetics of laboratory experimental evolution, despite the likely disparities between that laboratory system and others in nature.

Recent work by Burke *et al.* (2010) highlights the utility of the experimental evolution strategy for genetic analysis, despite these limitations. This study examined whole-genome sequence data from populations ($N_e \sim 10^3$) of Drosophila that had experienced over 600 generations of selection for accelerated development, as well as their ancestral or control populations. Flies in the strongly selected populations studied by Burke *et al.* (2010) develop ~20% faster than control flies and have evolved correlated phenotypic differences including smaller size, decreased stress resistance, and shorter mean life span. The primary goal of this study was to identify SNPs with significantly different allele frequencies in the experimental and control populations, as such loci can reasonably be associated with the aforementioned phenotypes. Burke *et al.* (2010) identified ~24,000 such SNPs, and since linkage disequilibrium extends up to 30–100 kb in these populations (see Teotonio *et al.* 2009), these SNPs localized to several dozen genomic regions that responded strongly to selection.

The observation that >20,000 SNPs significantly change in frequency suggests a large and complex genetic network underlying the response to selection for accelerated development. Perhaps more interesting is that Burke *et al.* (2010) found no evidence for the complete fixation of any of these alleles. Although local losses in heterozygosity were observed in the same areas of the genome at which there was significant differentiation in allele frequencies, in no region did heterozygosity come close to zero. The failure to observe the traditional signature of a selective sweep in this study is not necessarily unanticipated, given that 600 generations might not be enough time for newly arisen mutations to fix. On the other hand, very little allele-frequency differentiation was observed between replicate populations experiencing the same selection treatment; in other words, it is unlikely that beneficial new mutations arose independently in replicate populations and are currently in the process of fixing. The major conclusion to be drawn from this work is that, unlike microbial evolution experiments, selection acts primarily on standing variation, and not on new mutations, in sexually reproducing systems undergoing experimental evolution.

These experimental evolutionary genomic results ostensibly create a genetic load paradox. How can these laboratory Drosophila populations sustain so much genetic variation for fitness-related characters such as developmental speed, early fecundity, etc? But the genetic load paradox may be more apparent than real. Consider a simple case of balancing selection. For a locus at an overdominant equilibrium, the mean population fitness will be lower than the fitness of the heterozygote. This difference has been referred to as the segregational genetic load (Ewens 1979). More formally, genetic load, L, is defined as,

$$L = \left(w_{max} - \bar{w}\right)/\bar{w}, \tag{1}$$

where $\bar{w}$ is the mean fitness and $w_{\max}$ is the maximum fitness among all genotypes in the population. If the action of natural selection is assumed to act through strict viability selection, then a large genetic load has been taken to indicate that a population may be unable to numerically replace itself, which leads to extinction. Consequently, some (*e.g.*, Kimura 1968) have used this argument to suggest that there is a limit to how much genetic variation can be maintained by natural selection. These arguments have been countered by pointing out that natural selection does not always act through the deaths of individuals; rather, processes like frequency-dependent selection will ameliorate the impact of selection on population viability. Indeed, in Drosophila it has been well documented that adaptation in many severe larval environments results in changes in larval feeding rates (see review in Mueller *et al.* 2005), which in turn affects competitive ability for food, which is a frequency-dependent process (Mueller 1988).

Another reason that genetic load calculations are likely to overstate the negative impact of selection is that the most-fit genotype may be vanishingly rare, in which case the difference between the average and most-fit genotype may be dramatically smaller. For example, consider a single locus with two alleles, A_1 and A_2. Suppose the fitnesses for genotypes A_1A_1, A_1A_2, and A_2A_2 are $1 - s_1$, 1, and $1 - s_2$, respectively. If s_1, $s_2 > 0$, then the equilibrium frequency of the A_1

TABLE 1

Theoretical load (L), observed load ($\widehat{L}$), and mean fitness ($\widehat{w}$) in samples of 100 individuals

	$\widehat{p}$				
	0.05	0.1	0.2	0.4	0.5
s_1	0.038	0.018	0.008	0.003	0.002
L	1.6	1.5	1.2	0.82	0.65
$(\widehat{L})$	0.075	0.093	0.09	0.069	0.057
$(\widehat{w})$	0.39	0.41	0.45	0.55	0.61

The column heading $\widehat{p}$ and the row heading s_1 refer to the equation of overdominance outlined in the text. In each case, $s_2 = 0.002$. The value of s_1 was chosen to give the equilibrium allele frequency indicated in the table.

allele is $\widehat{p} = s_2/(s_1 + s_2)$. Now suppose we have 500 independent loci maintained by heterozygote advantage. The most-fit genotype is the multiple heterozygote with net fitness of $1^{500} = 1$. The mean fitness over all 500 loci is $[\widehat{p}^2(1-s_1)+2\widehat{p}(1-\widehat{p})+(1-\widehat{p})^2(1-s_2)]^{500}$. If s_1, $s_2 = 0.02$, then from Equation 1 we calculate the load as 150. However, the likelihood of seeing the heterozygote at all 500 loci in a finite population is very small.

To quantify this problem, we have created finite populations of 100 individuals and assigned them genotypes randomly at each of 500 loci, assuming that each locus was at the same overdominant equilibrium. We then calculated the load of this population of 100 individuals using Equation 1, but w_{max} was set to the highest fitness among the 100 individuals. We repeated this process 1000 times and computed the mean genetic load in these populations of 100 individuals and have contrasted this to the maximum expected load (Table 1). The results in Table 1 show that the load in small samples is 10–20 times less than the maximum load.

The significance of this finding is that outbred laboratory populations maintained under long-sustained selection regimes nonetheless can retain abundant genetic variation of evolutionary genetic and functional significance. This offers the prospect of using the genome-wide response to changes in the phenotypic focus of selection as a general-purpose tool for resolving questions of importance for both Mendelian and population genetics.

FROM GENOME-WIDE ANALYSIS BACK TO INDIVIDUAL GENES?

A natural goal for genetic research would be to proceed from genome-wide analysis to the functional dissection of individual genetic variants. This is perhaps most obvious in the case of QTL mapping, in that this experimental strategy is expressly focused on identifying specific regions of the genome that have a measureable effect on specific phenotypes. To the extent that QTL mapping identifies causally important specific quantitative trait nucleotides (QTNs), then such polymorphisms could be subjected to functional genetic validation using complementation tests among other standard Mendelian procedures. In the research of Burke *et al.* (2010), reasonably narrow areas of genomic differentiation were identified; with greater resequencing coverage, genome-wide sequencing of such experimentally evolved populations might provide candidate QTNs of comparable value to those of QTL analysis. Likewise, large-scale GWA research might also provide candidate SNPs that could be examined using more functional genetic studies. Alternatively, candidate SNPs identified by any of the three types of genome-wide scan could be subjected to such molecular genetic assays as RNA interference, overexpression in transgenic constructs, or allele replacement, at least with model organisms.

A significant problem impinging on such explicitly causal research is that the effect sizes of these candidate SNPs might be so small as to vitiate such direct assays of their phenotypic effects. The analysis of the Burke *et al.* (2010) data provided in its supplementary files suggests fairly stringent limits on the effect sizes of the SNPs involved, given that none of them showed evidence of a selective sweep during >600 generations of selection. Likewise, the small proportion of heritability accounted for by some GWA studies (Molino 2009) suggests that the candidate SNPs identified by such work often have a quantitatively small effect. Many inferred QTL may arise from local linkage disequilibrium among multiple QTNs, each of them having small effects individually. Andolfatto (2005) has estimated the likely selection coefficients of SNPs involved in selective sweeps and finds that they are quite small, suggesting that they have relatively small quantitative effects on measureable phenotypes. However, this is a question that is best resolved by experimental attempts to establish the magnitude of the phenotypic effects of SNPs identified by the kind of genome-wide experiments discussed here.

CONCLUSION

In an undivided experimental genetics, we should be free to take our experimental refutations and corroborations wherever we can, without regard to convention or habit. There are still good experiments to be done within the limited confines of both Mendelian genetics and experimental population genetics, as conventionally practiced. But there are now other experimental methods that have a good claim on our attention, particularly in an era when genomic technologies offer possibilities that were barely conceived of in the 1930s and 1940s, when tho two wings of experimental genetics discussed here first started to drift apart.

We provided a brief sketch of some of the possibilities for genomically founded biology in an earlier article (Rose and Oakley 2007), but it might be useful to

mention a few specifically promising lines of research for the undivided genetics that are now emerging:

Whole-genome resequencing of individuals obtained directly from wild populations should reveal the extent to which populations in nature sustain genetic variation and linkage disequilibrium at the nucleotide level, and what kind of structural and nucleotide-level variation is present in such populations.

Trajectories of genome-wide variation *among* replicated populations undergoing identical culture regimes in parallel over many generations at different population sizes should reveal the relative importance of genetic drift *vs.* deterministic patterns of selection in shaping genetic change over moderate evolutionary time periods.

Reverse experimental evolution (see Teotonio and Rose 2000, 2001) combined with whole-genome resequencing (*cf.* Teotonio *et al.* 2009) should reveal how common antagonistic pleiotropy is across entire genomes and thus indicate the potential for selection to sustain genetic polymorphisms.

It should be possible to use experimental evolution of replicated populations to test the relative importance of *de novo* mutation producing selective sweeps *vs.* shifting selectively sustained polymorphisms in the evolution of outbred sexual populations, provided enough generations of selection are monitored using whole-genome resequencing.

As discussed above, genome-wide scans followed by causal tests of the phenotypic effects of individual QTNs should test whether such QTNs have measureable phenotypic effects.

When QTNs with measureable phenotypic effects are identified, they could be tested for their pleiotropic and epistatic effects, and when these experiments are done over a variety of environments, they could also delimit patterns of genotype-by-environment interaction.

But as we move through the phenomena of genetics with the freedom to perform new kinds of experiments, we should be wary of supposing that the specifics of what we find in a particular laboratory or a particular set of samples from the wild will inductively generalize. Fortunately, although science grows out of a substratum of particulars, fruitful scientific debate is about general rules and theories, not about the artifacts generated by specific experimental methods. In the end, all experimental strategies are subject to difficulties that may render their conclusions suspect, suggesting that there is all the more reason to unburden ourselves of Procrustean methodologies and narrow views of the possibilities for genetic experimentation.

We are grateful to the anonymous referees, Robert J. S. Reis, Kevin R. Thornton, and Adam Wilkins for their comments on earlier drafts of this manuscript. The research of the authors has been supported by grants from the National Institutes of Health, National Science Foundation, and the University of California.

LITERATURE CITED

Altmüller, J., L. J. Palmer, G. Fischer, H. Scherb and M. Wjst, 2001 Genome-wide scans of complex human diseases. Am. J. Hum. Genet. **69**(5)**:** 936–950.

Andolfatto, P., 2005 Adaptive evolution of non-coding DNA in *Drosophila*. Nature **437:** 1149–1152.

Ayyadevara, S., R. Ayyadevara, S. Hou, J. J. Thaden and R. J. Shmookler Reis, 2001 Genetic mapping of quantitative trait loci governing longevity of *Caenorhabditis elegans* in recombinant-inbred progeny of a Bergerac-BO × RC301 interstrain cross. Genetics **157:** 655–666.

Ayyadevara, S., R. Ayyadevara, A. Vertino, A. Galecki, J. J. Thaden *et al.*, 2003 Genetic loci modulating fitness and life span in *Caenorhabditis elegans*: categorical trait interval mapping in CL2a × Bergerac-BO recombinant-inbred worms. Genetics **163:** 557–570.

Burke, M. K., J. P. Dunham, P. Shahrestani, K. R. Thornton, M. R. Rose *et al.*, 2010 Genome-wide analysis of a long-term evolution experiment with *Drosophila*. Nature **467:** 587–590.

Collins, F. S., M. S. Guyer and A. Chakravarti, 1997 Variations on a theme: cataloging human DNA sequence variation. Science **278**(5343)**:** 1580–1581.

Curtsinger, J. W., 2002 Sex specificity, life-span QTLs, and statistical power. J. Gerontol. **57A**(12)**:** B409–B414.

Curtsinger, J. W., and A. A. Khazaeli, 2002 Lifespan, QTLs, age-specificity, and pleiotropy in *Drosophila*. Mech. Ageing Dev. **123:** 81–93.

Dilda, C. L., and T. F. C. Mackay, 2002 The genetic architecture of Drosophila sensory bristle number. Genetics **162:** 1655–1674.

Dykhuizen, D. E., and A. M. Dean, 2009 Experimental evolution from the bottom up, pp. 67–89 in *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*, edited by T. Garland and M. R. Rose. University of California Press, Berkeley.

Ewens, W. J., 1979 *Mathematical Population Genetics*. Spinger-Verlag, New York.

Forbes, S. N., R. K. Valenzuela, P. Keim and P. M. Service, 2004 Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. I. Composite interval mapping. Genetics **168:** 301–311.

Garland, T., and M. R. Rose, 2009 *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*. University of California Press, Berkeley.

Hernandez, R. D., J. L. Kelley, E. Elyashiv, S. C. Melton and A. Auton *et al.*, 2011 Classic selective sweeps were rare in recent human evolution. Science **331**(6019)**:** 920–924.

Hirschhorn, J. N., and M. J. Daly, 2005 Genome-wide association studies for common diseases and complex traits. Nat. Rev. Genet. **6**(2)**:** 95–108.

Houle, D., 2010 Numbering the hairs on our heads: the shared challenge and promise of phenomics. Proc. Natl. Acad. Sci. USA **107**(Supp. 1)**:** 1793–1799.

Huey, R. B., and F. Rosenzweig, 2009 Laboratory evolution meets catch-22: balancing simplicity and realism, pp. 671–702 in *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*, edited by T. Garland and M. R. Rose. University of California Press, Berkeley.

Khazaeli, A. A., W. Van Voorhies and J. W. Curtsinger, 2005 The relationship between life span and adult body size is highly strain-specific in *Drosophila melanogaster*. Exp. Gerontol. **40:** 377–385.

Kimura, M., 1968 Evolutionary rate at the molecular level. Nature **217:** 624–626.

Leips, J., P. Gilligan and T. F. C. Mackay, 2006 Quantitative trait loci with age-specific effects on fecundity in *Drosophila melanogaster*. Genetics **172:** 1595–1605.

Leroi, A. M., A. K. Chippindale and M. R. Rose, 1994 Long-term laboratory evolution of a genetic trade-off in *Drosophila melanogaster*. I. The role of genotype × environment interaction. Evolution **48:** 1244–1257.

Lewontin, R. C., 1974 *The Genetic Basis of Evolutionary Change*. Columbia University Press, New York.

Lindeberg, S., 2010 *Food and Western Disease: Health and Nutrition from an Evolutionary Perspective*. Wiley-Blackwell, New York.

Luckinbill, L. S., and E. M. Golenberg, 2002 Genes affecting aging: mapping quantitative trait loci in *Drosophila melanogaster*

using amplified fragment length polymorphisms (AFLPs). Genetica **114**(2)**:** 147–156.

Macdonald, S. J., and A. D. Long, 2004 A potential regulatory polymorphism upstream of *hairy* is not associated with bristle-number variation in wild-caught Drosophila. Genetics **167:** 2127–2131.

Mackay, T. F. C., 1996 The nature of quantitative genetic variation revisited: lessons from Drosophila bristles. Bioessays **18:** 113–121.

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

Mackay, T. F. C., 2004 The genetic architecture of quantitative traits: lessons from Drosophila. Curr. Opin. Genet. Dev. **14:** 253–257.

Manolio, T. A., F. S. Collins, N. J. Cox, D. B. Goldstein, L. A. Hindorff *et al.*, 2009 Finding the missing heritability of complex diseases. Nature **461:** 747–753.

McCarthy, M. I., G. R. Abecasis, L. R. Cardon, D. B. Goldstein, J. Little *et al.*, 2008 Genome-wide association studies for compled traits: consensus, uncertainty and challenges. Nat. Rev. Genet. **9:** 356–369.

Mueller, L. D., 1988 Density-dependent population growth and natural selection in food limited environments: the *Drosophila* model. Am. Nat. **132:** 786–809.

Mueller, L. D., D. G. Folk, N. Nguyen, P. Nguyen, P. Lam *et al.*, 2005 Evolution of larval foraging behaviour in *Drosophila* and its effects on growth and metabolic rate. Physiol. Entomol. **30:** 262–269.

Nielsen, R., 2009 Adaptationism: 30 years after Gould and Lewontin. Evolution **63:** 2487–2490.

Nuzhdin, S. V., E. G. Pasyukova, C. L. Dilda, Z. Zeng and T. F. C. Mackay, 1997 Proc. Natl. Acad. Sci. USA **94:** 9734–9739.

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.

Provine, W. B., 1986 *Science and Its Conceptual Foundations: Sewall Wright and Evolutionary Biology*. University of Chicago Press, Chicago.

Riehle, M. M., A. F. Bennett and A. D. Long, 2001 Genetic architecture of thermal adaptation in Escherichia coli. Proc. Natl. Acad. Sci. USA **98**(2)**:** 525–530.

Rockman, M. V., S. S. Skrovanek and L. Kruglyak, 2010 Selection at linked sites shapes heritable phenotypic variation in *C. elegans*. Science **372:** 372–376.

Rose, M. R., 1991 *Evolutionary Biology of Aging*. Oxford University Press, Oxford.

Rose, M. R., and T. H. Oakley, 2007 The new biology: beyond the Modern Synthesis. Biol. Direct **2:** 30.

Shmookler Reis, R. J., P. Kang and S. Ayyadevara, 2006 Quantitative trait loci define genes and pathways underlying genetic variation in longevity. Exp. Gerontol. **41:** 1046–1054.

Simões, P., J. Santos and M. Matos, 2009 *Experimental evolutionary domestication*, pp. 89–110 in *Experimental Evolution: Concepts, Methods, and Applications of Selection Experiments*, edited by T. Garland and M. R. Rose. University of California Press, Berkeley.

Stern, D. L., 2000 Perspective: evolutionary developmental biology and the problem of variation. Evolution **54**(4)**:** 1079–1091.

Stern, D. L., 2010 *Evolution, Development, and the Predictable Genome*. Roberts & Company Publishers, Greenwood Village, CO.

Teotónio, H., and M. R. Rose, 2000 Variation in the reversibility of evolution. Nature **408:** 463–466.

Teotónio, H., and M. R. Rose, 2001 Perspective: reverse evolution. Evolution **55:** 653–660.

Teotónio, H., I. M. Chelo, M. Bradic, M. R. Rose and A. D. Long, 2009 Experimental evolution reveals natural selection on standing genetic variation. Nat. Genet. **41:** 251–257.

Thornton, K. R., and P. Andolfatto, 2006 Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaser*. Genetics **172:** 1607–1619.

Valenzuela, R. K., S. N. Forbes, P. Keim and P. M. Service, 2004 Quantitative trait loci affecting life span in replicated populations of *Drosophila melanogaster*. II. Response to selection. Genetics **168:** 313–324.

Van Voorhies, W. A., J. W. Curtsinger and M. R. Rose, 2006 Do longevity mutants always show trade-offs? Exp. Gerontol. **41:** 1055–1058.

Vieira, C., E. G. Pasyukova, Z. Zeng, J. Brant Hackett, R. F. Lyman *et al.*, 2000 Genotype-environment interaction for quantitative trait loci affecting life span in *Drosophila melanogaster*. Genetics **154:** 213–227.

Wellcome Trust Case Control Consortium, 2007 Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature **447**(7145)**:** 661–678.

Yang, J., B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders *et al.*, 2010 Common SNPs explain a large proportion of the heritability for human height. Nat. Genet. **42**(7)**:** 565–569.