Is Life Law-Like?

Kenneth M. Weiss[1] and Anne V. Buchanan

Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania 16802

Abstract Genes are generally assumed to be primary biological causes of biological phenotypes and their evolution. In just over a century, a research agenda that has built on Mendel's experiments and on Darwin's theory of natural selection as a law of nature has had unprecedented scientific success in isolating and characterizing many aspects of genetic causation. We revel in these successes, and yet the story is not quite so simple. The complex cooperative nature of genetic architecture and its evolution include teasingly tractable components, but much remains elusive. The proliferation of data generated in our "omics" age raises the question of whether we even have (or need) a unified theory or "law" of life, or even clear standards of inference by which to answer the question. If not, this not only has implications for the widely promulgated belief that we will soon be able to predict phenotypes like disease risk from genes, but also speaks to the limitations in the underlying science itself. Much of life seems to be characterized by *ad hoc*, ephemeral, contextual probabilism without proper underlying distributions. To the extent that this is true, causal effects are not asymptotically predictable, and new ways of understanding life may be required.

Perhaps the correct way of viewing the whole subject would be to look at the inheritance of every character whatever as the rule, and the non-inheritance as the anomaly.

Charles Darwin (1859)

WE are in a period in the history of science with important roots in the era known as the Enlightenment. About three centuries ago, empiricism replaced more casual and speculative approaches to understanding nature. Today, nature is assumed to be entirely physical and governed by universal, unexceptioned laws. A law of nature is a process, or mechanism, of cause and effect, and we pursue the laws of nature through what has become known as "the scientific method." The essence of the scientific method is systematic, controlled observation.

Among the implications of empiricism are reductionism, replication, prediction, and the ability to deduce new facts. Nature is complex, but if causation is law-like, then we should be able to reduce complex observations to simpler generalizations and to isolate the relevant variables in any situation. The same conditions should always lead to the same outcome so that experimental results should be entirely replicable, and if we understand the relevant "law," then from a given set of conditions, outcomes should be specifically predictable, with asymptotic precision, limited only by measurement error. Predictability was extended, especially in the 20th century, to include the law-like distributions of probabilistic processes. In this sense, nature is deterministic.

From an empiricist's perspective, there would seem to be no reason to think that life is exempt from law-like behavior. In the 19th century, catalyzed both iconically and in substance by the work of Charles Darwin, the study of life was brought under the Enlightenment tent. In the 152 years since the *Origin of Species* was published, progress in biology has been unmatched in human history. We now know a tremendous amount about the biochemistry of inheritance, the processes of development by which organisms are made, and so forth, but the more we know, the more complex the picture becomes, not less. The legacy of Enlightenment-derived science has enabled us to amass large amounts of detail about how life works, but in genetics as well as in many other fields, while pragmatic objectives have been met in many areas (such as experimental and agricultural genetics), phenogenetic prediction of complex traits and their evolutionary reconstruction from genomic data remain daunting challenges. This makes questions such as "what is life, beyond its biochemical core?" or "is there a 'law' of

life?" more than philosophical musings. The answers, or even how to answer them, are not clear.

Because Darwin's work is so iconic, and because much in biology is still seen by many to make sense only in light of Darwinism, we will rhetorically frame what follows around selections of Darwin's own words to represent the widely if often informally held view that genes are deterministic, traits are "adaptive," and science can uncover the cause-and-effect relationship between them. Certainly, progress in biology in the years since Darwin has included the recognition that genotype–phenotype relationships may not be as straightforward as it seemed from Mendel's purposefully designed studies with peas, and biologists and population geneticists have addressed this issue at least since R. A. Fisher's benchmark paper about polygenic causation in 1918 (Fisher 1918). But, perhaps because Mendelian traits are so much easier to parse than more complex traits, the desire persists to find the genes for almost any trait of interest, as does the belief that our current methods can identify them, making life inherently predictable from genotypic data.

Darwin repeatedly likened the laws of life, centered around inheritance of fundamental causal units, to Newton's universal, deterministic laws of physics. He even concluded the *Origin* with the image of evolution proceeding along with the earth's cycling "according to the fixed law of gravity" (Darwin 1859). Inheritance is essential to the development of organisms from seeds, the division of organisms into differentiated cells, and the origins and diversity of species. Darwin's triumph was to argue that the law-like process of inheritance connected all of life back to its common origin in "one grand natural system" (Darwin 1859).

In the *Origin*, Darwin was clear:

> Every one who believes, as I do, that all the corporeal and mental organs (excepting those which are neither advantageous nor disadvantageous to the possessor) of all beings have been developed through natural selection, or the survival of the fittest, together with use or habit, will admit that these organs have been formed so that their possessors may compete successfully with other beings, and thus increase in number (Darwin 1887).

The celebration of this great work, and the powerful potential of Darwin's theory compared to other explanations of life in his time, meant that his ideas became canonized, casting a long shadow over subsequent thinking. That persists today, when most of us are trained to be Darwinians, perhaps without realizing just what that means in light of what we now actually know, and in particular, as probabilism has become a fundamental aspect both of practical empiricism and of our understanding of nature herself.

The Darwinian View: "Everything in Nature Is the Result of Fixed Laws" (Darwin 1887)

On the Origin of Species is one of the most comprehensively nuanced and thoughtful books ever written. But overall Darwin characterized natural selection as a Newtonian kind of infinitesimally continuous force that detects "the smallest grain in the balance" [Darwin 1859, including chapter on natural selection, read to Linnaean Society in 1858] of relative fitness. It is the force that connects the units (of inheritance) of all life, much as gravity connects all matter.

Ironically, Darwin's own theory of inheritance, his "provisional hypothesis of pangenesis" (Darwin 1859), was in a sense antithetical to his own concept of evolution by natural selection. His units of inheritance, "gemmules," were deterministically causal because their form was molded by the organism's life experience, and then transmitted to its offspring. He suggested that "every separate part of the whole organization [of an organism] reproduces itself" (Darwin 1859). Gemmules were trait-specific, and rather than a unified structure, "An organic being is a microcosm— a little universe, formed of a host of self-propagating organisms inconceivably minute and numerous as the stars in heaven" (Darwin 1859). Translated into the genetic language of today, one might say "a host of self-propagating genetic elements inconceivably as minute and numerous as the nucleotides in the genome."

In Darwin's theory, ironically there is perhaps no need for natural selection, since any individual could behaviorally develop adapted gemmules. By contrast, modern views of natural selection are based on competitive screening of individuals bearing randomly arisen variation. Nonetheless, Darwin's universal adaptive determinism persisted into the intellectual gestalt when the modern synthesis melded single-gene Mendelian with Darwinian adaptive concepts. This is why genetic drift and non-Darwinian neutral evolution are still controversial in some circles.

The same worldview underlies the subjective expectation that genome-wide mapping and searches for genomic signatures of natural selection will yield tractably few meaningful, enumerable, and practically useful causal genes—Darwin's modified gemmules recast as Mendel's deterministic elements. This conceptual worldview explains the intense defense of genome-wide association study (GWAS), biobanks, personalized genomic medicine, and whole-genome sequencing on a population scale by the human genetics research community, despite the fact that expectations that these methods will identify many single genes with strong effects have not been met. The rhetoric tells the tale, as it is common to see reference to genes for complex traits such as heart disease and individual alleles driven by selective sweeps. This is Darwinian melodrama at its telegenic best. But is it accurate?

"Laws of Variation" (Darwin 1859): Life Is Cooperative by Nature

The mantra of Darwinian evolution is competition, with the fittest alleles advancing in frequency (or in heterosis being maintained) by selection. This seems certainly to be an important aspect of change and a major mechanism for the
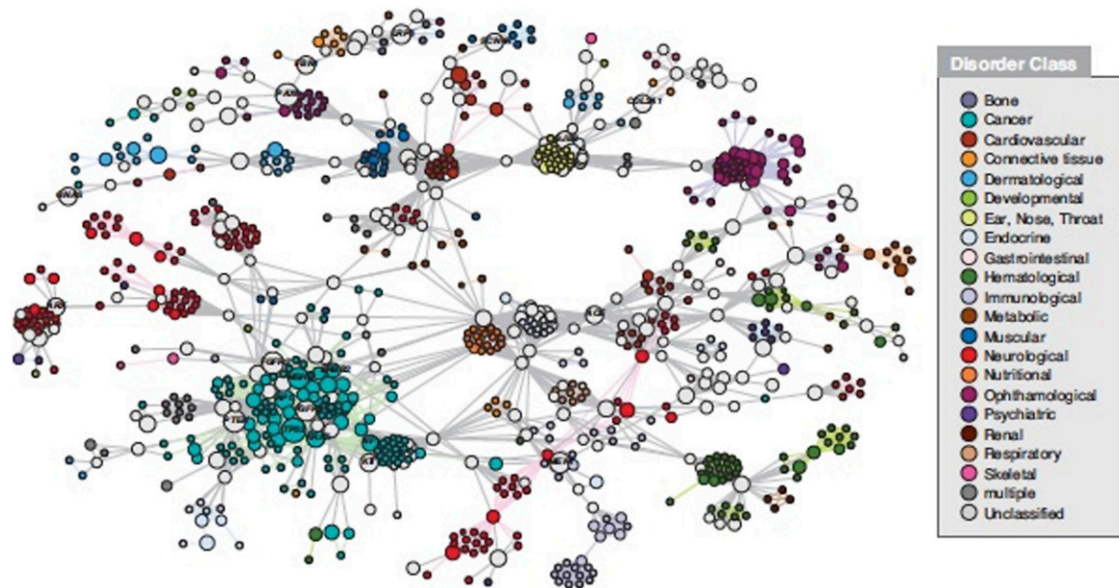
Figure 1 The polygenic, pleiotropic, and syndromic nature of life. Large networks of factors in a greater network, as exemplified by genes that are associated with human disease classes, are coded by color. Circle size reflects the number of disorders to which a gene contributes; links connect gene–gene associations; and light-gray nodes are involved in more than one disorder class, with gene names given for genes involved in more than five different disorders [from Goh *et al.* (2007) and Barabasi *et al.* (2011), where details can be found].

evolution of highly organized traits. But a focus on competition can draw attention from the fundamental role of cooperation among the countless interacting units that are responsible for organismal complexity. Cooperation—that is, co-operation—is a fundamental principle of life at all levels, including within cells, between cells, between tissues and organs, and between organisms as well as species sharing ecosystems (Weiss and Buchanan 2009a,b). Instead of wished-for simplicity, the complexity of life as demonstrated by mapping makes it clear that for many, if not most, traits a large number of cooperatively interacting genetic elements contribute to the trait's construction in an individual and its variation in a species.

These interacting factors correspond to Darwin's elements "as numerous as the stars in heaven." Yet he understood that, despite his theory of pangenesis, traits are not phenotypically or evolutionarily independent. He described the laws of "correlation of growth" such that "the whole organization is so tied together during its growth and development, that when slight variations in any one part occur, and are accumulated through natural selection, other parts become modified" (Darwin 1859).

In modern terms, we would say that life is largely *polygenic*, *pleiotropic*, and *syndromic*. That is, traits and organisms are made by overlapping, partly sequestered, and partly interacting processes. Life's cooperative complexity is what has enabled organized traits to evolve piecemeal from simpler beginnings. All one need to do is reflect on the way in which developmental genetics works through complex signaling and other interactions. A plethora of genome-wide mapping results clearly support this view, to the point that it is almost meaningless to ask how many loci are

reallyinvolved or to debate how many individuals must be sampled to identify all contributing loci by some statistical criterion.

It is also no surprise that there are widespread correlations among traits like diseases that can involve clinical variation in one or multiple organ systems and that genetic correlations are related to these phenotypes (Figure 1; Barabasi *et al.* 2011). The complex interactions from which biological traits are built shows clearly that life is far more about cooperative interaction than it is about competition and will be better understood in that way than by focusing on individual genes.

Current catchwords for cooperative interaction include "network" or "systems" biology. Often treated at least informally as higher-level causal units, genetic networks are identified by patterns of gene-product interaction that can be experimentally demonstrated or empirically observed in terms of correlated sequence variation or expression levels (*e.g.*, Barabasi *et al.* 2011).

Perhaps in part for subjective cultural reasons, current network approaches tend to search for hierarchy rather than community. Ranking criteria involve the number of associated traits or the number of regulatory network connections in which each gene participates. Genes involved in many connections are identified as "major," "hub," or "master genes" in their respective networks (networks are typically even named after such genes, as in "Wnt signaling") although the nomenclature may as often reflect the history of discovery as it does functional mastery (*e.g.*, Mattick *et al.* 2010). But these genes are typically pleiotropic, interacting with multiple other genes and expressed in multiple tissues and contexts.

On first thought, one might expect that such important genes would be more critical to normal function and so more likely to be disease genes when mutated. However, the empirical evidence suggests that hub genes are statistically *less* likely to be involved in disease (Barabasi *et al.* 2011). This really is not so surprising. Variation in critical genes is more likely to be quickly, even embryologically, lethal.

In the biomedical context, perhaps more important is that most common diseases arise in previously healthy persons, often mature adults, with clearly viable genotypes, implying that the majority of genetic factors responsible for disease are minor members of their respective networks and pathogenic only via interaction with their environmental or genomic context. The degree to which phenotypes depend upon such context is not, we think, nearly sufficiently appreciated (or, perhaps, too often appreciated only in passing), but has clearly been demonstrated in the biomedical arena (Sing *et al.* 2003, 2004; Dyson *et al.* 2007). Indeed, SNP alleles that are severely deleterious in one species can become normal in other species due, for example, to compensatory variation elsewhere in the genome (Kondrashov *et al.* 2002; Gao and Zhang 2003). Since minor genes are much more numerous than hubs, disease causation can be expected to be more polygenic than monogenic. This is what we observe.

These statements may seem delusional given the countless Mendelian traits, including diseases, that are observed in every species. However, those are the strong signals, many if not most being measured very close to the gene-product level, and while they seem clearly to be causally interpreted correctly, upon close inspection they often turn out not to be so simply Mendelian as was thought: for example, many alleles are usually found at a locus, with a range of penetrance effects and variable severity even for the "signal" alleles by which a Mendelian pattern was discovered.

"A Very Important Subject, Most Imperfectly Understood" (Darwin 1859)

Similar statements apply to evolution, the process responsible for building the genetic architecture that we observe today. Most mutations in hub genes or vital parts of any gene may be nonviable or seriously deleterious, removed from the population and explaining their high level of sequence conservation. This is interesting because in a sense such purifying selection, apparently the most prevalent form of selection, is non-Darwinian. The reason that hub variants are removed is not because their bearers cannot compete in a Malthusian world, but because at the molecular level they do not cooperate with other gene products during development: the embryo simply fails. Such selective screening is for cooperation rather than competition.

These findings reflect an essential fact of polygenic (multilocus) causation: that many different genotypes can yield similar phenotypes. But an implication of the purifying removal of vital genetic variation is that positive adaptation over time, like genetic architecture at any given time, may largely be due to variation in the more numerous minor rather than hub genes, because they are the genes that can sustain viable mutation and whose bearers can live to compete. These alleles make little individual contribution to the phenotype, their selective coefficients are very small, and their frequency change will be affected largely by genetic drift (Hartl and Clark 2007; Lynch 2007a,b). This may be why searches for signatures of selection in terms of comparative DNA sequence analysis (*e.g.*, Lopez Herraez *et al.* 2009; Pritchard *et al.* 2010; Weiss 2011), as well as genome-wide association studies (GWAS) searching for causal genes for complex traits (Weiss and Terwilliger 2000; Weiss 2008; Gibson 2010; Manolio 2010), find relatively few easily replicable signals, and these usually account for only a fraction of the observed change.

Selection systematically raises the frequency of an allele with major effect in the favored direction. But alleles at all other genes across the genome that happen to nudge the trait gently in the same direction will also be favored, even if only slightly and in aggregate. The more strongly favored alleles may advance more rapidly, leaving stronger population genetic signatures of selection in terms of various measures such as reduced haplotype variation (Sabeti *et al.* 2007; Lopez Herraez *et al.* 2009; Pritchard *et al.* 2010; Weiss 2011). But the preponderance of the phenotypic response could in fact be due to the much larger set of individually minor contributors that leave no detectable signal (*e.g.*, Burke *et al.* 2010; Pritchard *et al.* 2010). After the fact, the impression could be that the adaptation was based on the detected gene for the trait, when in fact it is a net, genome-wide holistic background effect as well. That is why transplanting the gene into different experimental animal strains often has highly differing effects, and why effects of the same nominal allele may be population-specific in humans. Indeed, over time the genomic "drag" of selection may lead to the muting of the initially strong net effects of major alleles. In that sense, the Mendelian alleles of our dreams may usually be rare and evolutionarily recent.

"Unless Profitable Variations Do Occur" (Darwin 1859): Life Is Probabilistic

As science progressed through the 19th and 20th centuries, there was an increasing recognition of the role of probability in how we understand the world. Two aspects of probabilism are important to both genetics and its evolution. One concerns repeatable outcomes, as in sampling and gambling. In such situations there are alternative, but deterministic and enumerable, outcomes that have regular parametric distributional properties, such as Mendelian segregation. At least statistically, such outcomes are predictable and replicable, two central criteria of Enlightenment-derived science. Related to this in a sense, sampling theory was developed to test hypotheses and thus formalize the scientific method.

Darwin did not have these methods at his disposal, but this is the kind of probabilism that Gregor Mendel and T. H. Morgan studied: genetic elements were causally deterministic but probabilistically transmitted.

There is also causal probabilism. Even before the probabilism introduced into physics by quantum mechanics, probabilism had arisen in physics in natural laws or principles, such as in statistical mechanics, like thermodynamics, by which order decayed to disorder: each atom moves probabilistically, but their aggregate behavior can be described in distributional terms, such as by the ideal gas law. The behavior is a strictly emergent property that can be understood and evaluated in aggregate, while the properties of the individual atoms are not enumerable—nor is that necessary to understanding.

Polygenic theory is an ideal gas law for biology. When enough genes contribute to a state, each treated as a replicate miniscule causal unit, their aggregate combination generates a net phenotype. By analogy with temperature, each organism is a "gas" produced by its set of genomic variants (and similarly at higher levels, are the phenotype distributions among individuals in a population or of species in an ecosystem).

However, this analogy is misleading for biology because, unlike a gas, the individual actors—*e.g.*, allelic variants among individuals—do not have inherent, much less uniform, causal properties of their own. All functional genomic elements may be structurally alike in that they involve DNA sequence but vary as to their specific effects due to intrinsic as well as contextual differences. Every gene, and the phenogenetics of every trait, is different. Even within the lifetime of each individual organism, somatic mutation introduces a phylogeny of genome-wide variation among cells. An uncritically applied polygenic model makes a remarkable assumption of context-free or even deterministic point causation by uniform causal elements.

DNA is basically an inert molecule, and in addition to the structural heterogeneity of its functional elements, each element works only in its contemporary environment, but that had to be established previously: a 4-billion-year regress of inheritance from earlier cells that can play havoc with Enlightenment, indeed Aristotelian, concepts of causation. Whether the purpose of science is to understand life or to pragmatically manipulate nature with predictable effects based on that understanding, it is clear that the intrinsic uncertainty of nature can stymie either endeavor.

"Organs Now of Trifling Importance" (Darwin 1859): Must Causation Be Significant?

If we want to understand something today in development, in homeostasis, or in inferring evolution, we have to cut the causal chain and identify some chosen starting point—a state at which a cause exists (or existed) whose effect we want to understand. But to demonstrate that, the cause must be—or must be assumed to be—replicable.

There are various ways in which genes (functional spots in the genome) become candidates for causal evaluation. Sometimes knowledge of the biological nature of a phenotype suggests a particular protein that may play a role in the trait's development, maintenance, or evolution. Alternatively, these days agnostic omics (fishing expedition, or hypothesis free) approaches are commonly taken; such approaches deliberately and nondiscriminately identify all proteins expressed in a given tissue, or use genome-spanning polymorphic markers to find regions of the genome whose variation is statistically, and presumably causally, associated with variation in some disease or other phenotype of interest.

The challenging epistemology of omics inference can be exemplified by the problem of identifying the causal basis of traits in the GWAS context. Implicit in GWAS approaches is the idea that the genomic architecture of a trait involves some unspecified but tractable number of genomic regions (coding, regulatory, or whatever). The evidence is probabilistic in several senses. The data inevitably compose only a small fraction of individuals in a population or species. Each contributing factor is assumed to have its own inherent causal effect, treated as a risk, penetrance probability, or an effect expected in some meaningfully probabilistic way. One must select some significance criterion for making inferential decisions.

In this kind of complex probabilistic environment, the scientific method makes inferential decisions that are statistical. Debates about the best kind of statistics (likelihood, frequentist, Bayesian) notwithstanding, the decision is ultimately a subjective one based on controlled ignorance: what we think is good-enough evidence. That we have conventions, like the $P = 0.05$ threshold, helps keep order in the henhouse. But it does not keep the fox out.

In GWAS and other similar omics approaches, we essentially confess *a priori* agnosticism about the specifics. However, the very reason for the move to agnostic science is that life is, in fact, complex. If, for the moment, we suspend skepticism and accept the notion that genomic elements have inherent causal effects and the empiricist belief that such things, if real, must be ascertainable, then it requires only altered study design to get the job of identifying these genomic elements done. But this puts us, not the fox, in a trap.

We can adjust for multiple testing by inventing more stringent significance cutoff criteria (Lander and Kruglyak 1995; Storey and Tibshirani 2003) so that what we detect is in some statistical sense more likely to be "real." But these methods do not enable us to account for all the genetic effects as reflected, for example, by heritability estimates. So a next-stage response is to increase sample size, and in this spirit it is becoming common to extrapolate from the distributions of allele frequency and allelic effects in observed samples to estimate the number of contributing loci that would be detected and the fraction of the phenotypic variance or heritability for which they would account (*e.g.*,

Pawitan *et al.* 2009; Lango Allen *et al.* 2010; Park *et al.* 2010; Speliotes *et al.* 2010).

The results consistently suggest that huge sample sizes in the hundreds of thousands would statistically identify hundreds of genome locations—and yet even this may account for only a fraction of the heritability of the trait. Similar findings have been made in animals, plants, and even unicellular organisms. And it is no surprise that we will be moving not just to much larger samples but from a mere million SNP markers to whole-genome sequence data on the basis of the implicit assumption of infinitesimal force-like causation, and that signal-to-noise ratios will be well behaved. This is pure Enlightenment-derived thinking based on significance criteria. But what if we used some other criterion?

A technique that is becoming popular for optimizing classification in a sample is to estimate the receiver operator characteristics (ROC) (Hanley and McNeil 1982) that, rather than using an *a priori* significance cutoff, systematically adjusts some classifying cutoff criterion until a set of observed individuals are optimally assigned to their proper category (such as case or control), based on a judgment as to an acceptable trade-off between false-positive and false-negative classifications. The trade-off distribution is called the area under the ROC curve and shows the efficacy of each possible value of the classificatory criterion, relative to the measure providing no information (Janssens *et al.* 2006; Evans *et al.* 2009; Wray *et al.* 2010). A conceptually similar regression approach to quantitative traits can be used to identify the set of mapping markers that maximizes predictive power (Lango Allen *et al.* 2010; Yang *et al.* 2010). Yet the choice of an appropriate cutoff value remains a subjective judgment.

At its root, this approach is an effort to stay in the shadow of standard Enlightenment-based concepts and reductionism, with the idea that we will eventually be able to enumerate all the contributing causes, with a tacit underlying assumption that this gets asymptotically to the biological truth, not just a description of our sample. That assumption is required if the result is going to be deductive or predictive with knowable precision.

This is the Stubbornist school of statistical inference, but it may not outfox the fox, because even this fix is largely illusory. At every SNP location, in any sample, the alternative alleles will be associated with different mean trait values—even if just by chance. But what does "just by chance" mean in actual causal terms? Indeed, our idea of whether an effect is real or not is rather circular if we define "real" in *P*-value terms.

Unless our understanding of genetics is much worse than it seems, the heritability in any given data set reflects the genome-wide allelic contributions *in those data*. Appropriately assessed genome-wide sequence or genotype data from the same individuals may plausibly be able to account for most of that heritability (*e.g.*, Pawitan *et al.* 2009). At the same time, if the bulk of heritability seems to be due to

polygenic effects, with many different genotypes able to generate equivalent genotypes, then prediction from most individual alleles will be essentially as useless as predicting gas pressure from a given atom. Whatever their individual effect in this sample, if most effects are trivially small and rare, then the set that will be found in the *next* sample will be different in presence, frequency, and effect of alleles. In this context, the criterion of replicability is problematic, except on an approximate basis for the fraction of larger, sufficiently frequent effects.

Because genetic architecture today is the product of yesterday, the same issues apply with even greater cogency to understanding, not to mention reconstructing evolution. We have plenty of data to support Darwin's general assertion that "We see nothing of these slow changes in progress, until the hand of time has marked the long lapse of ages, and then so imperfect is our view into long past geological ages, that we only see that the forms of life are now different from what they formerly were" (Darwin 1859).

Allelic effects on the phenotype that we have just seen are elusive and more numerous than the stars in heaven and have evolutionary importance only if they have individual effects on fitness, which is a degree of removal from merely having phenotypic effects. In this connection, even Darwin was a more perceptive triple-helix (Lewontin 2000) man than many are today: "When a variation is of the slightest use to a being, we cannot tell how much of it to attribute to the accumulative action of natural selection, and how much to the conditions of life." (Darwin 1859).

This suggests selective neutrality, but even that baseline notion of non-Darwinian evolution—s=0.00000, unvarying from zero—is a rather subjective if not metaphysical cutoff value when it comes to actual testing, which once again essentially defines reality on the basis of significance criteria. But even under strong selection, fitness may be distributed across so many contributing sites in a genome that it cannot be detected statistically unless we simply define any allele-frequency change to be due to selection.

The complexity of nature and the preponderance of very small, ephemeral, and context-specific causal effects raise a disturbing question: Need causation in life be accessible by significance-testing criteria? Small or rare effects might, in natural populations or samples, simply not be able to achieve persuasive statistical significance at any time or even at every time, even in principle. Their effects may be inherently probabilistic, and nonrepeatable context dependency may be their most important characteristic. These various facts point to the limits of our ability to connect genotypes to phenotypes (Clark 2000; Weiss 2008), a lesson that we think needs to be more soberly absorbed. Suppose that we were to recognize more formally that significance is simply not always a meaningful criterion for truth, but rather an illusion driven by Enlightenment-derived concepts of infinitesimal force-like causation.

According to Herodotus, before invading Greece in 480 BC, the Persian King Xerxes looked upon his magnificently

massed army of 100,000. Tears fell from his eye, as he observed that "of all this host, so numerous as it is, not one will be alive when a hundred years are gone by" to remember it. Yet there will still be a Persian army. Lives, genes, cells, and species alike are temporary embodiments of a more fluid continuity. Ideas of significance or enumerability in genomic causation and evolution are rather static, but life is perhaps better understood in terms of dynamic flux (or should we say *fluxion*, in Newton's honor?). During its lifetime, an organism is composed of a flux of biomolecules (sugars, proteins, even DNA) entering, changing, and leaving its various cells. Likewise, evolution is a flux of genomic causal elements entering by way of mutation and changing via population genetic processes. Individuals are connected by the genomic flux from parent to offspring or horizontally by infection. Flux involves individual molecules but life is an aggregate phenomenon, and this is why heritability may be in some ways as useful as enumerative genomic prediction (Aulchenko *et al.* 2009).

In this sense, phenotypes may be more causally real than their ephemeral genomic make-up (Weiss and Fullerton 2000), whose individual elements may be essentially inaccessible by current scientific criteria on the basis of criteria such as statistical significance. This makes it difficult, at the most fundamental causal level, to test data against what we think might be the laws of life.

But *are* there laws of life?

Ad hoc Ephemeral Probabilism? "An Extremely Intricate Subject" (Darwin 1859)

In the Enlightenment spirit, Darwin recognized that "all living things have much in common, in their chemical composition" (Darwin 1859). His contemporary Herbert Spencer, the polymath responsible for the phrase "survival of the fittest," believed that the universe followed one set of laws of matter and energy, which he extended to life and, using thermodynamic principles, even to the mind (Elliot 1917; Weiss 2010). In that sense, we are all just the Krebs cycle writ large (Morowitz *et al.* 2000; Smith and Morowitz 2004). But Darwinian thinking went beyond that because shared metabolism does not account for the diversity of life and its evolution, which create many emergent levels of organization that, beyond the assemblage of their parts, one might quip, have a life of their own.

In modern terminology, living diversity is based on information carried in heterogeneous polymeric molecules (Kay 2000; Morowitz *et al.* 2000; Weiss 2004; Weiss and Buchanan 2009b). The polymeric nature of nucleic acids and polypeptides enables function in DNA, RNA, and proteins to depend fundamentally and with little prior constraint on the number, identities, order, spacing, and arrangement of the components. Life is a Boolean ("logical") emergent elaboration of core metabolism whose key feature is combinatorial interaction of elements in time and space from the order of nucleotides along the genome to signaling by combinations of signals and receptors, in which, in a very nonlinear way, the information often has meaning only far from the genome, even extracellularly. That is, the various signaling elements interacting at a given location may have been produced by sets of other cells near and far.

The Darwinian law is that life is a process of divergence from common ancestry—that is, descent with modification, which is based on differential proliferation of heritable elements because of their relative ability to compete for limited resources. Both Darwin and Wallace used Malthus as a rationale for their theory of natural selection, a view that widely persists today, but does Malthusian population pressure resemble an ineluctable law of nature, or is it rather a frequently observed condition? Even if overpopulation is a law, it would enable, but not mandate, some genotypes to do better than others for selective reasons. Our lives may be nasty, brutish, and short but one of nature's cruel jests is that most of this may just be due to bad luck. Chance is an ever-present factor in reproductive success and is non-Malthusian. We noted above that purifying selection may often be due to failure to cooperate that need have nothing to do with competition for limited resources.

This is not the usual concept of evolution and suggests that natural selection is not a law the way gravitation is, Darwin's metaphoric analogy at the end of the *Origin* notwithstanding. As we noted earlier, the functional elements in DNA are not identical replicates like oxygen molecules in an ideal gas. It is not clear what is universal about them, beyond the fact that sequence order carries information in *a priori* unspecifiable numbers of ways, which differentially proliferate from a common origin.

Many argue that the Enlightenment idea of the laws of nature is outdated and that modern science no longer deals in such terms. We do not have to believe in natural law just because the founders of the Enlightenment, or Darwin, did. Progress has shown that the physical and living worlds are much more complex than was understood back then. But if the term is no longer fashionable, have we really abandoned the idea?

The resistance to selective neutrality is one of many ways in which laws, including ideas of universal adaptation and genetic determinism, are still widely found as default assumptions. Yet what we are learning is that, on many scales of observation at least, life manifests *ad hoc*, contextual, and ephemeral probabilism without clear-cut underlying distributions. And, if correct, this is not trivial because if there are no universal laws or properties of life—call them what you will—how can that be? Is that not even more mystic than the belief in laws? Indeed, we may not clearly know what the relevant laws are, but in a profound sense, denying their existence is tantamount to believing in miracles—effects without orderly causes.

In a sense similar to the ideal gas law, if the power to predict emergent higher orders of living organization from the fundamentals of life is "essentially zero" (Krakauer

et al. 2011), then what is the value of our relentless reductionism? One can answer this in many ways, but a fundamental reasoning is that because, unlike an ideal gas, a biological trait cannot be explained strictly in terms of homogenous units of quantitative genetics, or by population genetics theory, we use reductionism to seek its *ad hoc* nonhomogenous basis. But that does not imply that the latter can uniquely predict the former, leaving the emergent phenomena in Dante's circle of limbo: virtuous perhaps, but sighing without hope of salvation.

This raises a disturbing question. If life is a phenomenon that obediently follows the physical laws of chemistry, yet there is not a law of life, and if life is a set of statistical principles, but its dynamics need not be statistically significant, then how is life to be understood? Or, put a more practical way, if life is usually causally complex and contextual and does not follow a law, then against what baseline, or what criteria should we anchor our evidence, conclusions, or generalizations?

"How Far More Interesting. . .Will the Study of Natural History Become!" (Darwin 1859): End of an Enlightened Era?

Darwin was not dogmatic. He acknowledged that provisional hypotheses such as his pangenesis theory of inheritance "may often be of service" to science even if they are wrong (Darwin 1859). Darwin himself belongs to the ages, but the long shadow of his Newtonian determinism widely affects biological thinking today, even though we know better and can see why Darwin was right in those specific ways in which he clearly was presciently right. The same applies to Mendel, the other pillar of the modern synthesis: alleles, but not traits, segregate, and allelic effects are usually far from deterministic. The paucity of blockbuster findings from GWAS or searches for signatures of selection are often viewed with some angst or dismissed as being perplexingly negative (or due to as-yet inadequately Herculean data sets), but while this is disappointing relative to hopes, it is in fact very positive evidence about the nature of life.

Current debates over whether phenogenetic control and evolution are due to common or rare variants or how to enumerate them seem fruitless. Moving up one level to networks has been seen as a better way to recognize causal emergence and synthesize complex data. But even this may be substantially illusory. First, the criteria on which networks such as those shown in Figure 1 and countless others are built are vulnerable to the same epistemological issues of significance, replicability, and enumerability: they involve statistical expression cluster analysis, epidemiologic association, heterogeneous animal and cell-culture models, text-association informatic searches, and the like.

In addition, the evidence seems clear that genes (and their effects) are used in multiple partially overlapping networks or that only parts of networks may be involved
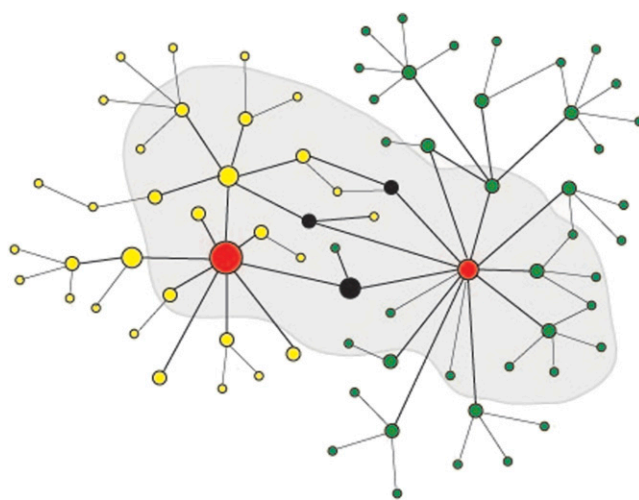

Figure 2 Gene networks are not distinct, discrete causal units. Two networks (yellow and green) are shown schematically, with genes participating in both shown in black. The red genes denote hub genes. The gray region shows the genes that are involved in one or more correlated phenotypes, such as multiple diseases. For a discussion of relevant recent data, see Barabasi *et al.* (2011).

in a given trait or that networks can rewire depending on circumstances (Figure 2; Bennett and Hasty 2008; Kim *et al.* 2010; Barabasi *et al.* 2011). Networks are not discrete causal units after all. Systems thinking may provide sets of genes to use in a kind of Bayesian way to modify or help interpret mapping (*e.g.*, Emily *et al.* 2009) and to help account explicitly for how evolution works to generate complex traits in ways that have been long understood in principle (Bard 2010). But even experimentally controlled approaches have not simplified matters (Chen *et al.* 2008; Emilsson *et al.* 2008; Keller *et al.* 2008; Eleftherohorinou *et al.* 2009; Li *et al.* 2010) in what might appear to be simple classically adaptive traits in flies (Ayroles *et al.* 2009; Harbison *et al.* 2009; Mackay *et al.* 2009) or even in yeast (Zhu *et al.* 2008).

In the long lead-in to announcing his theory, Darwin realized from his study of barnacles that there was "variability of every part in some slight degree of every species" and that "systematic work would be easy were it not for this confounded variation" (Darwin 1850). Today we know that this variation includes DNA sequences.

We earlier alluded briefly to the rather open-ended source of nongenetic contextual effects known generically as "environment." This variation includes lifestyle and other factors, often difficult to measure with satisfactory precision. More seriously, their relationship to genomic effects is necessarily assessed retrospectively, while useful prediction of fitness or personalized genomic medicine is prospective, but future environments are often inherently unpredictable (Buchanan *et al.* 2006; Weiss 2008).

A number of investigators have advocated what is essentially aggregate rather than reductive approaches to life and its complexities, recognizing that natural phenotypes are more layered than in the common gene-centered

view. Although the ideas (like most ideas!) go back to Aristotle, the modern recognition of these issues is usually attributed to the insightful maverick C. H. Waddington, who coined the term "epigenetics" (Waddington 1942; Krakauer *et al.* 2011). The term has taken on new meanings in some usages today, referring to DNA modification, but its broader sense involving genomes and environments is as cogent as ever.

Many other authors, current and past, have attempted to incorporate more ecologically interactive or cooperative systems and/or phenotype- rather than genotype-based views of life (a representative if somewhat arbitrary set of authors is Simon 1962; Riedl 1977; Laubichler 2000; Margulis 2000; Oyama *et al.* 2001; Jablonka and Lamb 2002; Newman and Mueller 2003; Olding-Smee *et al.* 2003; West-Eberhard 2003; Alonso and Wilkins 2005; Newman 2005; Jablonka and Lamb 2006; Leigh 2007; Wagner 2007; Wilkins 2007; Noble 2008; Gilbert and Epel 2009; Mustonen and Lassig 2009; Stadler *et al.* 2009; Nowak and Highfield 2011). There is also an active "complexity" school of thought that tries to relate reductionism to emergent self-organizing higher-order traits (*e.g.*, Krakauer *et al.* 2011).

Even the early geneticists such as Morgan and S. Wright, unfairly to name but two of them, were aware of genetic interactions and phenogenetic complexity. And as had these and other earlier authors, prominent geneticists were raising some of our same kinds of points even when the ability to document genetic complexity in molecular terms was just bursting on the scene (*e.g.*, Monod 1971; Lewontin 1974), the latter invoking a kind of self-organizing, cooperative complexity theory of its time. However, while it is not a problem *per se*, even these more holistic authors provide a variegated and often *ad hoc* picture, rather than a single, over-arching theory—or law—of life.

In the *ad hoc* realities of the biosphere, if our only test is a subjective statistical "has an effect" *vs.* 'has no effect," the result is a grab bag in which one can find whatever one wants. Regardless of one's opinion as to how full or empty the glass of our knowledge is, it is entirely consistent with what we know. We know that there is a fraction of the major effects on traits or fitness that stick out like Mendel's thumb. These causal elements behave very well, such as the genetic basis of pea color or Huntington's disease that do seem to be causal and to segregate in the usual senses. For many purposes, such as drug or crop development, these are "causes that make a difference" (a thoughtful discussion is by Waters 2007), and we can approach them as engineers, with approximate understanding not dependent on theory that may be good enough for many immediate pragmatic objectives.

There are many focal areas in the life sciences, from the cellular to the social, in which a variety of experimental or theoretical modeling approaches, built upon the epistemological kernel of our legacy from the Enlightenment, illuminate some major pattern features of complex systems (Krakauer *et al.* 2011). Such approximations do not get us

entirely out of emergence limbo, but they do satisfy Darwin's pragmatism that "they may often be of service."

Despite these successes, the reality to date does not entirely justify the celebratory hoopla surrounding omics. The fluxion of causal elements is heterogeneous, with some flowing more rapidly than others through evolution or through an organism's life, as they come and go by mutation, duplication, loss, and fixation. Classical genetics took an aggregate rather than enumerative view, although in a rather homogeneous ideal gas-like way that has not generally attempted to deal with the heterogeneous mix of functions and effects that we now know exist. And short-term extrapolation and replication of the stronger factors that are reliably or experimentally identified guide what has for a century been the most successful aspects of our approaches to understanding biological causation (Waters 2007), and good enough for many purposes. Yet "good-enough" is not the same as understanding nature, presumably an important goal of science, because most elements in the flux of life involve small or rare context-specific combinatorial effects that are historically *ad hoc* and heterogeneous or even unenumerable. The omics age is clearly demonstrating that even the narrowly focused pragmatic engineering approach is up against serious barriers, as we have tried to discuss.

The Enlightenment was triggered by technologies like optics that allowed the scale of observation to change, enticing us to empirical observation and controlled experiment, to identify the laws of nature. The 19th and 20th centuries formalized the scientific epistemology of our age in terms of quantitative and, in particular, statistical criteria for assessing evidence in the context of a belief that nature can be explained by infinitesimally assessable laws. We are still being aided (and baited) by continuing technological development, but this has enabled us to see that, in important ways, life might not be law-like in the Enlightenment sense, or even that we may not know when or if we have found such laws (Stanford 2006; Waters 2007). Such realization does not challenge empiricism, but questions current empirical approaches to causation and inference, especially in regard to understanding of the phenomenon of emergence. And that leads to a radical thought: Could it be that, as a result of what technology is revealing, the sun is setting on three centuries of Enlightenment-based science—with genetics leading the way to some new kind of empiricism?

Acknowledgments

Literature Cited

Alonso, C. R., and A. S. Wilkins, 2005 The molecular elements that underlie developmental evolution. Nat. Rev. Genet. 6: 709–715.

Aulchenko, Y. S., M. V. Struchalin, N. M. Belonogova, T. I. Axenovich, M. N. Weedon, et al., 2009 Predicting human height by Victorian and genomic methods. Eur. J. Hum. Genet. 17: 1070–1075.

Ayroles, J. F., M. A. Carbone, E. A. Stone, K. W. Jordan, R. F. Lyman, et al., 2009 Systems genetics of complex traits in Drosophila melanogaster. Nat. Genet. 41: 299–307.

Barabasi, A. L., N. Gulbahce, and J. Loscalzo, 2011 Network medicine: a network-based approach to human disease. Nat. Rev. Genet. 12: 56–68.

Bard, J., 2010 A systems biology view of evolutionary genetics: network-driven processes incorporate much more variation than evolutionary genetics can handle. This variation is hard to formalise but allows fast change. Bioessays 32: 559–563.

Bennett, M. R., and J. Hasty, 2008 Systems biology: genome re-wired. Nature 452: 824–825.

Buchanan, A. V., K. M. Weiss, and S. M. Fullerton, 2006 Dissecting complex disease: The quest for the Philosopher's Stone? Int. J. Epidemiol. 35: 562–571.

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.

Chen, Y., J. Zhu, P. Y. Lum, X. Yang, S. Pinto, et al., 2008 Variations in DNA elucidate molecular networks that cause disease. Nature 452: 429–435.

Clark, A. G., 2000 Limits to prediction of phenotypes from knowledge of genotypes, in Limits to Knowledge in Evolutionary Genetics (Evolutionary Biology, Vol. 32), edited by M. T. Clegg. Springer-Verlag, Berlin.

Darwin, Charles, 1859 On the Origin of Species. John Murray, London.

Darwin, C., 1887 The Autobiography of Charles Darwin. Darwin, F., and N. Barlow (Editors). Collins Clear-Type Press, London.

Darwin, Charles, 1850 Letter from Darwin to J. D. Hooker, June 12, 1850.

Dyson, G., R. Frikke-Schmidt, B. G. Nordestgaard, A. Tybjaerg-Hansen, and C. F. Sing, 2007 An application of the patient rule-induction method for evaluating the contribution of the apolipoprotein E and lipoprotein lipase genes to predicting ischemic heart disease. Genet. Epidemiol. 31: 515–527.

Eleftherohorinou, H., V. Wright, C. Hoggart, A. L. Hartikainen, M. R. Jarvelin, et al., 2009 Pathway analysis of GWAS provides new insights into genetic susceptibility to 3 inflammatory diseases. PLoS ONE 4: e8068.

Elliot, H. S. R., 1917 Herbert Spencer. Books for Libraries Press, Freeport, NY.

Emilsson, V., G. Thorleifsson, B. Zhang, A. S. Leonardson, F. Zink et al., 2008 Genetics of gene expression and its effect on disease. Nature 452: 423–428.

Emily, M., T. Mailund, J. Hein, L. Schauser, and M. H. Schierup, 2009 Using biological networks to search for interacting loci in genome-wide association studies. Eur. J. Hum. Genet. 17: 1231–1240.

Evans, D. M., P. M. Visscher, and N. R. Wray, 2009 Harnessing the information contained within genome-wide association studies to improve individual prediction of complex disease risk. Hum. Mol. Genet. 18: 3525–3531.

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

Gao, L., and J. Zhang, 2003 Why are some human disease-associated mutations fixed in mice? Trends Genet. 19: 678–681.

Gibson, G., 2010 Hints of hidden heritability in GWAS. Nat. Genet. 42: 558–560.

Gilbert, S. F., and D. Epel, 2009 Ecological Developmental Biology: Integrating Epigeneitcs, Medicine, and Evolution. Sinauer Associates, Sunderland, MA.

Goh, K. I., M. E. Cusick, D. Valle, B. Childs, M. Vidal, et al., 2007 The human disease network. Proc. Natl. Acad. Sci. USA 104: 8685–8690.

Hanley, J. A., and B. J. McNeil, 1982 The meaning and use of the area under a Receiver Operating Characteristic (ROC) curve. Radiology 143: 29–36.

Harbison, S. T., M. A. Carbone, J. F. Ayroles, E. A. Stone, R. F. Lyman, et al., 2009 Co-regulated transcriptional networks contribute to natural genetic variation in Drosophila sleep. Nat. Genet. 41: 371–375.

Hartl, D., and A. G. Clark, 2007 Principles of Population Genetics. Sinauer Associates, Sunderland, MA.

Jablonka, E., and M. J. Lamb, 2002 The changing concept of epigenetics. Ann. N. Y. Acad. Sci. 981: 82–96.

Jablonka, E., and M. Lamb, 2006 Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic: Variation in the History of Life. MIT Press, Cambridge, MA.

Janssens, A. C., Y. S. Aulchenko, S. Elefante, G. J. Borsboom, E. W. Steyerberg, et al., 2006 Predictive testing for complex diseases using multiple genes: Fact or fiction? Genet. Med. 8: 395–400.

Kay, L., 2000 Who Wrote the Book of Life? Stanford University Press, Stanford, CA.

Keller, M. P., Y. Choi, P. Wang, D. Belt Davis, M. E. Rabaglia, et al., 2008 A gene expression network model of type 2 diabetes links cell cycle regulation in islets with diabetes susceptibility. Genome Res. 18: 706–716.

Kim, M. S., J. R. Kim, and K. H. Cho, 2010 Dynamic network rewiring determines temporal regulatory functions in Drosophila melanogaster development processes. BioEssays 32: 505–513.

Kondrashov, A. S., S. Sunyaev, and F. A. Kondrashov, 2002 Dobzhansky-Muller incompatibilities in protein evolution. Proc. Natl. Acad. Sci. USA 99: 14878–14883.

Krakauer, D. C., J. P. Collins, D. Erwin, J. C. Flack, W. Fontana, et al., 2011 The challenges and scope of theoretical biology. J. Theor. Biol. 276: 269–276.

Lander, E., and L. Kruglyak, 1995 Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. Nat. Genet. 11: 241–247.

Lango Allen, H., K. Estrada, G. Lettre, S. I. Berndt, M. N. Weedon, et al., 2010 Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467: 832–838.

Laubichler, M. D., 2000 The organism is dead. Long live the organism! Perspect. Sci. 8: 286–315.

Leigh, E. G. Jr, 2007 Neutral theory: a historical perspective. J. Evol. Biol. 20: 2075–2091.

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

Lewontin, R. C., 2000 The Triple Helix. Harvard University Press, Cambridge, MA.

Li, P., L. Ponnala, N. Gandotra, L. Wang, Y. Si, et al., 2010 The developmental dynamics of the maize leaf transcriptome. Nat. Genet. 42: 1060–1067.

Lopez Herraez, D., M. Bauchet, K. Tang, C. Theunert, I. Pugach, et al., 2009 Genetic variation and recent positive selection in worldwide human populations: evidence from nearly 1 million SNPs. PLoS ONE 4: e7888.

Lynch, M., 2007a The frailty of adaptive hypotheses for the origins of organismal complexity. Proc. Natl. Acad. Sci. USA 104 (Suppl. 1): 8597–8604.

Lynch, M., 2007b The Origin of Genome Architecture. Sinauer Associates, Sunderland, MA.

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

Manolio, T. A., 2010   Genomewide association studies and assessment of the risk of disease. N. Engl. J. Med. 363: 166–176.

Margulis, L., 2000   *Symbiotic Planet: A New Look at Evolution*. Basic Books, New York.

Mattick, J. S., R. J. Taft, and G. J. Faulkner, 2010   A global view of genomic information: moving beyond the gene and the master regulator. Trends Genet. 26: 21–28.

Monod, J., 1971   *Chance and Necessity*. Alfred A. Knopf, New York.

Morowitz, H. J., J. D. Kostelnik, J. Yang, and G. D. Cody, 2000   The origin of intermediary metabolism. Proc. Natl. Acad. Sci. USA 97: 7704–7708.

Mustonen, V., and M. Lassig, 2009   From fitness landscapes to seascapes: non-equilibrium dynamics of selection and adaptation. Trends Genet. 25: 111–119.

Newman, S. A., 2005   The pre-Mendelian, pre-Darwinian world: shifting relations between genetic and epigenetic mechanisms in early multicellular evolution. J. Biosci. 30: 75–85.

Newman, S. A., and G. B. Mueller, 2003   *Origination of Organismal Form: Beyond the Gene in Developmental and Evolutionary Biology*. MIT Press, Cambridge, MA.

Noble, D., 2008   Genes and causation. Philos. Transact. A Math. Phys. Eng. Sci. 366: 3001–3015.

Nowak, M. A., and R. Highfield, 2011   *SuperCooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. Free Press, New York.

Olding-Smee, F. J., M. W. Feldman, and K. N. Laland, 2003   *Niche Construction: The Neglected Process in Evolution*. Princeton University Press, Princeton, NJ.

Oyama, S., P. E. Griffiths, and R. D. Gray (Editors), 2001   *Cycles of Contingency: Developmental Systems and Evolution*. MIT Press, Cambridge, MA.

Park, J. H., S. Wacholder, M. H. Gail, U. Peters, K. B. Jacobs, *et al.*, 2010   Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. Nat. Genet. 42: 570–575.

Pawitan, Y., K. C. Seng, and P. K. Magnusson, 2009   How many genetic variants remain to be discovered? PLoS ONE 4: e7969.

Pritchard, J. K., J. K. Pickrell and G. Coop, 2010   The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. Curr. Biol. 20: R208–R215.

Riedl, R., 1977   A systems-analytical approach to macro-evolutionary phenomena. Q. Rev. Biol. 52: 351–370.

Sabeti, P. C., P. Varilly, B. Fry, J. Lohmueller, E. Hostetter, *et al.*, 2007   Genome-wide detection and characterization of positive selection in human populations. Nature 449: 913–918.

Simon, H., 1962   The architecture of complexity. Proc. Am. Philos. Soc. 106: 467–482.

Sing, C. F., J. H. Stengard, and S. L. Kardia, 2003   Genes, environment, and cardiovascular disease. Arterioscler. Thromb. Vasc. Biol. 23: 1190–1196.

Sing, C. F., J. H. Stengard, and S. L. Kardia, 2004   Dynamic relationships between the genome and exposures to environments as causes of common human diseases. World Rev. Nutr. Diet. 93: 77–91.

Smith, E., and H. J. Morowitz, 2004   Universality in intermediary metabolism. Proc. Natl. Acad. Sci. USA 101: 13168–13173.

Speliotes, E. K., C. J. Willer, S. I. Berndt, K. L. Monda, G. Thorleifsson, *et al.*, 2010   Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. Nat. Genet. 42: 937–948.

Stadler, P. F., S. J. Prohaska, C. V. Forst, and D. C. Krakauer, 2009   Defining genes: a computational framework. Theory Biosci. 128: 165–170.

Stanford, P. K., 2006   *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press, New York.

Storey, J. D., and R. Tibshirani, 2003   Statistical significance for genomewide studies. Proc. Natl. Acad. Sci. USA 100: 9440–9445.

Waddington, C. H., 1942   Canalization of development and the inheritance of acquired characters. Nature 150: 563–565.

Wagner, G. P., 2007   The developmental genetics of homology. Nat. Rev. Genet. 8: 473–479.

Waters, C. K., 2007   Causes that make a difference. J. Philos. 104: 551–579.

Weiss, K. M., 2004   The phenogenetic logic of life. Nat. Rev. Genet. 6: 36–45.

Weiss, K. M., 2008   Tilting at quixotic trait loci (QTL): an evolutionary perspective on genetic causation. Genetics 179: 1741–1756.

Weiss, K. M., 2010   Shopping for answers from Marx and Spencer. Evol. Anthropol. 19: 168–173.

Weiss, K. M., 2011   Seeing the forest through the gene-trees. Evol. Anthropol. 19: 210–221.

Weiss, K. M., and A. V. Buchanan, 2009a   The cooperative genome: organisms as social contracts. Int. J. Dev. Biol. 53: 753–763.

Weiss, K. M., and A. V. Buchanan, 2009b   *The Mermaid's Tale: Four Billion Years of Cooperation in the Making of Living Things*. Harvard University Press, Cambridge, MA.

Weiss, K. M., and S. M. Fullerton, 2000   Phenogenetic drift and the evolution of genotype-phenotype relationships. Theor. Popul. Biol. 57: 187–195.

Weiss, K. M., and J. D. Terwilliger, 2000   How many diseases does it take to map a gene with SNPs? Nat. Genet. 26: 151–157.

West-Eberhard, M. J., 2003   *Developmental and Evolutionary Plasticity*. Oxford University Press, New York.

Wilkins, A. S., 2007   Between "design" and "bricolage": genetic networks, levels of selection, and adaptive evolution. Proc. Natl. Acad. Sci. USA 104(Suppl. 1): 8590–8596.

Wray, N. R., J. Yang, M. E. Goddard, and P. M. Visscher, 2010   The genetic interpretation of area under the ROC curve in genomic profiling. PLoS Genet. 6: e1000864.

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: 565–569.

Zhu, J., B. Zhang, E. N. Smith, B. Drees, R. B. Brem, *et al.*, 2008   Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. Nat. Genet. 40: 854–861.