Cell
PRESS

The many faces of pleiotropy

Annalise B. Paaby and Matthew V. Rockman

Department of Biology and Center for Genomics & Systems Biology, New York University, 12 Waverly Place, New York, NY 10003, USA

Pleiotropy is the well-established phenomenon of a single gene affecting multiple traits. It has long played a central role in theoretical, experimental, and clinical research in genetics, development, molecular biology, evolution, and medicine. In recent years, genomic techniques have brought data to bear on fundamental questions about the nature and extent of pleiotropy. However, these efforts are plagued by conceptual difficulties derived from disparate meanings and interpretations of pleiotropy. Here, we describe distinct uses of the pleiotropy concept and explain the pitfalls associated with applying empirical data to them. We conclude that, for any question about the nature or extent of pleiotropy, the appropriate answer is always 'What do you mean?'.

Multiple traits, multiple meanings

Pleiotropy refers to the phenomenon of a single gene affecting multiple traits and, as that manifestly ambiguous definition reveals, pleiotropy is a single word with multiple meanings. As biologists in different disciplines turn new genomic techniques toward questions of pleiotropy, the need to discriminate carefully between different meanings of the concept has become imperative.

Here, we first review the alternative meanings of pleiotropy and point to situations where they define unrelated phenomena. We then discuss the nature and quality of empirical data brought to bear on questions about pleiotropy in its many guises. Finally, we focus on how better to frame pleiotropy research, underscoring the continuing importance of this slippery concept.

What is pleiotropy?

At its essence, pleiotropy implies a mapping from one thing at the genetic level to multiple things at a phenotypic level. The natures of the things differ in different contexts. We identify three broad classes of question that use the word 'pleiotropy' to refer to different kinds of one-to-many maps. The classifications are not mutually exclusive (indeed, they typically overlap), but they do represent different and potentially contradictory perspectives.

In 'molecular gene pleiotropy', the question is about the number of functions a molecular gene has. These functions can be defined not only genetically, by the number of measured traits affected by a knockout, but also biochemically, for example, by the number of protein–protein interactors a gene has or the number of reactions it catalyzes.

In 'developmental pleiotropy', mutations rather than molecular genes are the relevant units. Developmental

pleiotropy is a feature of the genotype–phenotype map that defines the genetic and evolutionary autonomy of aspects of phenotype, independent of fitness. This is the mutational pleiotropy underlying the diverse manifestations of syndromic diseases, the ontogenetic pleiotropy that underlies classical questions about allometry and heterochrony, and the molecular pleiotropy that underlies questions about relative importance of *cis*-regulatory versus protein-coding variants.

In 'selectional pleiotropy', the question is about the number of separate components of fitness a mutation affects. In some cases, the multiple fitness components are life-history traits of a single individual, which is at the heart of the antagonistic pleiotropy model for the evolution of aging [1]. In other cases, the mutational effects manifest in different individuals in a population, which is the basis for sexually antagonistic pleiotropy and pleiotropic trade-offs underlying local adaptation. A key feature of selectional pleiotropy is that traits are defined by the action of selection and not by the intrinsic attributes of the organism.

Recent years have seen a surge of genome-scale analyses of empirical data designed to address questions about the scope of pleiotropy, and particularly to test hypotheses of universal pleiotropy (Boxes 1–3). However, given the incongruence between molecular gene, developmental, and selectional pleiotropy, data that reflect on pleiotropy in one setting may be uninformative or positively misleading in another. When considering the relevance of data to each of these classes of pleiotropy, four issues are critical.

First, are we discussing the genotype–phenotype map, or the genotype–fitness map? Instances of selectional pleiotropy can involve mutations that have only one molecular function or phenotypic effect but impact multiple aspects of fitness. For example, a mutation that changes the rate of organism metabolism may affect early- and late-life reproduction differently despite the unitary molecular mechanism, and a mutation that changes expression of a transporter on a yeast cell may be beneficial in one environment and deleterious in another, despite the total absence of any multiplicity in the genotype–phenotype map (Figure 1). A mutation with identical molecular and developmental effects in males and females is a typical cause of sexually antagonistic pleiotropy, yet it is the antithesis of molecular gene pleiotropy.

Second, are we discussing a molecular gene or a mutation? Failure to distinguish between effects of a mutation and functions of a gene is at the heart of much confusion [2]. From an evolutionary perspective, the molecular gene

---

Box 1. What, if anything, is universal pleiotropy?

The Hypothesis of Universal Pleiotropy has several different meanings. Neither Fisher [55] nor Wright [13], who are typically cited as its advocates (Box 2), actually endorsed the most literal interpretation, that is, that every mutation affects every trait [the Strong Hypothesis of Universal Pleiotropy (SHUP)]. The SHUP draws credence from consideration of the unified network architecture of the cell: all molecular genes compete for a common pool of polymerase, ribosomes, and nucleotides, among other things. Mutations that alter gene activity (or even just sequence) impinge on this central nexus and so alter activities of all other genes. This principle finds its clearest endorsement in the writing of Kacser and Burns in the paper that defines much of modern systems biology (p. 649, [56]):

> We have already remarked that very low coefficients will often be attached to enzymes kinetically ''distant'' from the particular flux considered. This is the reason for the well known observation that mutations in a different part of the map may have no detectable effect on one character even when another character is seriously affected. This apparent independence of most characters makes simple Mendelian genetics possible, but conceals the fact that there is universal pleiotropy. All characters should be viewed as ''quantitative'' since, in principle, variation anywhere in the genome affects every character.

The SHUP resembles models of diffuse coevolution or competition, wherein phenomena arise because everything affects everything, but not much and not to the exclusion of a smaller number of larger and more individually significant relations.

Critiques of the SHUP point out that trivial effects mediated by shared metabolite pools are biologically meaningless, and detectable pleiotropy is empirically restricted and perhaps highly structured, as proposed under the Hypothesis of Modular Pleiotropy (HMP) [57]. Under the HMP, mutations are generally constrained to affect traits that are part of the same functional complex, and this hypothesis has been invoked to explain modular relations observed between sets of genes and sets of traits [15]. As Kacser and Burns' description of universal pleiotropy makes clear, however, SHUP is a logical deduction consistent with the absence of detectable effects. It can be both true and inconsequential. The SHUP may be paraphrased: 'there is a single gene network'.

Box 2. Wright, Fisher, and the Weak Hypothesis of Universal Pleiotropy

Sewall Wright, who approached pleiotropy from the dual perspectives of population genetics and physiological genetics, drew on network architecture to advocate what we might call the Weak Hypothesis of Universal Pleiotropy (WHUP). The text of the locus classicus is worth quoting (pp. 59–60 [13]):

> There are a number of broad generalizations that follow from this netlike relationship between genome and complex characters. These are all fairly obvious but it may be well to state them explicitly.
> 1. The variations of most characters are affected by a great many loci (the multiple factor hypothesis).2. In general, each gene replacement has effects on many characters (the principle of universal pleiotropy).

In his other writings, Wright was careful to qualify universal pleiotropy, as in [58], which refers variously to 'practically', 'nearly', and 'almost universal pleiotropy'. Consequently, Wright should be cited as an advocate for the importance and ubiquity of pleiotropy, but not in any literal sense for its universality.

The other common source for claims of universal pleiotropy is R.A. Fisher's geometric model of adaptation [55], although this toy phenotype-fitness model makes no strong claim about the empirical structure of the genotype–phenotype map; it is simply a model to motivate analysis. As usually explicated, the model holds that mutations are random vectors in a phenotype space, which means that they may affect multiple phenotypes without excluding the possibility that they affect only subsets; indeed, Fisher's WHUP is sometimes described as the notion that a mutation 'can potentially affect all phenotypic traits' (e.g., [59,60]). Most importantly, the axes of Fisher's model are orthogonal traits under selection, with distances defined in fitness units, not simple morphological or life-history traits. Connecting Fisher's model to developmental pleiotropy depends on mapping the axes of selection onto measurable phenotypes, which is not trivial. Most such efforts have suggested that there are just a handful of selectable axes defined by linear combinations of measureable traits [42,49,61,62]. This is why efforts to estimate the dimensionality of organisms (i.e., the number of axes in Fisher's model) yield unintuitive results, on the order of two for worms, flies, and cows [47,63]. In other words, there is no intrinsic relation between the axes of Fisher's model and the genotype–phenotype map for measurable characters. Thus, Fisher's WHUP is the notion that mutations usually affect multiple aspects of fitness, without any claims about the number of affected developmental phenotypes.

---

is inconsequential. Mutations may affect only specific aspects of the functions of a gene, so that pleiotropic molecular genes may be subject to less-pleiotropic mutations [2]. At the other end of the spectrum, mutations that alter multiple molecular genes, whether by deleting large regions, hitting shared regulatory regions, or altering operon structure, are authentic sources of developmental and selectional pleiotropy. Even mutations that do not directly alter any molecular genes can be pleiotropic, by titrating limiting regulators; such mutations could be changes in DNA copy number of protein-binding motifs, for example, or structural changes that alter chromatin architecture. Thus, molecular gene pleiotropy may be uninformative about developmental and selectional pleiotropy, and vice versa.

Third, how are we enumerating traits? Mutations can have 'the same' or 'different' effects at the level of individual cells, tissues, organs, individuals, or populations. Is a mutation pleiotropic if it acts in a tissue that occurs in multiple parts of the body, but does so in the same molecular way in each part? Answers to this kind of question form the basis for elaborate taxonomies of molecular gene pleiotropy [3–5]. Things get further complicated when we consider traits that cannot be enumerated, such as function-valued traits,

particularly ontogenetic trajectories or traits defined as relations between measureable traits. Is a mutation more pleiotropic if it affects two traits uniformly, preserving the proportions between them, or if it affects one trait only, altering their relative sizes? In other realms of evolutionary biology, 'multiple traits' has yet another dimension, as we consider the effects of a mutation in different individuals or populations, including different sexes, environments, and genetic backgrounds. These phenomena are commonly treated as pleiotropy, but they are also instances of genotype-by-sex, -environment, or -genotype interactions, which are the subject of an independent theoretical tradition (Figure 1e). Moreover, pleiotropic effects within individuals have different evolutionary implications compared with pleiotropic effects among individuals (e.g., [6]).

Finally, what do we mean when we say that a gene or mutation 'affects' multiple traits? Here, discussion has centered on the molecular basis for pleiotropy and its biological implications. Many authors [3–5,7–11] have distinguished between multiple independent effects of a mutation and multiple effects that depend on one another

---

Box 3. How can we test the hypothesis of universal pleiotropy?

Given its slipperiness, universal pleiotropy poses problems for the conventional hypothesis-testing framework. Advocates of the HMP (Box 1) [64] argue that universal pleiotropy is untestable and that a hypothesis of 'no gene effect' is the appropriate experimental null model. That the physiological SHUP of Kacser and Burns (Box 1) is untestable is clear from its description; it is an ontological assertion and not a model at all. However, a WHUP (Box 2) tailored to measureable traits (i.e., for questions about developmental pleiotropy), with each mutation potentially affecting each trait according to a distribution of effect sizes (including zero), can provide quantitative predictions against which alternative models, such as the HMP, can be tested. Differences between the models in predictions about the multivariate distribution of effect sizes across traits may provide a statistical toehold [65]. In this framework, because the HMP can be treated as a submodel of the WHUP, likelihood ratio tests hold promise, as do Bayesian frameworks [66].

At the same time, alternative models of pleiotropy make predictions beyond effect detection patterns. Given the acknowledged deficiencies of QTL mapping, particularly its low power, quantitative genetic approaches that are not limited to large-effect loci might be more informative about the global structure of pleiotropy. These include the artificial selection methods described elsewhere in the paper [42,46]. Particularly in the evolutionary context that Fisher and Wright were working in, these selection-based tests for the extent of pleiotropy are better tests of the WHUP than are underpowered QTL studies of arbitrary morphological traits.

---

in a cascade (Figure 2). In the functional view of molecular gene pleiotropy, only the former is meaningful, whereas the latter has been termed 'spurious' [3]. Other authors have named these distinctions 'horizontal' and 'vertical' pleiotropy [9], or 'mosaic' and 'relational' [4] (see [10]). In the context of statistical inference, these modes are 'independent' and 'reactive' [12]. Wright [13] noted that 'from the standpoint of population genetics, these distinctions are, however, unimportant. There is pleiotropy if two effects contribute differently to selective value' (p. 61). Yet, the distinction between authentic/horizontal/mosaic/ independent pleiotropy on the one hand and spurious/ vertical/relational/reactive pleiotropy on the other hand may hold evolutionary relevance, insofar as the latter scenario produces phenotypic correlations between traits even in the absence of genetic variation (e.g., [14]).

The data: how much pleiotropy is there and how do we know?

Over the past few years, it has become possible to generate empirical data sets that can, in principle, estimate genome-wide levels of pleiotropy. The major sources of data are reverse-genetic screens and quantitative trait locus (QTL) mapping studies, and the results have been synthesized recently with the interpretation that 'in all cases the pleiotropy distribution is largely L-shaped, with a low median degree of pleiotropy' [11] (Figure 3). Even though these findings are surprisingly consistent [10], their interpretation is contentious (Box 3). Major concerns include statistical power (how do we declare that a gene has no effect on a trait?), trait delineation (what counts as a separate trait?), mutation spectrum (are we considering spontaneous mutations, gene deletions, or segregating variation?), and model structure (what should we expect to see under different hypotheses?).

Reverse genetics

Genome-wide reverse-genetic screens, in which phenotypes are measured in single-gene null mutants one by one, form the major class of experimental result. In an influential systematic analysis of published large-scale data sets from yeast and worms (and one less systematic data set from mice), a recent study [15] concluded that relatively low proportions of genes, when disrupted, affected multiple phenotypes, and even then, the number of affected traits was small. Here, we briefly consider each of the data sets.

The most comprehensive surveys were performed in yeast, in which a deletion library was used to measure both morphological traits (developmental pleiotropy) and growth rate phenotypes (selectional pleiotropy). Of 4718 gene deletion mutants tested for 254 morphological characters, more than one-third had statistically significant effects on two or more traits, which the original authors interpreted as widespread pleiotropy [16]. When assayed for growth rates across 21 different environments, however, most mutants showed phenotypes in zero or one environment, and only 5% produced phenotypes in three or more environments [17]. Notably, only differences in growth rate between normal and experimental conditions counted as phenotypes, such that a mutant growing slowly in all environments except the control was highly pleiotropic, but a universally slow-growing mutant had zero pleiotropy. This method is literally scoring gene-by-environment interactions for growth rate. These interactions underlie antagonistic pleiotropy, but synergistic pleiotropy (as in the case of a mutation that is deleterious in every environment) does not register. In addition, it is worth re-emphasizing that pleiotropy measured across environments can result from mutations that affect just a single mechanism or phenotype at the molecular or developmental levels (Figure 1f).

It is also important to note that these screens excluded lethal mutations, of which there are 1105 in the yeast deletion collection [18]. If molecular genes are the unit of analysis, logic requires including essential genes and, under any definition of developmental or selectional pleiotropy, these are maximally pleiotropic: any two traits will both differ between wild type and mutant. Inclusion of this large class of genes substantially alters the distribution of pleiotropy from L shaped to U shaped (Figure 3). It might be argued that lethal alleles are irrelevant to developmental or selectional pleiotropy, but the contribution of small-effect mutations in essential genes to evolution and development is simply unknown; the relevant data do not exist. At the same time, some of the best-known mutations underlying evolution of development are mutations in essential genes that only affect subsets of the activities of the gene; in sticklebacks, paired-like homeodomain 1 (*pitx-1*) and *Kit ligand* (*kitlg*) are essential genes that underlie the evolution of pelvic reduction and pigmentation [19,20], and in *Drosophila*, mutations in *ultrabithorax* underlie the evolution of trichome patterns [21].

Also excluded from these pleiotropy estimates were 'genes that do not affect any trait and traits that are not affected by any gene' [15]. From the molecular gene pleiotropy perspective, genes that have no detectable effect on
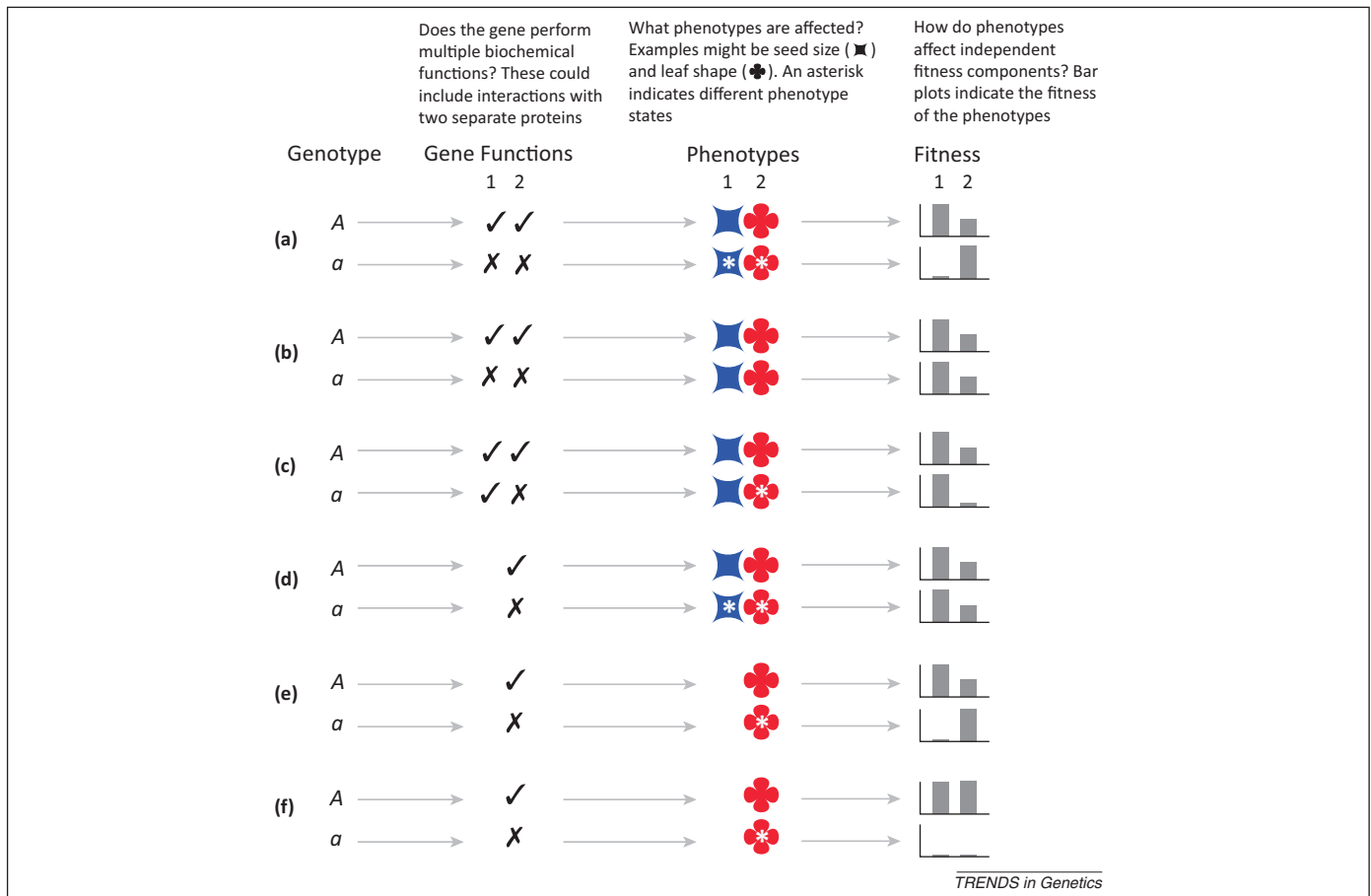
Figure 1. Different meanings of pleiotropy denote unrelated biological phenomena. Here, considering a single locus with alternate alleles (*A* and *a*) in a haploid organism, we represent genotype–gene function–phenotype–fitness mappings for alternate alleles that illustrate a few of the possible relations. (a) Pleiotropy by all accounts. A loss-of-function mutation in a multifunctional protein abolishes two molecular activities that affect two different morphological phenotypes that in turn affect two independent aspects of fitness. (b) Molecular gene pleiotropy occurs when a gene product carries out multiple independent biochemical functions. A gene deletion may abolish both functions, but if they are both redundant there may be no measureable phenotype (in this genetic background). (c) Genes versus mutations: a mutation in a multifunctional gene may affect only one function, in which case this instance of molecular gene pleiotropy does not correspond to any developmental or selectional pleiotropy. (d) Developmental pleiotropy occurs when a mutation affects two aspects of phenotype. This phenomenon does not require any multifunctionality at the level of the molecular gene or any particular pattern of impact on fitness. (e) A mutation that affects only a single phenotype may nevertheless yield selectional pleiotropy if the phenotype has varied effects on fitness in different circumstances (different environments or sexes, for example). (f) Perhaps the least-intuitive case is a mutation that affects only a single phenotype and has identical effects on orthogonal fitness components. Because these fitness components are independent traits in the selectional context, a mutation with uniform effects on more than one is necessarily pleiotropic.

any measured phenotype are an interesting class with undefined pleiotropy. From the developmental and selectional pleiotropy perspectives, these genes simply do not register as mutations. At the same time, their abundance
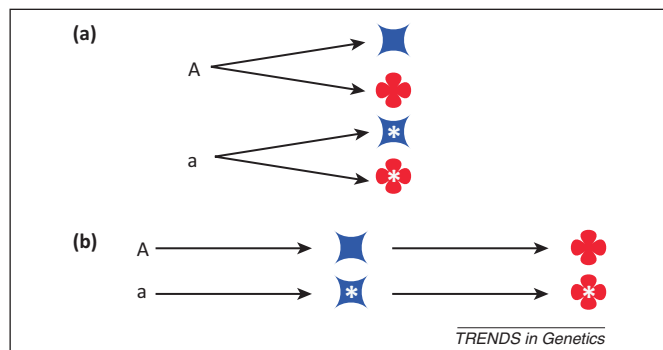

Figure 2. Two ways in which mutations can affect multiple traits. (a) An allelic difference (A/a) independently affects the blue trait and the red trait. (b) An allelic difference affects the blue trait, which in turn influences the red trait. The former case may be more susceptible to genetic modifiers of pleiotropy. In the latter case, independent environmental and genetic perturbations that affect the blue trait also affect the red trait.

(Figure 3) raises questions about the power of the experimental data to detect effects; to a first approximation, every gene should affect fitness, or it would become a pseudogene [22].

Pleiotropy estimates from the nematode data set are no less ambiguous. The original RNAi screen identified hundreds of genes that induced at least one of 45 phenotypic defects in the first four embryonic cell divisions [23]. In the systematic analysis of pleiotropy [15], the average gene was associated with 4.6 defects, which was described as consistent with low pleiotropy (Figure 3c). However, another group examined the same data and found that the average gene induced seven defects, and concluded that pleiotropy 'occurs extensively among genes' [24]. Clearly, there is room for interpretation not only in what is 'low' or 'high,' but also in how to do the counting. Procedures that require counting significant effects from vast data sets pose serious statistical concerns, and a recent critique of the systematic analysis shows that its statistical assumptions preclude any possibility of finding evidence for high pleiotropy, because patterns of restricted pleiotropy can be
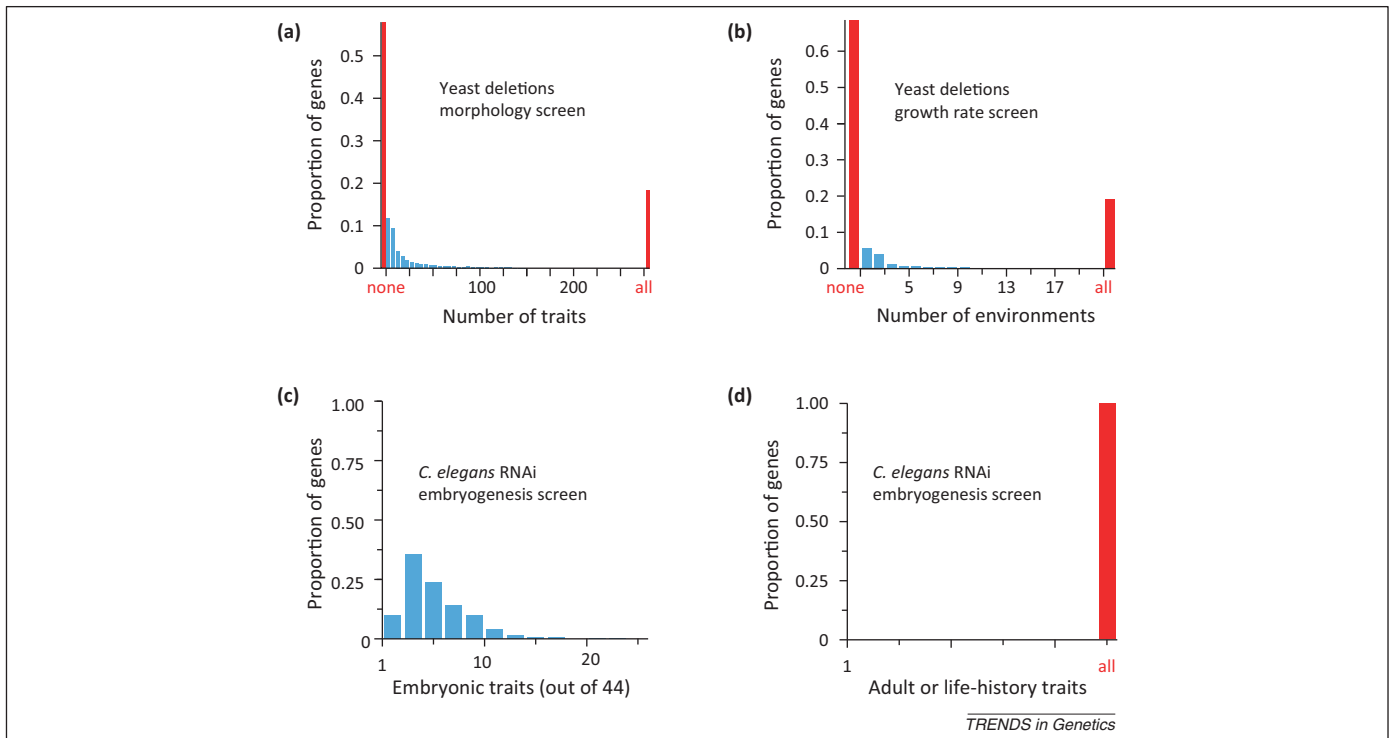
Figure 3. Distributions of pleiotropy for single-gene deletions. Previous estimates of pleiotropy have suggested that most genes have few effects, and the distribution of effects per gene is L shaped. The data responsible for those estimates [15] are replotted here in blue: most deletions affect few traits and a few deletions affect many. Data excluded from those estimates, plotted in red, show that the distribution of genic pleiotropy is actually U shaped, with a second mode at maximal pleiotropy. **(a)** Yeast deletions affect a fraction of the 253 measured morphological traits, if one excludes those deletions that affect all 253 traits (lethals, red bar at 'all'). Deletions with no detectable effects constitute an even larger fraction of genes (red bar at 'none'). **(b)** Yeast deletions affect growth rate in only a small fraction of environments, excluding lethals, which affect growth rate in all 21 studied environments ('all'). In this data set, the large category of genes with no detectable effect ('none') includes genes that uniformly decrease growth rate in all environments; only deletions exhibiting gene-by-environment interactions were counted as pleiotropic. **(c)** A live-imaging study of *Caenorhabditis elegans* embryos in which 661 essential genes had been knocked down by RNAi found that each knockdown resulted in only a subset of 44 possible embryonic phenotypes. **(d)** If any number of adult traits or life-history traits is included in the analysis of the 661 genes plotted in (c), each gene is maximally pleiotropic (all traits are affected) because every gene results in embryonic lethality.

inferred from data simulated with pervasive pleiotropy [25]. Finally, because the genes in this case are embryonic lethal, from a developmental or selectional pleiotropy perspective, these data are consistent with the most extreme form of universal pleiotropy (Figure 3d).

A final caveat about reverse-genetics approaches is whether molecular gene null mutations permit meaningful estimates of developmental and selectional pleiotropy, which draw from the spectrum of spontaneous mutations. Whole-gene deletions are expected to be more pleiotropic than most point mutations [2], upwardly biasing estimates of pleiotropy [11]. In some cases, however, natural mutations will be more pleiotropic, as when deletions affect multiple genes or large-scale chromatin structure. In other cases, null mutations will be caused by transposon insertions that may have their own phenotypic effects beyond abrogating gene function [26]. Finally, a significant fraction of natural mutations are gain-of-function, and these are often dominant or semidominant and, hence, immediately visible to selection. Positively selected gain-of-function mutations include those responsible for solitary foraging in *Caenorhabditis elegans* [27] and for lactase persistence in humans [28]. The former exhibits extensive developmental pleiotropy (e.g., [29]) and the latter arguably as well (e.g., on the viability of pastoralism). Mutations responsible for protein-aggregation diseases (e.g., Huntington's) are also

pleiotropic gain-of-function mutations that belie the simple assumption that null mutations are maximally pleiotropic.

QTL approaches
QTL are another category of empirical data that have been used to estimate genome-wide pleiotropy. These data fit squarely in the context of developmental pleiotropy and represent loci that harbor phenotypically penetrant alleles segregating in populations, which may be quite different in their effects from deletion alleles or spontaneous mutations. A QTL detected for multiple traits is said to be pleiotropic. A recent summary of QTL pleiotropy data concluded that most loci affect only a few traits, in a pattern similar to that observed in reverse genetic screens [11].

Probably the most comprehensive QTL survey for pleiotropy involved mapping 70 skeletal traits in recombinant inbred mice, which identified a total of 102 QTL and for which the maximum number of traits affected was 30 and the average 7.8 [30]. Similar work on 54 body-shape traits in sticklebacks [31] identified 26 QTL, approximately half of which only influenced one trait, and the average number of traits affected was 3.5.

Gene expression QTL (eQTL) have also been used to infer levels of genome-wide pleiotropy. Most eQTL appear to affect a small number of gene expression traits, but typically a handful of eQTL hotspots affect abundances of

hundreds to thousands of transcripts. Such patterns have been observed in yeast [32], mice [33], humans [34], nematodes [35], and *Arabidopsis* [36].

Despite the abundance of QTL data, their utility for questions about the global pattern of pleiotropy is doubtful. The gravest problem is power; in almost every study, the majority of heritable variation is unexplained. The eQTL data in particular imply that an enormous number of gene–phenotype links are simply undetected because of inadequate statistical power [37]. In this line-cross experimental setting, the only explanation is a large number of variants with pleiotropic but modest effects. Genome-wide association studies exhibit similar patterns and probably yield to the same explanation [38]. Adoption of arbitrary thresholds for significance is a viable method for evaluating pleiotropy only if we are prepared to stipulate that effects beneath the detection limits are inconsequential; however, ample evidence suggests that most causal genetic variation is beneath practical detection limits [39]. The power issue certainly dwarfs any bias from interpreting linked loci as single pleiotropic loci, which has been invoked as a source of upward bias in pleiotropy estimates [11]. Furthermore, some genetic architectures will cause this bias to be downward, for example, if highly pleiotropic alleles tend to accumulate linked suppressors of their pleiotropic effects. In most developmental and selectional pleiotropy contexts, the distinction between linkage and pleiotropy is not fundamental (Box 4). Finally, effect size has important implications for evolution. Genetic architectures with many highly pleiotropic small-effect loci may be unconstrained because the pleiotropic effects cancel out, whereas those with small numbers of large-effect pleiotropic loci typically induce strong genetic correlations (e.g., [40]).

The deletion and QTL data sets measure fundamentally unrelated phenomena, in neither case with any great efficacy. How, then, can they all produce similar L-shaped distributions (excluding lethals)? One possibility is that these data are revealing a fundamental unity of pleiotropy. An alternative is that these L-shaped patterns are artifacts

of limited statistical power, a result that is well established for QTL data [41] and now for reverse genetics [25].

How else can we estimate pleiotropy?
Classical quantitative genetic methods provide evidence of extensive pleiotropy without relying on detection of single locus effects. One striking observation is that the rate of accumulation of phenotypic variance by mutation for a single arbitrary metric trait implies a mutation rate per trait of approximately one-tenth of the total genomic mutation rate [42,43]. Consequently, there cannot be a large number of mutationally independent traits. This inference lends credence to the commonplace observation that there are a finite number of molecular genes and an effectively infinite number of traits. Statistical genetic analyses supporting extensive polygeny (thousands of loci affecting individual traits [38,44]) reinforce the notion that the mutational independence of traits must be limited.

Predictions about pleiotropy can also be tested by artificial selection. These methods are predicated on the simple assumption that a locus controlling some trait under selection will have other pleiotropic effects that reduce fitness [45]. The expected response to selection is simple to predict from the heritability of the focal trait, so the degree to which the response falls short is a measure of the extent of pleiotropy [46,47]. This principle has been demonstrated in *Drosophila* [48]. In some cases, pleiotropy will simply prevent any response to selection [49,50].

Evolutionary predictions can also facilitate improved single-locus estimates of pleiotropy. If pleiotropy tethers positive and negative fitness effects together across multiple traits, then one method for screening for pleiotropic loci would be to look for gene disruptions that increase some measure of fitness. Assuming gene disruption is an easy genetic change to acquire, and that natural selection would fix a disrupted allele if it had strictly beneficial effects, then the persistence of the wild type allele must indicate the existence of some trade-off between traits. This has essentially been done exhaustively for lifespan, in which systems such as flies and worms have been screened for mutations that increase longevity. The result is that, under careful examination, longevity alleles routinely show pleiotropy, most commonly with reproductive and stress tolerance phenotypes (e.g., [51], but see [52]).

The road forward
Given its diverse meanings, pleiotropy cannot be treated as a unitary concept with a definable prevalence. It must instead be treated as a suite of conceptually related but empirically independent phenomena, whose extent and consequences depend on the precise definition. The extent of molecular gene pleiotropy (how many reactions does a protein catalyze, for example) is not informative about the evolutionary independence of morphological traits or of life-history components. To be useful, pleiotropy must be precisely defined. We point to two examples of empirically successful hypotheses that invoke precise definitions of pleiotropy.

First, the hypothesis that *cis*-regulatory variants contribute more to the evolution of animal morphology than do protein-coding variants invokes the greater pleiotropy of

---

Box 4. Does the distinction between pleiotropy and close linkage matter?

Substantial effort in the QTL field is devoted to distinguishing between pleiotropy and tight linkage as the cause of QTL colocalization (e.g., [67–71]). In agricultural contexts, the goal is to determine whether genetic correlations between traits of agronomic interest are either breakable, as expected in the case of linkage, or inherent, as expected in the case of pleiotropy. However, both sources of correlation are subject to modification, the former by recombination and the latter by segregating modifiers [60,72] or mutation, including the well-studied case of duplication and subfunctionalization [73,74]. If the distinction between linkage and pleiotropy matters evolutionarily, it will be because these occur at different rates. But do they? Modifiers of pleiotropy are not rare [7,13,72,75]. By contrast, linkage can be so tight as to be immune to recombination, as in some chromosomal rearrangement, organellar genomes, or Y chromosomes. In those cases, genetic correlations caused by linked loci may be more likely to be altered by epistatic modifiers than by recombination. Ultimately, the evolutionary features that matter are the relative rates of recombination and modifier mutation (or the segregating frequencies of modifiers), which will depend on the details of the trait and the genetic distance separating the loci. That is, the evolutionary implications of linkage versus pleiotropy vary case by case.

the latter as a constraint [2,53]. Because *cis*-regulatory mutations often affect only a subset of the expression domains of a gene, they are less pleiotropic than are protein-coding mutations, which affect the protein everywhere it is expressed. The hypothesis is focused specifically on quantifying developmental pleiotropy by assessing the number (or fraction) of different gene expression domains affected by a mutation.

A second example draws on the theory of antagonistic pleiotropy for life-history traits. This theory describes how senescence can result from evolutionary pressure on selectionally pleiotropic alleles. If an allele increases fitness early in life by improving reproductive success, but decreases fitness later by promoting senescence, then natural selection may favor it and result in an organism with a prescribed lifespan [1]. This hypothesis yields a vast array of testable predictions, and a major focus in biology of aging research now includes mapping the effects of mutations onto physiology and life history [54].

The concept of pleiotropy has generated hypotheses that touch upon every aspect of the relation between genes and phenotypes. Testing these can address important outstanding questions: the road ahead is long and wide, and we are only becoming more equipped to travel it. However, even well-defined hypotheses about the nature and extent of pleiotropy are challenged by empirical limitations. The most informative work will test alternatives while avoiding the limits imposed by experimental power.

Acknowledgments

References

1 Williams, G.C. (1957) Pleiotropy, natural selection, and the evolution of senescence. *Evolution* 11, 398–411

2 Stern, D.L. (2000) Evolutionary developmental biology and the problem of variation. *Evolution* 54, 1079–1091

3 Gruneberg, H. (1938) An analysis of the 'pleiotropic' effects of a new lethal mutation in the rat (*Mus norvegicus*). *Proc. R. Soc. Lond. B: Biol. Sci.* 125, 123–144

4 Hadorn, E. (1961) *Developmental Genetics and Lethal Factors*, Methuen and Co

5 Hodgkin, J. (1998) Seven types of pleiotropy. *Int. J. Dev. Biol.* 42, 501–505

6 Turelli, M. and Barton, N.H. (2004) Polygenic variation maintained by balancing selection: pleiotropy, sex-dependent allelic effects and GxE interactions. *Genetics* 166, 1053–1079

7 Pyeritz, R.E. (1989) Pleiotropy revisited: molecular explanations of a classic concept. *Am. J. Med. Genet.* 34, 124–134

8 Emerson, K.J. *et al.* (2009) Complications of complexity: integrating environmental, genetic and hormonal control of insect diapause. *Trends Genet.* 25, 217–225

9 Tyler, A.L. *et al.* (2009) Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *Bioessays* 31, 220–227

10 Stearns, F.W. (2010) One hundred years of pleiotropy: a retrospective. *Genetics* 186, 767–773

11 Wagner, G.P. and Zhang, J. (2011) The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* 12, 204–213

12 Schadt, E.E. *et al.* (2005) An integrative genomics approach to infer causal associations between gene expression and disease. *Nat. Genet.* 37, 710–717

13 Wright, S. (1968) In *Evolution and the Genetics of Populations* (Vol. 1), University of Chicago Press

14 Gove, R.P. *et al.* (2012) Effects of causal networks on the structure and stability of resource allocation trait correlations. *J. Theor. Biol.* 293, 1–14

15 Wang, Z. *et al.* (2010) Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18034–18039

16 Ohya, Y. *et al.* (2005) High-dimensional and large-scale phenotyping of yeast mutants. *Proc. Natl. Acad. Sci. U.S.A.* 102, 19015–19020

17 Dudley, A.M. *et al.* (2005) A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol. Syst. Biol.* http://dx.doi.org/10.1038/msb4100004

18 Giaever, G. *et al.* (2002) Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* 418, 387–391

19 Shapiro, M.D. *et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428, 717–723

20 Miller, C.T. *et al.* (2007) *cis*-Regulatory changes in kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* 131, 1179–1189

21 Stern, D.L. (1998) A role of Ultrabithorax in morphological differences between *Drosophila* species. *Nature* 396, 463–466

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

23 Sonnichsen, B. *et al.* (2005) Full-genome RNAi profiling of early embryogenesis in *Caenorhabditis elegans*. *Nature* 434, 462–469

24 Zou, L. *et al.* (2008) Systematic analysis of pleiotropy in *C. elegans* early embryogenesis. *PLoS Comput. Biol.* 4, e1000003

25 Hill, W.G. and Zhang, X.S. (2012) On the pleiotropic structure of the genotype–phenotype map and the evolvability of complex organisms. *Genetics* 190, 1131–1137

26 Dennis, S. *et al.* (2012) *C. elegans* germ cells show temperature and age-dependent expression of Cer1, a Gypsy/Ty3-related retrotransposon. *PLoS Pathog.* 8, e1002591

27 McGrath, P.T. *et al.* (2009) Quantitative mapping of a digenic behavioral trait implicates globin variation in *C. elegans* sensory behaviors. *Neuron* 61, 692–699

28 Tishkoff, S.A. *et al.* (2007) Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39, 31–40

29 Reddy, K.C. *et al.* (2009) A polymorphism in npr-1 is a behavioral determinant of pathogen susceptibility in *C. elegans*. *Science* 323, 382–384

30 Wagner, G.P. *et al.* (2008) Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* 452, U470–U479

31 Albert, A.Y. *et al.* (2008) The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* 62, 76–85

32 Brem, R.B. *et al.* (2002) Genetic dissection of transcriptional regulation in budding yeast. *Science* 296, 752–755

33 Schadt, E.E. *et al.* (2003) Genetics of gene expression surveyed in maize, mouse and man. *Nature* 422, 297–302

34 Morley, M. *et al.* (2004) Genetic analysis of genome-wide variation in human gene expression. *Nature* 430, 743–747

35 Li, Y. *et al.* (2006) Mapping determinants of gene expression plasticity by genetical genomics in *C. elegans*. *PLoS Genet.* 2, e222

36 West, M.A. *et al.* (2007) Global eQTL mapping reveals the complex genetic architecture of transcript-level variation in *Arabidopsis*. *Genetics* 175, 1441–1450

37 Brem, R.B. and Kruglyak, L. (2005) The landscape of genetic complexity across 5,700 gene expression traits in yeast. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1572–1577

38 Purcell, S.M. *et al.* (2009) Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 460, 748–752

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

40 Pavlicev, M. and Hansen, T.F. (2011) Genotype–phenotype maps maximizing evolvability: modularity revisited. *Evol. Biol.* 38, 371–389

41 Bost, B. *et al.* (2001) Genetic and nongenetic bases for the L-shaped distribution of quantitative trait loci effects. *Genetics* 157, 1773–1787

42 Barton, N.H. (1990) Pleiotropic models of quantitative variation. *Genetics* 124, 773–782

43 Johnson, T. and Barton, N. (2005) Theoretical models of selection and mutation on quantitative traits. *Philos. Trans. R. Soc. Lond. B: Biol. Sci.* 360, 1411–1425

44 VanRaden, P.M. *et al.* (2009) Invited review: reliability of genomic predictions for North American Holstein bulls. *J. Dairy Sci.* 92, 16–24

45 Keightley, P.D. and Hill, W.G. (1990) Variation maintained in quantitative traits with mutation–selection balance: pleiotropic side-effects on fitness traits. *Proc. R. Soc. Lond. B: Biol. Sci.* 242, 95–100

46 Otto, S.P. (2004) Two steps forward, one step back: the pleiotropic effects of favoured alleles. *Proc. R. Soc. Lond. B: Biol. Sci.* 271, 705–714

47 Kirkpatrick, M. (2009) Patterns of quantitative genetic variation in multiple dimensions. *Genetica* 136, 271–284

48 Clayton, G.A. *et al.* (1956) An experimental check on quantitative genetical theory. I. Short–term responses to selection. *J. Genet.* 55, 131–151

49 McGuigan, K. *et al.* (2011) Pleiotropy, apparent stabilizing selection and uncovering fitness optima. *Trends Ecol. Evol.* 26, 22–29

50 Walsh, B. and Blows, M.W. (2009) Abundant genetic variation plus strong selection = multivariate genetic constraints: a geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* 40, 41–59

51 Paaby, A.B. and Schmidt, P.S. (2009) Dissecting the genetics of longevity in *Drosophila melanogaster*. *Fly* 3, 29–38

52 Partridge, L. *et al.* (2005) Sex and death: what is the connection? *Cell* 120, 461–472

53 Carroll, S.B. (2008) Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* 134, 25–36

54 Flatt, T. (2011) Survival costs of reproduction in *Drosophila*. *Exp. Gerontol.* 46, 369–375

55 Fisher, R.A. (1930) *The Genetical theory of Natural Selection*, Oxford University Press

56 Kacser, H. and Burns, J.A. (1981) The molecular basis of dominance. *Genetics* 97, 639–666

57 Wagner, G.P. *et al.* (2007) The road to modularity. *Nat. Rev. Genet.* 8, 921–931

58 Wright, S. (1964) Pleiotropy in the evolution of structural reduction and of dominance. *Am. Nat.* 98, 65–69

59 Su, Z. *et al.* (2010) A preliminary analysis of gene pleiotropy estimated from protein sequences. *J. Exp. Zool. B: Mol. Dev. Evol.* 314, 115–122

60 Pavlicev, M. and Wagner, G.P. (2012) A model of developmental evolution: selection, pleiotropy and compensation. *Trends Ecol. Evol.* 27, 316–322

61 Zhang, X.S. and Hill, W.G. (2003) Multivariate stabilizing selection and pleiotropy in the maintenance of quantitative genetic variation. *Evolution* 57, 1761–1775

62 Waxman, D. and Welch, J.J. (2005) Fisher's microscope and Haldane's ellipse. *Am. Nat.* 166, 447–457

63 Martin, G. and Lenormand, T. (2006) A general multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* 60, 893–907

64 Wagner, G.P.Z.J. (2012) Universal pleiotropy is not a valid null hypothesis: reply to Hill and Zhang. *Nat. Rev. Genet.* 13, 296

65 Hill, W.G. and Zhang, X.S. (2012) Assessing pleiotropy and its evolutionary consequences: pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. *Nat. Rev. Genet.* 13, 296

66 Balestre, M. *et al.* (2012) Bayesian mapping of multiple traits in maize: the importance of pleiotropic effects in studying the inheritance of quantitative traits. *Theor. Appl. Genet.* 125, 479–493

67 Jiang, C.J. and Zeng, Z.B. (1995) Multiple-Trait Analysis of Genetic-Mapping for Quantitative Trait loci. *Genetics* 140, 1111–1127

68 Almasy, L. *et al.* (1997) Bivariate quantitative trait linkage analysis: pleiotropy versus co-incident linkages. *Genet. Epidemiol.* 14, 953–958

69 Lebreton, C.M. *et al.* (1998) A nonparametric bootstrap method for testing close linkage vs. pleiotropy of coincident quantitative trait loci. *Genetics* 150, 931–943

70 Knott, S.A. and Haley, C.S. (2000) Multitrait least squares for quantitative trait loci detection. *Genetics* 156, 899–911

71 Varona, L. *et al.* (2004) Derivation of a Bayes factor to distinguish between linked or pleiotropic quantitative trait loci. *Genetics* 166, 1025–1035

72 Pavlicev, M. *et al.* (2008) Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution* 62, 199–213

73 Lynch, M. and Force, A. (2000) The probability of duplicate gene preservation by subfunctionalization. *Genetics* 154, 459–473

74 Des Marais, D.L. and Rausher, M.D. (2008) Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* 454, 762–765

75 Wolf, J.B. *et al.* (2006) The contribution of epistatic pleiotropy to the genetic architecture of covariation among polygenic traits in mice. *Evol. Dev.* 8, 468–476

Cell
PRESS

On the definition and measurement of pleiotropy

Jianzhi Zhang[1] and Gunter P. Wagner[2]

[1] Department of Ecology and Evolutionary Biology, University of Michigan, 830 North University Avenue, Ann Arbor, MI 48109-1048, USA
[2] Yale Systems Biology Institute and Department of Ecology and Evolutionary Biology, Yale University, 165 Prospect Street, POB 208106, New Haven, CT 06405-8106, USA

Pleiotropy, a 103-year-old concept [1], refers to the phenomenon of one mutation affecting multiple traits. Compared with the fundamental importance of pleiotropy in genetics, development, evolution, and medicine, our empirical knowledge of the prevalence and mechanism of pleiotropy is scarce [2]. Given this context, Paaby and Rockman's recent Opinion article on pleiotropy and its empirical estimation [3] is timely, but several of their claims are debatable.

First, based on the type of traits concerned, Paaby and Rockman proposed three kinds of pleiotropy, namely 'molecular gene pleiotropy', 'developmental pleiotropy', and 'selectional pleiotropy'. Although there is no doubt that the pleiotropic level of a mutation can vary depending on the types of trait considered, this classification scheme is fundamentally a categorization of traits rather than of pleiotropy. As an analogy, consider the common terms 'dominance' and 'recessiveness'. Whether a particular allele is dominant over another depends on the trait considered. For example, missense mutations in the gene encoding hexosaminidase A (*HEXA*) often cause the recessive Tay–Sachs disease (OMIM #272800) [4], but at the level of mRNA, the mutant alleles and wild type alleles are codominant, because the mutations do not affect *HEXA* gene expression. Despite the dependence of dominance on the trait considered, dominance or recessiveness is not classified, as Paaby and Rockman proposed for pleiotropy, as expression dominance or developmental dominance. We believe that, as long as the trait under discussion is clear, there is no need to classify dominance, recessiveness, or pleiotropy.

Second, Paaby and Rockman contended that a previous summary of the reverse genetic data on yeast gene pleiotropy [5] is biased because of the exclusion of essential genes, which are defined as genes that cause lethality or sterility when deleted. Note that gene pleiotropy is a simplified phrase for the pleiotropy of null mutations in a gene, because, strictly speaking, pleiotropy is a property of mutations rather than of genes. Paaby and Rockman considered essential genes to affect all traits of an organism and to be maximally pleiotropic. Although this is a potential way of looking at the situation, in most contexts lethal and/or sterile mutations are not relevant to start with. The pleiotropy of lethal and/or sterile mutations does not inform about the evolvability of a genotype, because such mutations will never compete for representation in the next generation. Furthermore, if null mutations of essential genes are specifically investigated for certain traits, their levels of pleiotropy are not maximally high [5,6].

Third, Paaby and Rockman argued that the pleiotropic level has been grossly underestimated from recent reverse genetic and quantitative trait locus mapping data [2,5,7–9], because they believe that these methods are not sensitive enough to detect most phenotypic effects of mutations. Although increasing the sample size and the precision of measurement is expected to boost the estimated degree of pleiotropy, the argument misses two critical points. First, the key question is not how many traits are significantly affected by a mutation statistically, but biologically. Of course, we do not have scalable methods to estimate fitness effects of trait differences, but in principle we should not mistake statistical issues for biological insights. Second, even if more phenotypic effects are detected upon experimental improvement, their effect sizes are expected to be smaller than those that have been detected. The majority of these small effects may be biologically irrelevant.

Finally, as mentioned previously [2], a measure of pleiotropy that is not sensitive to the statistical power of the experiment is urgently needed; for instance, methods describing the distribution of estimated effect sizes across traits, rather than only counting the statistically significant effects.

Although the total amount of empirical data for studying pleiotropy is still limited, the past few years have seen a rapid growth of suitable data, which has led to significant expansion of our knowledge of the prevalence and mechanism of pleiotropy [2,10,11]. We believe that the most productive way forward is to expand and improve attempts to estimate pleiotropy empirically. Paaby and Rockman highlighted two pleiotropy-related hypotheses on the genetic mechanisms of phenotypic evolution and aging that they suggest are well suited to exploring this phenomenon [3]. Although they mentioned the argument that *cis*-regulatory mutations are less pleiotropic than are coding sequence mutations as an example of a testable hypothesis, to our knowledge, there is only anecdotal evidence to support this. An empirical demonstration of the lower pleiotropy of *cis*-regulatory mutations than of coding sequence mutations will only be possible if the methodological problems discussed above are addressed. Similarly, the genes and mutations that antagonize development and/or reproduction and life span, and the mechanism of their trade-off have not been identified [12]. Although many basic questions on pleiotropy remain, we are hopeful that, with rapid advancement of genetic and genomic technologies and data-driven theory development, these questions will be addressed in the near future by empirical characterization of pleiotropy at a high resolution, on a large scale, and for many species.

Corresponding author: Zhang, J. (jianzhi@umich.edu); Wagner, G.P. (gunter.wagner@yale.edu).

CrossMark

References

1 Plate, L. (1910) Genetics and evolution. In *Festschrift zum sechzigsten Geburtstag Richard Hertwigs*. pp. 536–610, (in German), Fischer
2 Wagner, G.P. and Zhang, J. (2011) The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. *Nat. Rev. Genet.* 12, 204–213
3 Paaby, A.B. and Rockman, M.V. (2013) The many faces of pleiotropy. *Trends Genet.* 29, 66–73
4 Myerowitz, R. (1997) Tay–Sachs disease-causing mutations and neutral polymorphisms in the Hex A gene. *Hum. Mutat.* 9, 195–208
5 Wang, Z. *et al.* (2010) Genomic patterns of pleiotropy and the evolution of complexity. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18034–18039
6 Liao, B.Y. *et al.* (2010) Contrasting genetic paths to morphological and physiological evolution. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7353–7358
7 Wagner, G.P. *et al.* (2008) Pleiotropic scaling of gene effects and the 'cost of complexity'. *Nature* 452, 470–472
8 Albert, A.Y. *et al.* (2008) The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* 62, 76–85
9 Dudley, A.M. *et al.* (2005) A global view of pleiotropy and phenotypically derived gene function in yeast. *Mol. Syst. Biol.* 1, 2005.0001, (http://dx.doi.org/10.1038/msb4100004)
10 He, X. and Zhang, J. (2006) Toward a molecular understanding of pleiotropy. *Genetics* 173, 1885–1891
11 Qian, W. *et al.* (2012) The genomic landscape and evolutionary resolution of antagonistic pleiotropy in yeast. *Cell Rep.* 2, 1399–1410
12 Flatt, T. (2011) Survival costs of reproduction in *Drosophila*. *Exp. Gerontol.* 46, 369–375

Pleiotropy: what do you mean? Reply to Zhang and Wagner

Annalise B. Paaby and Matthew V. Rockman

Department of Biology and Center for Genomics and Systems Biology, New York University, New York, NY, USA

It is a testament to the conceptual slipperiness of pleiotropy that thoughtful people can reasonably disagree about its meaning and extent. Zhang and Wagner [1] raise important issues in response to our recent Opinion article [2], and readers will do well to consider their critique. Here, we offer a brief counterpoint to the objections that they raise.

First, Zhang and Wagner argue against our categorization of pleiotropy into three distinct forms. At some level, this point is semantic, because we could as well have described a single kind of pleiotropy in three contexts: molecular genetics, developmental genetics, and evolutionary genetics. Our central goal, however, is to discourage inappropriate applications of data from one context to theory from another. Zhang and Wagner's analogy to dominance illustrates our concern. As they note, a single mutation can have different dominance with respect to different phenotypes. Where the term 'dominance' appears in evolutionary population genetics, however, it refers exclusively to dominance with respect to fitness. Given that phenotypes may vary nonlinearly with fitness, data on dominance in the context of development may be misleading about the role of dominance in the evolutionary fate of mutations. The same holds for pleiotropy.

Next, Zhang and Wagner argue that lethal gene deletions are uninformative because they cannot contribute to evolution. This is our point. If the goal is to estimate the pleiotropy of null mutations to answer questions about evolution, then those mutations too pleiotropic to ever contribute to evolution are exceedingly relevant. They demonstrate that pleiotropy constrains evolution. Zhang and Wagner suggest that the patterns of pre-lethality

developmental pleiotropy exhibited by such mutations are informative, but we reiterate the importance of distinguishing developmental from selectional pleiotropy in such cases. Embryonic lethals all have the same pattern of selectional pleiotropy.

The third point of disagreement is whether subtle allelic effects that lie below thresholds for statistical detection are biologically important. We believe that the weight of empirical data decisively favors an affirmative answer, and we invite readers to revisit our notes on this point in the original Opinion [2]. We share with Zhang and Wagner the view that the development of methods to measure pleiotropy without dependence on statistical thresholds is an important goal.

Finally, Zhang and Wagner suggest that the way forward is more data. As empiricists, we echo this point. However, data alone are not sufficient. We need data that are relevant to our hypotheses. Zhang and Wagner suggest that the examples we offer of successes in the study of pleiotropy (the *cis*-regulatory theory of morphological evolution and the antagonistic pleiotropy model of aging) are themselves in dire need of such data, and we do not doubt that they remain hypotheses. Their virtue, from our perspective, is that they are well-posed hypotheses, such that the relevant data are precisely defined. In our Opinion, we argue that many questions about pleiotropy are ill posed, and that greater attention to conceptual foundations will improve our empirical progress.

References

1 Zhang, J. and Wagner, G.P. (2013) On the definition and measurement of pleiotropy. *Trends Genet.* 29, 383–384
2 Paaby, A.B. and Rockman, M.V. (2013) The many faces of pleiotropy. *Trends Genet.* 29, 66–73