4-1-2017

Trait Correlations in the Genomics Era

Julia B. Saltz  
*Rice University*

Frances C. Hessel  
*Baylor College of Medicine*

Morgan W. Kelly  
*Louisiana State University*

Follow this and additional works at: [https://repository.lsu.edu/biosci_pubs](https://repository.lsu.edu/biosci_pubs)

**Recommended Citation**


This Article is brought to you for free and open access by the Department of Biological Sciences at LSU Scholarly Repository. It has been accepted for inclusion in Faculty Publications by an authorized administrator of LSU Scholarly Repository. For more information, please contact ir@lsu.edu.
Review
Trait Correlations in the Genomics Era

Julia B. Saltz,1,* Frances C. Hessel,2 and Morgan W. Kelly3

Thinking about the evolutionary causes and consequences of trait correlations has been dominated by quantitative genetics theory that is focused on hypothetical loci. Since this theory was initially developed, technology has enabled the identification of specific genetic variants that contribute to trait correlations. Here, we review studies of the genetic basis of trait correlations to ask: What has this new information taught us? We find that causal variants can be pleiotropic and/or linked in different ways, indicating that pleiotropy and linkage are not alternative genetic mechanisms. Further, many trait correlations have a polygenic basis, suggesting that both pleiotropy and linkage likely contribute. We discuss implications of these findings for the evolutionary causes and consequences of trait correlations.

The Theory: Why Do We Care about the Genetic Basis of Trait Correlations?

Trait correlations (see Glossary) describe the situation in which two or more traits covary among individuals within a population (see Figure I in Box 1). Trait correlations are widespread among traits and taxa [1], so understanding how they evolve, both neutrally and under different types of selection, is essential to predicting their evolutionary effects [2]. Important work has demonstrated that many trait correlations do indeed have a genetic basis [3–5]. However, predictions about the evolutionary dynamics of trait correlations go beyond heritability: often we are interested in why traits are correlated, whether the correlation evolved under selection, and whether it is possible for selection or drift to change the magnitude or direction of trait correlations.

Quantitative genetics theory suggests that identifying genetic loci underlying trait correlations can contribute to answering these questions. Specifically, quantitative trait loci can produce trait correlations through pleiotropy, in which a single locus causally affects two or more traits, or through linkage, in which two or more loci each affect different traits, but are in linkage disequilibrium (LD) and therefore are inherited together (Box 1). This distinction is important because trait correlations caused by pleiotropy are expected to evolve (adaptively and by drift) differently than trait correlations caused by LD. In general, because LD is expected to erode through recombination, trait correlations generated by LD are expected to be transient and therefore have a limited scope for contributing to evolutionary change. Similarly, if a trait correlation is caused by LD, each trait might be produced by distinct functional mechanisms. By contrast, trait correlations caused by pleiotropy arise because a common mechanism contributes to the production of the correlated traits. Trait correlations caused by pleiotropy are not expected to break down simply through neutral processes, and in some cases can persist even if they are maladaptive because they reflect trade-offs [2]. Trade-offs occur when functional relationships among traits prevent evolution of optimum values for all traits simultaneously. Specifically, organisms are prevented from achieving maximum fitness because finite resources (which could include energy, time, specific molecules, etc.) allocated to one trait necessarily reduce investment in other traits. As a corollary, any variant that influences
investment of a finite resource in one trait must necessarily have pleiotropic consequences for other traits that require that resource [6].

What Types of Causal Variants Produce Trait Correlations?
Causal Genetic Variants for Trait Correlations Do Not Always Look Like Quantitative Trait Loci
Despite robust theory contrasting the evolutionary causes and consequences of pleiotropy and linkage, casual variants underlying trait correlations often fail to fall neatly into this “pleiotropy or LD” framework.

Box 1. Primer on Trait Correlations, Their Genetic Basis, and How Correlated Traits Evolve
Trait correlations describe phenotypic covariation between two or more traits among individuals in a population. For quantitative traits, the strength of the correlation between traits x and y can be measured as Pearson correlation coefficient, ρ. A trait correlation occurs when traits vary among individuals, and some combinations of traits are more common than others (illustrated in Figure A, left and center). When trait correlations are absent, all trait combinations occur at similar frequencies (Figure A, right).

Figure IB illustrates how a large-effect locus can generate trait correlations. Individuals with Variant 1 – which can be a single pleiotropic variant, or a set of linked variants – have high values of trait X and trait Y. Individuals with Variant 2 have low values of trait X and trait Y. (Variation among individuals with the same variant is due to effects of other variants, environmental variation, and/or measurement error.) When both variants segregate in a population, traits X and Y are correlated.

Evolution will proceed most rapidly when selection aligns with the direction of the trait correlation. For example, in Figure IC, if X and Y are positively correlated, selection for an increase in X and Y simultaneously will achieve a larger response to selection than selection for increased values of X but decreased values of Y (Figure IC).

Figure I. Primer on Trait Correlations. See text.

<table>
<thead>
<tr>
<th>(A)</th>
<th>ρ_{x,y} &lt; 0</th>
<th>ρ_{x,y} = 0</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait Y</td>
<td>Trait X</td>
<td>Trait Y</td>
</tr>
<tr>
<td>μ_x</td>
<td>μ_y</td>
<td>μ_y</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>(B)</th>
<th>(C)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Trait Y</td>
<td>Trait Y</td>
</tr>
<tr>
<td>μ_x</td>
<td>μ_y</td>
</tr>
</tbody>
</table>

Glossary
- **Causal variants**: “causal” genetic variants refer to those genetic differences between individuals that have been directly demonstrated (ideally using transgenics) to produce a trait correlation, that is, beyond merely showing a statistical association with the trait correlation.
- **Genetic variants**: variation in DNA sequence at a single genomic locus segregating in a population, including single nucleotide polymorphisms (SNPs), indels, copy number variants, etc.
- **Genome rearrangements**: segregating variation in chromosome structure, such as inversions, sex chromosomes, and translocations.
- **Genome-wide association study (GWAS)**: a method of identifying the genetic basis of variation in a trait. Numerous (typically thousands or more) individuals are measured for the trait(s) and genotyped; an association would be detected if the presence or absence of a particular variant is correlated with variation in the trait.
- **Linkage disequilibrium (LD)**: physical association between genetic variants, typically on the same chromosome, causing the variants to be co-inherited. If two or more linked variants influence different traits, then individuals inheriting different versions of the chromosome are expected to differ in multiple traits, producing trait correlations.
- **Long-range LD**: co-inheritance of variants that are relatively distant from each other in the genome and separated by regions of unlinked variants.
- **Pleiotropy**: a single genetic variant influences two or more phenotypic traits. Individuals with different versions of the variant (e.g., different alleles of a gene) are expected to differ in multiple traits, producing trait correlations.
- **Trade-offs**: because of competing demands on an organism, a change in the value of a particular trait would result in an increase in one aspect of fitness, but a decrease in another. Trade-offs occur when selection acts in opposing directions on the same trait, or when two or more traits are correlated in a direction that opposes the direction of selection.
- **Trait correlations (or ‘correlated characters’) [23]**: a phenotypic association between two or more
For instance, genome rearrangements have been well studied for their role in trait correlations [7,8]. Because a chromosomal rearrangement is inherited as a single unit, it acts as a single quantitative trait locus (QTL); when these genome rearrangements segregate within populations, they can produce trait correlations. For example, white-throated sparrows show discrete ‘tan-striped’ and ‘white-striped’ morphs (see Figure I in Box 2). The morphs differ in multiple behavioral traits, including mating preferences, courtship behavior, aggression, and investment in parental care, representing a highly multivariate set of trait combinations [3,9,10]. The genetic variants causing morph differences are located within a very large (approximately 100 Mb) pericentric inversion [11,12]. Expression of genes in the inversion differ between the brains of tan-striped and white-striped birds, and expression levels are correlated with song behaviors [13]. Therefore, the cause of this trait correlation has elements of both pleiotropy and LD; pleiotropy, because the inversion itself acts like a single pleiotropic QTL, influencing multiple correlated characters without being disassembled by recombination, and LD, because the individual casual variants are co-inherited in an inversion.

Conversely, a gene that appears to be pleiotropic — because allelic differences influence multiple phenotypes — might in fact represent multiple independent quantitative trait loci. For example, Carbone et al. [14] found in Drosophila that the gene Catecholamines up (Catsup) affects longevity, locomotor behavior, sensory bristle number, and starvation resistance, producing genetic correlations among these traits. However, fine mapping revealed that independent, unlinked mutations within catsup were associated with effects on different phenotypes [14]. These findings imply that the trait correlations generated by catsup represent a chance effect of sampling and are expected to be different even in a second sample of the same population [14].

Therefore, ‘gene-level’ pleiotropy and pleiotropic effects of individual genetic variants can result in different evolutionary fates for the trait correlations they produce [15].

The divergent predictions for trait correlations caused by pleiotropy and by LD assume that pleiotropy can be long-lasting, and that LD is ephemeral. However, cases (such as those highlighted earlier) where individual causal mutations have been identified demonstrate that the opposite can be true. We emphasize that this evidence is fully compatible with quantitative
(conspecific) morphs. Gene expression data suggested that mutations in different regions of doublesex, corresponding to different isoforms, independently control sex determination, sex-limited mimicry, and morph type [67].

(v) Seed beetles (Callosobruchus maculatus) lose performance on their preferred host (mung bean) when forced to evolve for multiple generations on its nonpreferred host (lentil), suggesting trade-offs in host performance. Genome scans reveal loci targeted by opposing selection during adaptation to each host [58].

(vi) Domestic dogs (Canis familiaris): The T gene controls tail length with deleterious pleiotropic effects in some dog breeds. The short-tailed mutation disrupts the ability of the T protein to bind to DNA, and in the 17 breeds segregating the variant in T, no dogs were found to be homozygous for the short-tailed variant, suggesting that it interferes with reproduction and/or embryo viability [69]. This example illustrates how strong artificial selection can produce maladaptive trait correlations.

Figure 1. Examples of Trait Correlations whose Genetic Basis has been identified (see text). (A) White-striped (left) and tan-striped morphs of the white-throated sparrow (photo: [80]). (B) Stickleback fish show multivariate differences between freshwater and marine populations (photo: https://commons.wikimedia.org/wiki/File:PICT0246-1.JPG). (C) Multivariate differences between wild-type (left) and DOG1 mutant A. thaliana (photo: [81]). (D) Polymorphic mimicry in butterflies (photo: [82]). (E) The seed beetle Callosobruchus maculatus (see Box 1; photo: [83]). (F) The T gene affects tail length with deleterious pleiotropic effects in dogs (photo: https://commons.wikimedia.org/wiki/Welsh_Corgi_Pembroke#/media/File:Pembroke_Welsh_Corgi.jpg).
genetics theory, but that understanding how specific causal genetic variants fit within a quantitative genetics framework is not always straightforward.

Pleiotropy and LD Are Not Mutually Exclusive
Some variants blur the distinction between pleiotropy and LD. In other cases, both pleiotropic and linked variants contribute to a trait correlation. For example, variation in the *Ectodysplasin* (*Eda*) gene in sticklebacks (*Gasterosteus aculeatus*) is associated with adaptation to freshwater environments and affects armor plating, growth rate (relative to the fully armored form, in fresh water) [16], the distribution of neuromasts along the lateral line [17], and behavior [18]. Within and among populations of freshwater and oceanic fish, variants in *Eda* are in LD with both nearby and long-distance variants [19]. Indeed, linkage between pleiotropic variants and other variants is expected to result from divergent multivariate selection, such as selection for adaptation to freshwater and oceanic habitats (Box 4).

Effect Sizes of Causal Variants Underlying Trait Correlations
Contrasting evolutionary predictions for trait correlations produced by pleiotropy and by LD are most applicable to trait correlations whose genetic architecture is dominated by a single large-effect locus: that locus is either a single variant that is pleiotropic, or a cluster of variants in LD. These predictions are less applicable to trait correlations caused by many loci, each with moderate or small effects, because some of these loci might correspond to pleiotropic variants, others might correspond to multiple variants in LD.

Examples of causal variants for trait correlations are dominated by variants of large effect, because small-effect variants are very challenging to conclusively link to phenotypes. Currently,
there is insufficient evidence to draw conclusions about the shape of the effect-size distribution for trait correlations [20]; however, at minimum, it is clear that trait correlations rarely show Mendelian genetic architecture, as we review in the following section.

Single Traits: An Exponential Distribution of Effect Sizes
The fact that most traits show polygenic, not Mendelian, inheritance is evident from pedigree studies and spurred the development of quantitative genetics. Orr [21] proposed that the effect sizes of loci underlying variation in most traits should follow an exponential distribution, with a few large-, and many small-effect variants. This view has been widely supported by empirical data (reviewed in [22]). In particular, genome-wide association studies (GWASs) frequently find no or few large-effect variants, even for highly heritable traits (reviewed in [23,24]).

Numerous, Highly Pleiotropic Variants
There are a large number of quantitative traits, and a very large, but finite, number of genetic variants. If most traits have a highly polygenic basis, then many of those variants must influence more than one trait [25,26]. This idea has been supported empirically [27], including with QTL analysis [28] and in mutation accumulation lines [29]. The same studies, as well as theoretical analysis [30] fail to support the concept of ‘universal pleiotropy’, that is, the idea that every...

Box 4. Environment-Dependent Pleiotropy
Just as the expression of univariate traits can be environment-dependent, environmental effects can alter the strength, or even the direction of trait correlations across environments (Figure I). Environment-dependent pleiotropy occurs when a variant affects the expression of two or more traits in one environment, but only one of these traits in another environment (Figure I). This phenomenon has been most commonly observed for life history traits [76]: for example, in the seed beetle Callosobruchus maculatus the sign of the correlation between fecundity and longevity is altered by food availability, potentially due to a shift between environments in the relative effect sizes of different causal loci [77].

A special case of environment-dependent pleiotropy arises in alleles contributing to local adaptation, which occurs when resident genotypes have higher fitness than those originating from other environments. Under conditions of moderate to high gene flow, maintenance of variation requires that alleles involved in local adaptation experience positive selection in one environment and negative selection in the other [78]. As a result, alleles involved in local adaptation might be especially likely to have environment-dependent pleiotropic effects on life history traits related to fitness. An increasing number of studies have identified alleles contributing to locally adapted traits [79], and many of these have turned out to be environment-dependent pleiotropic effects. For example, the low-armor allele of Eda (Box 2, and main text) has environment-dependent pleiotropic effects on growth rate: the allele affects the armor-plating phenotype in both freshwater and marine environments, but affects growth rate only in freshwater environments [16]. In Arabidopsis, the gene "Delay of Germination 1" (Dog1) contributes to local adaptation [65] and has environment-dependent effects on germination time [66], which in turn has downstream pleiotropic effects on the timing of other life history traits.

Figure I. Environment-dependent Pleiotropy. In Environment 1 (left), individuals with Variant 1 differ from individuals with Variant 2 in traits X and Y, producing a trait correlation. In Environment 2 (right), individuals with different variants differ in trait X, but not trait Y. Therefore, the pleiotropic effect of variation at this locus is environment dependent.
mutation affects all phenotypes equally, which is implied by Fisher’s geometric model [15,27,30]. Instead, evidence supports the view that pleiotropic variants typically have measurable effects on five to ten traits [15,27–29,31].

Genetic Architecture of Trait Correlations: Direct Evidence for Polygenic Architecture
In theory, the exponential distribution of effect sizes should apply both to univariate traits and to fitness, which is inherently multivariate [21]; currently, all we know is that the genetic architecture of trait correlations is commonly polygenic. Human GWASs have found that that comorbid diseases typically share 2 to 19 quantitative trait loci [32]. Similarly, in a review of QTL for trait correlations, Gardner and Latta [33] found that the average number of QTL shared between correlated traits was two. Both of these methods have significant limitations, including that they rely on LD among causal and non-causal variants to detect associations, and therefore cannot directly establish the number of causal variants or whether any variant has pleiotropic effects [34]. These limitations are expected to dramatically underestimate the degree of polygenicity, further underscoring the complex genetic architecture expected for most trait correlations.

Complementary evidence comes from mutation accumulation experiments, which have found that trait correlations can have a large mutational target [29,35], and from artificial selection experiments, which find that dozens of loci evolve in response to selection on trait correlations [36].

Genetic Architecture Is Expected to Depend on the Evolutionary Forces Maintaining the Trait Correlation
One of the least-understood aspects of trait correlations is how they are adaptively maintained, that is, why variation in the correlated traits persists despite selection and drift [20]. This question is important not only to understand the evolutionary maintenance of trait correlations but also because we expect trait correlations that are maintained by different evolutionary processes to have different genetic architectures.

Trade-offs, which occur when selection opposes the direction of the trait correlation, are expected to have complex polygenic architecture, with many variants of small effect. Trade-offs can be visualized as a multivariate fitness ridge, rather than a peak, in which many different combinations of trait values have equivalent, high fitness [20]. Any large-effect variant that mitigates the trade-off – that is, that increases one fitness component more than it decreases another – would be expected to rapidly evolve to fixation [1]. Variants of any effect-size that alter relative investment in different fitness components without improving overall fitness would be effectively neutral and would evolve by drift. Variants that decrease overall investment might be maintained by mutation-selection balance; such variants are expected to be abundant because organismal ‘quality’ or ‘condition’, which reflect the relative size of an individual’s total pool of resources, should be influenced by a large number of genes [37]. Since mutation-selection balance is biased toward maintaining small-effect variants [38], standing variation for trade-offs should rarely include large-effect variants.

By contrast, disruptive correlational selection – that is, two or more distinct fitness peaks in a multivariate adaptive landscape – might favor large-effect ‘master regulatory’ alleles that produce qualitatively distinct morphotypes with no intermediates [39]. For the trait correlation to persist, the morphotypes must also be subject to spatially varying selection and/or negative frequency-dependent selection [40], as in the white-throated sparrows described in an earlier section.

Implications for the Evolutionary Genetics of Trait Correlations
It Is Probably Pleiotropy and Linkage, So Focus on the Consequences of These
Evidence reviewed earlier suggests that it is not useful to describe pleiotropy as a property of a single gene and LD as linkage among genes, or to consider pleiotropy and LD to be mutually
exclusive. As reviewed earlier, pleiotropy and linkage are not defined by gene boundaries, and pleiotropic variants can be in LD with other pleiotropic or non-pleiotropic variants.

The underlying importance of distinguishing between LD and pleiotropy is that variants in LD are co-inherited, and pleiotropic variants have functions that influence the production of two or more traits. We suggest that each causal variant can be quantitatively evaluated for these two properties.

LD is already quantified as a spectrum from linkage equilibrium (i.e., no co-inheritance) to full co-inheritance. In this framework, the classical notion of a trait correlation caused by ‘pleiotropy’ – that is, one in which a pleiotropic variant contributes to a bivariate trait correlation – can be thought of as two quantitative trait loci, one for each correlated trait, that are always co-inherited [33].

This spectrum describes how trait correlations might be inherited, but ignores the mechanisms by which they are functionally produced. Therefore, causal variants underlying trait correlations can also be characterized based on the degree to which they contribute to variation via shared functional pathways. In a classical scenario of a trait correlation caused by ‘pleiotropy’, only a single large-effect variant is considered, so of course this variant must contribute to the production of all correlated traits. For polygenic trait correlations, the degree of shared function among different causal variants can vary from complete functional overlap (e.g., one variant regulates the expression of another variant) to complete functional independence (e.g., the variants are expressed at different times, in different tissues, and in different coexpression pathways). The classical idea of a trait correlation caused by ‘LD’ would be represented by two or more co-inherited variants characterized by functional independence. In practice, characterizing functional overlap or its absence will be challenging, because there are many different ways that variants can functionally contribute to the production of traits; we discuss some methodologies in Box 3.

Adaptive Hypotheses from Polygenic Trait Correlations
Pleiotropy is often considered an indication that a trait correlation is biologically and evolutionarily relevant, while LD is not, because pleiotropy indicates shared underlying mechanisms but LD can occur just by chance [41]. For example, a population of *Mimulus guttatus*, adapted to copper mine soils, also shows hybrid lethality when crossed to other populations. This pattern was originally interpreted as an example of ecological speciation, in which reproductive isolation evolved as a side effect of the evolution of copper tolerance. Recently, Wright et al. [42] found that loci conferring copper tolerance and hybrid sterility were in LD and that neither locus was pleiotropic for these two traits, suggesting that that adaptation to copper does not invariably produce reproductive isolation.

LD among any ‘two’ variants can easily occur by chance; indeed, every new variant arises on a chromosome containing other variants. However, polygenic trait correlations are unlikely to arise from chance LD events. Randomly arising LD should produce some linkage among variants that increase the values of both traits, some linkage among variants that decrease the values of both traits, and some linkage among variants that increase the value of one trait while decreasing the value of the other trait. If LD arises randomly, all outcomes of linkage should be equally common, producing no trait correlation at the phenotypic level. Trait correlations with a highly polygenic basis only arise when there is an overabundance of linked variants producing positive (or negative) correlations, which can occur because selection has removed deleterious instances of LD, and/or due to population admixture [43,44]. Similarly, trait correlations produced by long-range LD are only predicted to arise through selection or population admixture [43].
A version of this logic also applies for pleiotropy: although all pleiotropy indicates shared function between the traits, any particular pleiotropic variant might produce a positive or a negative trait correlation. One large-effect pleiotropic variant might produce a trait correlation in an arbitrary direction; but a highly polygenic trait correlation, generated by numerous small-effect pleiotropic variants, will manifest only when there is an overabundance of variants producing correlations in the same direction (positive or negative). Such a pattern might indicate physiological constraints and/or correlational selection.

Implications for Evolutionary Constraint or Its Absence
Since Darwin, theorists have argued that trade-offs must exist among life history traits, or else selection would drive any trait associated with fitness to its absolute maximum [45]. Since most values of those traits exist well below those bounds, their evolutionary potential must be constrained by trade-offs. And yet, artificial selection experiments that directly test these predictions typically discover rapid evolutionary responses in all directions [46]. For example, selection for increased egg size in a field-collected population of Drosophila melanogaster failed to find a negative genetic correlation with egg number, although these two traits are clearly functionally related, and require the same underlying resource [47].

Similarly, evolutionary change should be constrained to proceed more rapidly in the multivariate trait space for which genetic variation is available. Heritable trait correlations exist because variants producing certain trait combinations are more common than variants producing alternate combinations; therefore, maximum genetic variation is available when selection acts in the direction of the correlation (Box 1). Yet, artificial selection on genetically correlated traits has rarely shown evidence for any absolute constraint [48,49], and only a few cases have suggested quantitative constraints, that is, slower evolutionary response when selection acts to remove the correlation, relative to selection in the same direction as the correlation [50]. Even in cases where the trait correlation seems to be due to a single, large-effect, pleiotropic variant, evolutionary responses are typically rapid in response to any selection [48].

Constraints Might Result from Multivariate But Not Bivariate Correlations
As reviewed earlier, evidence suggests that pleiotropic variants typically influence more than two traits. Some of the best-studied causal variants for trait correlations control variation in numerous traits, such as the Eda allele in sticklebacks (Box 2). This evidence complements recent quantitative genetics theory pointing out that selection in nature is typically not bivariate, but rather, highly multivariate [51]. Therefore, evolution should depend on the degree to which genetic variation is aligned with the multivariate direction of selection.

However, direct tests for evolutionary constraints typically consider only bivariate genetic correlations. Expanding this perspective to consider highly multivariate trait correlations and their evolution is important to understanding evolutionary constraints: while genetic variants that alter bivariate trait correlations might be abundant (or at least, sufficient; see above), mutations altering many traits simultaneously in adaptive directions might be exceedingly unlikely [51]. This suggests a more limited potential for multivariate trait correlations to respond to selection.

To test this, Hine et al. [52] selected eight different directions (each of eight eigenvectors) on eight genetically correlated cuticular hydrocarbons in Drosophila serrata. Although each hydrocarbon was individually capable of responding to selection, the authors observed little response in three of the eight directions of selection characterized by low multivariate genetic variance [52].
Conclusions

The locations of genes relative to each other, and the local and global rates of recombination, influence the probability that a new functional variant will arise on a chromosome already containing another nearby functional variant, generating LD and potentially a trait correlation. Each of these processes can evolve. For example, recombination hotspots describe regions of the genome with elevated recombination rates, representing genomic locations where LD contributing to trait correlations is likely to decay through recombination more rapidly than expected. In primates, the location and ‘heat’ of hotspots are rapidly evolving [54]. By contrast, inversions, sex chromosomes, and other genomic regions of low recombination can facilitate the evolution and maintenance of trait correlations by allowing multiple variants to remain linked across longer timescales [7]. Finally, there is some evidence that clusters of genes conferring local adaptation have evolved through rearrangements, increasing the opportunity for LD to produce adaptive trait correlations even without mechanisms that suppress recombination [55,56].

Mechanisms that facilitate pleiotropy can also evolve. Genes contribute to the production of more than one trait if they produce a single product that affects more than one trait (e.g., through expression in different tissues and/or at different times, or by influencing one trait that affects other traits), or if they produce two or more distinct products each affecting the production of different traits. Changes in these gene-level properties can provide or limit the opportunity for variants in the coding or regulatory sequences of these genes to have pleiotropic effects. Evidence for the evolution of these properties comes from studies of macroevolutionary changes in gene regulatory elements that can influence a gene’s intensity, timing, and spatial extent of expression, as well as its downstream targets [57]. Similarly, evolutionary changes in factors that determine splicing can alter the number of functional products that a gene can produce [58,59]. Within species, the effect sizes and pleiotropic consequences of variants often depend on genetic background [60,61], indicating that there is standing genetic variation for the genetic architecture of trait correlations.

Importantly, these findings have been accompanied by emerging theories about how gene regulatory networks might evolve under drift [62,63]. For example, neutral forces are expected to increase the complexity of some gene networks in the absence of selection or under stabilizing selection, highlighting the need for caution in interpreting the role of adaptation in the evolution of complex gene networks [63].

Outstanding Questions

To understand trait correlations, we must not only understand why some trait combinations are present in a population but also, fundamentally, why others are absent. Have these missing trait combinations been removed by natural selection? Are they mutationally impossible to achieve (though they might be favored if present)? Or is their absence due to chance alone? Because trait correlations are an outcome of the joint effects of multivariate selection and the underlying genetic architecture of the traits under selection, we believe that the greatest insights will be gained through a close integration between descriptions of the genetic basis of trait correlations, and multivariate tests for selection. We list several important directions for future research below:

(i) Genetic architecture and selection: What is the distribution of effect sizes for trait correlations maintained by different adaptive or non-adaptive mechanisms? Does the distribution of effect sizes for trait correlations differ fundamentally from the distribution of effect sizes for single traits?
(ii) Fine-scale mechanisms: How often is gene-level pleiotropy the result of nonpleiotropic genetic variants that are linked within or among genes?
(iii) Constraint: Which portions of favorable trait space are unoccupied? What mechanisms prevent organisms from achieving these trait combinations?
(iv) Persistence of trait correlations: How persistent are trait correlations across deeper evolutionary time? Do trait correlations show a phylogenetic signal?
(v) Regulatory networks: What is the role of regulatory networks in producing trait correlations? How rapidly do these types of trait correlations evolve?
(vi) Shared function: How often are negative trait correlations generated through competition for a shared underlying resource? Do these types of correlations respond more slowly to selection?

All of these questions can be addressed with the technologies and data currently being employed to study trait correlations (Box 3). However new data, and the development of a larger number of organismal case studies will allow for a more robust test of the framework originally outlined by classical quantitative genetics theory.

Concluding Remarks

Because trait correlations are relevant to many ecologically and evolutionarily important traits, it is important to understand why some traits are correlated (and others are not), whether and how trait correlations affect trait evolution (relative to uncorrelated traits), and how trait correlations themselves evolve. By highlighting the ways that molecular genetic information has contributed to these goals, and conceptual and methodological challenges that have emerged in the process, we hope to foster continued interest in studying the causes of trait correlations and their evolutionary implications (see Outstanding Questions).

Acknowledgments

We thank the editor and four reviewers for constructive comments. J.B.S. was supported by the Dunn Foundation and Rice University startup funds. M.W.K. was supported by LSU startup funds, and the Louisiana Board of Regents.
References