Correlational selection in the age of genomics

Authors: Erik I. Svensson[1*], Stevan J. Arnold[2], Reinhard Bürger[3], Katalin Csilléry[4], Jeremy Draghi[5], Jonathan M. Henshaw[6,7], Adam G. Jones[6], Stephen De Lisle[1,8], David A. Marques[9], Katrina McGuigan[10], Monique N. Simon[2,11] and Anna Runemark[1].

Affiliations:

[1]Department of Biology, Lund University, SE-223 62 Lund, SWEDEN

[2]Department of Integrative Biology, Oregon State University, Corvallis, Oregon, USA

[3]Faculty of Mathematics, University of Vienna, Vienna, AUSTRIA

[4] Swiss Federal Research Institute WSL, Zürcherstrasse 111, 8903 Birmensdorf, SWITZERLAND

[5]Department of Biological Sciences, Virginia Tech, USA

[6]Department of Biological Sciences, University of Idaho, Moscow, Idaho, USA

[7]Institute of Biology I (Zoology), University of Freiburg, GERMANY

[8]Department of Ecology & Evolutionary Biology, University of Connecticut, Storrs, USA

[9]Eawag, Seestrasse 79, 6047 Kastanienbaum & University of Bern, Baltzerstasse 6, 3012 Bern, SWITZERLAND

[10]School of Biological Sciences, The University of Queensland, AUSTRALIA

[11]Department of Genetics and Evolutionary Biology, University of Sao Paulo, BRAZIL

*Correspondence to: *erik.svensson@biol.lu.se*

Svensson, E. I., Arnold, S. J., Bürger, R., Csilléry, K., Draghi, J., Henshaw, J. M., … Runemark, A. (2021). Correlational selection in the age of genomics. Nature Ecology & Evolution, 5, 562-573. https://doi.org/10.1038/s41559-021-01413-3

Abstract:

Ecologists and evolutionary biologists are well aware that natural and sexual selection do not operate on traits in isolation, but instead act on combinations of traits. This long-recognized and pervasive phenomenon is known as multivariate selection, or – in the particular case where it favours correlations between interacting traits – as correlational selection. Despite broad acknowledgement of correlational selection, the relevant theory has often been overlooked in genomic research. Here, we discuss theory and empirical findings from ecological, quantitative genetic and genomic research, linking key insights from different fields. Correlational selection can operate on both discrete trait combinations and on quantitative characters, with profound implications for genomic architecture, linkage, pleiotropy, evolvability, modularity, phenotypic integration and phenotypic plasticity. We synthesize current knowledge and discuss promising research approaches that will enable us to understand how correlational selection shapes genomic architecture, thereby linking quantitative genetic approaches with emerging genomic methods. We suggest that research on correlational selection has great potential to integrate multiple fields in evolutionary biology, including developmental and functional biology, ecology, quantitative genetics, phenotypic polymorphisms, hybrid zones and speciation processes.

Organisms are functionally integrated adaptive systems, where interactions among traits make the whole more than the sum of its parts. How and why did such functional integration evolve, and what are the evolutionary consequences of genetic correlations between traits? These questions have occupied evolutionary biologists for decades, resulting in a rich but scattered scientific literature on topics such as modularity[1], evolvability[1–3], multivariate selection on trait combinations[4–8] and the evolution of genetic correlation structure[9–12]. Early theoretical work by Cheverud[4] and Lande[13] predicted that genetic correlations between traits should become aligned with the direction of selection on trait combinations. This important insight made it possible to connect correlational selection (selection on trait combinations rather than traits in isolation; see formal definition in Box 1) to the field of evolutionary quantitative genetics, with its focus on genetic correlation structures. A central testable prediction was adaptive alignment between genetic correlations and the direction of correlational selection, although genetic correlations will also be influenced by other evolutionary forces (e.g. mutation and genetic drift) and ecological factors (e.g. fluctuating environmental conditions)[9,10].

Correlational selection forms a nexus between several traditionally separate research fields, including ecology and developmental biology (Fig. 1). Correlational selection links organismal level features, such as function and development, both to population phenomena such as modularity and genetic correlation structure and to underlying processes such as natural and sexual selection, which typically arise from interactions with mates, predators, mutualists or the abiotic environment (Fig. 2). These connections have not always been developed explicitly, with the result that whole research fields have largely remained separate, partly due to different terminologies. For instance, in a highly influential review about the evolution of modularity[1], correlational selection was not explicitly mentioned, and instead the authors used the terms modular selection

and a modular trait architecture as an expected outcome of selection. Correlational selection can either strengthen or weaken correlations between traits, depending on ecological context. For instance, plant evolutionary biologists studying floral pollination syndromes have noted that mutualistic interactions between pollinators and plants may lead to adaptive de-coupling between vegetative and floral parts, resulting in strong intramodule correlations but weak correlations between modules[14]. Similarly, antagonistic interactions like predation can impose strong correlational selection on behavioural traits, leading to tighter phenotypic integration and adaptive multivariate phenotypic plasticity in stickleback fish[15,16]. Studies of the outcomes of artificial selection and domestication processes have also revealed that correlations between animal personality traits have sometimes become decoupled, compared to the ancestors where these traits were more strongly genetically correlated[17].

In light of the genomic revolution, time is now ripe to evaluate Cheverud[4] and Lande[13]'s predictions about the evolution of genetic architecture and to ask: have they been confirmed or overturned by recent findings? In particular, are molecular signatures consistent with correlational selection having shaped the genomic architecture of organisms[6,18] and promoting functional integration, e.g. through linkage or pleiotropy? Here, we review quantitative genetic theory and data on correlational selection and link these to the partly separate literatures on modularity and evolvability, as well as to recent genomic research. Our aim is to synthesize insights from these different fields and to point out new directions for future research at their intersections.

Quantification and visualization of correlational selection

The first quantitative treatment of correlational selection was provided by Lande and Arnold[19] (Box 1). These pioneers introduced statistical tools to measure selection on continuously distributed phenotypic traits by estimating selection coefficients that could be incorporated into

96 the equations for predicting evolutionary responses. Below we discuss the interpretations of those
97 coefficients, and review the methods to estimate them.
98
99 Individual fitness surfaces are often complex, but can be analyzed to reveal the operation of
100 correlational selection (see definition in Box 1). Correlational selection is particularly likely when
101 the fitness surface resembles a ridge that is not parallel to either trait axis (Fig. 3A), as this form
102 of selection favors particular combinations of trait values over others and thereby selects for a non-
103 zero correlation between traits (Box 1). Correlational selection can also arise alongside disruptive
104 selection, for example when the fitness surface resembles a valley which is not parallel with either
105 trait axis (Box 1; Fig. 3).
106
107 The measurement of correlational selection requires data on the fitness and trait values of multiple
108 individuals (Fig. 3B). The major goals of such analyses are to visualize the fitness surface and
109 estimate coefficients that describe it[5]. In empirical studies, the true surface is unknown, but we can
110 deduce its properties by approximating the surface with simple functions. Quadratic surfaces are
111 often used to estimate coefficients corresponding to linear selection (β) and nonlinear selection (γ)
112 (Box 1)[19]. Unfortunately, it is difficult to visualize the fitness surface from the γ-coefficients alone.
113 However, the surface can be visualized by plotting it (Box 1) or by conducting a canonical analysis
114 that estimates the principal components (eigenvectors) of the surface (Box 1)[5,7,8]. Despite their
115 simplicity, quadratic coefficients can describe a wide variety of surfaces[5].
116
117 When a quadratic surface does a poor job of approximating the actual fitness surface, the surface
118 can be visualized using non-parametric methods[20]. These techniques can reveal multiple peaks and
119 valleys in the fitness surface (Fig. 3), if they exist, whereas the quadratic approaches will always
120 depict a smooth and simple relationship, regardless of the ruggedness of the underlying fitness

surface. However, non-parametric approaches have the shortcoming that they usually do not produce coefficients that are well-integrated into the equations of evolutionary change.

Our understanding of the empirical importance of correlational selection has lagged behind our understanding of the prevalence and consequences of directional selection[21], with only one meta-analysis of correlational selection published to date[22]. There are good reasons to expect correlational selection in a wide variety of ecological circumstances, and it might be particularly strong when fitness is affected by biotic interactions, which can generate strong and chronic selection on trait combinations[6]. Intraspecific interactions that have been shown to result in correlational selection often involve sexual or social selection[6]. Prime examples include selection on signaling traits such as colour[8,23,24] as well as selection on territorial behaviours, which can favor genetic coupling between traits like aggression, dispersal and colonization ability[25](Fig. 2). Interspecific interactions linked to correlational selection include predation based on colouration, morphology and behaviour traits[15,26], herbivory on plants[27] and mutualistic interactions between plants and their pollinators[28]. In many cases, the fitness surfaces are simple ridges or saddles, but sometimes the surface is more complex. Indeed, complex fitness surfaces could be common[20]. *A priori* we might expect to see multiple fitness peaks in organisms with discrete sympatric morphs[6,8,26] or between ecotypes[29] or newly formed species[30].

Evolution of genetic architecture in response to correlational selection

Correlational selection is central to our understanding of how genetic architecture evolves. Correlational selection is also closely connected, albeit not identical, to the concept of fitness epistasis in evolutionary genetics[31](Box 1). Importantly, although the single-generation effects of correlational selection on the genetic and phenotypic composition are readily understood, the transmission of these changes across generations is a complex theoretical and empirical issue.

147

148 To address how the effects of correlational selection are transmitted across generations, we must
149 first define two parameters. The first is the additive genetic variance-covariance matrix (**G**),
150 summarizing additive genetic variance for a set of traits[9,10]. The diagonal elements of **G** are the
151 additive genetic variances, and the off-diagonal elements are additive genetic covariances (Fig.
152 3C; see also Section 1 in Supplementary Material). Additive genetic variances and covariances
153 describe patterns of trait inheritance, and depend on the frequency and effects of alleles. The
154 additive genetic covariances are critical from a multivariate standpoint, because they describe the
155 extent to which inheritance of different traits tends to be non-independent. In the bivariate case, **G**
156 can be represented as an ellipse containing 95% of the genetic values of the individuals in a
157 population[32] (Fig. 3C). If two traits are strongly genetically correlated, the ellipse will be eccentric
158 and oriented such that it is not parallel to either trait axis. That is, genetic covariances between
159 traits result in directions of multivariate trait space with high (major axis of the correlation) and
160 low (minor axis of the correlation) genetic variance, even if genetic variance is high in all
161 individual traits[33] (Fig 3). Importantly, the long-axis of the **G**-matrix (g_{max}) represents a genetic
162 line of least resistance[34], the direction in phenotypic space which harbors the most genetic variance
163 and along which the population most easily evolves (see *"Consequences for pleiotropy,*
164 *evolvability, modularity and phenotypic plasticity"*).
165

166 Multivariate phenotypic effects of new mutations constitute a second set of key parameters, which
167 are summarized in the mutational variance-covariance matrix (**M**)[11,12]. Theory often assumes that
168 mutational effects are normally distributed. In the univariate case, when a locus affects only one
169 trait, this distribution can be described by a mean and a variance, and if mutations are unbiased,
170 the mean will be zero. In the multivariate case, some loci might be pleiotropic[13], meaning that they
171 affect more than one trait. In this case, the mutational effects are modeled as draws from a
172 multivariate normal distribution. This distribution is described by mutational variances for each

trait (diagonal elements of **M**) and mutational covariances between traits (off-diagonal elements of **M**). Positive mutational covariances mean that a mutation tends to affect both traits in the same direction, whereas negative mutational covariances indicate that mutations tend to affect traits in opposite directions.

Our analytical understanding of how correlational selection shapes the evolution of genetic variances and covariances comes from evolutionary quantitative genetic theory, particularly from the pioneering work by Russell Lande[13,35,36], and Wagner and Altenberg's[2] ideas about how selection on pleiotropic patterns could lead to parcelation or integration between traits. Lande's work suggested that inheritance should become aligned with the shape of the selection surface in well-adapted populations. Later, Cheverud used Lande's model of selection on pleiotropic mutations to predict that genetic correlations should match functional interactions among traits[4]. Recently, this suggestion was extended to predict a three-way alignment among selection, inheritance and mutation[12].

How short term responses to correlational selection are transmitted across generations depends on the distribution of allelic effects and the persistence of selection. Correlational selection can create genetic correlation by promoting linkage disequilibrium between alleles that affect two different traits[6]. However, such changes are expected to be eroded rapidly due to recombination if selection is relaxed in subsequent generations[6], suggesting that changes in genetic architecture due to this kind of correlational selection may be transient[37], unless correlational selection is persistent[6]. More realistically, if correlational selection acts on traits whose expression is affected by alleles with pleiotropic effects, then correlational selection will alter the frequencies of those pleiotropic alleles. Therefore, the distribution of mutational effects has important consequences for the efficacy of selection on genetic covariances.

Two recent advances have increased our general understanding of the evolution of genetic architecture. First, increasingly powerful computer simulations have enabled researchers to explore the long-term effects of correlational selection and mutation on the evolution of genetic covariances[9–12], expanding our knowledge beyond the case of mutation-selection balance under the classical infinitesimal model[35,38]. Second, a rapid increase in genomic data has provided insights into the empirical distributions of allelic effects in real populations. Combining both approaches provides exciting opportunities to understand how selection and genetics jointly shape the evolution of trait variation (see next section).

Simulation-based studies have verified the prediction by Lande[13] and Cheverud[4] that selection will cause standing genetic variation to become aligned with the fitness surface[9]. For instance, if the fitness surface is ridge-shaped, then populations will tend to harbor more variation in the direction of phenotypic space aligned along the crest of the ridge and less variation perpendicular to the ridge. However, other factors also influence the genetic architecture of traits: genetic drift can cause **G** to fluctuate over evolutionary time[9], a moving optimum stretches **G** in the direction of the movement[10], Migration also increases the genetic variance in the direction of phenotypic space pointing toward the mean of the migrant source population in an island-mainland model[39]. Recently, it has also been emphasized that the mutational variance is aligned with the direction of phenotypic plasticity[3,40], affecting both **G** and **M**. One interpretation of such alignment between plasticity and mutational variance is that developmental systems might respond similarly to environmental novelty as they do to genetic mutation[40]. Moreover, all else being equal, correlations in **M** should generate correlations in **G**, because standing genetic variation ultimately arises via mutation.

Interestingly, influences between the fitness surface, **G** and **M** can flow in both directions. While **M** can influence the shape of **G**, the fitness surface in turn can shape both **G** and **M**[11,12]. Thus, if

225 the fitness surface is a ridge in phenotypic space (Fig. 3), selection will cause the long axis of **G**
226 to align with the ridge. If such a selective regime is stable over evolutionary time, selection can
227 cause alignment between the fitness surface, **M** and **G**[12,41,42]. Simulations show that evolution of
228 the mutational distribution is especially plausible when different loci interact epistatically[12].
229 Recent progress in molecular biology, development and genomics suggests that such epistatic
230 interactions are extremely common[43]. Epistasis can therefore permit the evolution of the
231 mutational architecture because selection maintains variation at loci that have favorable
232 interactions under the prevailing selection regime.
233
234 A growing number of studies suggest that **G** can or has evolved in response to correlational
235 selection (Fig. 2). For instance, Delph et al.[44] imposed artificial correlational selection on
236 combinations of male and female floral traits in the dioecious flower *Silene latifolia* (Fig. 2I) to
237 test whether the between-sex genetic correlation was evolvable. High between-sex genetic
238 correlations would potentially constrain the evolution of sexual dimorphism. Between-sex genetic
239 correlations broke down after a few generations of selection[44], however, suggesting that these
240 correlations are due to linkage disequilibrium which is expected to break down rapidly under
241 artificial correlational selection or recombination. In another plant study, however, genetic
242 correlations were remarkably stable across several generations, suggesting that pleiotropy caused
243 these correlations[45].
244
245 **Genomic architecture of traits and consequences for multi-character evolution**
246
247 The development of next-generation sequencing (NGS) provides new opportunities to investigate
248 correlational selection beyond what has been possible with classical quantitative genetics.
249 Genomic data has allowed us to pinpoint the genetic basis and architecture of traits, to estimate

empirical distributions of allelic effects in real populations, to reconstruct the evolution of genome architecture relevant for trait evolution and to detect correlational selection from molecular footprints (Box 2).

Recent studies using quantitative trait loci (QTL) mapping, genome-wide association studies (GWAS), and whole genome sequencing of population samples (Box 2), have revealed that most genotype-phenotype maps[46] are complex. Most traits are determined by a large number of genes of small effect, consistent with the so-called 'polygenic model' of inheritance[38] allowing efficient quantitative genetics modelling ignoring details of multilocus inheritance by assuming the infinitesimal model[47]. However, empirical effect sizes distributions are often exponentially distributed[48,49] , with a few genes of major effect controlling a minority of traits for which the infinitesimal model is violated[50] and which often have an important role in adaptation and speciation[51,52]. Molecular studies have further revealed that many functional genetic variants are pleiotropic and affect multiple traits[53]. Multiple-mapping approaches, enabling joint estimation of effects on multiple traits, hold great promise to further improve our understanding of pleiotropy[54]. Molecular studies have also revealed that epistasis is common[55], with genotype-phenotype maps typically being highly nonlinear[56], suggesting pervasive epistasis in genotype networks[43]. The importance of epistasis is controversial because linear quantitative genetic models are rarely improved by the addition of interaction terms[57]. However, genomic quantitative genetic studies that incorporate a more precise estimate of the shared proportion of genome have revealed that higher order variance components are not negligible[58]. Interestingly, a recent study of *Timema* stick insects has shown that correlational selection arose from fitness epistasis due to ecological factors (predation), in spite of the underlying traits (colour) having an additive genetic basis[59].

These insights about the genetic architecture of traits have implications for the evolutionary response to correlational selection. One emerging insight from experimental evolution studies is that evolutionary changes in traits can often be achieved via many alternative "genomic solutions", suggesting important roles for redundancy and historical contingency in evolution[60]. Further, parallel evolution is often frequent for fitness itself, but less common for phenotypes, and less common still at the levels of genes or individual mutations[61]. For example, Therkildsen et al.[62] sequenced genomes of Atlantic silverside fish (*Menidia menidia*) selected for small and large size. Despite highly parallel phenotypic changes and several parallel allele frequency shifts in growth-related genes, genomic adaptation in one line was contingent on the presence of a large inversion with moderate phenotypic effect[62]. On the other hand, pleiotropy, functional constraints and the presence of major effect loci may limit the number of redundant genomic solutions in response to correlational selection[63]. For example, threespine stickleback adapting to freshwater habitats show highly repeatable evolution at a pleiotropic major effect locus[64,65].

Correlational selection changes the genetic covariances among traits and thereby ultimately shapes the evolution of genome architecture. Although many different mechanisms underly genetic covariances[66], the two basic causes are linkage disequilibrium and pleiotropy[38]. Linkage disequilibrium captured by physical linkage is one genomic cause of trait correlations, in which recombination is suppressed in heterochromatic regions, in genomic rearrangements, on sex chromosomes, or due to a high density of transposable elements[67]. For example, Choudhury et al.[68] sequenced 304 *Arabis alpina* genomes, and found that the S-locus supergene responsible for strict outcrossing was in a linkage disequilibrium block that included high levels of polymorphic transposable elements. Genomic rearrangements such as gene duplications, translocations,

297 chromosomal fusions or inversions can also maintain linkage disequilibrium, due to disrupted
298 meiotic chromosome pairing reducing recombination or the joint forces of physical linkage and
299 selection[69–71]. Linkage disequilibrium may be preserved in deep evolutionary time, forming so-
300 called ‘supergenes', some of which might resemble sex chromosomes[69,72–74] (Fig. 4A). There are
301 several empirical examples of how co-selected complex trait combinations, bound to supergenes,
302 cause behavioural and morphological differences between discrete morphs with different
303 reproductive tactics (Fig. 4B)[72–74]. For example, a recent study in the heterostylic plant genus
304 *Primula* revealed the build-up of an S-locus supergene controlling style, anther and pollen grains
305 via gene duplications and neofunctionalization[75].
306

307 Even in the absence of physical recombination suppression, genetic covariances among traits can
308 arise through linkage disequilibrium between loci[6]. Such linkage disequilibrium can potentially be
309 maintained by assortative mating among individuals with the same correlated trait combinations[76]
310 and by strong divergent or disruptive selection favoring several correlated trait optima within[6,77]
311 or between populations[78]. Correlational selection can also theoretically lead to speciation through
312 reinforcement of assortative mating by the evolution of genomic coupling between preference and
313 trait loci, even if they are initially unlinked[6,76,79]. There are several examples of ecotype or species
314 pairs where inter-chromosomal linkage disequilibrium is maintained either by strong selection[80]
315 or a combination of strong selection and assortative mating[81], with some studies demonstrating
316 genomic coupling between unlinked preference and sexually-selected trait loci[76].
317

Genomic tools in combination with quantitative genetic approaches also enable us to obtain better estimates of **G** (Box 2)[58,82]. In addition, comparisons between the genetic variances and covariances of **M** (the mutation matrix) and **G** (the matrix of standing genetic variation) can reveal the presence of correlational selection and how it operates during the organismal life-cycle and shapes both mutational pleiotropy and pleiotropy of the standing genetic variation[83–85] (Box 2). Mutation accumulation experiments (MA) have revealed strong mutational pleiotropy[85–87] and have indicated that correlational selection on such pleiotropy leads to a reduction in the corresponding genetic correlations in **G**[83](Box 2). Thus, correlational selection might operate against mutational pleiotropy, resulting in a discrepancy between **M** and **G**. Consequently, spontaneous and positive mutational correlations among traits could largely be maladaptive and reflect the input from mutation-selection balance[86]. This contrived example underscores the point that correlational selection can not only strengthen adaptive genetic correlations among traits, it can also weaken and break up maladaptive genetic correlations[44,83]

Finally, genomic information from several populations can be used to address how multiple traits co-evolve, using information from a coancestry matrix, which can be estimated with a handful of marker loci[88]. Csilléry et al.[89] recently used such an approach and found evidence for correlated character evolution in the timing and growth rate across 16 silver fir (*Abies alba*) populations. While this methodology is limited to describing the average effect across all causal loci, such approaches could enable us to describing the genomic architecture of trait correlations in terms of individual loci, their physical distribution across the genome, and their effect sizes[90].

Consequences for pleiotropy, evolvability, modularity and phenotypic plasticity

The intuition that pleiotropy slows and constrains evolution was well articulated by Orr[91], who updated Fisher's classical geometric model[92] to show that phenotypic complexity slows adaptation

when pleiotropy is universal. However, this link between pleiotropy and constraints on evolvability - the ability of a population to respond to selection[93,94] - has recently been challenged. First, pleiotropy may be largely confined within functional, integrated trait modules[46], allowing traits in separate modules to adapt semi-independently[2]. That is, as predicted by Cheverud, correlational selection will select for congruent phenotypic covariances[4,13]. Moreover, individual-based simulations with divergent multivariate directional selection, pushing groups of traits in opposite directions, showed that phenotypic variation can indeed evolve to become more modular[95]. Increased modularity may also evolve when environmental fluctuations favour new combinations of conserved functions[96] or when selection across multiple environments favors the expression of partially overlapping sets of genes[97] . While these studies suggest that circumstances favoring high evolvability can drive the evolution of modularity, theory has also shown that highly integrated and pleiotropic genetic architectures can have high evolvability[93]. There is still considerable room for development of theory to predict when we expect modularity to emerge as a solution to adaptive challenges (Section 2 in Supplementary Material).

A common feature for the evolutionary origin of modularity is directional selection, though modularity can also evolve as a consequence of selection for robustness to environmental perturbations[98], and merely adding selection to models with universal pleiotropy does not produce stable variational modules[99]. The responsiveness of modularity to directional selection also limits its use as a predictor of long-term evolutionary responses, perhaps explaining why functional modularity is only a modest predictor of co-evolutionary rates of evolution among genes[100]. Another potential explanation for this pattern is that functional and variational modularity only partially overlap. Empirical evidence of co-expression of genes is strong[101], but whether these co-adapted gene modules are organized as variational modules is controversial. Some studies using transcriptional data showed that genetically correlated transcripts tend to share developmental pathways, reflecting transcriptional modules that are mostly enriched with functionally related

genes[102,103], whereas other studies could not find significant overlap between gene expression and functional groupings[104]. Modular functional capacities do not require structural modularity[105], and modularity at the level of gene regulation may better predict evolvability[106]. The mismatch between variational modules and functional gene groupings can complicate the semi-independent evolution of phenotypic modularity.

Multivariate perspectives show that additive genetic variation within populations is distributed very unevenly across traits, with some linear combinations of traits accounting for most of the variance (i.e., g_{max}), while other trait combinations are associated with very little variance[33]. This pattern can stem from genetic variation being funneled through a few central developmental pathways, mediated by few developmental genes of large effect[107]. There has been considerable interest in g_{max}, (see section *"Evolution of genetic architecture in response to correlational selection"*) because it can either facilitate or bias evolutionary responses to selection depending on its alignment to the selective surface[108]. Additive genetic variance is determined by the effects and frequencies of contributing alleles, and at the genomic level, the initial response to selection should be dominated by loci with relatively high intra-locus variance and large effect. Although genomic studies, such as GWAS, have a tendency to detect loci with high frequencies of minor alleles and larger effects[109], empirical evidence of the contribution of variants to additive genetic variance points to mostly rare variants with mainly small but highly pleiotropic effects[110]. If g_{max} reflects the most common empirical pattern, selection aligned with g_{max} should promote adaptation through minor allele frequency changes at many loci[111].

The mere presence of additive genetic variation is not sufficient to predict evolutionary outcomes, as the response to selection depends on the orientation of selection relative to the distribution of genetic variation[94,112]. Only a few studies have measured both multivariate linear and quadratic selection and the distribution of genetic variation for those phenotypes. These studies typically

demonstrate relatively low genetic variance in the multivariate trait combinations associated with fitness variation, which slows down phenotypic evolution[113]. However, the causes of low genetic variance for multivariate phenotypes currently under selection, and thus the longer-term consequences of the covariance patterns, remain poorly resolved[33].

Because there is substantial genetic variance in other directions of trait space besides g_{max}, changes in the orientation of selection could result in relatively unbiased, rapid, adaptation[94]. Moreover, pleiotropy can be context-dependent[101,114], meaning that apparent pleiotropic constraints may shift in novel environments or evolve through epistatic interactions. For instance, changes in the selective environment could remove bias by changing the orientation of genetic variation[40], potentially through context-dependent pleiotropic effects of alleles[114], or through rapid evolution of **G**, which might be particularly likely if trait covariances are generated through opposing pleiotropic effects across contributing loci[93]. For example, a recent experimental study in yeast demonstrated that while alleles had pleiotropic effects on two life-history traits, variation in effects across environments resulted in genetic correlations ranging from -0.5 to +0.5[115]. Finally, mutational pleiotropy is only one of several factors influencing standing genetic variation, which also depends on multivariate selection and linkage disequilibrium.

The relationship between how genetic variance changes across contexts and how phenotypes respond to the environment directly (i.e., phenotypic plasticity) can determine the longer-term outcomes of correlational selection. A recent meta-analysis of published estimates of **G** and plastic responses to novel environments suggested that multivariate phenotypic plasticity might correspond to axes of genetic variation associated with substantial standing genetic variation[40]. This study and theory[3] suggest that bias in evolutionary response generated through **G** can become recapitulated through phenotypic plasticity. Clearly, more work is needed to understand how

environmental and genetic information are interpreted through the developmental systems (Section 2 in Supplementary Material).

Conclusions

Here, we have re-visited the early suggestions by Cheverud and Lande that correlational selection can shape genetic and phenotypic architecture in light of the recent genomic revolution. These early insights are consistent with increasing empirical evidence of genomic coupling and recombination suppression that could have arisen by correlational selection, although direct evidence for this process, in most cases, lacking. A remaining challenge is therefore to integrate organismal-level research on correlational selection on phenotypes with genomics and developmental biology. Below, we point to some promising new avenues for future integrative research in this exciting area.

First, despite empirical evidence that correlational selection can build up or eliminate genetic correlations between co-selected traits (Fig. 2), our knowledge of the mechanistic (i.e. genomic and developmental) underpinnings of such changes is still limited. To what extent are such changes caused by transient changes in linkage disequilibrium or the evolution of adaptive pleiotropy, and what is the relationship between modularity and correlational selection? We are only just beginning to understand the genomic mechanisms involved in adaptive recombination suppression caused by correlational selection, including the roles of supergenes[69,71], structural genomic rearrangements[70] such as gene duplications, chromosomal fusions or inversions and other mechanisms including TEs[67,68]. Promising future research directions in the study of the genomic consequences of correlational selection are to use genomic tools to study how correlational selection might lead to the gradual buildup of supergenes[75] and how such selection might operate on mutational pleiotropy across the organismal life-cycle, using a combination of mutation

accumulation (MA) experiments, quantitative genetics and quantification of gene expression changes during ontogeny[83,83,85–87]

Second, the relationship between phenotypic plasticity and correlational selection is largely unknown. The traditional perspective has been that correlational selection would primarily shape genetic correlation structure, by either strengthening or weakening genetic correlations between traits[4,6,44]. Research on stickleback fish has found that predation results in changed phenotypic correlation structures[15], but some of these phenotypic changes might reflect multivariate phenotypic plasticity rather than changes in genetic integration[16]. How genetic covariances and multivariate phenotypic plasticity jointly evolve under correlational selection is therefore a largely unexplored research area with great potential[16,40]. More generally, since correlational selection in the past might have shaped either phenotypic or genetic correlations (or both), it leaves an evolutionary "memory" of past selective environments[116] which can reveal itself in the form of alignment between the selective surface, **P**, **G** and **M**[12,41,42].

Third, the importance of correlational selection in speciation and macroevolution is largely unknown, despite early work on evolutionary allometry and the idea of evolution along "lines of least resistance"[34,117]. Recent research on shape-size allometry[118], brain-body size allometry[119] and metabolic allometry[120] have revealed that allometric relationships are not static evolutionary constraints, but can be altered by selection. Specifically, correlational selection could maintain adaptive allometric slopes, either due to internal causes related to deleterious pleiotropy[118] or because external ecological factors make certain slopes more beneficial than others[119,120].

Finally, we also see a great potential for research on the genomic consequences of correlational selection in the fields of animal and plant domestication[17], and in the context of dispersal strategies, social behaviours and personalities[16,25,121]. Humans might have consciously or unconsciously

either eliminated or strengthened genetic correlations between traits during domestication of plants and animals, through artificial correlational selection on suites of traits, which in some cases has led to adaptive introgression back into wild relatives[122]. One result of domestication is the formation of suites of co-inherited traits with distinct genomic signatures[17]. In natural populations, co-adaptation between social behaviours and dispersal[121] could frequently have been driven by correlational selection, resulting in increased genetic integration[25]. Artificial correlational selection to either strengthen[118] or eliminate genetic correlations[44] is a promising experimental approach in this context.

References

1. Wagner, G. P., Pavlicev, M. & Cheverud, J. M. The road to modularity. *Nat. Rev. Genet.* **8**, 921–931 (2007).
2. Wagner, G. P. & Altenberg, L. Perspective: Complex adaptations and the evolution of evolvability. *Evolution* **50**, 967–976 (1996).
3. Draghi, J. A. & Whitlock, M. C. Phenotypic plasticity facilitates mutational variance, genetic variance, and evolvability along the major axis of environmental variation. *Evolution* **66**, 2891–2902 (2012).
4. Cheverud, J. M. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* **110**, 155–171 (1984).
5. Phillips, P. C. & Arnold, S. J. Visualizing multivariate selection. *Evolution* **43**, 1209–1266 (1989).
6. Sinervo, B. & Svensson, E. Correlational selection and the evolution of genomic architecture. *Heredity* **16**, 948–955 (2002).
7. Blows, M. W. & Brooks, R. Measuring nonlinear selection. *Am. Nat.* **162**, 815–820 (2003).
8. Blows, M. W., Brooks, R. & Kraft, P. G. Exploring complex fitness surfaces: multiple ornamentation and polymorphism in male guppies. *Evolution* **57**, 1622–1630 (2003).
9. Jones, A. G., Arnold, S. J. & Bürger, R. Stability of the G-matrix in a population experiencing pleiotropic mutation, stabilizing selection, and genetic drift. *Evolution* **57**, 1747–1760 (2003).
10. Jones, A. G., Arnold, S. J. & Bürger, R. Evolution and stability of the G-matrix on a landscape with a moving optimum. *Evolution* **58**, 1639–1654 (2004).

11. Jones, A. G., Arnold, S. J. & Bürger, R. The mutation matrix and the evolution of evolvability. *Evolution* **61**, 727–745 (2007).

12. Jones, A. G., Bürger, R. & Arnold, S. J. Epistasis and natural selection shape the mutational architecture of complex traits. *Nat. Commun.* **5**, 3709 (2014).

13. Lande, R. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**, 203–215 (1980).

14. Armbruster, W. S., Pélabon, C., Hansen, T. F. & Mulder, C. P. H. Floral integration, modularity and accuracy: distinguishing complex adaptations from genetic constraints. in *Phenotypic integration: studying the ecology and evolution of complex phenotypes* (eds. Pigliucci, M. & Preston, K.) 23–49 (Oxford University Press, 2004).

15. Bell, A. M. & Sih, A. Exposure to predation generates personality in threespined sticklebacks (Gasterosteus aculeatus). *Ecol. Lett.* **10**, 828–834 (2007).

16. Dingemanse, N. J., Barber, I. & Dochtermann, N. A. Non-consumptive effects of predation: does perceived risk strengthen the genetic integration of behaviour and morphology in stickleback? *Ecol. Lett.* **23**, 107-118.

17. Hansen Wheat, C., Fitzpatrick, J. L., Rogell, B. & Temrin, H. Behavioural correlations of the domestication syndrome are decoupled in modern dog breeds. *Nat. Commun.* **10**, 1–9 (2019).

18. Hurst, L. D., Pál, C. & Lercher, M. J. The evolutionary dynamics of eukaryotic gene order. *Nat. Rev. Genet.* **5**, 299 (2004).

19. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).

20. Schluter, D. & Nychka, D. Exploring fitness surfaces. *Am. Nat.* **143**, 597–616 (1994).

21. Siepielski, A. M. *et al.* Precipitation drives global variation in natural selection. *Science* **355**, 959–962 (2017).

22. Roff, D. A. & Fairbairn, D. J. A test of the hypothesis that correlational selection generates genetic correlations. *Evolution* **66**, 2953–2960 (2012).

23. Svensson, E. I., McAdam, A. G. & Sinervo, B. Intralocus sexual conflict over immune defense, gender load, and sex-specific signaling in a natural lizard population. *Evolution* **63**, 3124–3135 (2009).

24. McGlothlin, J. W., Parker, P. G., Nolan, V. & Ketterson, E. D. Correlational selection leads to genetic integration of body size and an attractive plumage trait in dark-eyed juncos. *Evolution* **59**, 658–671 (2005).

25. Duckworth, R. A. & Kruuk, L. E. B. Evolution of genetic integration between dispersal and colonization ability in a bird. *Evolution* **63**, 968–977 (2009).

26. Brodie III, E. D. Correlational selection for color pattern and antipredator behavior in the garter snake Thamnophis ordinoides. *Evolution* **46**, 1284–1298 (1992).

27. Wise, M. J. & Rausher, M. D. Costs of resistance and correlational selection in the multiple-herbivore community of Solanum carolinense. *Evolution* **70**, 2411–2420 (2016).

28. Fenster, C. B., Reynolds, R. J., Williams, C. W., Makowsky, R. & Dudash, M. R. Quantifying hummingbird preference for floral trait combinations: The role of selection on trait interactions in the evolution of pollination syndromes. *Evolution* **69**, 1113–1127 (2015).

29. Arnegard, M. E. *et al.* Genetics of ecological divergence during speciation. *Nature* **511**, 307–311 (2014).

30. Martin, CH. & Wainwright, P. C. Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* **339**, 208–211 (2013).
31. Phillips, P. C. Epistasis - the essential role of gene interactions in the structure and evolution of genetic systems. *Nat.Rev.Gen.* **9**, 855–867 (2008).
32. Steppan, S. J., Phillips, P. C. & Houle, D. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* **17**, 320–327 (2002).
33. Blows, M. W. & McGuigan, K. The distribution of genetic variance across phenotypic space and the response to selection. *Mol. Ecol.* **24**, 2056–2072 (2015).
34. Schluter, D. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766–1774 (1996).
35. Lande, R. The maintenance of genetic variability by mutation in a polygenic character with linked loci. *Genet. Res.* **26**, 221–235 (1976).
36. Lande, R. The genetic correlation between characters maintained by selection, linkage and inbreeding. *Genet. Res.* **44**, 309–320 (1984).
37. Bulmer, M. G. The effect of selection on genetic variability: a simulation study. *Genet. Res.* **28**, 101–117 (1976).
38. Lynch, M. & Walsh, B. *Genetics and analysis of quantitative traits*. (Sinauer Associates, Inc., 1998).
39. Guillaume, F. & Whitlock, M. C. Effects of migration on the genetic covariance matrix. *Evolution* **61**, 2398–2409 (2007).
40. Noble, D. W. A., Radersma, R. & Uller, T. Plastic responses to novel environments are biased towards phenotype dimensions with high additive genetic variation. *Proc. Natl. Acad. Sci.* **116**, 13452–13461 (2019).

590 41. Houle, D., Bolstad, G. H., van der Linde, K. & Hansen, T. F. Mutation predicts 40 million
591 years of fly wing evolution. *Nature* **548**, 447–450 (2017).
592 42. Svensson, E. I. & Berger, D. The role of mutation bias in adaptive evolution. *Trends Ecol.*
593 *Evol.* **34**, 422–434 (2019).
594 43. Schweizer, G. & Wagner, A. Genotype networks of 80 quantitative Arabidopsis thaliana
595 phenotypes reveal phenotypic evolvability despite pervasive epistasis. *bioRxiv*
596 2020.02.26.966390 (2020) doi:10.1101/2020.02.26.966390.
597 44. Delph, L. F., Steven, J. C., Anderson, I. A., Herlihy, C. R. & Brodie, E. D., III. Elimination
598 of a genetic correlation between the sexes via artificial correlational selection. *Evolution* **65**,
599 2872–2880 (2011).
600 45. Conner, J. K. Genetic mechanisms of floral trait correlations in a natural population. *Nature*
601 **420**, 407–410 (2002).
602 46. Wagner, G. P. & Zhang, J. The pleiotropic structure of the genotype–phenotype map: the
603 evolvability of complex organisms. *Nat. Rev. Genet.* **12**, 204–213 (2011).
604 47. Barton, N. H., Etheridge, A. M. & Véber, A. The infinitesimal model: Definition,
605 derivation, and implications. *Theor. Popul. Biol.* **118**, 50–73 (2017).
606 48. Orr, H. A. The population genetics of adaptation: the distribution of factors fixed during
607 adaptive evolution. *Evolution* **52**, 935–948 (1998).
608 49. Flint, J. & Mackay, T. F. C. Genetic architecture of quantitative traits in mice, flies, and
609 humans. *Genome Res.* **19**, 723–733 (2009).
610 50. Stinchcombe, J. R., Weinig, C., Heath, K. D., Brock, M. T. & Schmitt, J. Polymorphic
611 genes of major effect: Consequences for variation, selection and evolution in *Arabidopsis*
612 *thaliana*. *Genetics* **182**, 911–922 (2009).

51. Orr, H. A. The genetics of species differences. *Trends Ecol. Evol.* **16**, 343–350 (2001).
52. Nadeau, N. J. *et al.* The gene cortex controls mimicry and crypsis in butterflies and moths. *Nature* **534**, 106–110 (2016).
53. Visscher, P. M. *et al.* 10 years of GWAS discovery: Biology, function, and translation. *Am. J. Hum. Genet.* **101**, 5–22 (2017).
54. Pitchers, W. *et al.* A multivariate genome-wide association study of wing shape in *Drosophila melanogaster*. *Genetics* **211**, 1429–1447 (2019).
55. Mackay, T. F. C. Epistasis and quantitative traits: using model organisms to study gene–gene interactions. *Nat. Rev. Genet.* **15**, 22–33 (2014).
56. Sailer, Z. R. & Harms, M. J. Detecting high-order epistasis in nonlinear genotype-phenotype maps. *Genetics* **205**, 1079–1088 (2017).
57. Hill, W. G. "Conversion" of epistatic into additive genetic variance in finite populations and possible impact on long-term selection response. *J. Anim. Breed. Genet.* **134**, 196–201 (2017).
58. Gienapp, P. *et al.* Genomic quantitative genetics to study evolution in the wild. *Trends Ecol. Evol.* **32**, 897–908 (2017).
59. Nosil, P. *et al.* Ecology shapes epistasis in a genotype–phenotype–fitness map for stick insect colour. *Nat. Ecol. Evol.* 1–12 (2020) doi:10.1038/s41559-020-01305-y.
60. Blount, Z. D., Lenski, R. E. & Losos, J. B. Contingency and determinism in evolution: Replaying life's tape. *Science* **362**, (2018).
61. Bolnick, D. I., Barrett, R. D. H., Oke, K. B., Rennison, D. J. & Stuart, Y. E. (Non)Parallel Evolution. *Annu. Rev. Ecol. Evol. Syst.* **49**, 303–330 (2018).

62. Therkildsen, N. O. *et al.* Contrasting genomic shifts underlie parallel phenotypic evolution in response to fishing. *Science* **365**, 487–490 (2019).

63. Stern, D. L. & Orgogozo, V. Is genetic evolution predictable? *Science* **323**, 746–751 (2009).

64. Colosimo, P. F. *et al.* Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928–1933 (2005).

65. Archambeault, S. L., Bärtschi, L. R., Merminod, A. D. & Peichel, C. L. Adaptation via pleiotropy and linkage: Association mapping reveals a complex genetic architecture within the stickleback Eda locus. *Evol. Lett.* **4**, 282–301 (2020).

66. van Rheenen, W., Peyrot, W. J., Schork, A. J., Lee, S. H. & Wray, N. R. Genetic correlations of polygenic disease traits: from theory to practice. *Nat. Rev. Genet.* **20**, 567–581 (2019).

67. Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W. & Smadja, C. M. Variation in recombination frequency and distribution across eukaryotes: patterns and processes. *Philos. Trans. R. Soc. B Biol. Sci.* **372**, 20160455 (2017).

68. Choudhury, R. R., Rogivue, A., Gugerli, F. & Parisod, C. Impact of polymorphic transposable elements on linkage disequilibrium along chromosomes. *Mol. Ecol.* **28**, 1550–1562 (2019).

69. Thompson, M. J. & Jiggins, C. D. Supergenes and their role in evolution. *Heredity* **113**, 1–8 (2014).

70. Yeaman, S. Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E1743-1751 (2013).

71. Faria, R., Johannesson, K., Butlin, R. K. & Westram, A. M. Evolving inversions. *Trends Ecol. Evol.* **34**, 239–248 (2019).

72. Tuttle, E. M. *et al.* Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350 (2016).

73. Kupper, C. *et al.* A supergene determines highly divergent male reproductive morphs in the ruff. *Nat. Genet.* **48**, 79–83 (2016).

74. Lamichhaney, S. *et al.* Structural genomic changes underlie alternative reproductive strategies in the ruff (*Philomachus pugnax*). *Nat.Genet.* **48**, 84–88 (2016).

75. Huu, C. N., Keller, B., Conti, E., Kappel, C. & Lenhard, M. Supergene evolution via stepwise duplications and neofunctionalization of a floral-organ identity gene. *Proc. Natl. Acad. Sci.* **117**, 23148–23157 (2020).

76. Merrill, R. M. *et al.* Genetic dissection of assortative mating behavior. *PLOS Biol.* **17**, e2005902 (2019).

77. Whitlock, M. C., Phillips, P. C., Moore, F. B.-G. & Tonsor, S. J. Multiple fitness peaks and epistasis. *Annu. Rev. Ecol. Syst.* **26**, 601–629 (1995).

78. Dudley, S. A. The response to selection on plant physiological traits: evidence for local adaptation. *Evolution* **50**, 103–110 (1996).

79. Kirkpatrick, M. & Ravigné, V. Speciation by natural and sexual selection: Models and experiments. *Am. Nat.* **159**, S22–S35 (2002).

80. Hohenlohe Paul A., Bassham Susan, Currey Mark & Cresko William A. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philos. Trans. R. Soc. B Biol. Sci.* **367**, 395–408 (2012).

81. Hench, K., Vargas, M., Höppner, M. P., McMillan, W. O. & Puebla, O. Inter-chromosomal coupling between vision and pigmentation genes during genomic divergence. *Nat. Ecol. Evol.* **3**, 657 (2019).
82. Gienapp, P., Calus, M. P. L., Laine, V. N. & Visser, M. E. Genomic selection on breeding time in a wild bird population. *Evol. Lett.* **3**, 142–151 (2019).
83. McGuigan, K., Collet, J. M., Allen, S. L., Chenoweth, S. F. & Blows, M. W. Pleiotropic mutations are subject to strong stabilizing selection. *Genetics* **197**, 1051-+ (2014).
84. McGuigan, K. *et al.* The nature and extent of mutational pleiotropy in gene expression of male *Drosophila serrata*. *Genetics* **196**, 911-+ (2014).
85. Hine, E., Runcie, D. E., McGuigan, K. & Blows, M. W. Uneven distribution of mutational variance across the transcriptome of *Drosophila serrata* revealed by high-dimensional analysis of gene expression. *Genetics* (2018) doi:10.1534/genetics.118.300757.
86. Estes, S., Ajie, B. C., Lynch, M. & Phillips, P. C. Spontaneous mutational correlations for life-history, morphological and behavioral characters in *Caenorhabditis elegans*. *Genetics* **170**, 645–653 (2005).
87. Houle, D. & Fierst, J. Properties of spontaneous mutational variance and covariance for wing size and shape in *Drosophila melanogaster*. *Evolution* **67**, 1116–1130 (2013).
88. Ovaskainen, O., Karhunen, M., Zheng, C., Arias, J. M. C. & Merilä, J. A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics* **189**, 621–632 (2011).
89. Csilléry, K. *et al.* Adaptation to local climate in multi-trait space: evidence from silver fir (Abies alba Mill.) populations across a heterogeneous environment. *Heredity* **124**, 77–92 (2020).

90. Berg, J. J. & Coop, G. A Population genetic signal of polygenic adaptation. *PLOS Genet.* **10**, e1004412 (2014).
91. Orr, H. A. Adaptation and the cost of complexity. *Evolution* **54**, 13–20 (2000).
92. Fisher, R. A. *The Genetical Theory of Natural Selection*. (Clarendon Press, 1930).
93. Pavlicev, M. & Hansen, T. F. Genotype-phenotype maps maximizing evolvability: modularity revisited. *Evol. Biol.* **38**, 371–389 (2011).
94. Hine, E., McGuigan, K. & Blows, M. W. Evolutionary constraints in high-dimensional trait sets. *Am. Nat.* **184**, 119–131 (2014).
95. Melo, D. & Marroig, G. Directional selection can drive the evolution of modularity in complex traits. *Proc. Natl. Acad. Sci.* **112**, 470–475 (2015).
96. Kashtan, N. & Alon, U. Spontaneous evolution of modularity and network motifs. *Proc. Natl. Acad. Sci.* **102**, 13773–13778 (2005).
97. Espinosa-Soto, C. & Wagner, A. Specialization can drive the evolution of modularity. *PLoS Comput. Biol.* **6**, e1000719 (2010).
98. Ancel, L. W. & Fontana, W. Evolutionary lock-in and the origin of modularity in RNA structure. in *Modularity: Understanding the Development and Evolution of Natural Complex Systems* (eds. Callebaut, W. & Rasskin-Gutman, D.) 129–141 (The MIT Press, 2009).
99. Wagner, G. P. & Mezey, J. G. The role of genetic architecture constraints in the origin of variational modularity. in *Modularity in Development and Evolution* (eds. Schlosser, G. & Wagner, G. P.) 338–358 (University of Chicago Press, 2004).
100. Fokkens, L. & Snel, B. Cohesive versus flexible evolution of functional modules in Eukaryotes. *PLOS Comput. Biol.* **5**, e1000276 (2009).

101. Huang, W. *et al.* Genetic basis of transcriptome diversity in Drosophila melanogaster. *Proc. Natl. Acad. Sci.* **112**, E6010–E6019 (2015).

102. Schweizer, R. M. *et al.* Physiological and genomic evidence that selection on the transcription factor Epas1 has altered cardiovascular function in high-altitude deer mice. *PLOS Genet.* **15**, e1008420 (2019).

103. Hämälä, T. *et al.* Gene expression modularity reveals footprints of polygenic adaptation in *Theobroma cacao*. *Mol. Biol. Evol.* **37**, 110–123 (2020).

104. Collet, J. M., McGuigan, K., Allen, S. L., Chenoweth, S. F. & Blows, M. W. Mutational Pleiotropy and the strength of stabilizing selection within and between functional modules of gene expression. *Genetics* **208**, 1601–1616 (2018).

105. Jiménez, A., Cotterell, J., Munteanu, A. & Sharpe, J. A spectrum of modularity in multi-functional gene circuits. *Mol. Syst. Biol.* **13**, (2017).

106. Verd, B., Monk, N. A. & Jaeger, J. Modularity, criticality, and evolvability of a developmental gene regulatory network. *eLife* **8**, e42832 (2019).

107. Pallares, L. F. *et al.* Mapping of craniofacial traits in outbred mice identifies major developmental genes involved in shape determination. *PLOS Genet.* **11**, e1005607 (2015).

108. Arnold, S. J., Pfrender, M. E. & Jones, A. G. The adaptive landscape as a conceptual bridge between micro- and macroevolution. *Genetica* **112–113**, 9–32 (2001).

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

110. Shikov, A. E., Skitchenko, R. K., Predeus, A. V. & Barbitoff, Y. A. Phenome-wide functional dissection of pleiotropic effects highlights key molecular pathways for human complex traits. *Sci. Rep.* **10**, 1–10 (2020).

111. Sella, G. & Barton, N. H. Thinking About the evolution of complex traits in the era of genome-wide association studies. *Annu. Rev. Genomics Hum. Genet.* **20**, 461–493 (2019).
112. Walsh, B. & Blows, M. W. Abundant genetic variation plus strong selection = Multivariate genetic constraints: A geometric view of adaptation. *Annu. Rev. Ecol. Evol. Syst.* **40**, 41–59 (2009).
113. Teplitsky, C. *et al.* Assessing multivariate constraints to evolution across ten long-term avian studies. *PLOS ONE* **9**, e90444 (2014).
114. Pavlicev, M. & Cheverud, J. M. Constraints evolve: context dependency of gene effects allows evolution of pleiotropy. *Ann. Rev. Ecol. Evol. Syst.* **46,** 413–434 (2015).
115. Wei, X. & Zhang, J. Environment-dependent pleiotropic effects of mutations on the maximum growth rate r and carrying capacity K of population growth. *PLOS Biol.* **17**, e3000121 (2019).
116. Parter, M., Kashtan, N. & Alon, U. Facilitated variation: How evolution learns from past environments to generalize to new Environments. *PLOS Comput. Biol.* **4**, e1000206 (2008).
117. Lande, R. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* **33**, 402–416 (1979).
118. Bolstad, G. H. *et al.* Complex constraints on allometry revealed by artificial selection on the wing of Drosophila melanogaster. *Proc. Natl. Acad. Sci., USA.* **112**, 13284–13289 (2015).
119. Tsuboi, M. *et al.* Breakdown of brain–body allometry and the encephalization of birds and mammals. *Nat. Ecol. Evol.* **2**, 1492–1500 (2018).
120. White, C. R. *et al.* The origin and maintenance of metabolic allometry in animals. *Nat. Ecol. Evol.* **3**, 598 (2019).

121. Mullon, C., Keller, L. & Lehmann, L. Social polymorphism is favoured by the co-evolution of dispersal with social behaviour. *Nat. Ecol. Evol.* **2**, 132–140 (2018).
122. Schweizer, R. M. *et al.* Natural selection and origin of a melanistic allele in North American gray wolves. *Mol. Biol. Evol.* **35**, 1190–1209 (2018).
123. Brodie III, E. D. Genetic correlations between morphology and antipredator behaviour in natural populations of the garter snake *Thamnophis ordinoides*. *Nature* **342**, 542–543 (1989).
124. Auinger, H.-J. *et al.* Model training across multiple breeding cycles significantly improves genomic prediction accuracy in rye (Secale cereale L.). *Theor. Appl. Genet.* **129**, 2043–2053 (2016).
125. Xie, K. T. *et al.* DNA fragility in the parallel evolution of pelvic reduction in stickleback fish. *Science* **363**, 81–84 (2019).
126. Slate, J. INVITED REVIEW: Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol.* **14**, 363–379 (2005).
127. Brieuc, M. S. O., Waters, C. D., Drinan, D. P. & Naish, K. A. A practical introduction to Random Forest for genetic association studies in ecology and evolution. *Mol. Ecol. Resour.* **18**, 755–766 (2018).
128. Nielsen, R. Molecular Signatures of Natural Selection. *Annu. Rev. Genet.* **39**, 197–218 (2005).
129. Barghi, N., Hermisson, J. & Schlötterer, C. Polygenic adaptation: a unifying framework to understand positive selection. *Nat. Rev. Genet.* 1–13 (2020) doi:10.1038/s41576-020-0250-z.

130. Lemos, B., Araripe, L. O. & Hartl, D. L. Polymorphic Y chromosomes harbor cryptic variation with manifold functional consequences. *Science* **319**, 91–93 (2008).
131. Haddad, R., Meter, B. & Ross, J. A. The genetic architecture of intra-species hybrid mito-nuclear epistasis. *Front. Genet.* **9**, 481 (2018).
132. Long, A., Liti, G., Luptak, A. & Tenaillon, O. Elucidating the molecular architecture of adaptation via evolve and resequence experiments. *Nat. Rev. Genet.* **16**, 567–582 (2015).
133. Bollback, J. P., York, T. L. & Nielsen, R. Estimation of 2Nes from temporal allele frequency data. *Genetics* **179**, 497–502 (2008).
134. Marques, D. A., Jones, F. C., Di Palma, F., Kingsley, D. M. & Reimchen, T. E. Experimental evidence for rapid genomic adaptation to a new niche in an adaptive radiation. *Nat. Ecol. Evol.* **2**, 1128–1138 (2018).
135. Svensson, E. I. & Calsbeek, R. *The Adaptive Landscape in Evolutionary Biology*. (Oxford University Press, 2012).
136. Hämälä, T., Gorton, A. J., Moeller, D. A. & Tiffin, P. Pleiotropy facilitates local adaptation to distant optima in common ragweed (Ambrosia artemisiifolia). *PLOS Genet.* **16**, e1008707 (2020).
137. Roda, F., Walter, G. M., Nipper, R. & Ortiz-Barrientos, D. Genomic clustering of adaptive loci during parallel evolution of an Australian wildflower. *Mol. Ecol.* **26**, 3687–3699 (2017).

Acknowledgements

We are grateful to Debora Goedert for comments on the first draft of this manuscript. E.I.S. and A.R. were funded by grants from the Swedish Research Council (VR: grant no. 2016-03356 and 2018-04537, respectively). D. M. was supported by the Swiss National Science Foundation grant #31003A_163338 to Ole Seehausen, Laurent Excoffier and Rémy Bruggmann. J.A.D. acknowledges support by NSF 13-510 Systems & Synthetic Biology, award #1714550. J.M.H. is supported by the German Federal Ministry of Education and Research (BMBF). K.C. is supported by an SNF grant (CRSK-3_190288). K.M. was funded by the Australian Research Council (DP190101661). M.N.S. was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), projects 2015/19556-4 and 2016/22159-0. The authors also wish to thank Butch Brodie for kindly providing us with the photograph of the garter snakes in Fig. 2A, and the original figure of his classic fitness surface in Box 1.

Author contributions

E.I.S. and A.R. conceived of the paper, organized the writing and put together the first draft. All authors contributed to writing, improving and finalizing the manuscript.

Competing interests

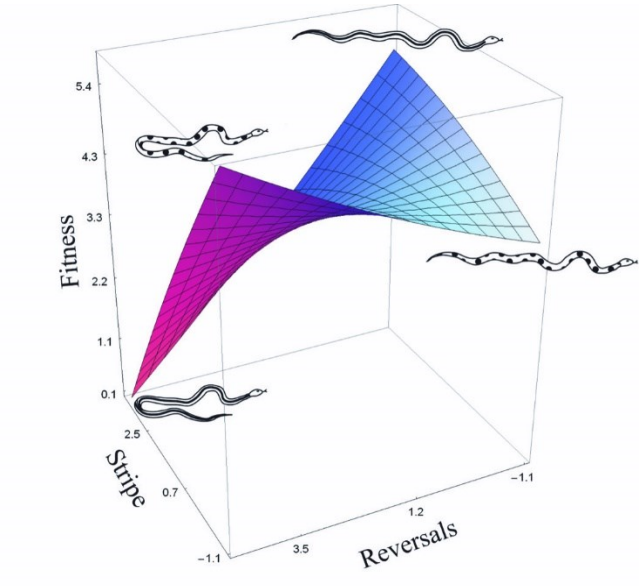
The authors declare no competing interests.

Box 1: What is correlational selection?

Correlational selection involves several interrelated concepts. Here, we define the most important terms.

The individual fitness surface

Imagine a function relating an individual's trait values to that individual's expected lifetime fitness (Fig. 3). Supposing fitness depends on two traits, we can depict the fitness function as a three-dimensional surface. The two horizontal axes represent trait values and elevation represents fitness. Fitness peaks and valleys represent regions in trait space with high and low fitness, respectively. Such a surface will reveal regions where favorable trait combinations produce high fitness, as well as unfavorable trait combinations that confer low fitness. Individual fitness surfaces can take almost any shape, including single-peaked surfaces, multi-peaked surfaces or ridges (Fig. 3), and can involve any number of traits.



Brodie's pioneering study[26] of coloration and behavior in garter snakes provided one of the first empirical examples of how such individual fitness surfaces can illustrate correlational selection in a natural population (see inset). A snake's color pattern could be either blotched or striped.

862 Moreover, snakes either crawled in a straight line or reversed directions repeatedly when evading
863 predators. Striped snakes had higher fitness when they fled predators in a straight line, whereas
864 blotched snakes had higher fitness when they reversed directions. Therefore, survival depended on
865 the interaction between two types of traits – colour and behavior – rather than on single traits.
866 Interestingly, colouration and behavior were also genetically correlated with each other[123],
867 providing empirical support for the prediction[4,13] that correlational selection can promote and
868 maintain genetic correlations.
869
870 *An operational definition of correlational selection*
871 Correlational selection occurs when the relationship between an individual's trait value and
872 expected fitness for one trait depends on that individual's trait values for other traits, and direct
873 selection acts in such a way as to establish or maintain genetic and hence phenotypic correlations
874 among traits[6]. One way to think about correlational selection is to imagine slicing the fitness
875 surface parallel to one of the trait axes. If the slices differ in shape as we proceed along the fitness
876 surface (Fig. 3A), then fitness is the result of interactions between traits.
877
878 Lande and Arnold[19] showed that correlational selection could be measured by using simple
879 regression approaches. If we assume that traits have a multivariate normal distribution, then the
880 fitness surface can be estimated by a regression of the form (in the bivariate case):
881
882 $$w(z_1, z_2) = \alpha + \beta_1 z_1 + \beta_2 z_2 + \tfrac{1}{2}\gamma_{11} z_1^2 + \tfrac{1}{2}\gamma_{22} z_2^2 + \gamma_{12} z_1 z_2 + \varepsilon, \qquad (1)$$
883
884 where α is an intercept, z_1 and z_2 are trait values, and ε is a residual term. The parameters β_1 and
885 β_2 are the *linear selection gradients*, which estimate directional selection on each trait. The matrix
886 of *quadratic selection coefficients* $\boldsymbol{\gamma} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{12} & \gamma_{22} \end{bmatrix}$ estimates stabilizing, disruptive and

887 correlational selection. The diagonal elements of $\boldsymbol{\gamma}$ measure quadratic selection on each trait (i.e.,
888 stabilizing or disruptive selection), whereas the off-diagonal elements represent correlational
889 selection. Thus, non-zero off-diagonal elements of $\boldsymbol{\gamma}$ constitute evidence of correlational selection.
890
891 *Fitness epistasis and epistatic selection*
892 Much of the fitness variation in complex phenotypic traits originates from allelic variation at
893 individual loci, each with small fitness effects. Favourable trait combinations at the organismal
894 level often also reflect favourable allelic combinations at separate sets of loci. At the genomic
895 level, correlational selection occurs when the fitness effects of a particular locus depend on the
896 genotype at another locus or, more generally, depend on the genetic background. This situation is
897 often described as fitness epistasis or epistatic selection and is likely to have a big impact on
898 genome evolution[6,18].
899
900
901
902
903
904
905
906
907
908
909
910
911

Box 2. Methods to study genomic signatures of correlational selection

Genomics can inform quantitative genetics

Due to the highly polygenic nature of most traits, quantitative genetics is a pragmatic method to predict short term evolutionary change in phenotypes[58]. Genomic tools can however be integrated with quantitative genetics methodology to expand our understanding[38,58,82]. For example, the so-called GBLUP approach[82] allows the pedigree-relatedness matrix of an "animal model" to be replaced by a marker-based relatedness matrix to infer genetic variances and covariances, i. e. **G**. By accurately determining the proportion of genome shared, such genomic approaches may improve the estimates of **G** compared to using pedigree data alone, where relatedness is based on a shallow pedigree[124].

Genomic approaches can also provide information about mutation rates of SNPs and indel variants, thereby improving our understanding of the role of mutation rates in evolution[125] and the importance of mutational pleiotropy and **M**-matrix evolution[83–85]. Of particular interest is the effects of new mutations on genetic variances and covariances, i. e. **M**[83–85]. A promising approach is the combination of mutation accumulation experiments (MA) with estimates of **M** and **G**[85–87]. Studies on MA-lines have revealed strong mutational pleiotropy across the transcriptome[85]. Such strong mutational pleiotropy in **M** contrasts with weaker pleiotropy in **G**, suggesting that correlational selection operates against maladaptive strong mutational covariance, which results in a weakening of pleiotropy during the course of the life cycle[83,84].

To quantify outcomes of correlational selection, we need to identify the genetic loci under such selection. Traditionally, this has been achieved by quantitative trait loci (QTL) mapping, admixture mapping and genome wide association studies (GWAS)[126] which have limited power to detect small effect size genes. Newer approaches map pleiotropy by simultaneously associating genomic loci with multiple traits[54] and can also detect epistatic interactions using machine learning algorithms[127].

Detecting the genomic signatures of correlational selection

Correlational selection could potentially be inferred from signatures of selective sweeps at loci under strong selection[128] or, for highly polygenic traits, allele frequency shifts that are not explainable by genetic drift[90,129]. Selection on polygenic traits often leads to small frequency changes in many genes, which is more difficult to detect[129]. Since correlational selection favors certain allele combinations, one outcome is deviations from Hardy-Weinberg equilibrium and the build-up of linkage disequilibrium between alleles at unlinked loci, detectable both across individuals and between age classes within populations. Genomic data may also indicate whether recombination suppression leading to trait correlations[67], such as in supergenes or genomic rearrangements, has been favored by correlational selection. On longer time scales, genomic data can reveal how such supergenes are gradually built up and assembled ~~via gene duplications and neofunctionalization~~ [75]. Experimental assays such as introgression lines[130] or reciprocal crosses of diverged lineages[131] can be used to confirm whether combinations of alleles or genomic regions are under correlational selection. Evolve and re-sequence experiments comparing populations

955 before and after selection[132], or studies of allele frequency time series during an experiment[133] can
956 give further, detailed insight into allelic interactions and both genomic and phenotypic responses
957 to experimentally imposed selection[62,134].

958

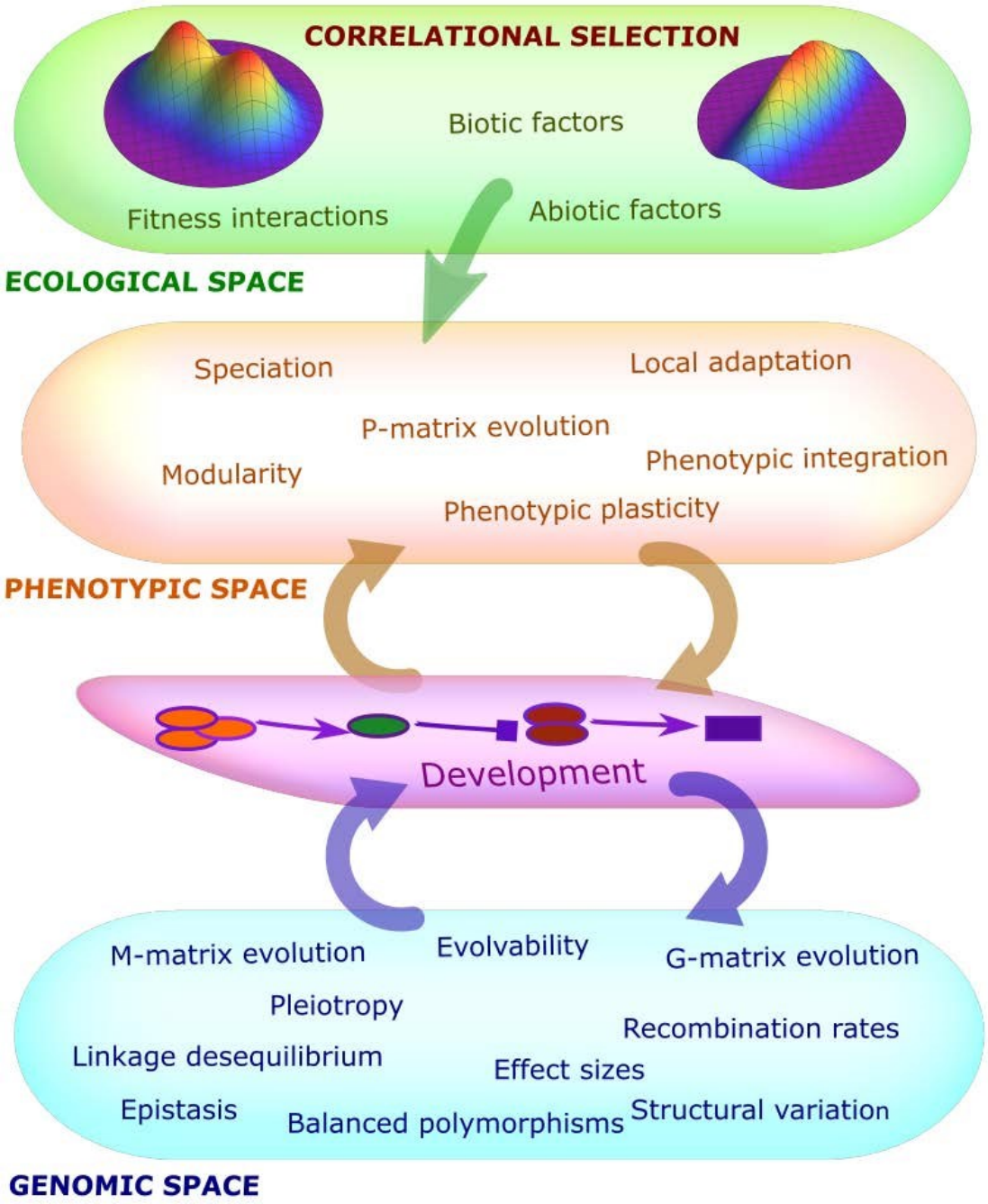
959 *Bridging the genotype-phenotype-fitness map*

960 Ideally, the genotypic and phenotypic levels should be studied alongside the adaptive
961 landscape[108,135] and integrated into a genotype-phenotype-fitness map. This integration has been
962 achieved for very few non-model organisms such as threespine stickleback[29], Bahama pupfish[30]
963 and *Tinema* stick insects[59] in which the fitness landscape was mapped experimentally with
964 information about the genomic architecture of traits. Experimental field studies on fitness epistasis
965 in natural populations combined with genomic data is a promising integrative approach to detect
966 the genomic consequences of correlational selection[59].

967

968

969 **Legends to figures**



970

971 **Figure 1. The scope of correlational selection and its links to different fields in evolutionary**
972 **biology.** Correlational selection is relatively well-understood statistically and theoretically (Box
973 1), but we still do not know its prevalence in natural populations and the extent to which it has
974 shaped genome evolution in diverse organisms. In a few cases, correlational selection has been
975 studied and documented in natural field populations and in laboratory artificial experimental
976 studies (Fig. 2). The main effects of correlational selection are to strengthen or reinforce
977 phenotypic and/or genetic correlations between traits[6,22,23,26], which may be governed by separate
978 sets of loci, or to break up non-adaptive or maladaptive genetic correlations, such as between the
979 sexes[44]. These effects of correlational selection on phenotypic and potentially also genetic
980 correlation structure have consequences for several organismal-level phenomena that are of great
981 interest in evolutionary genetics and developmental biology. These include **G**-matrix evolution,
982 phenotypic plasticity, modularity, evolvability and phenotypic integration (upper part of figure),
983 as discussed in this review. Theory suggests that correlational selection at the organismal level can
984 potentially drive the downstream evolution of genomic architecture[4,6,18] (lower part of figure),
985 although here our knowledge is more limited. Correlational selection could preserve adaptive
986 genetic correlations between traits that are governed by different sets of loci by suppressing
987 recombination rates, thereby maintaining inversion polymorphisms and other structural genomic
988 variation that is often associated with balanced genetic polymorphisms (Fig. 4). In addition,
989 correlational selection could lead to adaptive pleiotropy, such as during range expansions when
990 populations are far away from their adaptive peaks[136], and could shape patterns of epistasis
991 between loci[12]. Finally, correlational selection is likely to be involved in local adaptation, if
992 different sets of character combinations are favoured in different abiotic[137] or biotic

993 environments[27,28], but the consequences for speciation and other aspects of macroevolution remain
994 largely unexplored.

995

996

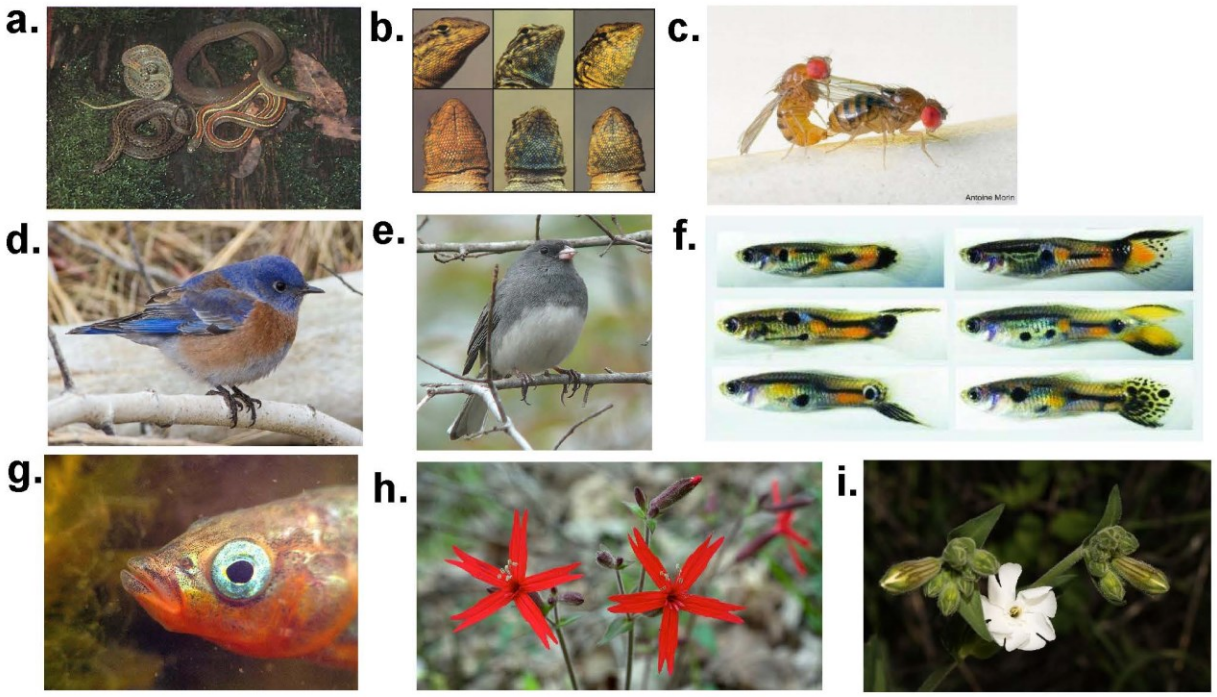
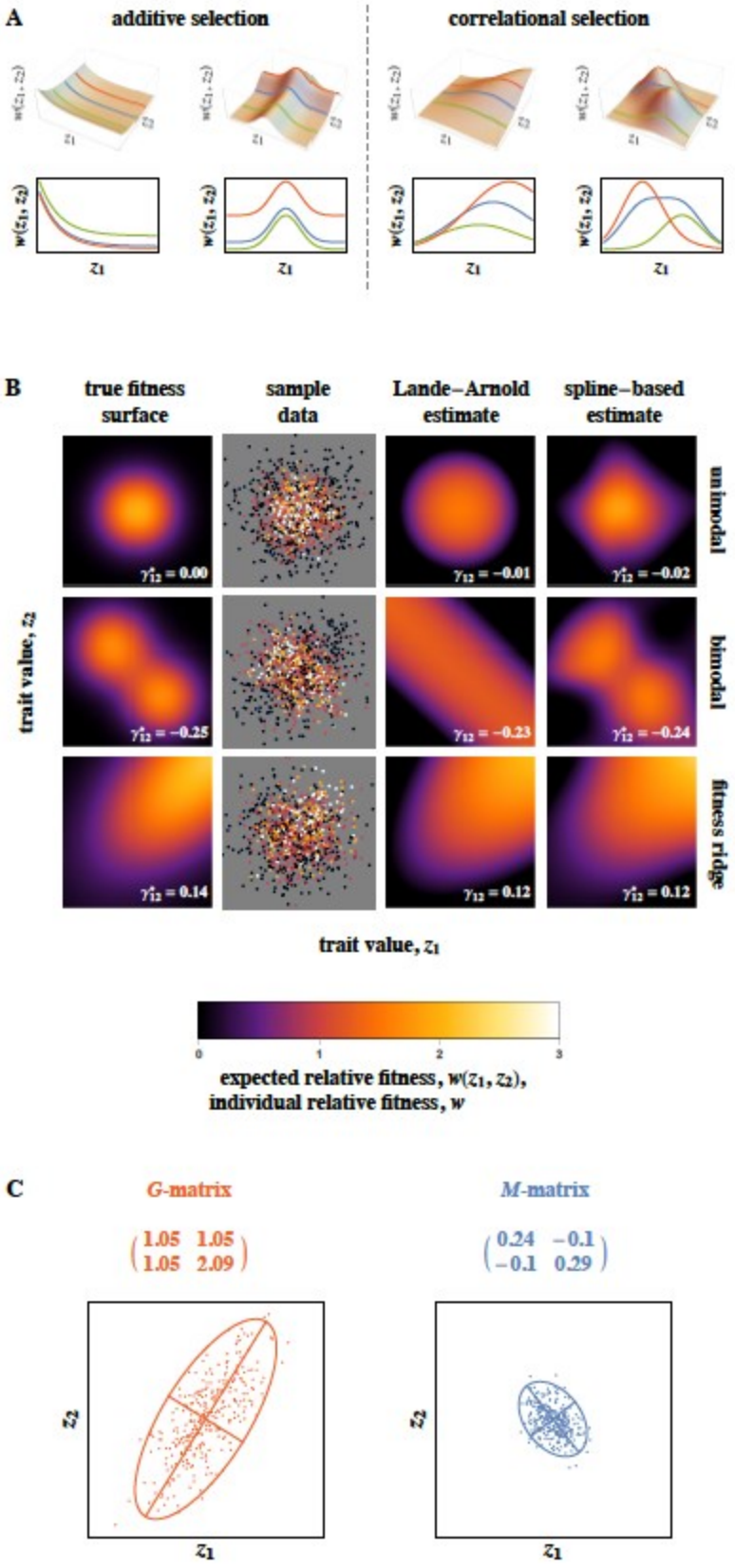


Figure 2. Phenotypic and quantitative genetics studies on organisms and traits in which correlational selection has experimentally been demonstrated or inferred in the field or in the laboratory. A. Northwestern garter snake (*Thamnophis ordinoides*). **B.** Side-blotched lizard (*Uta stansburiana*). **C.** Australian fruit fly (*Drosophila serrata*). **D.** Western bluebird (*Sialia mexicana*). **E.** Dark-eyed junco (*Junco hyemalis*). **F.** Guppy (*Poecilia reticulata*). **G.** Three-spined stickleback (*Gasterosteus aculeatus*). **H.** Fire pink (*Silene virginica*). **I.** White campon (*Silene latifolia*). Correlational selection has been demonstrated and quantified for a number of different traits, including both discrete colour polymorphisms[6,8,23,26] (**A**, **B**, **F**) and continuous, quantitative characters[15,24,25,28,44,94] (**C, D-E, G-I**), both in animals and in plants. The ecological causes and selective agents driving such correlational selection have been shown to be predators (**A,G**), interspecific mutualists such as pollinators (**H**) and conspecific interactions, especially under sexual selection (**B-C, E-F**). In some of these studies, the phenotypic traits that were found or implicated to be under correlational selection were also genetically or phenotypically correlated

with each other (**A-B, D-E**), suggesting that correlational selection can build up, promote or strengthen genetic integration between the traits in question. Conversely, artificial correlational selection has been demonstrated to be efficient in breaking up an intersexual genetic correlation in at least one case (**I**). Finally, traits that have been found to experience correlational selection include visual colouration traits (**A,B,E,F**), chemical communication traits (**C**), behavioural traits such as dispersal, aggression and personality (**D,G**) and structural traits such as size and shape (**H**). Photo credits: **A:** Butch Brodie. **B.** Barry Sinervo. **C-I:** Public domain. **C.** Antoine Morin: **https://www.eurekalert.org/multimedia/pub/94488.php**. **D.** Wikimedia Commons: **https://en.wikipedia.org/wiki/Dark-eyed_junco#/media/File:Dark-eyed_Junco-27527.jpg** **E.** Wikimedia Commons (Ken Thomas): **https://commons.wikimedia.org/wiki/File:Dark-eyed_Junco-27527.jpg** **F.** Kimberly Hughes/Nature: **https://www.nature.com/articles/nature12717/figures/1?draft=collection** **G.** Wikimedia Commons (Piet Spaans): **https://commons.wikimedia.org/wiki/File:GasterosteusAculeatusMaleHead.JPG** **H.** Wikimedia Commons (Eric Hunt): **https://en.wikipedia.org/wiki/Silene_virginica#/media/File:Silene_virginica_Arkansas.jpg** **I.** Wikimedia Commons (Walter Siegmund): **https://en.wikipedia.org/wiki/Silene_latifolia#/media/File:Silene_latifolia_9631.JPG**


A
additive selection
correlational selection
$w(z_1, z_2)$
z_1
z_2
B
true fitness surface
sample data
Lande–Arnold estimate
spline–based estimate
unimodal
bimodal
fitness ridge
trait value, z_2
trait value, z_1
$\gamma^*_{12} = 0.00$
$\gamma_{12} = -0.01$
$\gamma^*_{12} = -0.02$
$\gamma^*_{12} = -0.25$
$\gamma_{12} = -0.23$
$\gamma^*_{12} = -0.24$
$\gamma^*_{12} = 0.14$
$\gamma_{12} = 0.12$
$\gamma^*_{12} = 0.12$
0
1
2
3
expected relative fitness, $w(z_1, z_2)$,
individual relative fitness, w
C
G-matrix
M-matrix
$\begin{pmatrix} 1.05 & 1.05 \\ 1.05 & 2.09 \end{pmatrix}$
$\begin{pmatrix} 0.24 & -0.1 \\ -0.1 & 0.29 \end{pmatrix}$
z_2
z_1

1033

Figure 3. **Illustration of correlational selection, along with important parameters used to quantify it and determine how its effects are carried across generations. A.** Example fitness surfaces for hypothetical traits z_1 and z_2 (top row) and conditional fitness curves for z_1 given fixed values of z_2 (colored lines in both rows). **A.** When selection is additive, the fitness effects of z_1 are independent of the value of z_2. The conditional fitness curves are then identical aside from their height above the trait axes (bottom row, left of the dashed line). Under correlational selection, in contrast, the fitness effects of z_1 depend on z_2, and so the shape of the conditional fitness curves changes with value of z_2 (bottom row, right of the dashed line). **B**. Estimation of multivariate fitness surfaces from samples of individual trait values, z_1 and z_2, and relative fitness, w. The true fitness surface (first column) is unobservable but can be sampled by measuring the relative fitness of individuals in a population (second column). The true surface can then be estimated via quadratic regression (third column) or by non-parametric smooth splines (fourth column). See Section 3 in Supplementary Material for full details. **C**. The $\boldsymbol{G}$-matrix (orange, left) is the variance-covariance matrix of additive genetic effects (i.e. breeding values) for a multivariate phenotype. The $\boldsymbol{M}$-matrix (blue, right) is the variance-covariance matrix of additive mutational effects. Points represent individual breeding values (orange) and additive mutational effects (blue) respectively. If the distribution of point values is multivariate normal, it can be summarized via an ellipsoid. The principal axes of the ellipsoid (crossed lines) align with the eigenvectors and their lengths are proportional to the square roots of the eigenvalues. The major axis, associated with the largest eigenvalue, indicates the direction of maximum additive genetic or mutational variance.

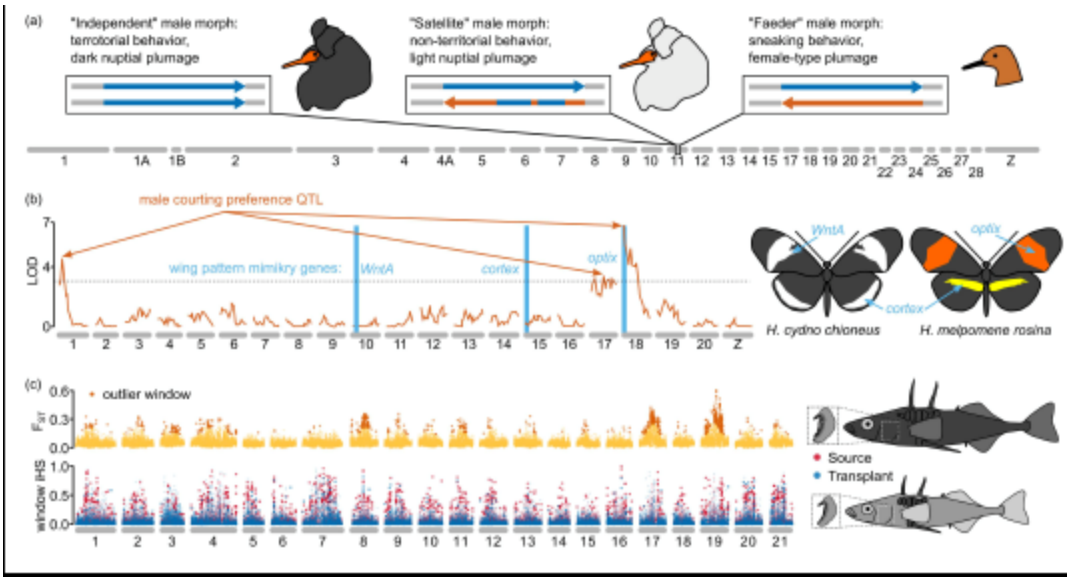


Figure 4. Examples of genomic trait architectures that might reflect past or ongoing correlational selection. We focus here on empirical examples where multiple loci are involved in the adaptive traits in question, as these reflect the most challenging situations to maintain adaptive genetic correlations between traits, due to the eroding effects of recombination when traits are governed by multiple unlinked loci. However, we underscore that correlational selection could equally well lead to the evolution of adaptive pleiotropy[136] as an alternative mechanism to maintain adaptive genetic correlations between traits. **A.** A complex mating polymorphism in male ruff (*Philomachus pugnax*) reproductive tactics involves multiple correlated morphological and behavioral traits, and the different character combinations in the male morphs are preserved because of the lack of recombination between different loci that are held together in a single large chromosomal inversion[73,74].**B.** Assortative mating maintains linkage disequilibrium between unlinked color pattern loci under correlational selection in *Heliconius* butterfly species, facilitated by tight linkage between preference and trait loci on one chromosome[76]. **C.** In a multifarious selection experiment on threespine sticklebacks (*Gasterosteus aculeatus*), the predicted phenotypic changes in multiple traits were caused by widespread underlying genomic changes that could potentially be attributed to correlational selection for different character combinations in the different phenotypes[134].