Evo Edu Outreach (2012) 5:231–244
DOI 10.1007/s12052-012-0430-1

EVO-DEVO

Coming to Grips with Evolvability

Mihaela Pavlicev · Günter P. Wagner

Published online: 7 July 2012

Abstract To explain the evolution of complex organisms by random mutation, drift, and selection is not a trivial task. This becomes obvious if we imagine an organism in which most genes affect most traits and all mutations are immediately expressed in the phenotype. Most of the mutations will be deleterious. Computer programmers experienced a similar problem when trying to evolve computer programs by introducing random changes to a conventional computer code, realizing that almost all random changes are "lethal." Everyone who has done any programming knows that conventional computer languages are very brittle! Real organisms are not organized in this way but rather involve mediation between the genes and the phenotypic traits, namely development, also sometimes called the genotype–phenotype map. This map of genetic effects is structured in a way that enables evolvability, that is, enhances the probability that mutations will improve the performance of the organism. Here we outline two properties of organismal development, namely modularity and robustness. Modularity refers to the situation in which genes affect a restricted number of functionally related phenotypic characters. Robustness describes a situation in which cryptic mutations can accumulate without effect on fitness but can become visible to selection in a new environment or genetic background. We discuss recent empirical evidence in support of both phenomena and their effect on evolvability and also briefly address their evolution.

Keywords Robustness · Modularity · Canalization · Genotype–phenotype map · Constraint · Development

M. Pavlicev (✉)
Department of Theoretical Biology, University of Vienna,
Althanstrasse 14,
1090 Vienna, Austria
e-mail: mihaela.pavlicev@univie.ac.at

M. Pavlicev
Konrad Lorenz Institute for Cognition and Evolution research,
Altenberg, Austria

G. P. Wagner
Department for Ecology and Evolutionary Biology and Yale
Systems Biology Institute, Yale University,
New Haven, CT, USA

Why is Evolution of Complex Organisms Not Intuitive?

The ability of a population of organisms to respond to a selective challenge caused by environmental change depends on the presence of individuals that are suited to survive and reproduce under the new circumstances. This ability of a population to cope with the changing environment by adaptation we call evolvability. In a simplified way, this process has two levels: selection acts at the phenotypic level—the individuals are selected according to their phenotypic differences. If this phenotypic difference is due to genetic difference, for example due to a gene variant with the slightly different sequence, selection will result in accumulating gene variants conferring the better suited phenotype and cause change at the population genetic level. Response to selection then means that the individuals with a gene variant (allele) conferring a phenotypic advantage, such as for example tolerance to heat, will contribute more offspring to the next generation than individuals with a less fit phenotype. As the gene variant becomes more common, the average phenotype of a population changes, resulting in adaptation. Different gene variants arise by random mutations, where *random* means that the effects of mutations are independent of the environmental challenge the organism is experiencing; thermal stress doesn't make it more likely that the particular mutations causing thermotolerance will arise with greater frequency.

This cycle of repeated mutation and selection of the gene variants via the phenotypes they generate is at the core of evolutionary theory. But the mutations can generate an

incredible number of effects on the phenotype, and most of them will be deleterious under any circumstances, if not lethal. How does such random genetic change produce the "right" kind of phenotypic deviation often enough? How is change possible where multiple mutations are necessary but intermediate steps have no apparent advantage? How probable is adaptation if only some of the traits should be changed, without affecting those that are already in place? The problem of evolvability is most evident when it comes to complex organisms, consisting of many parts, as sophisticatedly integrated as for example in a vertebrate eye. Darwin (1859) already recognized the difficulty when first proposing his theory:

> To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by (random mutation and (*added by authors*)) natural selection, seems, I freely confess, absurd in the highest possible degree.

Of course Darwin did not consider this problem insurmountable, and spent the subsequent paragraphs addressing it as well as he could with the knowledge available to him. Nevertheless, evolvability of complex traits is one of the most intriguing properties and at the same time one of the most enduring objections against the Darwinian theory of evolutionary change. It is picked up by researchers genuinely interested in understanding how evolution works, as well as by those who use its intricacy as a weapon against evolutionary biology, most notably even by those academics working for the creationist movement (e. g., Behe 1996). Because the problem of evolvability of complex organisms–and in a way all organisms are complex–represents a weakness in the public perception of evolutionary theory; and also because it is an intrinsically interesting problem of fundamental importance, we think that it should receive much more attention from mainstream evolutionary biologists and educators.

The answer to the problem of evolvability as presented above lies in the mechanisms that translate the genetic change into the phenotypic change. These mechanisms are development and physiology, and their mediating role is often referred to as a genotype–phenotype map (Fig. 1). Any evolutionary change of a trait requires a change in these mechanisms. For tractability, the genotype–phenotype map is summarized statistically rather than considering the details of all the developmental processes. For example, the genotype–phenotype map can be described in terms of the number of traits the mutations at a gene affect (pleiotropy), how many genes affect the same trait (polygeny), distribution of effect sizes, etc. We know that this is a simplification, a working tool that enables inferences about

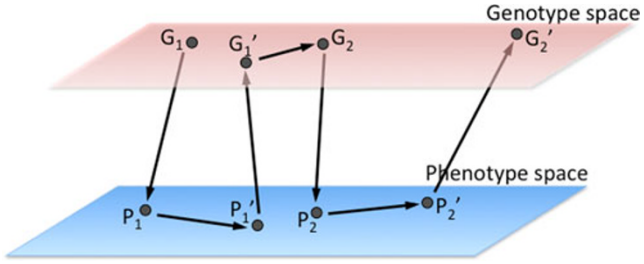

Fig. 1 Genotype-phenotype map. Genetic changes arising by mutation in the genotype space are translated into phenotypic variance (phenotype space) depending on the structure of the genotype-phenotype map, and become exposed to selection. Subsequently the population change in the phenotype due to selection causes the change in genetic composition and so forth. Figure modified after Houle et al. (2010)

certain specific structures (Pavlicev and Hansen 2011). The real maps are not simple, but they are also not random. Genetic changes often affect very specific sets of traits; some small genetic changes can have large and complex phenotypic effects, while others may have little or none. The effect on the phenotype thus depends on how the change percolates through development; that is, it depends on the structure of the genotype–phenotype map. The idea behind explaining the high evolvability of organisms is that the genotype–phenotype maps of extant organisms are such that the probability that random mutation will improve the phenotype has increased during evolution.

Here we want to summarize some important ideas about evolvability and track their maturation from the formulation of the evolutionary process as mainly a population genetic change in allele frequencies up to the current thinking within evolutionary biology where the organism and its genotype–phenotype map structures the availability of phenotypic variation. In this context, we will also present some recent results from experimental research on factors affecting evolvability.

Early Attempts to Understand Evolvability

Population geneticist Ronald A. Fisher illustrated the problem of evolvability with a metaphor in which he compared mutation and selection to the mechanical tuning of a microscope, arguing that every large change will have a very small probability of improving the image (Fisher 1930). Fisher also expressed this idea in a more mathematical language in the so-called geometric model (Fig. 2). To understand this model, consider for simplicity a phenotype consisting of only two traits. This *phenotypic space* representing all combinations of values of two traits can be represented as a two-dimensional plane. Each point in the plane represents

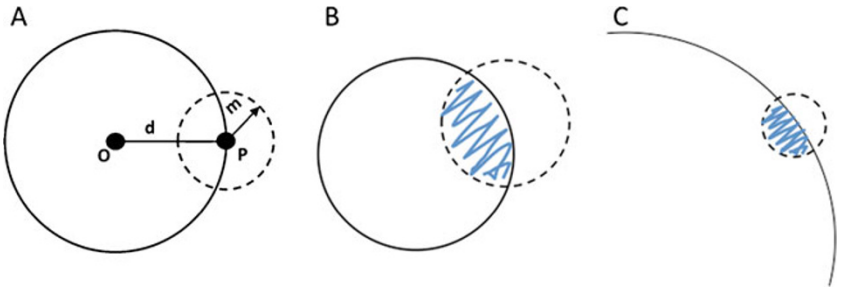

Fig. 2 Fisher's geometric model. **a** Optimum (O) can be envisioned as a point in the phenotypic space, with fitness uniformly decreasing in all directions. A phenotype P with a certain fitness *d* units lower than optimum can then be positioned on a circle with radius *d* and center at O. Mutational step that changes the phenotype P for *m* units can again be shown as a circle around the P, with radius *m*. **b** The overlap between the two circles represents the proportion of the mutations of the particular size, occurring in the phenotype P, that are advantageous. **c** The smaller the mutational step relative to the curvature (i.e. distance from the optimum), the greater will be the proportion of mutations that are advantageous

a unique combination of phenotypic trait values, and mutations can be imagined as a dislocation of an individual's phenotype from one point in this abstract plane to another. Let us consider a situation where an individual phenotype P is positioned at some distance *d* from the optimal phenotype O. Let us further assume that fitness (in terms of contribution of the individual with a certain phenotype to the next generation) is decreasing in all directions with the distance from the optimum O, so that all points with the same distance *d* from the optimum, which can be represented by a solid circle around the optimum (Fig. 2) have the same fitness. Any phenotypic change from the point P that brings the individual toward the optimal phenotype, i.e. inwards of the circle, is advantageous, and conversely, any change that positions it outside this circle is disadvantageous. Now consider all possible phenotypic changes of a certain step size, beginning from the phenotype P. These again constitute a circle (dashed line in Fig. 2), this time with P in the center and the radius of a step size *m*. The intersection of the two encircled areas is reached by the mutational steps of a certain size that are advantageous, i.e. lead to a phenotype with higher fitness than P. Fisher noted that the smaller the mutational step *m*, i.e. the smaller the radius of the dashed circle around P, the larger is the proportion of advantageous changes out of all possible ones, approaching 50% with infinitely small steps. As the mutational effects get larger, the proportion of advantageous mutations out of all possible ones decreases, and so does the probability of improvement. In complex organisms with more than two traits and thus also higher dimensionality of mutational effects, the probability of improvement by mutation decreases even more steeply with mutational size. Based on this model, Fisher proposed a solution to the question of how evolution in complex organisms nevertheless occurs, namely that the more traits the organism has, and therefore the more dimensional the mutations are, the smaller must be the mutational steps. As noted by Orr (2000), this leads to the so-called *cost of complexity*, an expectation that the more complex the adaptations, the lesser the adaptive step will become, and consequently the rate of evolution will decrease with increasing complexity, eventually coming to a halt at very high levels of complexity. This is an intriguing result. We know that even complex organisms respond to selection. The most apparent examples are the numerous sorts of domestic plants and animals that were selected for different traits for human consumption, use, or pleasure.

One blow to Fisher's theoretical solution came when Crow and Kimura (1970) noticed that mutations need to have a certain minimal effect on fitness to become selected, establishing a limit to evolution by very small steps. Another important finding stems from research into evolutionary optimization algorithms by a German engineer, Rechenberg (1973), who found that evolution by random mutation can be very efficient but requires a very precisely controlled mutational step size, depending on how the trait changes affect fitness.

If evolution cannot proceed by genetic changes with very small phenotypic effects, then how are the many potentially deleterious effects of large phenotypic changes avoided? In particular in complex organisms, where so many traits are integrated into a whole, how can some traits change significantly without rendering some other traits nonfunctional? In other words, if complex organisms evolve by mutation and selection, there must be a mechanism that makes mutations more likely to produce adaptive phenotypes than predicted by the Fisher's model.

The problems of the mutational step size and of the evolution of organismal complexity revealed the weaknesses of evolutionary biology. Whereas the main principles of evolutionary theory from the population genetic perspective were established early (Sewall Wright, Ronald A. Fisher, and their contemporaries in the 1930s), their integration with the organismal perspective, represented by comparative morphology

and embryology, was lacking and was even actively discouraged. Selection appeared almighty, but can the variation required for the selection to cause evolutionary change really be produced in any amount and of any sort? Can genetic change cause just any phenotype to arise, or are there limits? Why are some morphologies common, whereas others are thinkable but never realized? Why are some forms unchanged for long evolutionary times while others evolve rapidly? How do the effects of genetic mutations percolate through the physiology and development, from the molecules they affect, to the traits, and trait combinations, which are selected? How much genetic change is needed for certain phenotypic change?

Around this set of questions, the notion arose that evolutionary changes in organismal intricacies cannot be sufficiently understood by changes in the frequency of allele variants alone (Rensch 1959; Waddington 1957; Simpson 1953; Riedl 1978). Prominent controversies around this subject included the questions whether phenotypic evolution proceeds by punctuated or gradual phenotypic change (Eldredge and Gould 1972; Gould and Eldredge 1977), and to what extent adaptation alone can explain the phenotypes we observe (Gould and Lewontin 1979; Riedl 1978). These criticisms converged on the issue of how to explain evolution of complex organisms solely by random mutation and selection.

Due to the abstract nature of population genetic models, the problem of explaining evolvability is vulnerable to misunderstandings and misinterpretations. Frequently, evolvability is simply declared unexplainable by those who oppose evolutionary thinking in general, like the supporters of creationism (the issue is then referred to as "irreducible complexity"). Moreover, ignorance of the existing work on the issue has led others to refute the evolution by natural selection altogether (e.g., "What Darwin got wrong," Fodor and Piatelli-Palmarini 2010). See the critical review of this later position in Sober (2008) or Futuyma (2010).

Modern evolutionary biology came a long way towards understanding biological complexity and its evolution, if compared to the time when the problem was first raised. After some of the theoretical principles had been established, the details and consequences of which are still a topic of lively research, empirical work also began to show them at work. What started as a systems theoretical query was complemented by an empirical research program, which is greatly aided in the recent years by technological advances.

In the following we outline what we think are important components toward understanding the evolvability of complex organisms and emphasize recent developments. Our aim is to provide the reader with a general idea of the argument for evolvability, as well as with the ability to access more detailed studies on specific questions. We explain the historical context and issues from the perspective of, but not restricted to, population genetics. This is justified by the fact that population genetics has been a dominating branch of evolutionary biology for many decades. Other perspectives on evolvability have been published, most notable among them the contributions by Gerhart and Kirschner (2007). In short, they point out that many properties of the organismal structure facilitate the origination of viable phenotypic variation and enable populations to produce adaptive variants and thus survive selective challenges. They single out three principles: exploratory processes, weak linkage, and compartmentation. We will show that these principles describe similar properties of organisms, albeit at the different organismal level, as the ones discussed here from the population genetic perspective.

It is due to the very nature of this paper that many important detailed studies will remain unmentioned. We apologize to the numerous contributors to the field for that. We recommend that the interested reader turn to existing accounts (de Visser et al. 2003; Gerhart and Kirschner 2007; Kirschner and Gerhart 2010; Wagner 2012; Wagner et al. 2007; Wagner and Zhang 2011) for a more inclusive coverage of the research in this field.

Systems Conditions of Evolution

One of the biologists that explicitly addressed evolvability early on was Rupert Riedl. In short, Riedl recognized that morphological characters evolve at different rates. Consequently, some traits differ only slightly across taxa whereas others differ more. Variation is also structured within populations, with some traits varying more than others. He ascribed the hierarchy of variation to the notion of *burden*, meaning that the more burdened characters are those upon which many further characters depend in their development or function. More burdened characters will be less variable and evolve more slowly. Riedl didn't think that this pattern is just because selection was sorting the phenotypic variants, but that it mirrored the mutations, namely that mutational pattern was structured by the internal developmental constraints. Thus Riedl's theory suggested that newly arising variation is structured by the development and presented to selection in a nonrandom way. A mutation affecting a deeply engrained part of development, for instance the patterning of the body axis in a mammal, will have many deleterious effects and cannot be selected. This is the likely reason why mammals almost always have seven neck vertebrae, for example (Galis 1999; Narita and Kuratani 2005; Schoch 2010). While the mutation may be random with respect to whether it improves or reduces the fit to the environment, it is not random with respect to which traits it affects.

Importantly, Riedl furthermore suggested that this structuring mechanism, the development, itself evolves such that

it mimics the functional interdependencies. In other words, the arising mutational variation would be shared between the traits that are commonly selected together, but not between traits that require independence in their selection response. He specifically argued that these patterns are there because they increase evolvability, and thus that evolvability is a selectable trait. For more detail, consider Riedl's *Order in Living Organisms* (1978; in original *Die Ordnung des Lebendigen*, 1975), subtitled "A Systems Analysis of Evolution" (Die Systembedingungen der Evolution), as well as a summary of his ideas in Wagner and Laubichler (2004).

By conceptualizing the organisms in terms of patterns of variation, Riedl also created the much needed connection between organismal comparative biology and the variation-based Neo-Darwinian theory of evolution. In this way Riedl was not only one of the first to articulate the importance of development for understanding organismal evolution, he also provided a link between population genetic and developmental (organismal) approaches to evolution, a task still considered difficult by many (see Amundson 2005). Several lucid accounts of the central conceptual importance of evolvability in developmental evolution have been published (Brigandt 2007, 2012; Hendrikse et al. 2007).

However, in Riedl's work, the idea that organisms may have properties that would increase their ability to evolve was still articulated vaguely. Interestingly, the first clear formalization of evolvability stems from computer science. Lee Altenberg worked on optimization algorithms in genetic programming and defined evolvability as "the ability of a population to produce variants fitter than any yet existing" (Altenberg 1994). We will stick to this general meaning of evolvability for the purpose of this paper, in spite of later differentiation of the term (see Pigliucci 2008).

In the 1990s biologists and computer scientists thus set off in parallel to explore evolvability (Wagner and Altenberg 1996; Gerhart and Kirschner 1997; Kirschner and Gerhart 1998). One of the influential ideas was that particularly evolvable systems are likely to have a modular mapping of genetic effects on the phenotype. Modularity was defined as a structure of genotype–phenotype map where genes are shared only among the related traits (with common function or development) but not between unrelated traits (Fig. 3). Such a genotype–phenotype map was proposed to reduce interference between the characters when the characters are under conflicting selection pressures (Wagner and Altenberg 1996).

Another important contribution to understanding the importance of the structure of genotype–phenotype mapping for evolvability has arisen independently from the work of theoretical chemists. The group around Peter Schuster from the University of Vienna (Schuster et al. 1994) focused on the secondary structure of RNA. The RNA is a single-strand macromolecule generated by transcription from the DNA. It

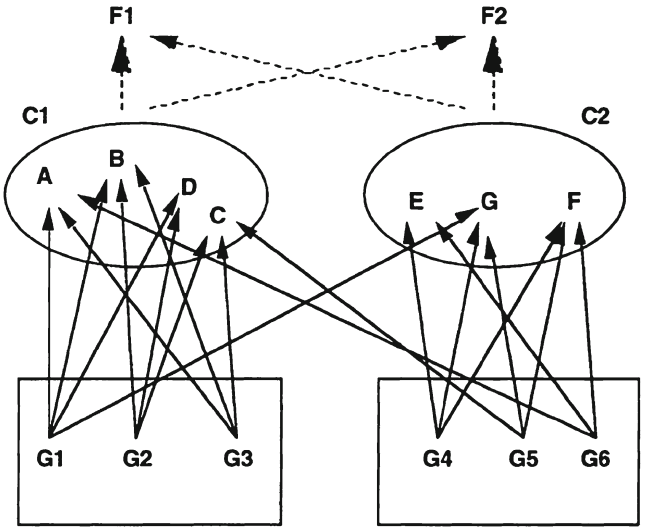

Fig. 3 Modular genotype–phenotype map. Effects of genes (G1–G6) are organized such that they preferably affect traits of one module (either C1 or C2) with the same function (F1 or F2), whereas the interference between modules is minimized. Figure from Wagner and Alteberg (1996)

serves as a template for translation of a sequence into protein but also serves many other functions in the cell, for example as an enzyme (aka ribozyme), regulatory molecule, or as a part of the ribosome. Several of its functions depend on a proper folding of the single-stranded molecule upon itself. This folding requires a certain pairing of the corresponding nucleotides in the strands that come to bind (Fig. 4a). Schuster et al. (1994) treated folding as the phenotypic trait. This phenotype level immediately follows the transcription from the sequence, and is therefore "developmentally" close to the genotype level, i.e., the genetic sequence. This so-called RNA secondary structure and its underlying sequence are therefore ideally suited for the study of immediate effects of mutational changes on the phenotypic level and exploring the simplest genotype–phenotype map. Schuster et al. (1994) observed at the phenotypic level that there are only few very common secondary structures and many very rare ones among the existing structures. Upon inspection of the underlying sequences, they found (a) that there are many-to-few mappings, that is, many different sequences fold into the same stable structure (Fig. 4b) and (b) that these many sequences that fold into same structures are similar. Schuster and colleagues represented the genetic sequences (i.e., genotypes) by nodes in a network, where nodes connected by an edge are one mutational change from each other. Groups of connected nodes that fold into the same phenotype are called neutral networks (Fig. 4c). Differences among genotypes within a neutral network represent cryptic genetic variation with no effect on the phenotype, and can be seen as the portion of *sequence space* that the individuals of a population can occupy without

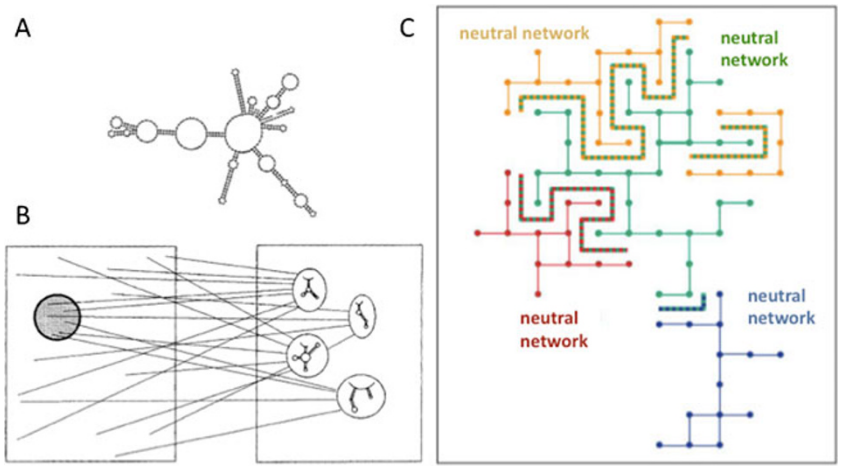

Fig. 4 Robustness of the RNA folding. **a** Single-stranded RNA molecule folds onto itself, to form stems and loops. **b** Mapping from the sequence space (*left*) to the phenotypic space (*right*) is many to few, meaning that many sequences fold into the same secondary structure. **c** Sequences that fold into the same structure and hence do not differ in their phenotype belong to neutral networks (*yellow*, *green*, *red*, *blue*), consisting of nodes connected by single mutational changes. Figures from Fontana (2002)

selective differences. Nevertheless, note that the greater the neutral network, the greater is its outer border. The outer border separates the neutral network from the genetic neighborhood, and when a mutation crosses this border, the phenotype changes. Large neutral networks with a large number of border genotypes also have a high potential to access new phenotypes by a single mutational step, as a population. By at the same time increasing the number of neutral changes and the number of different phenotypes a population can reach, mutational corridors without an effect on the phenotype make it significantly more likely that a random mutation will find a viable sequence. In other words, the genotype–phenotype maps enabling larger neutral networks are more evolvable (Fig. 4c). This early RNA work provided a basis for the intense theoretical exploration of the relationship between robustness and adaptation (Huynen 1996; Huynen et al. 1996) and between robustness and evolvability (Schuster and Fontana 1999; Wagner 2005; Draghi et al. 2010).

From the above work thus emerged two principles relevant to organismal evolvability. The first is modularity of genetic effects—a principle according to which not all genes affect all traits. The second is robustness (also referred to as canalization or capacitance)—an idea that not every mutation has an effect on the phenotype under all circumstances, or more generally, that the mutational effects change across environments and genetic backgrounds. In the following we briefly describe both concepts and illustrate them with some of the recent, predominantly empirical examples.

Modularity

The concept of modularity is used broadly in biology, generally referring to the idea that some parts of the organism or a system are more connected among themselves than with other parts. This can regard different levels of organization, such as the network of molecular interactions, temporal modularity in co-expression of genes, as well as the distribution of gene effects on the phenotype. We maintain our focus here on the modularity of gene effects on the phenotype, the genotype–phenotype map, due to its importance for evolvability as explained above (Fig. 3).

Modularity of gene effects means that pleiotropy (i.e., the number of traits that a gene affects) is restricted to relatively few traits per gene even in complex organisms. Furthermore, the traits sharing genes are suggested to have a related function. As mutations occur and cause trait variation, the traits that share genes will tend to change together. This enables them to maintain their functionality if they are functionally integrated, like the parts of the eye, and perhaps more importantly, this avoids side effects on functionally unrelated traits. In contrast, Fisher's geometric model described above assumes that each mutation potentially affects all traits, and hence the more traits an organism has (the higher its complexity), the higher is the potential dimensionality of a mutation. From this assumption of "universal pleiotropy" arises the "cost of complexity" in Fisher's geometric model.

Much evidence for a pattern of restricted pleiotropy stems from quantitative trait locus (QTL) studies. These studies map genomic locations at which alternative gene variants are associated with alternative phenotypic values. For example, if individuals with one gene variant are on average smaller than those with an alternative gene variant, everything being equal, we can conclude that the gene has something to do with the body size. Often these loci include many genes; therefore, it is rather a location, a region that is revealed in this general approach. However a variety of methods are used to increase the resolution and reliability of the approach. Quantitative or complex phenotypic traits refers to traits that are affected by many genes, and hence often (depending on the variation available in the population) many loci will be found for each complex trait. Scanning the genome for such associations thus allows us to determine at which locations the genes are found that affect the particular phenotypic trait. Using QTL mapping for many

traits simultaneously, one can locate genomic regions that are associated with many traits, and thus one can determine the pattern of pleiotropic effects. One such study was performed in the lab of Jim Cheverud in St. Louis, on mice skeletal traits. Jane Kenney-Hunt and colleagues have mapped the QTL that affect one or more of the 70 traits (Kenney-Hunt et al. 2008). A follow-up study of these data (Wagner et al. 2008) has shown that most of the QTL affected very few of these 70 traits (average five to six), and only a few QTL affected many traits (up to 38), but even these affected much less than 70. This study thus provided evidence in support of restricted pleiotropy. The study also explored whether the genes that affect a greater number of traits have a smaller effect per trait. This would be the case if there were a cost to complexity. The study rejected the notion that complexity leads to smaller mutational size and consequently to a cost for evolvability (Orr 2000). QTL studies were furthermore used not only to explore how restricted the pleiotropy may be (Albert et al. 2008; Wagner et al. 2008; Zou et al. 2008; Su et al. 2010) but also to show that the shared loci indeed often coincide with common development or function (Mezey et al. 2000; Ehrich et al. 2003; Albertson et al. 2005; Parsons et al. 2012).

Advanced molecular methods moreover enabled systematic assays of the patterns of pleiotropy on an array of organisms and phenotypes. For example, by targeted genetic perturbations of single genes, their effects can be recorded directly, and the ambiguity of how many genes may be responsible, inherent in QTL studies, can be avoided. Again, the studies exploring this question often find a surprisingly low number of traits affected by the same gene, compared to the number of traits tested. Moreover, there is strong support for the clustering of the detected effects into phenotypic modules for genomes of yeast, nematode, and mouse (Wang et al. 2010a).

Robustness (for a Recent Review, See Wagner 2012)

Robustness or canalization refers to the invariance of the phenotype in spite of genetic mutations or environmental perturbations. The general idea goes back to English biologist Conrad Waddington, who observed that the wild-type populations of fruit fly in the lab tend to show less variation (i. e., are more *canalized*) than populations carrying a large mutation, implying that a mutation changes (*decanalizes*) the effect of other mutations, i.e., exposes otherwise small or cryptic genetic variation (Waddington 1942).

The contribution of robustness to evolvability is counter-intuitive at the first glance: how is a system that is apparently resistant to change more evolvable? Yet the robust organisms are able to accumulate genetic variation precisely because this variation is hidden from selection, i.e. because it is cryptic. It is also well established that many kinds of environmental and genetic perturbations can lead to the release of this accumulated genetic variation in the phenotype, by which this variation is exposed to selection and may fuel adaptation. Thus robust populations may have a better shot at responding to new environmental challenges, as they contain potentially advantageous genetic variants upfront, rather than having to wait for new mutations to arise (Rutherford 2003).

There are two main ways that cryptic variation can become visible to selection. It can be released by the genetic or by the environmental change. If accumulated genetic variants that had no effect on the phenotype suddenly become manifested as phenotypic variation due to a single mutation elsewhere in the genome, we speak of a case of genetic interaction or epistasis. More generally, epistasis means that a mutation in one part of the genome changes the way an allele (genetic variants coexist in a population) at another locus is affecting phenotypic variation. Thus mutational effects on a trait cannot be simply added up across loci, rather their combinations are important due to this so-called interaction effect. In the case of release of variation (decanalization), a new mutation causes the previously neutral genetic polymorphisms to affect phenotypic variation and may become selectively relevant (Hermisson and Wagner 2004). Because this phenotypic variation is genetic and thus heritable, its release can contribute to a rapid burst of evolutionary change. In a similar way, environmental change, such as the introduction into the new environment, can cause decanalization. Also in this case, previously neutral genetic variation may become selectively relevant. The examples below will elucidate these phenomena.

A well-known example of genetic robustness comes from the study of fruit fly, *Drosophila*. Rutherford and Lindquist (1998) have studied the role of chaperones, the molecules that are involved in folding and stabilizing of other proteins. Stabilization of proteins is crucial for their proper functioning. Chaperones are especially important when an organism is under stress, such as for example, heat shock. Under stress, otherwise stable proteins also often fold incorrectly, which impedes their function. Rutherford and Lindquist (1998) observed that a mutation in one of the chaperones, the so-called heat-shock protein Hsp90 in *Drosophila melanogaster*, increases morphological variation in multiple traits, including wings, legs, eyes, and abdomen. By a selection experiment in which they selected two of the newly variable traits (wings and eyes), they have shown that this new variation is heritable. As this genetic variation unlikely has appeared solely by a sudden increase in mutagenesis, they reasoned that it must have been present previously as cryptic genetic variation without an effect on the traits, and was manifested due to the malfunctioning of the chaperone. Indeed, apparently, the mutation's destabilizing proteins were present in a population but could be buffered

due to the stabilizing action of a chaperone. When chaperone becomes scarce, either because of increased need due to stress or due to a mutation in a chaperone, the accumulated destabilizing mutations cannot be buffered and the malfunctioning of proteins manifests in the mutant phenotype. This work, independently from the RNA work described above, established the idea that robustness may increase evolvability and generated a further line of research.

Recently, an additional mechanism for the sudden increase in genetic variation due to Hsp90 was proposed: Specchia and colleagues (2010) have demonstrated that nonfunctional Hsp90 also affects the silencing of the transposable elements in the germ line. Transposable elements (TE) are DNA fragments that replicate independently from the rest of the DNA, and the replicates insert themselves at other places in the genome. By insertion, they can cause a genetic change that may have a phenotypic effect, for example when by insertion they introduce a regulatory element or impair one, thereby changing when and in which cells a certain gene is expressed. Mutation rate is thus increased by the presence of transposable elements. The activity of TE is normally suppressed in the germ line, which is the line of cells producing the sperm and egg cells, therefore reducing the potentially heritable genetic changes. Specchia et al. (2010) proposed that increased genetic variability following the mutation in Hsp90 may also be induced by the failure to suppress mutagenesis due to transposable elements. In this case, the newly arisen genetic variation would be due to increased mutation rate, rather than due to release of the already present but hidden genetic variation in the phenotype.

This shows that closer examination reveals multiple ways of increasing the ability to adapt. As will be shown below, research on robustness flourished in particular due to advantages of an experimental evolution approach.

Testing for Evolvability in the Laboratory

Modern molecular technology allows monitoring evolution while it is occurring in the laboratory. Thus we can move from asking the question whether modularity and robustness exist, to asking whether they also increase evolvability as predicted, and under what circumstances they themselves evolve. Bacteria and viruses are particularly suitable for this task because of their high reproduction rates, low maintenance effort, and easy manipulation (Colegrave and Collins 2008; Garland and Rose 2009). One example of using experimental evolution to study the relationship between robustness and evolvability comes from Paul Turner's lab at Yale University. McBride et al. (2008) founded populations of viruses from the clones with high, respectively low robustness (how these clones were generated will be described in a later section). The level of robustness was measured by observing the variation in survival under environmental challenge. Initially, when exposed to thermal stress, the brittle and robust groups of clones had equal mean survival. The experiment consisted of exposing both groups to periodic heat shock, and comparing their relative ability to adapt by evolving thermotolerance. Indeed the study found that the populations founded by the robust clones evolved greater resistance to heat shock than the populations founded by the brittle clones, even though their initial mean fitness in the heat-shock treatment was indistinguishable (McBride et al. 2008). This supports the notion that robustness confers advantage for evolvability.

In another recent experimental evolution study, Hayden and colleagues (2011) at the University of Zürich, Switzerland, studied the evolution of robustness and its effect on evolvability in a catalytic RNA enzyme. As mentioned above, RNA can serve as informational molecule (i.e., relating information encoded in a sequence), as well as a functional molecule with secondary structure, as in the case of RNA enzyme. RNA enzyme can be evolved in vitro, by the reverse transcription of the RNA sequence into DNA sequence, which can be reproduced via polymerase chain reaction. During reproduction, mutagenesis can be induced in order to enhance the accumulation of genetic variation. Hayden et al. (2011) used this approach and evolved two populations of RNA enzymes by mutagenesis. During mutagenesis, they applied stabilizing selection to maintain the native catalytic function of the enzyme. Thus during these ten generations of mutagenesis and stabilizing selection, the enzyme accumulated cryptic genetic mutations, that is, those which did not significantly affect its native catalytic function. Subsequently, the populations were challenged to adapt to a different substrate. Researchers recorded how fast the enzyme increased the efficiency of catalysis on the new substrate, therefore measuring the evolvability of the enzymes to the new environment. They compared the evolvability of an enzyme that was allowed to accumulate cryptic variation in the previous step, to the adaptability of a control that didn't accumulate cryptic variation. They found that the populations with cryptic variation adapted much faster, even though they started from the same mean value of catalytic efficiency as the control. Analysis of the mutations revealed that specific combinations of mutations were particularly important in conferring selective advantage. Genetic interaction (epistasis) between mutations is thus emphasized here, as often the adaptive phenotypes are more than a single mutation away, and if the intermediate step is deleterious, transversing to the potentially advantageous genotype can be impossible with single-step mutations. Robustness allows these adaptive genotypes to be reached while they are not "seen" by selection.

Another remarkable example of evolvability research is a long-term experimental evolution study of bacteria *Escherichia coli*, conducted in Rick Lenski's lab at Michigan State University. *E. coli* has been evolving in Lenski's lab for over 55,000 generations, with populations collected regularly as evolutionary snapshots and frozen for later research. Bacteria can easily be revived and used in a further study at later time points. In a recent study, the authors focused on such earlier population (500th generation of the experiment) and compared the fitness of two clones with gene variants that eventually took over the population, with the fitness of two other contemporary clones that lacked the particular variants and went extinct in the original experiment (Woods et al. 2011). Upon competing them against each other to compare their fitness, the eventual losers surprisingly showed higher competitive fitness. So how come did they go extinct in the long term in the original experiment? Authors repeated the long-term evolution experiment from these clones many times, thereby excluding an explanation by pure chance. They showed that the outcome was mostly the same: the winners from the original experiments eventually win in most replicates (Woods et al. 2011). The mutations characterizing the two types of clones thus presumably conferred the long-term evolvability by affecting subsequent mutations. This difference decided the long-term outcome of evolution and overcame the effect of short-term competitive superiority of the alternative clones.

Empirical studies thus not only find differences in robustness, but also demonstrate that these differences indeed affect the ability of population to evolve, such that the more robust populations adapt faster to new challenges than the less robust ones.

Is robustness always evolutionarily advantageous? Or is the intuition that strong robustness can also impede evolution correct? There is indeed support for this intuition. Draghi and colleagues (2010) explored the relationship between robustness and evolvability in a population genetic model and showed that robustness is not always advantageous. Indeed evolvability is reduced at very high and very low robustness. This is plausible as when robustness is very high, little variation becomes exposed to selection, and when robustness is very low, the conditionally adaptive mutations cannot accumulate because they are immediately exposed to selection.

The Origin of Evolvability

The examples above provide empirical evidence for the existence of robustness and modularity and for their contributions to evolvability. But how do the organisms attain these properties and with them evolvability? This question is yet unsolved and the subject of some controversy.

Why is this question so hard to answer? Evolvability per se does not directly contribute to the fitness of an individual under the present environment in the same way as, for example, larger leaves contribute to securing a greater exposure to sunlight. Therefore evolvability is not a direct target of selection. Not surprisingly then, its evolution is intriguing, even unthinkable to some. The types of models proposed to explain the evolution of evolvability differ mainly with respect to the role that natural selection plays in the process. They range from the models in which evolvability arises without natural selection to those that explicitly require the action of natural selection, whether direct or indirect (Hansen 2011).

An example of a model where evolvability arises due to nonadaptive increase in modularity involves gene duplication. Gene duplication is a common evolutionary event and leads to genes being present in the genome in multiple copies. The resulting paralogous genes can subsequently differentiate in terms of the functions they have inherited from the ancestral single-gene stage (Fig. 5; Force et al. 2005). This may happen by random mutations that cause some of the initially redundant functions to become lost. This situation enables subfunctionalization, a process during which each of the gene copies maintains only a complementary portion of the original functions, resulting in subdivision of functions, a decrease of pleiotropy, and potentially also in the modularization of gene effects. A nice example for evolutionary increase in evolvability due to gene duplication is the evolution of scaleless fish. Scale loss occurs frequently among cypriniform fishes, i.e., fishes related to carp and zebrafish, but is less common in other fishes. It turns out that one gene that is essential for scale development is the gene for FGF receptor 1, *fgfr1*. However, this gene also has many other vital functions, and therefore most mutations affecting *fgfr1* supposedly have many negative side effects. In cyprinids, however, there are two copies of *fgfr1*, one of which is exclusively dedicated to scale development, presumably due to duplication and subfunctionalization. Hence in these fish, the mutation in the scale-related copy of *fgfr1* and loss of scales occurs without strong negative side effects, and this is what has repeatedly happened in this group of fish but not in others where scale development is more integrated with the rest of the body. In cyprinid fishes the scaleless character is therefore more evolvable because of a past gene duplication, which led to a gene specifically dedicated to scale development (Rohner et al. 2009).

Gene and whole genome duplications are frequent events in evolution and are considered by some to be the main driving force responsible for the increase in complexity in general (Lynch 2007a, b; see Finnigan et al. 2012 for a recent empirical example). Using gene duplication to explain variational patterns by the action of mutation and drift

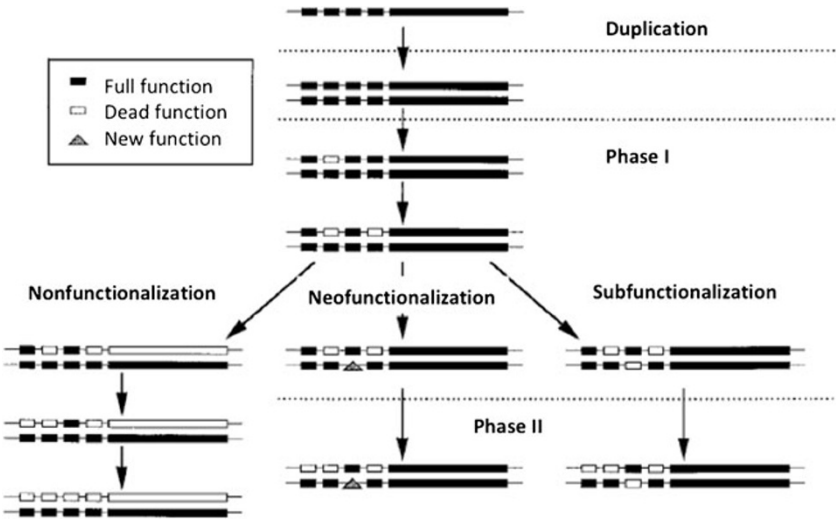

Fig. 5 Fates of duplicated genes. After gene duplication, the two copies of a gene diverge (*Phase I*) and subsequently one copy can degrade (nonfunctionalization), assume an entirely new function (neofunctionalization), or the two copies may reduce their functions to complementary subsets (subfunctionalization). Figure from Force et al. (1999)

alone, instead of also invoking selection, is tempting. However, whereas duplication/subfunctionalization process explains that the pleiotropy (and with it the correlations between sets of traits) undergoes evolutionary change, one may still ask why it is that the particular traits or sets of functions are separated from the others, instead of these combinations being random.

Rather, the traits that evolved lower correlation are usually those that have different functions, and those more closely correlated the ones that share a common function (Olson and Miller 1958). That this pattern is adaptive can be concluded from the observation that the correlation between parts decreases when the parts undergo evolutionary specialization for different functions. For example, fore and hind limbs in primates are parts that share much of development and are so to say duplications of a part within the body (so-called serially homologous traits). Fore and hind limbs have differentiated and fulfill distinct functions in apes and humans, but not in monkeys. Correspondingly, the correlation between the fore and hind limbs in apes and humans decreased, but not in apes (Young et al. 2010). The same pattern holds with bats, which have specialized fore and hind limbs compared to quadrupedal mammals (Wang et al. 2010b).

Similarly, parts of plants can become specialized for different functions. For example, flowers in plants that are more highly specialized for pollination are less correlated with their vegetative parts than the case in plants where flowers are less specialized for pollinators (Berg 1960; Pélabon et al. 2011).

Next, let us consider explicitly what it means that modularity and robustness evolve by selection. In the case of modularity, it means that among gene variants that affect different numbers of traits, those gene variants are selected which affect lower numbers of traits, as well as certain combinations of traits. Alternatively, the gene effect can be changed due to genetic background, as is the case due to epistasis, where the mutation at the locus A can change the effect size and also the number of traits that the locus B affects. Thus, given that the pleiotropy of a gene B variant changes dependently on the genetic background A, those variants in genetic background are selected in which the gene variant affects the preferred combinations of traits (Pavlicev et al. 2008; Pavlicev and Wagner 2012). What does evolution by selection mean in the case of robustness? It means that the genotypes are preferred in which mutations have little or no effect on fitness.

Support for the evolution of modularity by selection derives predominantly from theoretical studies (see review in Wagner et al. 2007; also see Hansen 2006). An interesting empirical result comes from the selection study on sexually dimorphic plant *Silene latifolia*. Delph and collaborators (2011) have successfully selected for reduced correlation between the flower sizes in male and female plants. The flower sizes in the two sexes can be treated as traits sharing many genes, and not surprisingly having a high correlation in the population. As genetic correlation is a manifestation of pleiotropic effects, the results suggest that evolution of modularity by selection had occurred. Note however that here the selection was for correlation and not for evolvability per se.

Again the experimental evolution work has revealed important insights about the evolution of evolvability. To be able to compare the evolvability of strains with different degree of robustness, the empirical studies mentioned above had to repeatedly select for robustness. Strong stabilizing selection on fitness, while simultaneously accumulating mutations, selects for neutral mutations. For example, Turner and Chao (1998) grew the virus RNA bacteriophage $\Phi 6$ on the host bacterium *Pseudomonas syringae* under high and low degrees of coinfection. Because viruses can compensate for the deleterious mutations by exchanging the superior protein

products between viruses within the same host, the viruses in the highly coinfected setting were under weaker stabilizing selection than those evolving under low coinfection that were unable to compensate. Montville et al. (2005) subsequently used the bacteriophage populations resulting from the above experiment and performed a mutation accumulation study. The results confirmed that the fitness consequences of the accumulated mutations were lower in the viruses historically evolving at low coinfection, meaning that these viruses were indeed more robust. In contrast, the viruses evolving under high levels of coinfection were less robust to accumulating mutations (Montville et al. 2005). These populations of brittle and robust viruses were used in the study by McBride et al. (2008) described in the previous section.

Another example for evolution of robustness under stabilizing selection was already discussed above (Hayden et al. 2011). As explained, the cryptic variation that led to increased evolvability accumulated under a combination of mutagenesis and the stabilizing selection for the native catalytic function in the RNA enzyme. For completeness, it should be noted that while the studies mentioned here demonstrate that robustness *can* evolve under selection in the lab, we do not know whether the same process also acts in nature and whether the dominant form of selection is selection for evolvability.

The Relationship Between Robustness and Modularity

So far we treated robustness and modularity as two independent concepts. They are often studied separately, however they are closely related. One reason is that modularity increases robustness, because it reduces the number of traits affected by mutations. This can be shown on the example of RNA folding. The secondary structure of RNA consists of folds in the form of double-stranded stems, where the matching between nucleotides is good, and loops, where the strand is unpaired. These structures can be relatively independently functioning modules, and then the RNA molecule is in its entirety (fitness) more robust to environmental and mutational perturbations than if the parts are nonmodular. This is because the perturbations only affect a single module at a time, maintaining the remaining functions intact (Ancel Meyers and Fontana 2005). A change in modularity, hence, causes a change in robustness at the same time. This doesn't mean that the increase in robustness requires modularity, however modularity is one way in which robustness can be realized.

Conclusion

Research in evolvability encompasses the integration of the mechanisms of population genetics and the mechanisms translating genetic variation into structured organismal variation. Evolvability is thus a central question for evolutionary developmental biology or evo-devo (Hendrikse et al. 2007), as well as for the field of evolutionary systems biology.

Since the time when the question of how complex organisms evolve first arose, evolutionary biologists have identified several principles of evolvability. Here we focused on modularity and robustness and the evidence that these indeed play a role in determining the evolvability of real organisms, as well as on how they themselves may have evolved. This is not the only possible perspective to approach evolvability. The principles that enhance variation have been identified at other levels of organization. Specifically, at the cellular, developmental, and physiological levels, the above-mentioned work by John Gerhart and Marc Kirschner is particularly important. In their theory of facilitated variation (Kirschner and Gerhart 2005, 2010; Gerhart and Kirschner 2007), the authors identify organismal system properties that "furnish a specific favorable kind of phenotypic variation." Even though the properties at the levels of development and physiology manifest in more complex ways, their insights can be related to the principles of robustness and modularity. A full comparison of the two perspectives exceeds the scope of this paper, and is complicated by the fact that the authors inherently focus on processes generating variation rather than variational patterns reviewed here, but some will make the point. One property that Gerhart and Kirschner identified is the compartmentalization of phenotypic effects to avoid negative side effects of genes. This property is clearly related to modularity as described here, a separation of gene effects on different unrelated traits. Similar is the case with exploratory processes. Exploratory processes are those that are not rigidly determined and invariantly executed, but are rather context dependent, such as development of blood vessels and muscles that follows the development of the limb skeletal elements, even if the limb is built in the wrong place. To some extent this property is a manifestation of the high integration of traits within modules; in the above example, the module is the limb, and the integrated traits are the bones, vessels, and muscles. Robustness could be associated with this property and with the fact that mutational changes in one trait do not render the module nonfunctional but are rather accommodated. A third systems property identified by Gerhart and Kirschner is the so-called weak linkage and refers to the situation in which a small and even unspecific change triggers a complex system of responses. This again can be related to the effect of the pleiotropic gene, which has evolved to affect a nonrandom set of related functions that are integrated. In this way, the mutation itself may be random, but it affects a nonrandom combination of traits, increasing the potential for complex adaptation. More work is needed to connect the processes with the variational patterns they generate.

Overall, we think that an important insight required for understanding the evolution of evolvability at the genetic level is the role of context dependency of the mutational change. The effect of a mutation on the phenotype is not a property of a gene but rather differs across the different genetic backgrounds in which it occurs, depending on what other gene variants are present in the genotype. This is not surprising, as gene effects are defined due to correlation between genotype and phenotype, regardless what the mechanisms mediating these effects are, and this includes many interactions between genes. Context dependency furthermore arises due to the environment in which mutation occurs. We discussed above that this effect contributes to the evolvability of gene effects.

This paper should not be understood as a claim that the recent advances discussed here are completely undisputed. The research in this field is very dynamic and hence also generates friction and vigorous discussions. For example, with respect to one aspect of modularity, namely restricted pleiotropy, it has been argued that reducing the number of traits affected by a gene will also reduce the mutational target for traits, that is, the number of genes available to affect each trait will decrease (given a constant gene number). As a lesser number of genes affect a trait, the potential for generating genetic variation may be reduced, revealing a trade-off between positive and negative effects of modularity on evolvability (Hansen 2003, 2006; Pavlicev and Hansen 2011). And there are challenges to the validity of some evidence (Hill and Zhang 2012a, b; but see Wagner and Zhang 2012). This and many more open questions are still being explored.

Therefore, whereas the detailed mechanisms of evolvability are still a major research topic, naturally (and fortunately) generating much scientific discourse, these discussions should not be misunderstood as disagreement on the fact of evolution itself. As described above, organisms have the properties it takes to evolve complex adaptation, whether these properties themselves evolved by natural selection or not. Biologists agree that the question of evolution of complex organisms starts with *how* rather than *whether*.

Acknowledgments MP is supported by the postdoctoral fellowship at the Konrad Lorenz Institute for Evolution and Cognition Research. GPW gratefully acknowledges funding by the John Templeton Foundation (JTF grant #12793). The opinions expressed in this publication are not those of the John Templeton Foundation.

References

Albert AY, Sawaya S, Vines TH, Knecht AK, Miller CT, Summers BR, et al. The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. Evolution. 2008;62:76–85.

Albertson RC, Streelman JT, Kocher TD, Yelick PC. Integration and evolution of the cichlid mandible: the molecular basis of alternate feeding strategies. Proc Natl Acad Sci U S A. 2005;102:16287–92.

Altenberg L. The evolution of evolvability in genetic programming. In: Kinnear Jr KE, editor. Advances in genetic programming. Cambridge: MIT Press; 1994. p. 47–74.

Amundson R. The changing rule of the embryo in evolutionary thought: roots of evo-devo. Cambridge studies in philosophy and biology. Cambridge: Cambridge University Press; 2005.

Ancel Meyers LW, Fontana W. Evolutionary lock-in and the origin of modularity in RNA structure. In: Callebaut W, Rasskin-Gutman D, editors. Modularity: understanding the development and evolution of complex natural systems. Cambridge: MIT Press; 2005.

Behe M. Darwin's black box. New York: Free Press; 1996.

Berg R. The ecological significance of correlation pleiades. Evolution. 1960;17:171–80.

Brigandt I. Typology now: homology and developmental constraints explain evolvability. Biol Philos. 2007;22:709–25.

Brigandt I. Intelligent design and the nature of science: philosophical and pedagogical points. In: Kampourakis K, editor. Philosophical issues in biology education. Berlin: Springer; 2012.

Colegrave N, Collins S. Experimental evolution: experimental evolution and evolvability. Heredity (Edinb). 2008;100:464–70.

Crow JF, Kimura M. An introduction to population genetics theory. New York: Harper & Row; 1970.

Darwin C. On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life. London: John Murray; 1859.

de Visser JA, Hermisson J, Wagner GP, Ancel Meyers L, Bagheri-Chaichian H, Blanchard JL, et al. Perspective: evolution and detection of genetic robustness. Evolution. 2003;57:1959–72.

Delph LF, Steven JC, Anderson IA, Herlihy CR, Brodie 3rd ED. Elmination of a genetic correlation between the sexes via artificial correlational selection. Evolution. 2011;65:2872–80.

Draghi JA, Parsons TL, Wagner GP, Plotkin JB. Mutational robustness can facilitate adaptation. Nature. 2010;463:353–5.

Ehrich TH, Vaughn TT, Koreishi SF, Linsey RB, Pletscher LS, Cheverud JM. Pleiotropic effects on mandibular morphology I. Developmental morphological integration and differential dominance. J Exp Zool B Mol Dev Evol. 2003;296:58–79.

Eldredge N, Gould SJ. Punctuated equilibria: an alternative to phyletic gradualism. In: Schopf TJM, editor. Models in paleobiology. San Francisco: Freeman Cooper; 1972.

Finnigan GC, Hanson-Smith V, Stevens TH, Thornton JW. Evolution of increased complexity in a molecular machine. Nature. 2012;481:360–4.

Fisher RA. The genetical theory of natural selection. 2nd ed. New York: Dover; 1930.

Fodor J, Piatelli-Palmarini M. What Darwin got wrong. 1st ed. New York: Farrar, Straus and Giroux; 2010.

Fontana W. Modelling 'evo-devo' with RNA. Bioessays. 2002;24:1164–77.

Force A, Cresko WA, Pickett FB, Proulx SR, Amemiya C, Lynch M. The origin of subfunctions and modular gene regulation. Genetics. 2005;170:433–46.

Force A, Lynch M, Pickett FB, Amores A, Yan Y, Postlethwait J. Preservation of duplicate genes by complementary, degenerative mutations. Genetics. 1999;151:1531–45.

Futuyma DJ. Two critics without a clue. Science. 2010;328:692–3.

Galis F. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes, and cancer. J Exp Zool. 1999;285:19–26.

Garland T, Rose MR. Experimental evolution: concepts, methods, and applications of selection experiments. Berkeley: University of California Press; 2009.

Gerhart J, Kirschner M. Cells, embryos and evolution. Malden, MA: Blackwell Science; 1997.
Gerhart JC, Kirschner MW. The theory of facilitated variation. PNAS. 2007;104:8582–9.
Gould SJ, Eldredge N. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology. 1977;3:115–51.
Gould SJ, Lewontin RC. The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc R Soc Lond B Biol Sci. 1979;205:581–98.
Hansen TF. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. Biosystems. 2003;69:83–94.
Hansen TF. The evolution of genetic architecture. Annu Rev Ecol Evol Syst. 2006;37:123–57.
Hansen TF. Epigenetics: adaptation or contingency? In: Hallgrimsson B, Hall BK, editors. Linking genotype and phenotype in development and evolution. Berkeley: University of California Press; 2011.
Hayden EJ, Ferrada E, Wagner A. Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. Nature. 2011;474:92–5.
Hendrikse JL, Parsons TE, Hallgrimsson B. Evolvability as the proper focus of evolutionary developmental biology. Evol Dev. 2007;9:393–401.
Hermisson J, Wagner GP. The population genetic theory of hidden variation and genetic robustness. Genetics. 2004;168:2271–84.
Hill WG, Zhang X-S. On the pleiotropic structure of the genotype–phenotype map and the evolvability of complex organisms. Genetics. 2012. doi: 10.1534/genetics.111.135681.
Hill WG, Zhang X-S. Assessing pleiotropy and its evolutionary consequences: pleiotropy is not necessarily limited, nor need it hinder the evolution of complexity. Nat Rev Genet. 2012;13:296.
Houle D, Govindaraju DR, Omholt S. Phenomics: the next challenge. Nat Rev Genet. 2010;11:855–66.
Huynen MA. Exploring phenotype space through neutral evolution. J Mol Evol. 1996;43:165–9.
Huynen MA, Stadler PF, Fontana W. Smoothness within ruggedness: the role of neutrality in adaptation. Proc Natl Acad Sci U S A. 1996;93:397–401.
Kenney-Hunt JP, Wang B, Norgard EA, Fawcett G, Falk D, Pletscher LS, et al. Pleiotropic patterns of quantitative trait loci for 70 murine skeletal traits. Genetics. 2008;178:2275–88.
Kirschner M, Gerhart J. Evolvability. PNAS. 1998;95:8420–27.
Kirschner MW, Gerhart JC. The plausibility of life: resolving Darwin's dilemma. New Haven: Yale University Press; 2005.
Kirschner MW, Gerhart JC. Facilitated variation. In: Müller GB, Pigliucci M, editors. Evolution—the extended sythesis. Cambridge: MIT Press; 2010.
Lynch M. The evolution of genetic networks by non-adaptive processes. Nat Rev Genet. 2007a;8:803–13.
Lynch M. The frailty of adaptive hypotheses for the origins of organismal complexity. Proc Natl Acad Sci U S A. 2007b;104 Suppl 1:8597–604.
McBride RC, Ogbunugafor CB, Turner PE. Robustness promotes evolvability of thermotolerance in an RNA virus. BMC Evol Biol. 2008;8:231.
Mezey JG, Cheverud JM, Wagner GP. Is the genotype–phenotype map modular? A statistical approach using mouse quantitative trait loci data. Genetics. 2000;156:305–11.
Montville R, Froissart R, Remold SK, Tenaillon O, Turner PE. Evolution of mutational robustness in an RNA virus. PLoS Biol. 2005;3:e381.
Narita Y, Kuratani S. Evolution of the vertebral formulae in mammals: a perspective on developmental constraints. J Exp Zool B Mol Dev Evol. 2005;304:91–106.
Olson EC, Miller RL. Morphological integration. Chicago: University of Chicago Press; 1958.
Orr AH. Adaptation and the cost of complexity. Evolution. 2000;54:13–20.
Parsons KJ, Marquez E, Albertson RC. Constraint and opportunity: the genetic basis and evolution of modularity in the cichlid mandible. Am Nat. 2012;179:64–78.
Pavlicev M, Hansen TF. Genotype–phenotype maps maximizing evolvability: modularity revisited. Evol Biol. 2011;38:371–89.
Pavlicev M, Wagner GP. A model of developmental evolution: selection, pleiotropy and compensation. Tree. 2012;27:316–22.
Pavlicev M, Kenney-Hunt JP, Norgard EA, Roseman CC, Wolf JB, Cheverud JM. Genetic variation in pleiotropy: differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. Evolution. 2008;62:199–213.
Pelabon C, Armbruster S, Hansen TF. Experimental evidence for the Berg hypothesis: vegetative traits are more sensitive than pollination traits to environmental variation. Funct Ecol. 2011;25:247–57.
Pigliucci M. Is evolvability evolvable? Nat Rev Genet. 2008;9:75–82.
Rechenberg I. Evolutionsstrategie. Stuttgart: Friedrich Frommann Verlag; 1973.
Rensch B. Evolution above the species level. New York: Columbia University Press; 1959.
Riedl RJ. Order in living organisms: a systems analysis of evolution. New York: Wiley; 1978.
Rohner N, Bercsenyi M, Orban L, Kolanczyk ME, Linke D, Brand M, et al. Duplication of *fgfr1* permits Fgf signaling to serve as a target for selection during domestication. Curr Biol. 2009;19:1642–7.
Rutherford SL. Between genotype and phenotype: protein chaperones and evolvability. Nat Rev Genet. 2003;4:263–74.
Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. Nature. 1998;396:336–42.
Schoch RR. Riedl's burden and the body plan: selection, constraint, and deep time. J Exp Zool B Mol Dev Evol. 2010;314:1–10.
Schuster P, Fontana W. Chance and necessity in evolution: lessons from RNA. Physica D. 1999;133:427–52.
Schuster P, Fontana W, Stadler PF, Hofacker IL. From sequences to shapes and back: a case study in RNA secondary structures. Proc Biol Sci. 1994;255:279–84.
Simpson GG. The major features of evolution. New York: Simon and Schuster; 1953.
Sober E. Fodor's *Bubbe Meise* against Darwinism. Mind and Lang. 2008;23:42–9.
Specchia V, Piacentini L, Tritto P, Fanti L, D'Alessandro R, Palumbo G, et al. Hsp90 prevents phenotypic variation by suppressing the mutagenic activity of transposons. Nature. 2010;463:662–5.
Su Z, Zeng Y, Gu X. A preliminary analysis of gene pleiotropy estimated from protein sequences. J Exp Zool B Mol Dev Evol. 2010;314:115–22.
Turner PE, Chao L. Sex and the evolution of intrahost competition in RNA virus phi6. Genetics. 1998;150:523–32.
Waddington CH. Canalization of development and the inheritance of acquired characters. Nature. 1942;150:563–5.
Waddington CH. The strategy of the genes. New York: MacMillan Co.; 1957.
Wagner A. Robustness and evolvability in living systems. Princeton studies in complexity. Princeton: Princeton University Press; 2005.
Wagner A. The role of robustness in phenotypic adaptation and innovation. Proc Biol Sci. 2012. doi:10.1098/rspb.2011.2293. Published before print.
Wagner GP, Altenberg L. Complex adaptations and evolution of evolvability. Evolution. 1996;50:967–76.
Wagner GP, Laubichler MD. Rupert Riedl and the re-synthesis of evolutionary and developmental biology: body plans and evolvability. J Exp Zool B Mol Dev Evol. 2004;302:92–102.

Wagner GP, Zhang J. The pleiotropic structure of the genotype–phenotype map: the evolvability of complex organisms. Nat Rev Genet. 2011;12:204–13.

Wagner GP, Zhang J. Universal pleiotropy is not a valid null hypothesis: reply to Hill and Zhang. Nat Rev Genet. 2012;13:296.

Wagner GP, Pavlicev M, Cheverud JM. The road to modularity. Nat Rev Genet. 2007;8:921–31.

Wagner GP, Kenney-Hunt JP, Pavlicev M, Peck JR, Waxman D, Cheverud JM. Pleiotropic scaling of gene effects and the 'cost of complexity'. Nature. 2008;452:470–2.

Wang Z, Liao BY, Zhang J. Genomic patterns of pleiotropy and the evolution of complexity. Proc Natl Acad Sci U S A. 2010a;107:18034–9.

Wang Z, Dong D, Ru B, Young RL, Han N, Guo T, Zhang S. Digital gene expression tag profiling of bat digits provides robust candidates contributing to wing formation. BMC Genomics. 2010b;11:619.

Woods RJ, Barrick JE, Cooper TF, Shrestha U, Kauth MR, Lenski RE. Second-order selection for evolvability in a large *Escherichia coli* population. Science. 2011;331:1433–6.

Young NM, Wagner GP, Hallgrimsson B. Development and the evolvability of human limbs. Proc Natl Acad Sci U S A. 2010;107:3400–5.

Zou L, Sriswasdi S, Ross B, Missiuro PV, Liu J, Ge H. Systematic analysis of pleiotropy in *C. elegans* early embryogenesis. PLoS Comput Biol. 2008;4:e1000003.