Review

The role of robustness in phenotypic adaptation and innovation

Andreas Wagner*

*Institute of Evolutionary Biology and Environmental Studies, University of Zurich,
Y27-J-54 Winterthurerstrasse 190, 8057 Zurich, Switzerland*

Phenotypes that vary in response to DNA mutations are essential for evolutionary adaptation and innovation. Therefore, it seems that robustness, a lack of phenotypic variability, must hinder adaptation. The main purpose of this review is to show why this is not necessarily correct. There are two reasons. The first is that robustness causes the existence of genotype networks—large connected sets of genotypes with the same phenotype. I discuss why genotype networks facilitate phenotypic variability. The second reason emerges from the evolutionary dynamics of evolving populations on genotype networks. I discuss how these dynamics can render highly robust phenotypes more variable, using examples from protein and RNA macromolecules. In addition, robustness can help avoid an important evolutionary conflict between the interests of individuals and populations—a conflict that can impede evolutionary adaptation.

Keywords: genotype networks; genotype space; conflict

1. INTRODUCTION

A feature or *phenotype* of an organism is robust if it persists when perturbed. Phenotypes encompass a broad range of traits, from macroscopic, visible traits, to molecular traits, such as the expression level of a gene, or the three-dimensional conformation of a protein.

The perturbations that can affect a phenotype fall into two broad categories. The first comprises environmental perturbations. These include changes in an organism's exterior environment, such as changes in temperature, in available nutrients or in the abundance of other organisms, such as potential prey. They also include changes in an organism's internal environment, such as temporal fluctuations in gene-expression levels, which are caused by ubiquitous intracellular noise. The second kind of perturbations are of mutations, changes in an organism's DNA, its *genotype*. Mutations affect an organism more permanently than environmental change because the changes they cause are readily inherited from generation to generation. For this reason, they are especially important study objects for students of evolution. I will here focus on robustness to genetic mutations—mutational robustness. A huge body of literature shows that living systems on all levels of organization—from macromolecules to whole organisms—are to some extent robust to mutations [1–4]. I note that mutationally robust systems are often also robust to environmental change [4–9], even though exceptions may exist [10,11]. Thus, most observations I make here about mutational robustness apply to environmentally robust systems as well.

During a population's evolutionary adaptation to a new environment, which takes place over multiple generations, its members undergo mutations. Most of these mutations are detrimental, and only few change an existing phenotype into a new, better-adapted phenotype [12–15]. Very occasionally, one or more mutations may also bring forth evolutionary innovations—new phenotypes that are qualitatively different and superior to the existing phenotypes. The ability of mutations to bring forth new phenotypes is important to Darwinian evolution. I will refer to it as *phenotypic variability*.

Phenotypic variability and robustness may seem opposite properties, because in a robust system, mutations do not easily change a phenotype. The main purpose of this review is to show why this view is not necessarily correct, and why robustness can instead be a prerequisite for phenotypic variability.

To make this point beyond individual case studies—anecdotes of natural history—one needs to study the relationship between genotypic change and the resulting phenotypic change *systematically*, either experimentally, through comparative data, or through computational modelling. I will focus here on protein and RNA macromolecules, for which this has become possible in recent years, but similar principles also hold for very different classes of systems (see the electronic supplementary material and Wagner [16]).

Other recent reviews have focused on phenotypic variability and its relationship to recombination [17], enzyme promiscuity [18], commonalities among different system classes [19] and phenotypic constraints [20]. In contrast, this review's focus is the role of robustness in phenotypic variability. In §2, I introduce some concepts and discuss the most difficult problem organisms face in evolutionary adaptation and innovation. Sections 3 and 4 focus on the two main respects in which robustness affects phenotypic variability. Specifically, §3 shows that robustness influences how genotypes with the same phenotype are organized in a vast space of genotypes.

*andreas.wagner@ieu.uzh.ch

This *static* or *structural* influence of robustness contrasts with its influence on the evolutionary *dynamics* of populations, which is the focus of §4. A brief §5 discusses the evolutionary conflict between the interests of an individual and that of a population in producing phenotypic variation. It points out that robustness as a variability mechanism can avoid this conflict. I will not discuss the important but controversial claim that mechanisms endowing living systems with robustness originated in evolution *because* they favour variability, for lack of sufficient evidence [21].

2. EXPLORING THE NEW WHILE CONSERVING THE OLD

Proteins and RNA catalyse all chemical reactions, provide structural support, help cells and organisms move, guide cell communication and carry out many other functions. The *genotype* of each such molecule is a sequence of amino acids or RNA nucleotides. *Genotype space* comprises all sequences of a given length L. It is astronomically large, and comprises 20^L protein genotypes and 4^L RNA genotypes. A molecule's *phenotype* refers to the fold or conformation that it forms in space, and to the biochemical function that this fold makes possible. New phenotypes in molecules arise through genotypic changes that cause phenotypic change. Macromolecules are well-studied, with many known adaptations and innovations [16,18,22,23].

Evolutionary change takes place in populations of organisms. Each member of a population has some genotype. For the purpose of this review, it is useful to think of a population as a collection of genotypes in genotype space. The members of this population change their location in this space through mutations. An especially important class of mutations are point mutations, which transform a genotype into one of its neighbours—a genotype that differs in one amino acid or nucleotide. The individuals in an evolving population also have some phenotype. Natural selection preserves individuals with well-adapted phenotypes and eliminates mutants with poorly adapted phenotypes.

Somewhere in genotype space, a superior genotype may exist whose phenotype is better adapted to the current environment. The central problem in evolutionary adaptation is how a population can find such a superior genotype. The problem is difficult, because this genotype may exist far away from the population's current genotype, and because the vast majority of mutant phenotypes a population explores are inferior and not superior to the existing phenotypes [12–14]. What is more, during a population's 'search' for this genotype, an existing well-adapted phenotype must be preserved. Perhaps the most compact way to express this problem is with an analogy from politics: evolving populations need to be both 'conservative' and 'progressive' at the same time. A tall order indeed.

Two generic features of genotype space make it possible to reconcile these conflicting demands. The first is the existence of connected genotype networks: vast sets of genotypes whose members all have the same phenotype (figure 1). These sets extend far through genotype space, and can be traversed in many small steps of individual mutations with little or no phenotypic change. Their existence was first suggested by computational models of phenotype formation [24,25], but they do also exist
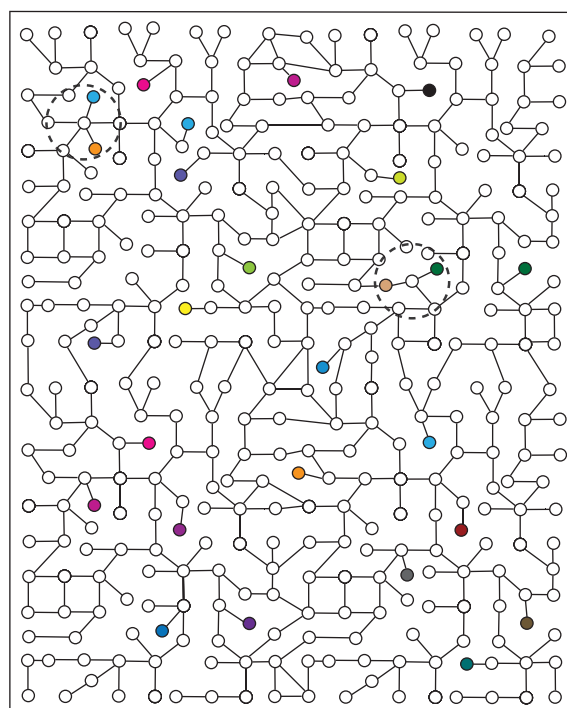

Figure 1. Connected genotype networks facilitate accessibility of diverse phenotypes. The figure schematically represents a hypothetical set of genotypes (small open black circles) in genotype space (rectangle) that share the same phenotype and form a genotype network; neighbouring genotypes are connected by black lines. Coloured circles indicate genotypes with different phenotypes. The two large dashed circles denote the neighbourhoods of two different genotypes on the genotype network. The upper left neighbourhood contains two novel phenotypes (blue and orange), whereas the lower right neighbourhood contains two different novel phenotypes (beige and green). The figure illustrates that many different novel phenotypes can be accessed from a connected genotype network that spreads far through genotype space. Note that a two-dimensional figure like this cannot capture many features of high-dimensional genotype spaces. These include the fact that individual genotypes can have thousands of neighbours, and not just the few shown here, and that the phenotypes shown in colour also have vast genotype networks that are not shown. Adapted from Wagner [16]. Reprinted with permission from Oxford University Press.

in real macromolecules. A paradigmatic example is the family of oxygen-binding globins. They comprise hundreds of known members that are connected through single amino acid changes to a common ancestor. They share a common fold and biochemical function—oxygen binding—but have diverged in more than 95 per cent of their amino acid residues [26,27]. Macromolecules like this are the rule rather than the exception [28–31].

Genotype networks are sometimes called neutral networks [25]. However, evolutionary change on such networks is usually all but neutral. That is, such change may affect fitness. For example, weakly deleterious mutations are more abundant than neutral mutations in most macromolecules, but they are often followed by compensatory changes that allow a preservation of phenotype. Similarly, in large populations, the simultaneous occurrence of multiple mutations can help a population 'tunnel' through a region of low fitness in genotype space, and thus help preserve a phenotype [13,32–36].

Because phenotype preservation does not require neutrality of individual mutations, I refrain from using the word 'neutrality' in this context.

The second central feature of genotype space regards the collection of those genotypes that can be reached from any one genotype through one or few mutations. This collection is also called a genotype's *neighbourhood*. Neighbourhoods are important, because the set of different phenotypes in a neighbourhood are easily accessible by mutation. The size of this set is thus a simple measure of how phenotypically variable a genotype is in response to mutations [37]. The second feature of genotype space is that neighbourhoods of different genotypes typically contain different novel phenotypes (figure 1; see also the electronic supplementary material).

The first feature, genotype networks, allows individuals in a population to preserve their phenotype while changing their genotype in many small steps that, cumulatively, can add up to substantial divergence. Because of the second feature, the different genotypes on a genotype network can explore different phenotypes, precisely because their neighbourhoods contain different novel phenotypes.

Thus far, I implicitly assumed that one genotype has one phenotype, a simplification that helps illustrate key concepts in simple terms. However, it is important to be aware that many molecules can form multiple folds and exert multiple functions [38,39]. Such multi-functionality can play an important role in the origin of novel phenotypes [18], and can further enhance the variability caused by genotype networks [16, ch. 13].

3. IN BRINGING FORTH GENOTYPE NETWORKS, ROBUSTNESS FACILITATES PHENOTYPIC VARIABILITY

To appreciate the role of robustness in phenotypic variability, it is useful to define robustness in the genotype space framework. Specifically, a genotype is to some extent robust to mutations if it has some neighbours with the same phenotype P as itself. One can show that robustness thus defined is both necessary and sufficient for the existence of genotype networks (see the electronic supplementary material).

Now compare, as a thought experiment, two kinds of genotypes. The first is a minimally robust genotype, that is, a genotype that has no neighbours with phenotype P. Figure 2a shows such a hypothetical, minimally robust genotype G (black circle) with eight neighbours (dashed black lines). The second genotype is a genotype with some robustness, as exemplified by the left-most hypothetical genotype G (black circle) in figure 2b. Half of the eight neighbours of this genotype have the same phenotype P as itself (solid lines), whereas the other half (dashed dark blue lines) have new phenotypes (not shown in the figure), all of which might be different from each other.

Which of these two genotypes is phenotypically more variable, in the sense that it can access a greater number of novel phenotypes through mutation? The answer is genotype G in figure 2a. Because it is minimally robust, all of its eight neighbours have a phenotype different from P. In contrast, the left-most genotype in figure 2b is more robust but less phenotypically variable, because only some of its neighbours have a phenotype different

from P. This is the core argument why robustness reduces phenotypic variability, cast in the abstract but precise language of genotype space.

Figure 2b also illustrates why this argument is flawed. The robust left-most genotype G has neighbours with the same phenotype P as itself, one of which, G', is shown as the middle circle in figure 2b. This neighbour G' itself has a neighbourhood containing five genotypes with new phenotypes (dashed medium blue lines) that may not occur in the neighbourhood of G itself. G' has three further neighbours (solid lines) that have the same phenotype P, one of which is shown as G''. The neighbourhood of G'' contains four genotypes with novel phenotypes (dashed light blue lines). In comparison with the non-robust genotype, from which up to eight different novel phenotypes are accessible, the robust genotype can—merely through the neighbours shown in the figure—access up to $4 + 5 + 4 = 13$ new phenotypes, more than if it were not robust. This argument takes only the neighbourhoods of G' and G'' into account, and not the neighbourhoods of several other neighbours of G with the same phenotype P. Thus, the actual number of different accessible phenotypes may be even higher for the robust genotype.

Three concrete examples show how much higher. These examples are based on three different natural RNA molecules, a guide RNA, a ribozyme and a telomerase (figure 2c, same colour coding as figure 2a,b). The phenotype in question is RNA secondary structure, a planar structure that arises when an RNA molecule folds onto itself through internal base pairing. Secondary structure is essential for the function of many RNA molecules, and thus in itself a phenotype worthy of study. It can be predicted computationally using known biophysical RNA folding principles [40].

Each of the three panels of figure 2c shows, in a black bar, the maximally possible number of novel secondary structure phenotypes accessible to an RNA genotype if its phenotype were minimally robust (as in figure 2a). This number equals the total number of neighbours of an RNA molecule, which equals three times the molecule's total length L in nucleotides, because every one of the molecule's nucleotides can mutate into three other nucleotides. For example, for the guide RNA with length $L = 40$, these would be $3 \times 40 = 120$ neighbours and novel phenotypes.

The dark blue bars indicate the *actual* number of accessible new phenotypes in the neighbourhood of a genotype. This number was obtained by computing the minimum free energy secondary structure phenotype of each neighbour of a genotype with an RNA folding algorithm [40]. For the guide RNA, this number is 40, many fewer than the maximally 120 new phenotypes without robustness, thus confirming the principle illustrated in figure 2a,b.

The medium and light blue bars indicate the total number of different phenotypes that are accessible up to two and three mutations away from the starting genotype, respectively. This number—obtained again by computing the structures for all genotypes in these neighbourhoods—is much larger than the 120 maximally attainable phenotypes in the absence of robustness. Specifically, for the guide RNA discussed here, 746 (medium blue bars) and 1174 (light blue bars) distinct new phenotypes
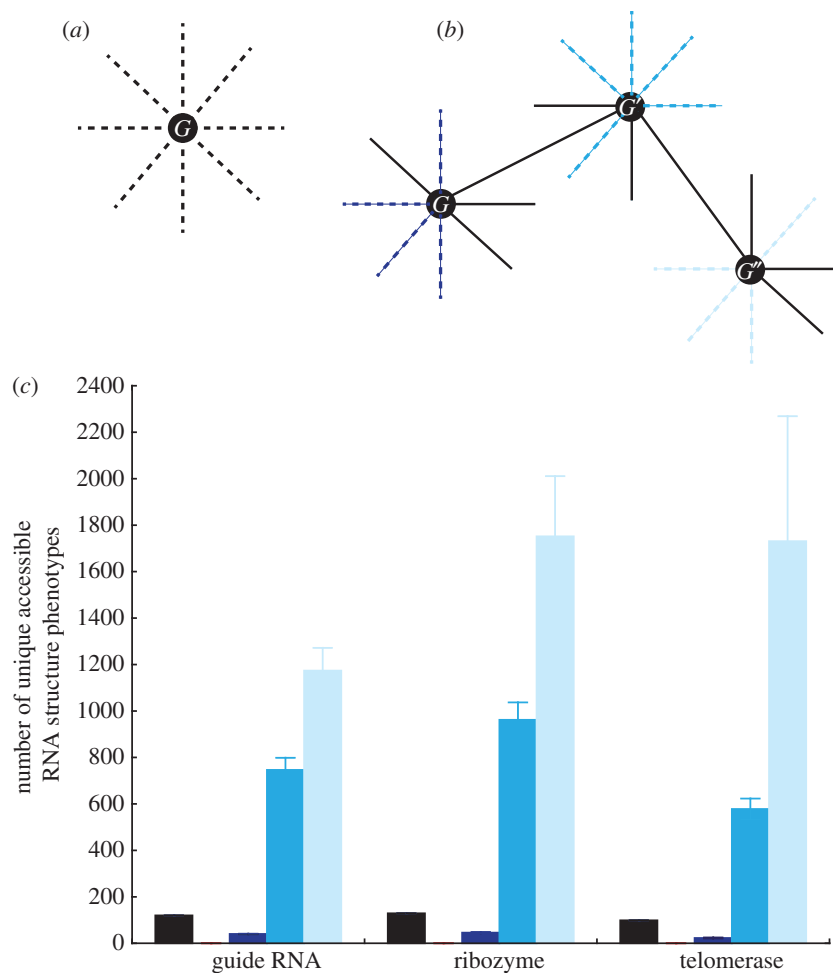
Figure 2. Robustness makes many phenotypic variants accessible to mutations. Circles in (*a,b*) represent genotypes with some hypothetical phenotype *P*. Straight lines connect a genotype to its neighbours (not all neighbours of a genotype are shown as circles), solid lines connect a genotype to a neighbour with the same phenotype *P* and dashed lines connect a genotype to a neighbour with a new phenotype. (*a*) The hypothetical genotype *G* shown here has no robustness, that is, no neighbours with the same phenotype. All eight of its neighbours have new phenotypes. It thus shows maximal phenotypic variability. (*b*) All three genotypes shown are to some extent robust, that is, they have neighbours with the same phenotype *P*. Dark, medium and light blue dashed lines point to genotypes with new phenotypes that are one, two and three mutations away from the left-most genotype in (*b*), respectively. Robustness makes more new genotypes accessible. See §3 for details. (*c*) The figure illustrates this principle through the actual number of new accessible phenotypes for three different natural RNA molecules (horizontal axis), and for computationally predicted [40] minimum free secondary structure phenotypes in their neighbourhood. Each of these molecules has some phenotype *P* (not shown). The black bar in each of the three panels indicates the maximally possible number of different phenotypes one mutation away from an RNA genotype *G*. This number is equal to 3*L*, where *L* is the number of nucleotides in a molecule. It would be attained only in the absence of robustness, as in (*a*). Dark, medium and light blue bars indicate, just as in (*b*), the number of distinct new phenotypes that are accessible in the neighbourhood of the molecule *G* ('1 mutation away'), in the neighbourhoods of all its neighbours *G'* with phenotype *P* ('2 mutations away') and in the neighbourhood of the neighbours *G''* of *G'* with phenotype *P* ('3 mutations away'). Data in (*c*) are averages (error bars: 1 s.e.m.) from 10 inversely folded [40] RNA genotypes per RNA secondary structure phenotype. The individual RNA molecules have been obtained from the functional RNA database (http://www.ncrna.org/frnadb) [41]. They include a guide RNA (*Trypanosoma brucei*, fRNAdb accession number L25590, *L* = 40 nucleotides), a hammerhead ribozyme (*Schistosoma mansoni*, accession number AF036740, *L* = 43) and a telomerase (*Moneuplotes crassa*, accession number AF061109; *L* = 33 nucleotides). See Jörg *et al.* ([42]; electronic supplementary material, table S1) for predicted secondary structure phenotypes *P* of these RNA molecules.

become accessible two and three mutations away. Thus, robustness allows access to many novel RNA phenotypes.

I note that the genotypes considered in this computation (up to two mutations away from *G* and with the same phenotype *P*) are a tiny fraction of the genotypes that form a typical genotype network. For example, a simple calculation shows that there are only 7.2×10^3 total genotypes that differ from the guide RNA with *L* = 40 in no more than two mutations. However, the size of this guide RNA's genotype network—which can be computed—equals approximately 9.1×10^{17} ($\pm 3.3 \times 10^{16}$) genotypes,

and is thus more than 14 orders of magnitude larger [42]. (Astronomically large genotype networks are typical for natural RNA molecules.) It is currently not feasible to compute the number of distinct phenotypes near a genotype network this large, but this number would also be surely astronomical.

Taken together, these observations mean that the mere existence of robustness makes a dramatic difference in phenotypic variability. It is the difference between the few novel phenotypes accessible in the immediate neighbourhood of a non-robust genotype, and the extremely

large number of new phenotypes accessible from the neighbours of a genotype network. By bringing genotype networks into existence, robustness makes vastly more new phenotypes accessible.

I will next discuss two lines of experimental evidence that indicate how important this principle is for the discovery of new molecular phenotypes.

The first experiment revolves around one natural and one synthetic ribozyme [43]. The natural ribozyme that catalyses its own cleavage is encoded by the human hepatitis delta virus. The synthetic RNA is the so-called class III self-ligating ribozyme, which joins an oligonucleotide substrate to its own 5′-end. The two ribozymes are unrelated in sequence and fold [43]. Schultes & Bartel [43] were able to design a mutational path through RNA genotype space that starts from either one of the ribozymes and leaves its native activity largely intact, until it reaches a hybrid ribozyme that is more than 40 mutational steps from each starting point. This hybrid can act both as a self-cleaving ribozyme and as a ligase, albeit with lower catalytic activity than the starting enzymes. By constructing a hybrid ribozyme and constructing a path through sequence space back to its ancestors, this work makes two key points. First, many consecutive changes in a genotype are possible that do not affect an RNA's (catalytic) phenotype. Without robustness, this would not be the case. Second, these changes can be very important intermediate steps to create a new catalytic function. Similar principles have been suggested for other ribozymes [44,45].

This experiment demonstrated the role of robustness in the origin of new phenotypes using an engineered path through genotype space. Biological evolution does not use such pre-meditated paths, but random changes in evolving populations. The next experiment shows that robustness is also highly relevant in such populations [46]. The experiment revolves around the concept of cryptic variation. Cryptic variation is genotypic variation in a population that is not visible on the level of phenotype [47]. An example is variation in genotypes on the same genotype network. Cryptic variation cannot exist without mutational robustness. The experiment asks whether cryptic variation can help a population find a new and superior genotype during an evolutionary search.

The study system was again an RNA ribozyme, the so-called *Azoarcus* ribozyme [48], which is a naturally occurring ribozyme that can ligate a short RNA fragment to its own end. One can subject populations of ribozymes like this to repeated cycles of mutagenesis, and to selection to maintain or to modify this catalytic activity [49].

The experiment in question compared two different kinds of populations—one that consisted mostly of identical or similar genotypes, all of them copies of a single ribozyme sequence, and another that consisted of many diverse genotypes (figure 3). Ribozymes in the two kinds of populations had similar catalytic activities on a specific RNA substrate, such that the average activities of the populations were indistinguishable. In other words, the first kind of populations contained little or no cryptic variation, whereas the second kind contained lots of it. The experiment then changed the chemical environments in which these populations existed. That is, it exposed both kinds of populations to a new, chemically modified RNA substrate, on which the starting

ribozyme has low catalytic activity. Both kinds of populations then experienced repeated rounds of mutagenesis and selection that favoured high activity on the new substrate [46].

Populations with much cryptic variation adapted up to six times faster to the new chemical environment (figure 3b) [46]. They did so through genetic changes that improved the ribozyme's catalytic activity on the new substrate. DNA sequencing of thousands of genotypes from evolving populations subsequently showed why: there exists a superior genotype, and populations with much cryptic variation discover this genotype faster because they are genotypically more diverse, and they contain genotypes that are already close in genotype space to the superior genotype. In sum, this experiment shows that cryptic variation—a consequence of robustness—can accelerate evolutionary adaptation to a new chemical environment. Similar phenomena are likely to exist in proteins [50,51], even though we do not yet have proof that cryptic variation can accelerate the rate of adaptation for them.

4. ROBUSTNESS CAN FACILITATE PHENOTYPIC VARIABILITY BY AFFECTING EVOLUTIONARY DYNAMICS ON LARGE GENOTYPE NETWORKS

Thus far, I have argued that robustness can facilitate variability through its *static, structural* role in organizing genotypes with the same phenotype into genotype networks. I will next turn to a second role of robustness, which builds on the first role: robustness can increase variability through its influence on the evolutionary *dynamics* of populations on genotype networks. To this end, I will first introduce some further terminology from the genotype space framework. After that, I will explain how robustness can affect the evolutionary dynamics of populations, and then discuss a mix of pertinent experimental, computational and comparative data.

Some phenotypes have very large associated genotype networks and are formed by many different genotypes. Others have much smaller genotype networks and are formed by fewer genotypes [25,42,52–55]. This difference in genotype network size is accompanied by a difference in the average robustness of genotypes encoding these phenotypes. Specifically, the larger a phenotype *P*'s genotype network is, the greater is also the average fraction of each genotype's neighbours with this phenotype *P*. In other words, genotypes on a large genotype network are more robust to mutations than genotypes on a small genotype network [37,56]. This observation allows one to extend the definition of robustness that I have used thus far—the number of a *genotype*'s neighbours with the same phenotype. Specifically, one can define the robustness of a *phenotype* as the average robustness of all genotypes encoding it. Phenotypes with large genotype networks are more robust.

Now consider a population of initially identical genotypes with the same phenotype *P*. Subject the population to repeated cycles ('generations') of mutations and selection that confines the population to the genotype network of the phenotype *P*. After a given number of generations, examine the neighbourhood of each individual in the population, and enumerate the number of *different* or *unique* phenotypes that occur in these neighbourhoods. That is, if the same phenotype is formed by two or more
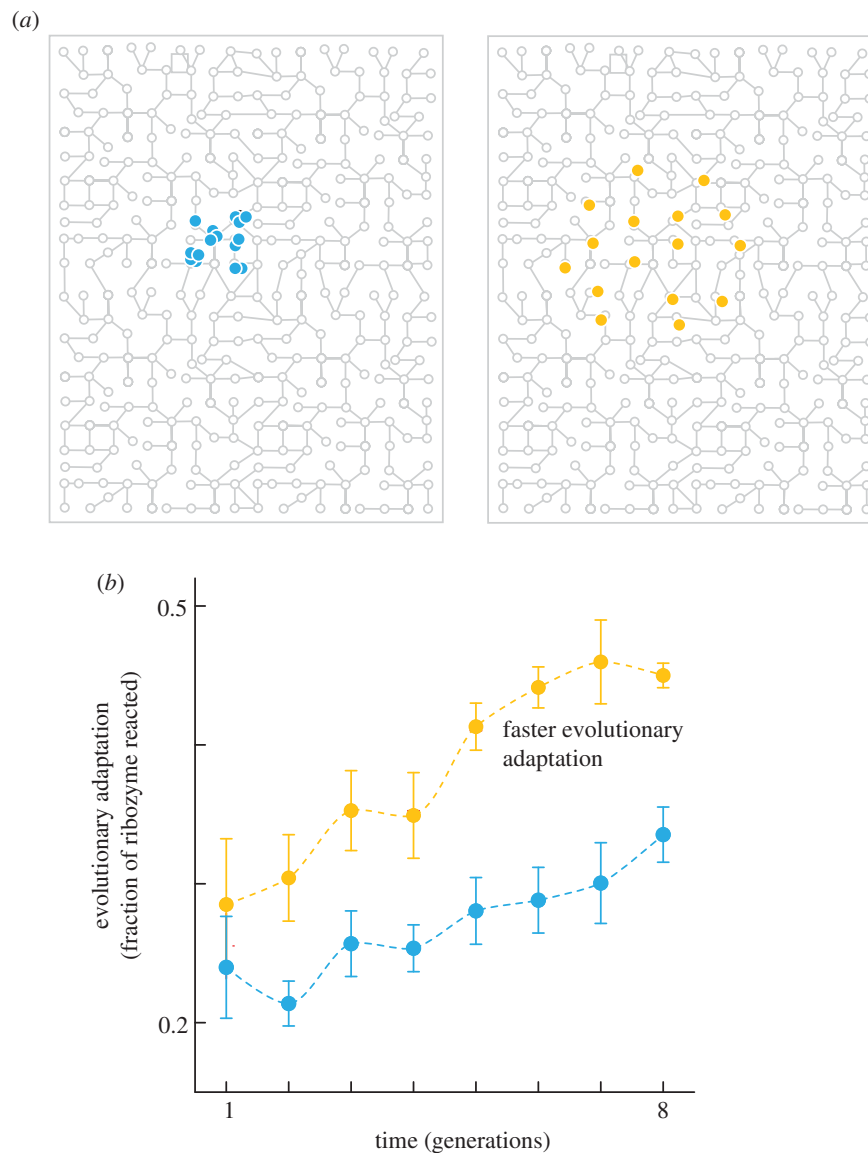
(a)

(b)

Figure 3. Cryptic variation can facilitate evolutionary adaptation. (*a*) Each large rectangle represents a genotype space into which a hypothetical genotype network is inscribed (grey open circles connected by grey lines). The coloured circles symbolize individuals in a population on this genotype network. The population on the left (blue circles) is genotypically less diverse, and thus contains less cryptic variation than the population on the right (yellow circles). (*b*) A laboratory evolution experiment showing how fast two populations of ribozymes with indistinguishable phenotype (catalytic activity on an RNA substrate) but different amounts of cryptic genotypic variation adapt evolutionarily to a new RNA substrate. As in (*a*), blue and yellow correspond to populations with little and much cryptic variation. The horizontal axis shows time in generations and the vertical axis shows a measure of the biochemical activity of each population on the new RNA substrate. The population with more cryptic variation adapts faster [46].

genotypes in these neighbourhoods, count it only once. This number is a measure of the phenotypic variability of an entire population, not just of a single individual. It encompasses all phenotypes that a population can access through a single nucleotide change in some individual.

To understand how phenotypic variability is affected by phenotypic robustness, it is necessary to examine how populations evolve on genotype networks that vary in size. Recent work on populations of evolving RNA molecules has done that for computationally predicted secondary structure phenotypes [37]. It found that populations whose phenotypes have greater robustness also show greater phenotypic variability. This observation is based on thousands of randomly sampled phenotypes, and is thus independent of any one particular phenotype. It is a generic feature of RNA genotype space.

To understand this observation, one needs to understand two different phenomena with opposite effects on phenotypic variability. The first of these is the number of different phenotypes in the neighbourhood of any one genotype. This number will be lower for highly robust phenotypes, because their genotypes have, on average, more neighbours with unchanged phenotype. The second phenomenon is the rate at which a population spreads through a genotype network. This rate is determined by the likelihood that a mutation is deleterious, that is, that it does not preserve the phenotype P. Individuals suffering deleterious mutations are eliminated from the population, which slows the population's diversification. The greater the incidence of such mutations, the more slowly a population spreads through genotype space. A lack of robustness thus acts like a brake on the

genotypic diversification process of a population. This diversification process is important, because the fraction of unique phenotypes in the neighbourhoods of two genotypes increases with the distance between them (figure 1). This means, as I discussed earlier, that populations with greater (cryptic) genotypic diversity can access more novel phenotypes through mutations. They have greater phenotypic variability.

In sum, considering only the first phenomenon, high phenotypic robustness entails low variability. In contrast, considering the second phenomenon, high robustness entails high variability. In evolving populations, these two phenomena have opposite effects on variability. For RNA secondary structure phenotypes, the second phenomenon—greater population diversity—exerts the dominant influence on phenotypic variability. This is why more robust phenotypes have higher phenotypic variability overall [37].

Observations from computational analyses like these can help us appreciate that we must study the dynamics of evolving populations—not just individual genotypes—to understand the quantitative link between robustness and phenotypic variability. A combination of experimental evolution work and comparative analyses further indicates that robustness also matters for real molecules.

One class of experiments worth highlighting regards chaperones, proteins that assist other proteins in folding and help to maintain their fold and function. Chaperones can reduce the effects of environmental stress, such as high temperature, and they can eliminate the deleterious effects of some mutations that reduce protein stability and abolish a protein's activity [57,58]. In the language of genotype space, one could say that a chaperone increases the size of the genotype network of a particular phenotype because it can render some mutations neutral that would otherwise be deleterious or lethal. In other words, a chaperone can make a phenotype more robust. Recent laboratory evolution experiments on four different enzymes expressed in *Escherichia coli* support this notion. Specifically, populations of these enzymes tolerated twice as many amino acid changes and evolved greater genotypic diversity when large amounts of a chaperone were present. One of these enzymes, a phosphotriesterase that can hydrolyse the pesticide paraoxon, was also subjected to laboratory evolution for activity on a new catalytic substrate, 2-naphtylhexanoate. Populations of this enzyme attained higher activity and specificity on the new substrate when the chaperone was overexpressed. In sum, high robustness—in this case induced by a chaperone—is associated with superior evolutionary adaptation [59].

Laboratory evolution experiments of enzymes also provide relevant evidence independent from that of chaperones [50,60,61]. A case in point is cytochrome P450, which belongs to an enzyme superfamily whose members hydroxylate many different substrate molecules. The relevant experiments mutagenized different variants of this enzyme that differed in their thermodynamic stability, and in their robustness to mutations. The stable and more robust variants of cytochrome P450 more readily evolved the ability to hydrolyse new substrates, such as the anti-inflammatory compound naproxen [60,61].

Experiments like these can show how robustness can facilitate evolutionary adaptation on short, laboratory time-scales. They are silent about how this relationship translates

onto the enormous time-scales on which proteins diversified in life history. Only a comparative analysis of the phenotypic diversity of proteins—a record of past evolutionary innovation—can answer this question. That is, it can answer whether highly robust protein phenotypes have adopted many different functions in their evolutionary history.

Such an analysis has become possible with the ability to estimate the robustness of protein folds (not just genotypes) to point mutations [62], and to estimate the functional diversity of proteins, for example through well-catalogued enzyme functions. A recent study of 112 ancient protein folds showed that highly robust folds have evolved greater functional diversity using different and complementary measures of functional diversity [63].

In sum, evidence that ranges from computational to comparative and experimental work suggests that more phenotypic robustness can increase the ability of RNA and protein molecules to adapt and diversify in evolution. The computational work i discussed earlier in this section helps explain why: phenotypic robustness accelerates the spreading of populations through a genotype network, makes a broader spectrum of phenotypes accessible through mutation and thus increases the likelihood of encountering a beneficial phenotype.

5. ROBUSTNESS CAN HELP AVOID CONFLICTS BETWEEN INDIVIDUALS AND POPULATIONS IN BRINGING FORTH PHENOTYPIC VARIATION

It is sometimes stated that biological systems bring forth novel features because this ability has been 'selected for'. This assertion is naive and problematic. To see why, consider mutator alleles, variants of genes that can increase an organism's mutation rate (and phenotypic variability) dramatically [64,65]. Mutators can be quite abundant in bacterial populations [65]. A facile explanation for their abundance resorts to the advantage they confer to a *population*: they help create many new phenotypes. Even though most of these new phenotypes may be deleterious, the few beneficial phenotypes may help the population survive in a challenging environment. However, this advantage is overshadowed by a great disadvantage to the *individual*—typically just one in a large population—who first acquires a mutator: because most mutations are deleterious, carrying the mutator genotype is detrimental to this individual and will thus often lead to its extinction [64]. A conflict thus exists between the interests of a population and that of an individual. How this conflict is resolved may depend on details of a population's life history and environment. Sometimes the conflict may be resolved in favour of the population, and at other times in favour of the individual. In the latter case, variability would be reduced. Thus, the emergence of phenotypic variability in evolution is not a foregone conclusion. Similar conflicts exist for other mechanisms that facilitate phenotypic variability [66].

Robustness as a variability principle, however, has a remarkable property: it can avoid this conflict. In RNA and proteins, where more robustness promotes greater variability, the interests of the individual and the lineage can be perfectly aligned. This is a simple consequence of how robustness influences the evolutionary dynamics of populations. Consider a population where stabilizing selection maintains a well-adapted phenotype. If this

phenotype is highly robust, it is not easily perturbed through mutation or environmental changes, because the two kinds of robustness are usually correlated [4–9]. Such robustness is advantageous for an individual that has this phenotype, because this individual experiences fewer perturbations with deleterious effects. At the same time, it is also advantageous for populations of such individuals. The reason lies in the evolutionary dynamics I discussed in the preceding section: robust phenotypes in both RNA and protein molecules show greater phenotypic variability, and can become phenotypically more diverse on evolutionary time-scales [37,63]. Robustness can thus benefit both an individual and its lineage. Evolutionary conflicts are among the most serious impediments to adaptation, which makes their avoidance here even more significant [67]. Their general role in the evolution of phenotypic variability needs further study.

6. SUMMARY AND OPEN QUESTIONS

In sum, I have distinguished between two roles of robustness in evolutionary adaptation and innovation: a structural and a dynamic role. First, robustness causes the existence of genotype networks, complex web-like structures formed by genotypes with the same phenotype, which facilitate phenotypic variability. Second, a robust phenotype can help the evolutionary exploration of new phenotypes in macromolecules by accelerating the dynamics of change in an evolving population.

Many open questions persist in this young research field. They fall into two broad classes. The first regards quantitative aspects of evolutionary dynamics. How do the sizes of evolving populations and their mutation rates interact with robustness to influence phenotypic variability? Do the principles I discuss here also apply in environments that change rapidly and continually, where populations always track a moving optimal phenotype? Do these principles apply to systems with extremely high or low robustness? Does robustness also accelerate the evolutionary exploration of new phenotypes in systems other than macromolecules, such as evolving regulatory circuits or metabolic networks? Population genetic models and computational analyses of genotype–phenotype relationships are beginning to tackle these questions [5,68–71]. However, we still lack a sufficient body of concordant evidence from different approaches to draw general conclusions.

A second class of questions regards evolutionary changes in robustness itself. Experimental and comparative work suggests that the robustness of macromolecules can change on evolutionary time-scales [72–74]. If robustness benefits both individuals and populations, then natural selection may favour robust phenotypes. If so, the robustness of phenotypes might increase over time. Only tentative evidence exists that naturally occurring phenotypes may be unusually robust [42,75]. We do not yet know the causes of this robustness, we do not yet have relevant evidence from other system classes and we are ignorant about the population genetic conditions under which such an increase would occur. Only with such evidence will we be able to answer a last and most fundamental question: does robustness evolve in a way that facilitates evolutionary adaptation and innovation?

REFERENCES

1 Kitano, H. 2004 Biological robustness. *Nat. Rev. Genet.* 5, 826–837. (doi:10.1038/nrg1471)

2 Stelling, J., Sauer, U., Szallasi, Z., Doyle, F. J. & Doyle, J. 2004 Robustness of cellular functions. *Cell* 118, 675–685. (doi:10.1016/j.cell.2004.09.008)

3 Wagner, A. 2005 Distributed robustness versus redundancy as causes of mutational robustness. *Bioessays* 27, 176–188. (doi:10.1002/bies.20170)

4 Wagner, A. 2005 *Robustness and evolvability in living systems.* Princeton, NJ: Princeton University Press.

5 Ancel, L. W. & Fontana, W. 2000 Plasticity, evolvability, and modularity in RNA. *J. Exp. Zool./Mol. Dev. Evol.* 288, 242–283. (doi:10.1002/1097-010X(20001015) 288:3<242::AID-JEZ5>3.0.CO;2-O)

6 Lehner, B. 2010 Genes confer similar robustness to environmental, stochastic, and genetic perturbations in yeast. *PLoS ONE* 5, e9035. (doi:10.1371/journal.pone.0009035)

7 Masel, J. & Siegal, M. L. 2009 Robustness: mechanisms and consequences. *Trends Genet.* 25, 395–403. (doi:10.1016/j.tig.2009.07.005)

8 Meiklejohn, C. & Hartl, D. 2002 A single mode of canalization. *Trends Ecol. Evol.* 17, 468–473. (doi:10.1016/S0169-5347(02)02596-X)

9 Szollosi, G. J. & Derenyi, I. 2009 Congruent evolution of genetic and environmental robustness in micro-RNA. *Mol. Biol. Evol.* 26, 867–874. (doi:10.1093/molbev/msp008)

10 Cooper, T. F., Morby, A. P., Gunn, A. & Schneider, D. 2006 Effect of random and hub gene disruptions on environmental and mutational robustness in *Escherichia coli. BMC Genomics* 7, 237. (doi:10.1186/1471-2164-7-237)

11 Milton, C. C., Huynh, B., Batterham, P., Rutherford, S. L. & Hoffmann, A. A. 2003 Quantitative trait symmetry independent of Hsp90 buffering: distinct modes of genetic canalization and developmental stability. *Proc. Natl Acad. Sci. USA* 100, 13 396–13 401. (doi:10.1073/pnas.1835613100)

12 Eyre-Walker, A. & Keightley, P. D. 2007 The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* 8, 610–618. (doi:10.1038/nrg2146)

13 Sawyer, S. A., Parsch, J., Zhang, Z. & Hartl, D. L. 2007 Prevalence of positive selection among nearly neutral amino acid replacements in *Drosophila. Proc. Natl Acad. Sci. USA* 104, 6504–6510. (doi:10.1073/pnas.0701572104)

14 Soskine, M. & Tawfik, D. S. 2010 Mutational effects and the evolution of new protein functions. *Nat. Rev. Genet.* 11, 572–582. (doi:10.1038/nrg2808)

15 Wloch, D. M., Szafraniec, K., Borts, R. H. & Korona, R. 2001 Direct estimate of the mutation rate and the distribution of fitness effects in the yeast *Saccharomyces cerevisiae. Genetics* 159, 441–452.

16 Wagner, A. 2011 *The origins of evolutionary innovations. A theory of transformative change in living systems.* Oxford, UK: Oxford University Press.

17 Masel, J. & Trotter, M. V. 2010 Robustness and evolvability. *Trends Genet.* 26, 406–414. (doi:10.1016/j.tig.2010.06.002)

18 Khersonsky, O. & Tawfik, D. S. 2010 Enzyme promiscuity: a mechanistic and evolutionary perspective. In *Annual Review of Biochemistry* (eds R. D. Kornberg, C. R. H. Raetz, J. E. Rothman & J. W. Thorner), vol. 79, pp. 471–505. Palo Alto, CA: Annual Reviews. (doi:10.1146/annurev-biochem-030409-143718)

19 Wagner, A. 2011 The molecular origins of evolutionary innovations. *Trends Genet.* 27, 397–410. (doi:10.1016/j.tig.2011.06.002)

20 Wagner, A. 2011 Genotype networks shed light on evolutionary constraints. *Trends Ecol. Evol.* **26**, 577–584. (doi:10.1016/j.tree.2011.07.001)

21 Pigliucci, M. 2008 Is evolvability evolvable? *Nat. Rev. Genet.* **9**, 75–82. (doi:10.1038/nrg2278)

22 Cheng, C. 2006 Old proteins and DNA, new tricks: molecular evolution of antifreeze proteins in polar fishes. *Comp. Biochem. Physiol. A-Mol. Integr. Physiol.* **143**, S61–S64. (doi:10.1016/j.cbpa.2006.01.059)

23 Li, H. 1997 *Molecular evolution*. Sunderland, MA: Sinauer.

24 Lipman, D. & Wilbur, W. 1991 Modeling neutral and selective evolution of protein folding. *Proc. R. Soc. Lond. B* **245**, 7–11. (doi:10.1098/rspb.1991.0081)

25 Schuster, P., Fontana, W., Stadler, P. & Hofacker, I. 1994 From sequences to shapes and back: a case-study in RNA secondary structures. *Proc. R. Soc. Lond. B* **255**, 279–284. (doi:10.1098/rspb.1994.0040)

26 Goodman, M., Pedwaydon, J., Czelusniak, J., Suzuki, T., Gotoh, T., Moens, L., Shishikura, F., Walz, D. & Vinogradov, S. 1988 An evolutionary tree for invertebrate globin sequences. *J. Mol. Evol.* **27**, 236–249. (doi:10.1007/BF02100080)

27 Hardison, R. C. 1996 A brief history of hemoglobins: plant, animal, protist, and bacteria. *Proc. Natl Acad. Sci. USA* **93**, 5675–5679. (doi:10.1073/pnas.93.12.5675)

28 Bastolla, U., Porto, M., Roman, H. E. & Vendruscolo, M. 2003 Connectivity of neutral networks, overdispersion, and structural conservation in protein evolution. *J. Mol. Evol.* **56**, 243–254. (doi:10.1007/s00239-002-2350-0)

29 Rost, B. 2002 Enzyme function less conserved than anticipated. *J. Mol. Biol.* **318**, 595–608. (doi:10.1016/S0022-2836(02)00016-5)

30 Thornton, J., Orengo, C., Todd, A. & Pearl, F. 1999 Protein folds, functions and evolution. *J. Mol. Biol.* **293**, 333–342. (doi:10.1006/jmbi.1999.3054)

31 Todd, A., Orengo, C. & Thornton, J. 1999 Evolution of protein function, from a structural perspective. *Curr. Opin. Chem. Biol.* **3**, 548–556. (doi:10.1016/S1367-5931(99)00007-1)

32 Eyre-Walker, A., Keightley, P. D., Smith, N. G. C. & Gaffney, D. 2002 Quantifying the slightly deleterious mutation model of molecular evolution. *Mol. Biol. Evol.* **19**, 2142–2149.

33 Hartl, D. & Clark, A. 2007 *Principles of population genetics*. Sunderland, MA: Sinauer Associates.

34 Kern, A. D. & Kondrashov, F. A. 2004 Mechanisms and convergence of compensatory evolution in mammalian mitochondrial tRNAs. *Nat. Genet.* **36**, 1207–1212. (doi:10.1038/ng1451)

35 Kulathinal, R. J., Bettencourt, B. R. & Hartl, D. L. 2004 Compensated deleterious mutations in insect genomes. *Science* **306**, 1553–1554. (doi:10.1126/science.1100522)

36 Weinreich, D. M. & Chao, L. 2005 Rapid evolutionary escape by large populations from local fitness peaks is likely in nature. *Evolution* **59**, 1175–1182. (doi:10.1554/04-392)

37 Wagner, A. 2008 Robustness and evolvability: a paradox resolved. *Proc. R. Soc. B* **275**, 91–100. (doi:10.1098/rspb.2007.1137)

38 O'Brien, P. J. & Herschlag, D. 1999 Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* **6**, R91–R105. (doi:10.1016/S1074-5521(99)80033-7)

39 Tokuriki, N. & Tawfik, D. S. 2009 Protein dynamism and evolvability. *Science* **324**, 203–207. (doi:10.1126/science.1169375)

40 Hofacker, I., Fontana, W., Stadler, P., Bonhoeffer, L., Tacker, M. & Schuster, P. 1994 Fast folding and comparison of RNA secondary structures. *Monatshefte fuer Chemie* **125**, 167–188. (doi:10.1007/BF00818163)

41 Kin, T. *et al.* 2007 fRNAdb: a platform for mining/annotating functional RNA candidates from non-coding RNA sequences. *Nucleic Acids Res.* (*Database Issue*) **35**, D145–D148. (doi:10.1093/nar/gkl837)

42 Jörg, T., Martin, O. & Wagner, A. 2008 Neutral network sizes of biological RNA molecules can be computed and are not atypically small. *BMC Bioinform.* **9**, 464. (doi:10.1186/1471-2105-9-464)

43 Schultes, E. & Bartel, D. 2000 One sequence, two ribozymes: implications for the emergence of new ribozyme folds. *Science* **289**, 448–452. (doi:10.1126/science.289.5478.448)

44 Beckert, B., Nielsen, H., Einvik, C., Johansen, S. D., Westhof, E. & Masquida, B. 2008 Molecular modelling of the GIR1 branching ribozyme gives new insight into evolution of structurally related ribozymes. *EMBO J.* **27**, 667–678. (doi:10.1038/emboj.2008.4)

45 Huang, Z. & Szostak, J. W. 2003 Evolution of aptamers with a new specificity and new secondary structures from an ATP aptamer. *RNA* **9**, 1456–1463. (doi:10.1261/rna.5990203)

46 Hayden, E., Ferrada, E. & Wagner, A. 2011 Cryptic genetic variation promotes rapid evolutionary adaptation in an RNA enzyme. *Nature* **474**, 92–95. (doi:10.1038/nature10083)

47 Gibson, G. & Reed, L. K. 2008 Cryptic genetic variation. *Curr. Biol.* **18**, R989–R990. (doi:10.1016/j.cub.2008.08.011)

48 Tanner, M. A. & Cech, T. R. 1996 Activity and thermostability of the small self-splicing group I intron in the pre-tRNA(IIe) of the purple bacterium *Azoarcus*. *RNA–A Publ. RNA Soc.* **2**, 74–83.

49 Beaudry, A. A. & Joyce, G. F. 1992 Directed evolution of an RNA enzyme. *Science* **257**, 635–641. (doi:10.1126/science.1496376)

50 Amitai, G., Gupta, R. D. & Tawfik, D. S. 2007 Latent evolutionary potentials under the neutral mutational drift of an enzyme. *HFSP J.* **1**, 67–78. (doi:10.2976/1.2739115)

51 Bloom, J. D., Romero, P. A., Lu, Z. & Arnold, F. H. 2007 Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution. *Biol. Direct* **2**, 17. (doi:10.1186/1745-6150-2-17)

52 Ciliberti, S., Martin, O. C. & Wagner, A. 2007 Circuit topology and the evolution of robustness in complex regulatory gene networks. *PLoS Comput. Biol.* **3**, e15. (doi:10.1371/journal.pcbi.0030015)

53 Ciliberti, S., Martin, O. C. & Wagner, A. 2007 Innovation and robustness in complex regulatory gene networks. *Proc. Natl Acad. Sci. USA* **104**, 13591–13596. (doi:10.1073/pnas.0705396104)

54 Li, H., Helling, R., Tang, C. & Wingreen, N. 1996 Emergence of preferred structures in a simple model of protein folding. *Science* **273**, 666–669. (doi:10.1126/science.273.5275.666)

55 Samal, A., Rodrigues, J. F. M., Jost, J., Martin, O. C. & Wagner, A. 2010 Genotype networks in metabolic reaction spaces. *BMC Syst. Biol.* **4**, 30. (doi:10.1186/1752-0509-4-30)

56 Reidys, C., Stadler, P. & Schuster, P. 1997 Generic properties of combinatory maps: neutral networks of RNA secondary structures. *Bull. Math. Biol.* **59**, 339–397. (doi:10.1007/BF02462007)

57 Fares, M. A., Ruiz-Gonzalez, M. X., Moya, A., Elena, S. F. & Barrio, E. 2002 Endosymbiotic bacteria: GroEL buffers against deleterious mutations. *Nature* **417**, 398. (doi:10.1038/417398a)

58 Hartl, F. U. & Hayer-Hartl, M. 2002 Protein folding: molecular chaperones in the cytosol from nascent chain to folded protein. *Science* **295**, 1852–1858. (doi:10.1126/science.1068408)

59 Tokuriki, N. & Tawfik, D. S. 2009 Chaperonin overexpression promotes genetic variation and enzyme evolution. *Nature* **459**, 668–673. (doi:10.1038/nature08009)

60 Bloom, J. D., Labthavikul, S. T., Otey, C. R. & Arnold, F. H. 2006 Protein stability promotes evolvability. *Proc. Natl Acad. Sci. USA* **103**, 5869–5874. (doi:10.1073/pnas.0510098103)

61 Bloom, J. D., Silberg, J. J., Wilke, C. O., Drummond, D. A., Adami, C. & Arnold, F. H. 2005 Thermodynamic prediction of protein neutrality. *Proc. Natl Acad. Sci. USA* **102**, 606–611. (doi:10.1073/pnas.0406744102)

62 England, J. L. & Shakhnovich, E. I. 2003 Structural determinant of protein designability. *Phys. Rev. Lett.* **90**, 218101. (doi:10.1103/PhysRevLett.90.218101)

63 Ferrada, E. & Wagner, A. 2008 Protein robustness promotes evolutionary innovations on large evolutionary time scales. *Proc. R. Soc. B* **275**, 1595–1602. (doi:10.1098/rspb.2007.1617)

64 Sniegowski, P., Gerrish, P., Johnson, T. & Shaver, A. 2000 The evolution of mutation rates: separating causes from consequences. *Bioessays* **22**, 1057–1066. (doi:10.1002/1521-1878(200012)22:12<1057::AID-BIES3>3.0.CO;2-W)

65 Taddei, F., Radman, M., Maynard-Smith, J., Toupance, B., Gouyon, P. & Godelle, B. 1997 Role of mutator alleles in adaptive evolution. *Nature* **387**, 700–702. (doi:10.1038/42696)

66 Kirschner, M. & Gerhart, J. 1998 Evolvability. *Proc. Natl Acad. Sci. USA* **95**, 8420–8427. (doi:10.1073/pnas.95.15.8420)

67 Futuyma, D. J. 2009 *Evolution*. Sunderland, MA: Sinauer.

68 Draghi, J., Parsons, T., Wagner, G. P. & Plotkin, J. 2010 Mutational robustness can facilitate adaptation. *Nature* **463**, 353–355. (doi:10.1038/nature08694)

69 Draghi, J. & Wagner, G. R. 2008 Evolution of evolvability in a developmental model. *Evolution* **62**, 301–315. (doi:10.1111/j.1558-5646.2007.00303.x)

70 Espinosa-Soto, C., Martin, O. C. M. & Wagner, A. 2010 Phenotypic robustness can increase phenotypic variability after non-genetic perturbations in gene regulatory circuits. *J. Evol. Biol.* **24**, 1284–1297. (doi:10.1111/j.1420-9101.2011.02261.x)

71 Rodrigues, J. F. & Wagner, A. 2011 Genotype networks in sulfur metabolism. *BMC Syst. Biol.* **5**, 39. (doi:10.1186/1752-0509-5-39)

72 Montville, R., Froissart, R., Remold, S. K., Tenaillon, O. & Turner, P. E. 2005 Evolution of mutational robustness in an RNA virus. *PLoS Biol.* **3**, 1939–1945. (doi:10.1371/journal.pbio.0030381)

73 Sanjuan, R., Forment, J. & Elena, S. F. 2006 *In silico* predicted robustness of viroids RNA secondary structures. I. The effect of single mutations. *Mol. Biol. Evol.* **23**, 1427–1436. (doi:10.1093/molbev/msl005)

74 Wagner, A. & Stadler, P. 1999 Viral RNA and evolved mutational robustness. *J. Exp. Zool./Mol. Dev. Evol.* **285**, 119–127. (doi:10.1002/(SICI)1097-010X(19990815)285:2<119::AID-JEZ4>3.0.CO;2-D)

75 Cowperthwaite, M. C., Economo, E. P., Harcombe, W. R., Miller, E. L. & Meyers, L. A. 2008 The ascent of the abundant: how mutational networks constrain evolution. *PLoS Comput. Biol.* **4**, e1000110. (doi:10.1371/journal.pcbi.1000110)