Evolution of Phenotypic Robustness

Joachim Hermisson and Günter P. Wagner

Department of Ecology and Evolutionary Biology, Yale University,
New Haven, CT 06520, USA

1 Introduction

Evolutionary biology, in the neo-Darwinian tradition, is based on the study of genetic and phenotypic variation and its fate in populations. Thus the observation that genetic variability of a trait is itself influenced by the genotype has obvious theoretical implications (Waddington, 1957; Stearns, 1994). It is in this context that the robustness of phenotypic traits was first conceptualized as *canalization* and became the focus of a significant research effort (reviewed in Scharloo, 1991). With increasing awareness of the intricate molecular mechanisms maintaining the life of cells, the ubiquity of buffering and compensatory mechanisms came into focus (Wilkins, 1997; Gerhart and Kirschner, 1997; Rutherford, 2000). Recent years have seen a confluence of the classical concept of canalization and new research in molecular biology that resulted in a sharp increase in the interest in canalization and related phenomena. One can speak of an emerging field of *biological robustness* research that is able to draw on a sophisticated arsenal of technical and theoretical tools that were developed over the last ten years (de Visser *et al.*, 2003).

The present chapter aims at summarizing the principal findings of the classical and the newer literature on robustness of phenotypes and to identify the issues that require attention in future research. We will start with an attempt to formalize the notion of phenotypic robustness to clarify the criteria that need to be met in order to experimentally demonstrate phenotypic robustness. Then we review the experimental evidence about environmental and genetic robustness and find that genetic robustness is particularly difficult to demonstrate. Finally an overview of theoretical models for the evolution of canalization is provided which shows that the existing literature is strongly biased towards a few scenarios. We conclude that there are major unsolved questions both in the experimental demonstration of canalization as well as in understanding the evolutionary dynamics of canalization.

2 Defining phenotypic robustness

Phenotypic robustness is about the sensitivity of the phenotype (e.g. some quantitative trait) with respect to changes in the underlying variables (genotype and environment) which determine its expression in an individual. Intuitively, this seems to be a clear notion. However, although the concept originated more than forty years ago there is a conspicuous lack of a formal definition. Such a definition should entail a clear criterion for the detection and characterization of robustness in empirical *and* theoretical study. For empirical work, robustness must be formalized as a operational concept. In order to be of conceptual use for an understanding of evolutionary processes, robustness should be classified into types that provide information about its evolutionary role and the circumstances of its origination. In this section, we will point out some key elements for such a formal definition and discuss demarcations to related concepts.

Our starting point is the formalization of robustness as a state of reduced impact of a given source of variation (such as mutations or environmental change) on the trait, i.e. as reduced variability due to that source (cf Wagner and Altenberg, 1996; Wagner *et al.*, 1997). Let us state this first in an informal way: *A character state that has evolved under natural selection is phenotypically robust if the variability of the character under a given source of variation is significantly reduced in this state as compared to a*

set of alternative states. Classification of different types of robustness now goes along with the formal clarification of three key notions of this definition, namely the character state, the source of variation, and the reference set of alternative states.

Genotypic and population states Since phenotypic robustness describes how variation in genotypic or environmental values is translated into variation on the level of the phenotype, a formal notion of the mapping between these two levels is needed. Each individual phenotype in a population, through development, is determined by a number of heritable and non-heritable factors, which we will refer to as its genotype and the environment experienced. Among the environmental factors, two types need to be distinguished in the following, commonly refered to as macro- and micro-environmental factors (cf Rutherford, 2000). Macro-environmental factors, such as temperature and light intensity, describe the identifiable environment external to the organism. They are, at least in principle, experimentally controllable. Micro-enviromental (or developmental) factors, in contrast, describe the uncontrollable developmental noise that is, to a large extent, internal to the organism (Gärtner, 1990). Imagine now the space G of all possible genotypes $\boldsymbol{g}$ in a population, and sets E and I of all experimentally controllable environmental and uncontrollable developmental conditions, $\boldsymbol{e}$ and $\boldsymbol{i}$. Each individual is then represented by the micro-state vector $\boldsymbol{y} := (\boldsymbol{g}, \boldsymbol{e}, \boldsymbol{i})$. Following e.g. Rice (1998), we may view a character ϕ as evolving on a phenotype landscape, defining phenotype as a function of the underlying variables, $\phi = \phi(\boldsymbol{g}, \boldsymbol{e}, \boldsymbol{i})$. For a given trait ϕ we now introduce the *genotypic character state* (or genotypic state) $(\boldsymbol{g}, \boldsymbol{e}, \rho_I)_\phi$ which describes the trait values of a genetically homogeneous population under controlled environmental conditions $\boldsymbol{e}$. The uncontrollable developmental factors are represented by a distribution ρ_I over I. For a natural (outbred) population with genotype distribution ρ_G, we define the *population (character) state*, $(\rho_G, \boldsymbol{e}, \rho_I)$, which describes the phenotype statistics of a population on the landscape in an averaged micro-environment.

Sources of variation and measures of variability Phenotypes can be buffered against variation from very different sources and reduced variability with respect to one source may or may not correlate with robustness under a different mode of variation. In order to explain what is actually buffered in a robust state, we will therefore need to specify the source of variation. The distinction with the most important evolutionary consequences is certainly the one between genetic and environmental types of robustness (Wagner *et al.*, 1997), depending on whether variation in heritable or non-heritable variables is buffered. Both types can assume a variety of forms. While genetic sources include different types of mutation and recombination, environmental sources may affect various macro-environmental factors. An important (and unavoidable) environmental source is further given by the uncertenties of development, namely developmental noise. Formally, any source S can be described by a mapping which assigns every micro-state $\boldsymbol{y}$ a distribution of variant micro-states $\mu^{(S)}_{\boldsymbol{y}}$. Usually, we can assume that genetic and macro-environmental sources and developmental noise are all independent of each other, and hence $\mu^{(S)}_{\boldsymbol{y}} = \mu^{(S)}_{\boldsymbol{g}} \mu^{(S)}_{\boldsymbol{e}} \mu^{(S)}_{\boldsymbol{i}}$. This does not imply that the *phenotypic* variation effects are independent. Since developmental variations are not heritable, the new variation per generation due to developmental noise is equal to the distribution of developmental factors in the population, $\mu^{(S)}_{\boldsymbol{i}} \equiv \rho_I$. It is important to note that this is in general not the case for new genetic variations, $\mu^{(S)}_{\boldsymbol{g}} \neq \rho_G$.

In order to quantify variability due to a source S, consider a trait ϕ equipped with some scale for the measurement of variation effects. (The choice of the appropriate scale is an important, but subtle issue, cf e.g. Lynch and Walsh, 1998, Chap. 11.) As long as the mean effect of the variations on the trait can be neglected, the variance of the variation effects is a valid measure of variability, i.e.

$$v^{(S)}_\phi(\boldsymbol{g}, \boldsymbol{e}) = \int d\boldsymbol{g}' d\boldsymbol{e}' d\boldsymbol{i} \Big(\phi(\boldsymbol{g}', \boldsymbol{e}', \boldsymbol{i}) - \bar{\phi}\Big)^2 \mu^{(S)}_{\boldsymbol{g}}(\boldsymbol{g}') \mu^{(S)}_{\boldsymbol{e}}(\boldsymbol{e}') \rho_I(\boldsymbol{i})$$

for the variability of a genotype state and $v^{(S)}_\phi(\rho_G, \boldsymbol{e}) = \int d\boldsymbol{g}\, v^{(S)}_\phi(\boldsymbol{g}, \boldsymbol{e}) \rho_G(\boldsymbol{g})$ on the population level.

If S is purely genetic or purely environmental, this is just the mutational or environmental variance, respectively. The mutational variance is an insufficient measure of variability if mutational effects are strongly asymmetric or even unidirectional. In this case, the mean square effect or the (absolute) mean are more appropriate.

Reference genotypes Defined as a state of reduced variability, robustness is a relative concept. Characterizing a character state as robust always presupposes that we compare it with a set of alternative states which are, on average, more variable. For a general notion of robustness, many choices of this reference are possible. If we are interested in robustness as a physiological property, we might e.g. compare the character states of various traits and ask for the elements of the genetic architecture that lead to observed differences in variability. From an *evolutionary* point of view, however, the alternative states should reveal the reduction of variability as a property that has evolved under natural selection. Consequently, the reference must be chosen from the states that have been (or still are) segregating in the population, i.e. from the naturally occuring variants or the ancestral genotypes. Within this set, the specific choice of the reference depends on a precise reformulation of the evolutionary question: In the last paragraph, we have distinguished different types of robustness according to the source of variation. In the following section, we will further refine our definition by asking whether, how and why robustness originated in evolution. Comparison with the properly chosen set of reference states then provides relevant statistical information for these questions. Note, that this choice of a reference does not depend on whether these states are actually available for variability measurements. For ancestral genotypes, this is usually not the case. One then must resort to indirect information, e.g. by comparing homologs across different species as a standard method to reconstruct the state of a common ancestor.

2.1 Adaptive or intrinsic?

It is a trivial truth that any cellular or organismal property is evolved in the sense that it is the outcome of the evolutionary process. Evolutionary biology, however, is concerned with the level of organization where it first arose, asking for the causes and circumstances of its origination. In the case of phenotypic buffering, this is made explicit in the distinction of *adaptive* and *intrinsic* forms of robustness (cf also Gibson and Wagner (2000); de Visser *et al.* (2003), who distinguish in a narrower sense "evolved" from intrinsic canalization).

Adaptive robustness We call robustness *adaptive* if the buffering of the trait with respect to some source of variation has been a *target of natural selection*, i.e. the robust character states have been selected because of their reduced variability. In order to come to an operational criterion for adaptive robustness, we test if the target state shows less variability than its mutational neigbors in genotype space that lead to the same phenotype. More precisely: we restrict our reference to mutational neighbors with a mean associated phenotype of equal or even higher fitness. This restriction ensures that the target state has not been selected because of a correlated direct effect on fitness. Note that the definition does not exclude the possibility that robustness with respect to a given source of variations evolves as correlated response to selection for reduced effects of variations from a different source. This may have important evolutionary consequences in particular if buffering for genetic and environmental variations is coupled (Wagner *et al.*, 1997). Phenomena of this kind have been observed in computational models of RNA secondary structure and have been called plasto-genetic congruence (Ancel and Fontana, 2000). In order to account for congruence effects, we may further refine our definition by restricting the reference set to states with equal variabilities from other sources.

In the following, we will use the term *canalization* as synonymous with adaptive robustness. This is in line with the classic literature on this subject: According to Schmalhausen (1949) and Waddington (1957), (see also Scharloo, 1991), canalization is a property that evolves for its own sake. The natural force assumed to be responsible for its evolution is stabilizing selection acting directly on the character or on

some highly correlated pleiotropic trait (in which case buffering of the correlated trait is the evolutionary target).

Within adaptive robustness, two main types may be distinguished, marking the endpoints of a scale. The first, *mechanistic* type corresponds to the classic view of canalization. Here, the trait and its buffering mechanism are genetically independent. Since the selective advantage of buffering depends on the primary trait, the evolution of the buffering mechanism is secondary to the character adaptation itself. Empirically, one might think of a high-level feedback mechanism or special "canalizing genes" (such as, perhaps, certain heat-shock proteins), in theoretical modeling this corresponds to approaches where variability is regulated by independent modifier loci. The other endpoint of the scale is the *cooperative* scenario. In this case, the primary trait and the variability are still uncorrelated in genotype space (structurally independent *sensu* Stadler *et al.*, 2001), but are influenced by the same genes and mutations will in general have effects on both. A distinction between primary and buffering genes is no longer possible. Buffering may then evolve much more in parallel with character adaptation, although there may still be primarily adaptive and buffering phases. Theoretical approaches to cooperative canalization include the UMF model (Wagner *et al.*, 1997) and the phenotypic landscape models by Rice (1998).

Intrinsic robustness refers to cases where evolution leads to states of reduced variability although buffering is not directly selected for. Robustness then arises as a by-product of selection for some other, correlated property of the state. The variability of intrinsically robust states is not lower that at its mutational neighbors with the same phenotype. It is only reduced in comparison with some ancestral states that exhibit different phenotypes (which then form the reference set). Therefore, in order to detect intrinsic robustness, the correlation of variability to some other trait must be demonstrated and an ancestral value of this trait must be known.

Intrinsic robustness is sometimes rather a system-level property than an attribute of a particular state. Robustness and buffering may be understood as emergent phenomena of regulatory networks. Here, a scenario favoring intrinsic robustness could be brought about if complex tasks for adaptation require fine-tuned regulation of gene expression which in turn require sufficiently complex gene networks – with robustness arising as a by-product of complexity or connectivity (see A. Wagner (1996) who shows that higher connectivity leads to higher robustness in certain networks). In this case, we would assume buffering to evolve for traits which use genes with multiple pleiotropic functions and in landscapes where adaptation is restricted by trade-offs and physical or biochemical constraints. Taking a wild-type pattern of gene-expression as the evolutionary target a mechanism of this kind has recently been described in a model for the segment polarity network in *Drosophila* (von Dassow *et al.*, 2000). Here, the first model network found to produce the desired expression pattern at all also led to significantly increased robustness with respect to variations in the network parameters and the initial conditions as a non-selected by-product. Note, however, that the phenotype itself is not contained in this model. Instead, selection is assumed to act directly on gene expression. Interpretation of the result as an example of intrinsic *phenotypic* robustness, therefore, still has to await further study.

Another potential example of intrinsic robustness is the diminishing returns function between enzyme activity of a given enzyme and metabolic flux. This relationship was first proposed as an explanation for the dominance of wild type genotypes against loss of function mutations by Wright (1929). Dominance can be seen as an indication of phenotypic robustness. In this case the phenotypic character is the steady state flux of a metabolic pathway. If selection favors increased flux it will reach a genotype where most variants have small effect on the character as a result of saturation phenomena. Experiments in *E. coli* have shown that this is in fact the case (e.g. Dykhuizen and Hartl, 1980). The famous controversy between R. A. Fisher, who thought that dominance is an adaptive phenomenon, and Wright, who thought it is a passive consequence of enzyme biochemistry is essentially a controversy over whether dominance is a consequence of canalization (adaptive robustness) or intrinsic robustness of states with maximal flux. Wright's position later got strong support from metabolic control theory (Kacser and Burns, 1981) and Fisher's adaptive hypothesis is commonly dismissed today. A close examination, however, shows that

this conclusion may not be as straightforward (cf Bagheri-Chaichian, 2001).

Summarizing our formal considerations, we have introduced phenotypic robustness for genotypic and population states, describing genetically homogeneous and outbred populations on a phenotypic landscape. Phenotypic robustness, however, can not be defined as a property of the phenotypic landscape alone (as it appears in Rice, 1998), but depends on a specific source of variation and the appropriate set of reference genotypes. Types of robustness are collected in the following table:

Typology of phenotypic robustness

Classification with respect to ...

1. the source of variation
   - genetic (mutational, recombinational, ...)
   - environmental (macro-env., micro-env. = developmental, ...)
2. the evolutionary origin
   - adaptive (robustness as evolutionary target = canalization)
   - intrinsic (robustness as correlated by-product)

2.2 Demarcations

Studies of phenotypic robustness or certain aspects of it have been pursued by many authors from various research traditions, not only evolutionary biology, with the focus on different problems. This entails the possibility for a fruitful exchange of ideas, but also led to a confusing variety of terminologies for similar and overlapping concepts. In order to sharpen the above notion of phenotypic robustness, we briefly discuss its relation to three similar concepts.

Robustness and stability In our formal definition of sources of variation, we have not required variations to be especially large or unusual. As a consequence, we do not distinguish robustness and stability (i.e. reduced variability under large and small perturbations) as is sometimes done in other fields, such as engineering and control theory (Slotine and Li, 1991). We feel that there are good biological reasons for this choice: Clearly, the restriction to unusual change (such as meteorites or genome duplications) would severely narrow the possibility of adaptive robustness from the outset. But also for the understanding of intrinsic robustness, the interrelation of buffering on different levels of organization and with respect to very different sources should play a key role. Within this larger scope, it is, of course, an interesting question to ask which type of robustness could also lead to reduced variability under major changes as a correlated by-product. If this is the case, and if robustness is needed to guarantee function and survival, major change could further advance the spread of robustness by means of population or species level selection. Note, however, that the *primary* cause for the evolution of robustness must always be found by selection among individuals and under more "usual" circumstances.

Robustness of the trait, genic buffering, and cryptic variation Genic buffering, i.e. a reduced level of expressed variation for single genes, is a widely observed phenomenon (Hartman IV *et al.*, 2001). In many cases double or multiple mutants produce much stronger phenotypic changes than each of the single mutants. Buffering of single gene functions in this sense and phenotypic robustness are closely related concepts, and the former has often been taken as evidence of the latter (cf Wilkins, 1997). It is important, however, to distinguish phenotypic robustness from cases where the variability of the contribution of a single gene to the trait is reduced, but not of the trait as a whole. Genic buffering that is not connected with phenotypic robustness may naturally arise as gene networks grow and become increasingly more complex: 1. Share of control between an increasing number of genes may reduce the effects of mutations in every single gene without changing the total impact of natural mutation on the trait. 2. Epistasis

and genotype × environment interactions make the effects of mutations in single genes dependent on the genetic background and the environmental conditions. If interactions are strongly variable without a net directional effect, as found e.g. for transposon insertions in *E. coli* (Elena and Lenski, 1997), the increase of mutational effects at some loci after the background change is compensated by a decrease at other loci. Evidence of single genes that are buffered under wild conditions therefore does not necessarily imply reduced variability of the trait as a whole. Under this scenario, large amounts of hidden ("cryptic") genetic variation can build up in mutation-selection equilibrium which can be expressed after a change in the environmental conditions or the genetic background. Clearly, this means that the release of cryptic variation can not be regarded a sufficient criterion for robustness or canalization (see below).

Robustness, neutrality, and the mutation rate Following our definition of genetic robustness as a state of reduced variability with respect to mutations, the evolution of complete invariability of the trait under some of these mutations is clearly a special form of robustness. Increased robustness, therefore, can manifest itself in an increasing fraction of neutral vs. non-neutral mutations on the molecular level or in a decrease of the mutational target size (i.e. number of genes that affect a trait) or the trait-specific deleterious mutation rate U on the level of gene-loci. This must also be taken into account if evolving systems with variable genome size are considered: Since decreasing the genome size may reduce the mutational target for a given trait, this may lead to an reduced overall impact of this source of genetic variation on the trait and therefore to increased robustness. This effect may, of course, be overcompensated if mutations in the remaining genome have more severe effects than before, e.g. due to a reduction of genetic redundancy.

Phenotypic robustness, however, is still about variation effects, not about rates of variation on the level of the underlying variables. We therefore do not consider decrease in the molecular mutation rate μ as a mode of genetic robustness or canalization. Genetic canalization and the evolution of mutation rates are "sister problems" dealing with indirect selection on the genotypic level that does not increase the maximum fitness in the population (cf Sniegowski *et al.*, 2000, for a recent review on mutation rate evolution). In both cases, mutational variation has a double function: it produces the genetic variation needed for any kind of selection, but also controls the selective advantage of mutants with lower mutation rate or higher robustness. Nevertheless, for phenomenological as well as conceptual reasons, both problems should be clearly distinguished. From a phenomenological point of view, entirely different mechanisms are responsible for these two processes. Mechanisms that regulate the copying fidelity and mutational repair act directly on the DNA level and may affect the genome as a whole. They have little in common with buffering mechanisms like genetic redundancy or feedback regulation which act on an intermediate level between genotype and phenotype and often are trait-specific. But also the evolutionary forces that favor the respective adaptations are different: Selection for modifications of the molecular mutation rate μ entirely depends on linkage disequilibria between the modifier locus and directly fitness related loci (Sniegowski *et al.*, 2000). Buffering mechanisms, on the other hand, may also be selected in linkage equilibrium since they increase the fitness of sub-optimal mutants and therefore have a higher marginal fitness (cf Wagner *et al.*, 1997).

3 Empirical evidence

Although environmental and genetic robustness of phenotypes seem to be palpable phenomena in nature, genetic robustness, in particular, is exceedingly difficult to measure and to prove in empirical work. In this section we will briefly discuss different ways robustness could be detected experimentally, also pointing out some of the practical problems. In our above reasoning we have singeled out two requirements for the detection of robustness, namely the measurement of the variability from a clearly distinguished source of variations, and the determination of an adequate reference point for this value. Since direct measurements of these quantities are often rather time-consuming and technically difficult, most empirical studies rely

on indirect evidence from measurements of associated quantities. In the second part of this section, we will present a short review of the experimental literature on phenotypic robustness and discuss the validity of these indirect estimates in the light of our formalization above. We find in particular that there is no convincing proof for adaptive genetic robustness.

3.1 Detecting robustness

Obviously the most direct approach is to measure the variability in wild types and mutants. Variability measurements are relatively straightforward for environmental sources. Here, superimposed variation from other sources (such as mutations) can be controlled or even ruled out. For the measurement of developmental noise this is most elegantly done by considering fluctuating asymmetry. The visible phenotypic variation in a population then provides a direct measure of variability. Measurement of genetic variability (i.e. the mutational variance), on the other hand, is connected with major technical difficulties (cf Gibson and Wagner (2000) for a discussion of some of the problems). There is, as yet, no data available that allows for the comparison of mutational variances in different genetic backgrounds. Conclusions in empirical studies therefore rely on indirect evidence (see below). A second practical problem for the detection of robustness lies in the difficulty of experimentally accessing the set of reference genotypes. Ideally, the decrease of variability would be observed during the evolution of robustness under natural or artificial selection. There are several such attempts in the literature. Convincing results, however, only exist for environmental robustness. In most cases, also the reference point for the variability has to be determined indirectly.

An alternative route for the detection of robustness is to identify particular buffering mechanisms first. If these mechanisms have no advantageous phenotypic effect themselves, but rather are physiologically costly, this would provide good evidence for evolved robustness. Intrinsic robustness can be demonstrated if a function of the mechanism is found that is directly related to fitness and *necessarily* leads to buffering as a by-product. Of course, only mechanistic forms of robustness (as opposed to cooperative forms, see above) can be detected that way.

3.2 Discussion of experimental results

Evolution of environmental robustness The classic example of natural selection for robustness against developmental noise comes from measurements of fluctuating asymmetry (FA) in Australian bowflies (Clark and McKenzie, 1987). Here, a mutation leading to insecticide resistance had deleterious pleiotropic effects on development when it first occurred, increasing FA. However, in subsequent generations modifiers reducing the pleiotropic effect were selected for, leading to an increase of developmental robustness. This clearly shows that environmental robustness can (and does) evolve in the wild, either for its own sake or as a correlated response to selection.

Heat-shock proteins Expression of Heat-shock proteins (Hsps) is an important mechanism to guarantee protein function under environmental stress. By impairing the function of Hsp90 by a heterozygous mutation or an inhibitor in *Drosophila*, Rutherford and Lindquist (1998) found that large amounts of previously silent polygenic variation with effects on many different traits were expressed. Hsp expression therefore seems to entail a multifunctional buffering mechanism with a clearly defined molecular basis. Conclusions about the evolutionary origin and maintenance of Hsp function remain nevertheless difficult. Many Hsps have multiple, often vital functions in the cell even under "ideal", unperturbed conditions which could be primary targets for selection. Certain levels and patterns of Hsp expression, however, seem to have evolved in response to the phenotypic effects of environmental variations, one example being Hsp70 expression in *Drosophila* (cf Feder and Hofmann, 1999).

Release of cryptic variation Following the original work of Waddington (1953), the increase of the phenotypic variation after a major mutation or exposure to an environmental challenge (e.g. heat shock) during development is the primary observation from which genetic canalization has been inferred. The classical examples are vibrissae number in mice (Dunn and Fraser, 1958) and scutellar bristles (Rendel, 1959) and cross-vein formation (Waddington, 1953) in *Drosophila*, all reviewed in Scharloo (1991). In all these experiments a character with almost vanishing variance in the wild-type showed significant variation after mutational or developmental perturbation. Much of the released variation was shown to be *genetic* (since it responds to artificial selection) and based on unexpressed (cryptic) variation already present in the base population (since inbred lines showed no selection response after a similar treatment). The increase in variance is then interpreted as reduced variability of the wild-type, hence canalization.

The model that has commonly been used for the quantitative analysis of these experiments has been developed by Rendel (cf Rendel, 1967). It assumes that the observed phenotypic value z is the function of some (unobserved) underlying character (sometimes called *liability*), $z = \phi(y)$. The *developmental map* ϕ rescales the underlying liability into the phenotype and determines the degree of canalization: flat regions of the developmental map correspond to so-called *zones of canalization*. Assuming that the liability is one-dimensional and normally distributed in the base population, the developmental map can be reconstructed. A measure for canalization is then given by the relative slope (wild-type/mutant) of the developmental map (cf Fig. 1).

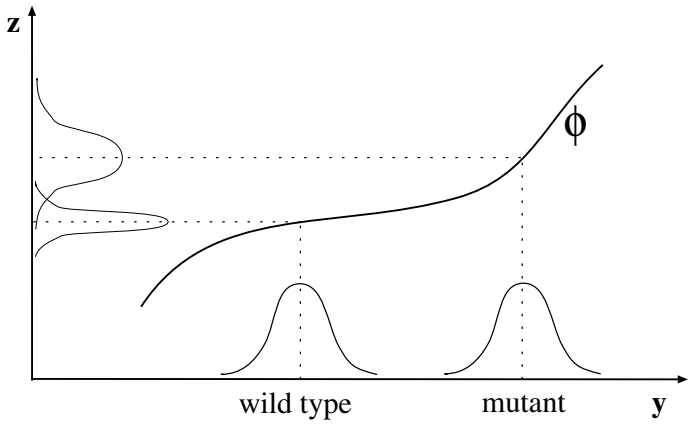

Figure 1: Rendel model. The distribution of the liability y in the wild population is shifted due to a background mutation (or environmental change). Since the developmental map ϕ is nonlinear, this results in a change of the genetic variance of the phenotype z. If the variance in the liability has a larger phenotypic effect in the mutant, cryptic variation is released and canalization of the wild type (relative to the mutant) can be concluded.

One may ask how these interpretations relate to the formal criteria for phenotypic robustness stated above. As far as the reference point is concerned, the mutant genotypes considered in the experiments are neither ancestral states, nor mutational neighbors with the same phenotype (the mean phenotype is changed in all cases). Nevertheless, since many perturbations of the wild conditions (with changes of the wild phenotype in both directions) lead to coherent results, it does not seem too far fetched to regard the mutant variabilities as typical for large regions of the genotype space. Problems, however, arise from the variability measurement itself: Because of the difficulties involved in determining *mutational* variances directly, the *genetic* variance in the population under wild and perturbed conditions is taken as a measure of genetic variability in all these experiments. As it turns out, however, this indirect measure is biased. This bias comes about because stabilizing selection during many generations has formed the genotype distribution in the wild population, reducing the genetic variance. After the experimental perturbation,

the population is not in mutation-selection equilibrium, but still carries the genetic variation shaped by selection in the old background. Since this shape is not adapted to the new background, it will in general be stronger expressed here, even if the mutational variance (which measures variability due to *new* mutations) may not be increased. Theoretical estimates of this effect in population genetic models show that this bias can easily account for all observed effects, if there is sufficiently strong variation in the epistatic interactions among loci (Hermisson and Wagner, in prep.). A detailed reading of the original experimental literature shows that variable epistatic interactions are indeed present in all cases where canalization has been suspected.

This problem is not apparent in the Rendel model where any release of variation translates into an increase in slope of the developmental map which again implies an increase in the mutational variance. This connection, however, is a result of the simple model assumption of epistasis acting on a one-dimensional underlying character. This assumption impies that either all mutational effects are increased, or all are decreased by the background shift. The argument breaks down if there is variance in the epistatic interactions. In this case, the underlying parameter space must be higher dimensional, with the slope of the developmental map increasing in some dimensions, but decreasing in others (cf Fig. 2). An increase in the expressed variation then also arises if the mutational variance remains unchanged.

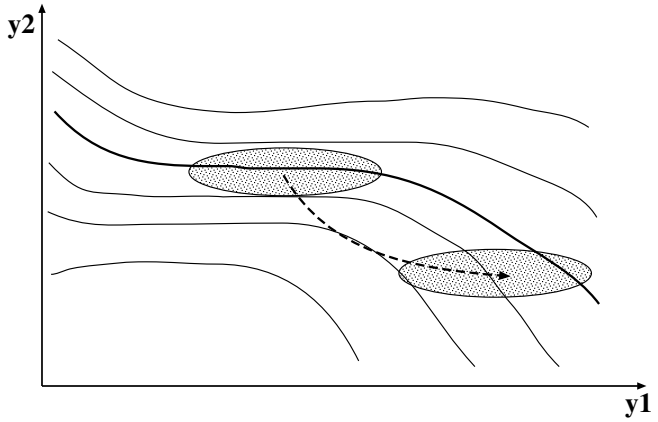

Figure 2: Release of cryptic variation and canalization for multidimensional liability. The figure shows a 2-dim liability map with iso-phenotype lines and wild and mutant liability distributions (shaded areas). The wild distribution is concentrated around the optimal phenotype (bold line). As in the one-dimensional case, the liability distribution is shifted by a background mutation. In the mutant background, the phenotypic variation is increased (distribution crosses more iso-phenotype lines). If the variance of new mutations in both liability directions is equal, however, the mutant background is gentically more robust (developmental map is less steep).

Artificial evolution Artifical evolution of genetic canalization has been reported in the above experimental settings by imposing mutant strains to stabilizing selection around a new phenotypic optimum (Rendel, 1967; Scharloo, 1988). The criterion for canalization, however, again relies on the assumptions of the Rendel model. Alternative explanations are possible. For example, the experiments can not exclude the possibility that selection created a leptocurtic distribution of the underlying variables. In a recent study of artifical *E. coli* evolution in which effects of new mutations in an ancestral and a derived background are directly measured, no evidence for the evolution of genetic robustness has been found (Elena and Lenski, 2001).

Introgression experiments An interesting extension of the above experimental set-up is to study the effects of introgression of mutations into a panel of inbred lines. With this technique, dramatic increases of the variances among mutant relative to wild-type lines have been found for *Drosophila* bristle number (Moreno, 1995) and photoreceptor differentiation (Polacysk *et al.*, 1998). Note, however, that interpretation of the results as robustness of the wild-type faces the same problems as discussed above since all wild-type lines have experienced the same selection pressure towards a common optimum. A notable exception is the study of Gibson and van Helden (1997) of *Drosophila* haltere characters in wild-type and *Ubx* mutant flies. Here, the variation (in shape and size) among the wild-type lines is rather due to drift or different optimal trait values in these lines which were taken from different populations. The combined distribution is therefore not constrained by selection towards a *common* optimal trait value and is thus not biased as compared with the variation among mutant strains. Interestingly, no increase in the phenotypic variance, and hence no canalization, is found in this study.

Comparing trait variabilities Stearns and co-workers have related the insertional variabilities to the fitness sensitivities of life-history traits in *Drosophila* by measuring the phenotypic effects of *P-element* insertions (Stearns and Kawecki, 1994; Stearns *et al.*, 1995). They find smaller variabilies in traits that are more strongly coupled to fitness and explain this result as evidence for canalization: traits under stronger selection are able to evolve higher levels of robustness (note that this expectation is not supported by theoretical work, see below). In contrast to the above cases, variability measurements are direct and unbiased here. The conclusion, however, has been challenged by Houle (1998), who suggests that the observed differences among traits are better explained by different numbers of genes affecting these traits. We have argued above that changes for a smaller mutational target size could well be a form of canalization. Nevertheless, the direct comparison of trait variabilities is biased if the mutational target sizes differ not only in the wild-type, but also on the set of reference genotypes. In particular, this will be the case if traits on a inherently different level of complexity are compared, such as, perhaps, adult size and fecundity. In general, the variability of one trait can not be used as a reference point for the variability of another trait without further argument.

RNA secondary structure Using program packages for RNA secondary structure prediction, A. Wagner and Stadler (1999) compared the stability of conserved and non-conserved elements in the secondary structure of RNA viruses with respect to point mutations. They find a trend towards a lower variability of conserved elements. Conserved elements probably have conserved functions and thus behave as 'traits' under stabilizing selection. Since the non-conserved elements, which are used as the reference set in this approach, do not show obvious differences in size or thermodynamic stability, adaptive genetic robustness is inferred. Open questions remain, however, with regard to the conditions under which robustness might have evolved: Since mutational stability seems to be correlated to stability of the structure with respect to sub-optimal folds (Wuchty *et al.*, 1999), either genetic or environmental variations may have been the driving force for selection. As for the level of organization, selection could have acted among structures with equal functional properties (favoring more robust ones) or among sequences that fold into a given structure. This last point could be clarified in a study comparing the mutational robustness of evolved 'character' states with sequences that fold into the same structure, as was done in Ancel and Fontana (2000).

4 Models for the evolution of robustness

In recent years a large number of papers have been published about models for the evolution of robustness. The models differ with respect to the kind of robustness considered and assumptions about the population genetic mechanisms involved. The consensus of the majority of papers is that, in principle, selection for robustness is possible by conventional natural selection, i.e. does not require exotic mechanisms like group

selection. There is, however, no consensus with regard to which of the proposed mechanisms is likely to be effective in natural populations (Gibson and Wagner, 2000; de Visser *et al.*, 2003). This conflict is somewhat hidden in the literature and is not explicitly discussed, in part because the literature is fragmented with studies pursued in disconnected fields ranging from physics to artificial life research. The purpose of the present section is to sketch the emerging consensus about the population genetic mechanisms for the evolution of robustness and to identify the remaining unresolved questions. Here, we will concentrate on models that explain the evolution of robustness for its own sake (hence on canalization according to our definition). This excludes intrinsic robustness, which needs to be explained by modeling the physiology of the trait rather than by population genetics. We will also omit congruence from our considerations. Models for plasto-genetic congruence will be discussed in the chapter by Fontana (this volume).

The standard model for the evolution of adaptive robustness is a quantitative trait under stabilizing selection (Waddington, 1957). The reason is that, under stabilizing selection and if the mean phenotype coincides with the optimal phenotypic state, any form of variation, genetic or environmental, decreases fitness. Any mechanism that buffers the trait by decreasing the phenotypic effects of the variations should thus be favored by natural selection. However, the dynamical properties of population genetic models for the evolution of environmental and genetic canalization are nevertheless quite different and will therefore be discussed in separate sections.

4.1 Models for environmental canalization

Since environmental canalization is the stability of the phenotype against environmental or developmental perturbations, it is the complement to phenotypic plasticity, i.e. the ability to realize different phenotypes with the same genotype (Schlichting and Pigliucci, 1998). In the literature on phenotypic plasticity, differnces between macro- and micro-environmental variations have been stressed (Scheiner, 1993). The need to distinguish these two forms of non-genetic variation comes from the observation that they are influenced by different genetic factors. As it turns out, the potential population genetic mechanisms for the evolution of macro- and micro-environmental canalization are also different.

Waddington (1957) proposed a model for the evolution of canalization against macro-environmental variation. He considered a scenario in which variation, across habitats or time, of an environmental factor changes the phenotype associated with a genotype but leaves the optimal phenotype unchanged. In this situation it would be beneficial to suppress phenotypic plasticity to this environmental factor and evolve a genotype that always realizes the optimal phenotype. This scenario is a special case of the evolution of phenotypic reaction norms. An extensive literature exists on the evolution of reaction norms (Schlichting and Pigliucci, 1998) that will not be reviewed here. Instead we will briefly discuss models about the evolution of micro-environmental (or developmental) canalization that are usually not considered in the literature on reaction norms. In quantitative genetic theory micro-environmental variation is usually thought of as the component of phenotypic variance that is not accounted for by genetic differences and interactions beween the genotype and the observed macro-environment. Assuming stabilizing selection where the fitness is a negative quadratic function of the phenotype $m(z) = m_{max} - kz^2$, the mean fitness in equilibrium is $\bar{m} = m_{max} - k(V_G + V_E)$. Here, the micro-environmental variance V_E may depend on the genotype, but not on the strength of selection, k. Any reduction in environmental variance increases mean fitness and will thus be favored by selection. With realistic estimates of V_E, the predicted selection coefficients in favor of environmental canalization are large enough for moderately strong selection, making this a plausible mechanism for the evolution of environmental canalization (Gavrilets and Hastings, 1994; Wagner *et al.*, 1997; Eshel and Matessi, 1998). Indeed, the main question is why V_E is still as large as it is for most characters. There has to be an as yet unidentified mechanism that limits the level of environmental canalization reached in natural populations. Several scenarios are discussed in Gibson and Wagner (2000).

4.2 Models for the evolution of genetic canalization

Models for the evolution of genetic canalization are more complicated and more ambiguous in their implications than those for environmental canalization. There are various selection scenarios that have been shown to favor genetic canalization, and there are different modeling approaches used with apparently different implications as to the mechanisms involved in the evolution of genetic canalization (e.g. compare Wagner *et al.* (1997) and van Nimwegen *et al.* (1999)). Most of the work has been done on models of mutation selection balance. Here, the vast majority of mutations that have a phenotypic effect on a trait are deleterious. Reduction of mutational effects should increase mean fitness and therefore be favored by selection independently of the relation of trait and fitness. In particular, also characters under directional selection should be subject to selection for canalization. But mutation selection balance is not the only scenario that may occur in nature and thus more work is needed to explore alternative situations. Several possible directions for further exploration are mentioned at the end of this section.

To our knowledge, the first study demonstrating the evolution of genetic robustness in a computational model is the paper by A. Wagner (1996). In this paper the phenotype is a gene activation state and evolution proceeds by changing the strength of regulatory interactions among the genes. Andreas Wagner demonstrated that networks acquire higher genetic robustness under stabilizing selection. He interpreted the results as an evolution of the population along neutral networks (fitness ridges) towards regions where the ridge is broader (see A. Wagner, 1996, : Fig. 11a). This interpretation of the evolution of genetic canalization was further formalized in the work of van Nimwegen *et al.* (1999) and anticipates the results of Bornholdt and Sneppen (2000). In a quantitative genetic context this is paralleled by the concept of genetic canalization as evolution towards flat regions on an iso-phenotype conture (Rice, 1998). Here, canalization may evolve whenever two or more underlying factors of a phenotype interact in a non-linear manner.

A population genetic theory of the selection for genetic canalization only exists for populations in or close to mutation selection balance (Wagner *et al.*, 1997). The analysis revealed a catch 22 for the selection for genetic canalization. The selection coefficient for a canalizing effect increases with the intensity of stabilizing selection, the degree of canalization caused by the allele and the amount of genetic variation affected by the canalizing effect, $s_c = kCV_G$. Since, however, the amount of genetic variation in mutation selection equilibrium is inversely related to the intensity of stabilizing selection the resulting selection coefficient can actually remain the same with increasing selection intensity. The reason is that stabilizing selection is eliminating genetic variation and thus eliminates the effects for which canalizing alleles are selected. The magnitude of the selection coefficient depends to a great degree on the mutation rate in these models. With per locus mutation rates of less than 10^{-4} the selection coefficients are very low. Only with high mutation rates can genetic canalization be effectively selected in mutation selection balance. The same conclusion is reached in similar models of genetic *redundancy* (Nowak *et al.*, 1997; Wagner, 1999). A model with fluctuating stabilizing selection also predicts low selection coefficients (Kawecki, 2000).

The model analysed in Wagner *et al.* (1997) assumes that the phenotype landscape allows for simultaneous buffering of all genes. This, however, is only possible at isolated points in genotype space which usually do not correspond to the optimal phenotype (Wagner and Mezey, 2000). A model that takes this into account was analyzed recently. Surprisingly, the equilibrium points in mutation-selection balance do not coincide with these maximally canalized points in many cases, but show rather high variability. Buffering, however, is found on the genic level for genes with high mutation rates (Hermisson *et al.*, 2003).

Most of the models cited above use the standard population genetic approach in which one considers the frequencies of genes influencing the trait of interest. A somewhat different model is discussed by van Nimwegen *et al.* (1999). In their model the elements are genotypes on a fitness landscape which change by mutation in the absence of recombination. As already observed by A. Wagner (1996), evolution of canalization can be described, in models with and without recombination, as the movement of the population along a neutral network towards regions of higher neutrality. Nimwegen *et al.* call their

model "neutral evolution" of robustness since movement on the ridge does not change fitness. However, the results presented in the paper show that genetic robustness is rather reached as a result of indirect natural selection against the effects of genetic variation. After all, the equilibrium frequencies clearly depend on selection and are distinct from a model without fitness differences among genotypes. Consistent with the population genetic models (Nowak *et al.*, 1997; Wagner *et al.*, 1997; Wagner, 1999), selection for robustness in the model is driven by mutation, with a selective advantage that independent of the strength of selection and proportional to the mutation rate. Consequently, the robust state breaks down if small fitness differences among states on the ridge occur.

Given these similarities of models with and without recombination, it is interesting to highlight some of the conceptual differences. In population genetic models with recombination, the unit of selection clearly is the gene. Selection for canalizing can be conceptualized as genic selection and is driven by the higher marginal fitness of alleles with specific (epistatic) gene effects (Wagner *et al.*, 1997). For clonal reproduction, the potential unit of selection is the genotype. If mutation can be neglected, selection among the genotypes only depends on differences in death and reproduction rates. Clearly, there is then no selective advantage among genotypes, canalized or not, on a neutral network. Including mutations in the model, the rate of growth of a genotype depends on the mutational loss in each generation, which again does not discriminate genotypes on the network, and on the gain due to *back mutations*. These depend in subtle ways on the fitness of the genotypes in a mutational neighborhood (Bürger, 2000; Hermisson *et al.*, 2002). If mutant genotypes are more fit, i.e. if the wild-type is more robust, the contribution of back mutations to the growth rate of the optimal genotype is higher than if the mutants are strongly selected against. Perhaps this is an example where we reach the limits of the unit of selection concept. What in fact is selected are population distributions not individual genotypes, i.e. quasi species sensu Eigen and Schuster (1979). Quasi species can increase their fitness by moving to a new, more robust quasi-equilibrium even if the optimal fitness in the population decreases (broader peaks can outcompete higher, but sharper peaks, cf Schuster and Swetina (1988) and Wilke *et al.* (2001)). The maximum gain in fitness, however, is small: as long as the trait specific mutation rate U_x is small against fitness differences, the selective advantage is of order U_x^2, and in general may not exceed U_x (cf Bürger, 2000; Hermisson *et al.*, 2002).

Mutation selection theory thus predicts that, in mutation-selection balance, genetic canalization is unlikely to evolve by selection directly for genetic canalization if the mutation rate is low. This can, in fact, easily be understood in terms of the genetic load. Since canalization does not increase the optimal fitness in the population, the selective advantage that any canalizing genetic effect may have is limited by the load component L_x of the character x under consideration. If the load is entirely due to deleterious mutations in mutation-selection balance, it is very well approximated by twice the trait-specific mutation rate, $L_x \approx 2U_x$ (resp. $L \approx U_x$ for haploids), largely independent of the fitness landscape and mutation schemes (cf Bürger, 2000). If mutation rates are not higher than thought, this can mean three things: 1. genetic canalization is rare or does not exist at all. In this case, genetic robustness could still be *intrinsic*. 2. genetic canalization evolves by congruence associated with environmental canalization (see Fontana, this volume), or 3. there are alternative mechanisms for the evolution of genetic canalization. According to the above, it is natural to look for alternative scenarios where the genetic load is higher. For stabilizing selection, and the optimum phenotype coinciding with the mean, the load is essentially proportional to the amount of genetic variation maintained in the population. There are several situations where more genetic variation is available. Already in 1960, Ernst Mayr proposed that canalization should be favored in spatially structured populations with gene flow (p 377 in Mayr, 1960). Indeed there is more genetic variation maintained in populations with clines in a polygenetic trait (Slatkin, 1978; Barton, 1999). Preliminary results show that selection for canalizing modifiers is predicted to be strong (Hermisson and Wagner, in prep.). The other alternative is to consider non-equilibrium situations in which genetic variation for a character under stabilizing selection can increase. This is for instance the case when genes have pleiotropic effects on a character that is under directional selection (Baatz and Wagner, 1997). Preliminary results show that selection for modifiers of pleiotropic effects can be quite strong (Wagner,

unpublished; Mezey, 2000).

5 Conclusions

1. The definition and the experimental demonstration of phenotypic robustness require the thorough determination of 1. the source of variation in order to define and measure variability and 2. a reference point with respect to which robustness of a character state is measured.

2. Environmental robustness is experimentally well established and theoretical models show that it should be easy to evolve.

3. Genetic robustness is difficult to demonstrate. We conclude from a review of the literature that there is no convincing proof for the existence of genetic canalization (adaptive genetic robustness).

4. Most of the published models for the evolution of genetic canalization assume mutation selection balance. These models show that genetic canalization is unlikely to be selected for in the wild.

5. Studies of alternative scenarios for the evolution of genetic canalization are severely lacking. Particularly promising directions are models with population structure, non-equilibrium models, and models including plasto-genetic congruence.

Acknowledgments

We thank Steve Stearns for valuable discussions and Homa Attar, Ashley Carter, Hans Larsson, Sean Rice, and Terri Williams for their helpful comments on the manuscript. J.H. gratefully acknowledges financial support by an Emmy Noether fellowship of the *Deutsche Forschungsgesellschaft* (DFG).

References

Ancel, L. W. and Fontana, W., 2000, Plasticity, evolvability and modularity in RNA, *J. Exp. Zool. (Mol Dev Evol)* **288**, 242–283.

Baatz, M. and Wagner, G. P., 1997, Adaptive inertia caused by hidden pleiotropic effects, *Theor. Pop. Biol.* **51**, 49–66.

Bagheri-Chaichian, H., 2001, Evolution of Mutational Effects in Metabolic Physiology, PhD Thesis, Department of Ecology and Evolutionary Biology, Yale University, New Haven.

Barton, N. H., 1999, Clines in polygenic traits, *Genet. Res. Camb.* **74**, 223–236.

Bornholdt, S. and Sneppen, K., 2000, Error thresholds on correlated fitness landscapes, *Proc. Roy. Soc. London* **B 267**, 2281–2286.

Bürger, R., 2000, The Mathematical Theory of Selection, Recombination, and Mutation, Wiley, Chichester.

Clark, G. M. and McKenzie, J. A., 1987, Developmental stability of insecticide resistance in the bowfly: a result of canalizing natural selection, *Nature* **325**, 345–346.

de Visser, J. A. G. M., Hermisson, J., Wagner, G. P., Ancel Meyers, L., Bagheri-Chaichian, H., Blanchard, J. L., Chao, L., Cheverud, J. M., Elena, S. F., Fontana, W., Gibson, G., Hansen, T. F., Krakauer, D., Lewontin, R. C., Ofria, C., Rice, S. H., von Dassow, G., Wagner, A., and Whitlock, M. C., 2003, Perspective: Evolution and detection of genetic robustness, *Evolution* **57**, ???–???

Dunn, R. B. and Fraser, A. S., 1958, Selection for an invariant character – 'vibrissae number' – in the house mouse, *Nature* **181**, 1018–1019.

Dykhuizen, D. and Hartl, D. L., 1980, Selective neutrality of 6pgd allozymes in *E. coli* and the effects of genetic background, *Genetics* **96**, 801–817.

Eigen, M. and Schuster, P., 1979, The Hypercycle, Springer, Berlin.

Elena, S. F. and Lenski, R. E., 1997, Test of synergistic interactions among deleterious mutations in bacteria, *Nature* **390**, 395–398.

Elena, S. F. and Lenski, R. E., 2001, Epistasis between new mutations and genetic background and a test of genetic canalization, *Evolution* **55**, 1746–1752.

Eshel, I. and Matessi, C., 1998, Canalization, genetic assimilation and preadaptation: a quantitative genetic model, *Genetics* **149**, 2119–2133.

Feder, M. E. and Hofmann, G. E., 1999, Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology, *Annu. Rev. Physiol.* **61**, 243–282.

Gärtner, K., 1990, A third component causing random variability beside environment and genotype. A reason for the limited succes of a 30 year long effort to standardize laboratory animals?, *Laboratory Animals* **24**, 71–77.

Gavrilets, S. and Hastings, A., 1994, A quantitative-genetic model for selection on developmental noise, *Evolution* **48**, 1478–1486.

Gerhart, J. and Kirschner, M., 1997, Cells, Embryos, and Evolution, Blackwell Scientific, Oxford.

Gibson, G. and van Helden, S., 1997, Is the function of the *Drosophila* homeotic gene *Ultrabithorax* canalized?, *Genetics* **147**, 1155–1168.

Gibson, G. and Wagner, G. P., 2000, Canalization in evolutionary genetics: a stabilizing theory?, *BioEssays* **22**, 372–380.

Hartman IV, J. L., Garvik, B., and Hartwell, L., 2001, Principles for the buffering of genetic variation, *Science* **291**, 1001–1004.

Hermisson, J., Redner, O., Wagner, H., and Baake, E., 2002, Mutation-selection balance: Ancestry, load, and maximum principle, *Theor. Pop. Biol.* **62**, 9–46.

Hermisson, J., Hansen, T. F., and Wagner, G. P., 2003, Epistasis in polygenic traits and the evolution of genetic architecture under stabilizing selection, *The American Naturalist* **161**, 708–734.

Houle, D., 1998, How should we explain variation in the genetic variance of traits?, *Genetica* **102/103**, 241–253.

Kacser, H. and Burns, J. A., 1981, The molecular basis of dominance, *Genetics* **97**, 6639–6666.

Kawecki, T. J., 2000, The evolution of genetic canalization under fluctuating selection, *Evolution* **54**, 1–12.

Lynch, M. and Walsh, J. B., 1998, Genetics and Analysis of Quantitative Traits, Sinauer, Sunderland.

Mayr, E., 1960, The emergence of evolutionary novelties, in Evolution after Darwin (S. Tax, ed.), Harvard University Press, Cambridge, MA, pp. 349–380.

Mezey, J., 2000, Pattern and evolution of pleiotropic effects: analysis of QTL data and an epistatic model, PhD Thesis, Department of Ecology and Evolutionary Biology, Yale University, New Haven.

Moreno, G., 1995, Genetic archtecture, genetic behavior, and character evolution, *Annu. Rev. Ecol. Syst.* **25**, 31–44.

Nowak, M. A., Boerlijst, M. C., Crooke, J., and Maynard Smith, J., 1997, Evolution of genetic redundancy, *Nature* **388**, 167–171.

Polacysk, P. J., Gasperini, R., and Gibson, G., 1998, Naturally occurring genetic variation affects *Drosophila* photoreceptor determination, *Dev. Genes Evol.* **207**, 462–470.

Rendel, J. M., 1959, Canalization of the scute phenotype of *Drosophila*, *Evolution* **13**, 425–439.

Rendel, J. M., 1967, Canalization and gene control, Logos Press, New York.

Rice, S. H., 1998, The evolution of canalization and the breaking of von Baer's laws: Modeling the evolution of development with epistasis, *Evolution* **52**, 647–656.

Rutherford, S. L., 2000, From genotype to phenotype: buffering mechanisms and the storage of genetic information, *BioEssays* **22**, 1095–1105.

Rutherford, S. L. and Lindquist, S., 1998, Hsp90 as a capacitor for morphological evolution, *Nature* **396**, 336–342.

Scharloo, W., 1988, Selection on morphological patterns, in Population Genetics and Evolution (G. de Jong, ed.), Springer, Berlin, pp. 230–250.

Scharloo, W., 1991, Canalization: Genetic and developmental aspects, *Annu. Rev. Ecol. Syst.* **22**, 65–93.

Scheiner, S. M., 1993, Genetics and the evolution of phenotypic plasticity, *Annu. Rev. Ecol. Syst.* **24**, 35–68.

Schlichting, C. and Pigliucci, M., 1998, Phenotypic Evolution: a reaction norm perspective, Sinauer, Sunderland.

Schmalhausen, I. I., 1949, Factors of evolution: the theory of stabilizing selection, Chicago University Press, Chicago, (reprinted edition 1986).

Schuster, P. and Swetina, J., 1988, Stationary mutant distributions and evolutionary optimization, *Bull. Math. Biol.* **50**, 635–660.

Slatkin, M., 1978, Spatial patterns in the distribution of polygenic characters, *J. theor. Biol.* **70**, 213–228.

Slotine, J.-J. and Li, W., 1991, Applied Nonlinear Control, Prentice Hall, Upper Saddle River, NJ.

Sniegowski, P., Gerrish, P., Johnson, T., and Shaver, A., 2000, The evolution of mutation rates: separating causes from consequences, *BioEssays* **22**, 1057–1066.

Stadler, B. M. R., Stadler, P., Wagner, G. P., and Fontana, W., 2001, The topology of the possible: Formal spaces underlying patterns of evolutionary change, *J. Theor. Biol.* (in press).

Stearns, S. C., 1994, The evolutionary links between fixed and variable traits, *Acta Pal. Pol.* **38**, 215–232.

Stearns, S. C. and Kawecki, T. J., 1994, Fitness sensitivity and the canalization of life-history traits, *Evolution* **48**, 1438–1450.

Stearns, S. C., Kaiser, M., and Kawecki, T. J., 1995, The differential genetic and environmental canalization of fitness components in *Drosophila melanogaster*, *J. Evol. Biol.* **8**, 539–557.

van Nimwegen, E., Crutchfield, J. P., and Huynen, M., 1999, Neutral evolution of mutatiuonal robustness, *Proc. Natl. Acad. Sci. USA* **96**, 9716–9720.

von Dassow, G., Meir, E., Munro, E. M., and Odell, G. M., 2000, The segment polarity network is a robust developmental module, *Nature* **406**, 188–192.

Waddington, C. H., 1953, The genetic assimilation of an aquired character, *Evolution* **7**, 118–126.

Waddington, C. H., 1957, The Strategy of the Genes, MacMillan, New York.

Wagner, A., 1996, Does evolutionary plasticity evolve?, *Evolution* **50**, 1008–1023.

Wagner, A., 1999, Redundant gene functions and natural selection, *J. Evol. Biol.* **12**, 1–16.

Wagner, A. and Stadler, P. F., 1999, Viral RNA and evolved mutational robustness, *J. Exp. Zool. (Mol Dev Evol)* **285**, 119–127.

Wagner, G. P. and Altenberg, L., 1996, Complex adaptations and the evolution of evolvability, *Evolution* **50**, 967–976.

Wagner, G. P. and Mezey, J., 2000, Modeling the evolution of genetic architecture: A continuum of alleles model with pairwise $A \times A$ epistasis, *J. theor. Biol.* **203**, 163–175.

Wagner, G. P., Booth, G., and Bagheri-Chaichian, H., 1997, A population genetic theory of canalization, *Evolution* **51**, 329–347.

Wilke, C. O., Wang, J. L., Ofria, C., Lenski, R. E., and Adami, C., 2001, Evolution of digital organisms at high mutation rates lead to survival of the flattest, *Nature* **412**, 331–333.

Wilkins, A. S., 1997, Canalization: a molecular genetic perspective, *BioEssays* **19**, 257–262.

Wright, S., 1929, The evolution of dominance, *Amer. Nat.* **63**, 556–561.

Wuchty, S., Fontana, W., Hofacker, I. L., and Schuster, P., 1999, Complete suboptimal folding of RNA and the stability of secondary structures, *Biopolymeres* **49**, 145–165.