CHAPTER 5 -- EVOLVABILITY, STABILIZING SELECTION, AND THE PROBLEM OF STASIS

BY THOMAS F. HANSEN & DAVID HOULE

"Organisms.. have not done nearly as much evolving as we should reasonably expect" Williams (1992, p 128).

"For a century we have been mesmerized by the successes of evolution. It is time now that we paid equal attention to its failures" Bradshaw(1991, p303).

Introduction

Evolutionary stasis has a paradoxical position in current evolutionary thinking. Williams (1992) argued convincingly that stasis is one of the most important unsolved problems of evolutionary biology (see also Bradshaw 1991). Stasis is arguably the predominant mode of evolution (Gould & Eldredge 1993); the cry 'stasis is data' (Gould & Eldredge 1977) echoes faintly in most neontologist's ears. Despite this, stasis is one of the most neglected theoretical problems in evolutionary biology. When population geneticists think of stasis at all, they usually regard it as an almost trivial consequence of stabilizing selection (e.g. Charlesworth et al. 1982; Maynard Smith 1983; Lande 1985, 1986). The fundamental problem of what may cause persistent stabilizing selection in changing environments is rarely addressed at length. It has been largely left to macroevolutionists to speculate on the microevolutionary underpinnings of stasis (e.g. Van Valen 1982; Wake et al. 1983; Williamson 1987; Hoffman 1989; Lieberman & Dudgeon 1996; Sheldon 1996; Eldredge1999; Gould 2002).

The paradox of stasis has it roots in quantitative genetics. It is the abundant variation in quantitative traits that makes stasis so difficult to explain on current thinking. Quantitative genetic experiments and simple population genetical theory seem to indicate that most characters should be very evolvable (i.e. have a high capacity to evolve), and there are many examples of rapid microevolutionary change that seem to confirm this ability (Hendry & Kinnison 1999, 2001). For example, the neutral theory for the evolution of quantitative genetic characters (Lynch & Hill 1986; Lynch 1993) shows that very rapid changes are expected on a geological time scale even in the

absence of any selection, simply due to drift and random fixation of new mutations. In fact, it is rare to find characters that evolve faster than the neutral expectation on macroevolutionary time scales, and very easy to find characters that are much too conservative for it (Lynch 1990). The need for phylogenetic comparative methods is another simple illustration of how conservative evolution is. Comparative studies need to deal with correlations stemming fromphylogenetic inertia.

This conflict between stasis and abundant genetic variation has an all-too-obvious solution in stabilizing selection. Alternatively, the abundance of genetic variation may be illusory, and some form of variational constraint limits the evolvability of characters, precluding them from tracking environmental changes. In this essay we discuss the logic of these two hypotheses. We argue that stabilizing selection is too readily accepted, while constraints are too readily dismissed. In particular, we suggest that conceiving of an organism as an integrated phenotypic and/or genotypic entity suggests two classes of variational explanation for stasis. Although there is often abundant genetic variation in quantitative characters, it is the quality, and not the quantity, of variation that is important for evolvability. The quality of variation may be reduced by either integration of characters (pleiotropic constraints) or by specific types of integration among genes (epistatic constraints). We suggest several ways to operationalize these notions.

Stasis in insect wings

Our thinking on this problem is best introduced with an example. Insect wings are conservative characters. Wing characters are usually reliable taxonomic indicators on the level of families, and qualitative differences are rare within genera, although coloration is sometimes a striking exception. Quantitative differences in shape occur, but are limited in extent. A good example is provided by variation in the genus *Drosophila*. Fig. 1 summarizes the size-adjusted positions of vein intersections in 22 species of *Drosophila*, plus one closely related genus in the family Drosophilidae (Galpern 2000). While the wings differ in length by an order of magnitude, the size-adjusted positions of the intersections are remarkably conservative. This conservatism is also readily detected in the behavior of *Drosophila* in flight. The slow, hovering “cargo-helicopter” flight mode is instantly recognizable to the experienced eye. Despite this conservatism, discriminant function analysis of wing shape shows that nearly all of the specimens in Fig. 1 can

be correctly identified to species, as shown in Table 1. This seeming paradox is resolved when we note that it is the result of low variation within species and not high variation among species.

The relative conservatism of wing shape in *Drosophila* is particularly remarkable because the genus is thought to be at least 50 million years old (Powell 1997). As expected in such an old group, there is substantial ecological diversity in the genus, with flies inhabiting both temperate and tropical habitats from rain forest to deserts. Their larvae feed on the microflora in a wide range of substrates, including decaying fruits, flowers, wood, leaves, fungi, or carrion. A few exceptional species mine leaves, feed on pollen or prey on other insects (Powell 1997; Kambysellis & Craddock 1997). As wings are often used in elaborate courtship displays, the diversity in mating systems, including lek breeding, solitary territoriality, assault mating and scrambles, is particularly noteworthy.

Even more remarkable is that wings throughout the Acalyptratae, the huge subordinal clade that includes *Drosophila*, are quite similar to those of the family Drosophilidae. The acalyptrates comprise over 22,000 species and 65 of the 113 families recognized in the Diptera (McAlpine 1989). There are a few key differences: the two costal breaks used as landmarks in Fig. 1 are missing in most other acalyptrate families, and *Drosophila* lack a basal crossvein that most other Acalyptrates have. Otherwise, the overall shape and placement of the common veins is usually conserved, despite the even greater range of ecological variation in the wider group.

Wing size and shape are highly heritable in *Drosophila* populations (e.g. Robertson & Reeve 1952; Robertson 1959; Cavicchi et al. 1981), making this a typical example of the paradox of stasis. Mutation produces novel phenotypic variation in wings at typical rates (Santiago et al. 1992; Houle, unpublished). Variation among populations has been demonstrated in both natural (e.g. Coyne & Beecham 1987; Gilchrist et al. 2000) and laboratory populations (Cavicchi et al. 1985), as has rapid divergence under natural selection (Huey et al. 2000; Gilchrist et al. 2001 for field studies and Cavicchi et al. 1985 for lab populations). Particularly compelling evidence for heritable variation in shape comes from the artificial selection experiments of Weber (1990; 1992; Weber et al. 1999), who obtained responses of up to 20 standard deviations in seven different arbitrary measures of wing shape in a relatively small number of generations.

Stabilizing selection as an explanation of stasis

Stabilizing selection has a somewhat odd position in evolutionary biology. On the one hand, there are few detailed empirical studies, and the direct evidence for stabilizing selection in the wild is far from overwhelming (Endler 1986; Travis 1989; Kingsolver et al. 2001). On the other hand, some of the most successful research programs in evolutionary biology are based on the assumption that traits are maintained at local optima by stabilizing selection (Mitchell & Valone 1990; Parker & Maynard Smith 1993). This is a fundamental assumption in behavioral ecology, life-history theory, and functional biology. These fields could not be as successful as they are if stabilizing selection is not the norm.

There appears to be fairly general acceptance of the idea that stabilizing selection is the cause of stasis. Even Gould, who may have been the most prominent champion of constraint hypotheses, recently acknowledged stabilizing selection based on niche tracking as a likely explanation of stasis (Gould 2002, pp. 880-885). From the theoretical point of view, stabilizing selection is a potent conservative force, but its relation to stasis is not simple. Explaining stasis with stabilizing selection requires not just that stabilizing selection is common, but that the selective optimum varies only within a narrow range. A stable optimum may seem plausible for a trait affected by a single selective factor, but the fitness functions of most quantitative traits are likely the result of a compromise among a large number of selective factors, any of which may be affected by changes in the environment. Dipteran wing shapes are certainly under stabilizing selection, but are the optimal shapes likely to be nearly the same in thousands of species of widely different size, living under widely different conditions with respect to temperature, humidity, and wind conditions? Why does not allocation to wing mass and muscle depend on the relative importance of flight to energetic constraints? Should not shape depend on this allocation? Should not the importance of wings for mate choice have substantial effects on their optimal shape? Should not males and females with differently shaped bodies have wings more different in shape? And if there really is one global optimum that fits all these conditions, why then would thousands of similarly-sized hymenopterans have such different wings?

Mammalian body temperatures (BTs) provide another good illustration of the problems involved in assuming a constant optimum (Williams 1992). Almost all placental mammals keep their operating BT between 37 and 38C, and equally puzzling, they keep their testicular temperatures 1C below that. If this is to be explained in terms of direct selection on BT we need to show that an arctic lemming and an African elephant have similar ecologically determined temperature optima. This seems next to impossible in view of the huge differences in ambient

temperature, heat exchange, metabolic needs and energetic constraints. An explanation based on internal selection where BT is "burdened" (Riedl 1978) by interactions with other traits is perhaps more plausible (see fuller discussion below), but we agree with Williams (1992) that mammalian BT seems inexplicably conservative under a stabilizing-selection hypothesis.

Thus, just evoking stabilizing selection to explain stasis is insufficient; we need to explain why selective optima themselves should be conservative. Surprisingly little work has been devoted to this problem. Only a few sketches of candidate hypotheses have been proposed, such as niche-tracking, population averaging, and ecological equilibration.

Niche tracking is perhaps the best-known mechanism proposed for maintaining stable selective environments (e.g. Eldredge 1999). To varying extents, all organisms are able to seek out favorable living conditions by behavioral means. In doing so, they will also stabilize many selective factors. More generally, Wake et al. (1983) suggested that any sort of plasticity (behavioral, physiological or developmental) in one set of traits will tend to allow stasis in the remaining set of characters. There are, however, problems with these propositions. Note that this explanation just shifts the problem from one set of traits to another, for example from morphology to behavior. Why should a given behavioral "habitat" preference remain optimal in a changing environment? Indeed the idea of the “Baldwin effect” is that plasticity, including behavioral plasticity, facilitates adaptive shifts (Baldwin 1896; Robinson and Dukas 1999), precisely the opposite of niche tracking. It is thus unclear if adaptive plastic responses will generally stabilize or disrupt selection pressures. Explicit models may help to clarify when plasticity helps or hinders further evolutionary change (e.g. Ancel 2000), but more work is needed in this area.

Lieberman & Dudgeon (1996) suggest that stasis is a result of averaging over many semi-independent populations that separately track fluctuating optima. This simply lifts the problem to a higher hierarchical level. Why should the separately fluctuating optima exactly cancel, and why would the many environmental conditions that undoubtedly affect the entire metapopulation remain constant?

Williams (1992) suggests that stasis could be explained by the existence of hyperstable niches. These are core sets of environmental conditions that are always present somewhere. Unstable "niches" come and go as the environment changes, and species that adapt to them will tend to go extinct with these shifts. What we observe in the fossil record are thus the forms that reside in these hyperstable niches, which also must be the common niches for the forms that reside in them to dominate the fossil record.

Sheldon (1996) suggests a similar lineage-selection mechanism. The species that survive environmental fluctuations are those that are least affected by them, typically generalists. But Sheldon's hypothesis does not explain why some species should be immune to the fluctuations. Similarly, Williams' hypothesis does not explain why we should expect niches to be hyperstable. How can we test the hypotheses that Dipteran wings and mammalian BT are conserved by stabilizing selection in a hyperstable niche? The ecological niche concept itself may not be sufficiently operational to allow a direct test, as this would require very precise descriptions of niches. At least these hypotheses predict that we should find numerous short-lived taxa that do deviate from the norm. We believe this can be rejected in the case of both Dipteran wings and mammalian BT, but Williams suggests that repeated fresh-water radiations of sticklebacks from a stable marine form (Bell 1989) may provide an example.

A theoretical mechanism that may favor niche conservation has been proposed by Holt & Gaines (1992; Holt 1996), and Kawecki (1995). The idea is that selection for adaptation to a core niche is stronger than selection for adaptation to any marginal niche or habitat, since individuals living under conditions to which they are well adapted have higher reproductive output than individuals living under marginal (sink) conditions. In other words, more individuals are affected by selection in the core habitat and selection in this habitat is therefore more important. This helps maintain adaptation to the core habitat and makes adaptation to alternative habitats more difficult. Although a significant theoretical observation, this does not solve the stasis problem. First, the argument still depends on the core niche conditions remaining constant, and second, the argument is only valid if the core habitat is more abundant than any marginal habitat the species may encounter. In fact, these models may as well predict that shifting habitat abundances should be a powerful driver of adaptive shifts.

Ecology is no doubt essential in understanding stasis. Stenseth & Maynard Smith (1984) developed a model of community dynamics that was able to predict stasis or gradual (Red Queen) evolution depending on the strength of ecological interactions. As with Sheldon's hypothesis, this is still no complete explanation of stasis within lineages, as it simply assumes that evolution will come to a halt in a stable environment.

A different type of explanation for stasis focuses on the complexity of forces that affect optima. Adaptation in one focal trait to one "primary" selective factor may be hindered by the need to stay adapted to a myriad of "secondary" factors including other traits that have been tuned to the previous state of the focal trait (Simpson 1944; Hansen 1997). Change is not

impossible in this scenario, but may be slow, as any new adaptation needs to be coordinated with a host of secondary adaptations. Consistent with this scenario, related species often deviate from an adaptive prediction in the same direction, leading to phylogenetic correlations. Although not stasis per se, this is direct evidence of evolutionary inertia. Hansen (1997) suggested that evolutionary inertia may be governed by the dynamics of the secondary factors, and developed a phylogenetic comparative method around this assumption. A skeptic may, however, argue that it is just as plausible that the secondary factors provide more opportunity for environmental changes to nudge the optimal state around. Theoretical work is necessary to determine whether, or under what conditions, a complex system of interrelated traits and selection forces will resist change.

In short, all proposed mechanisms for preserving optima are ultimately based on shifting the problem elsewhere, be it to other traits or to other levels of ecological organization. Thus, we join Arnold et al. (2001) in suggesting that work is urgently needed on the estimation and dynamics of adaptive landscapes. We simply do not have enough empirical evidence to conclude that landscapes are stable, nor do we have any compelling theoretical justification for assuming such stability. Given this situation, the alternative notion that stasis is due to constraints should be entertained as a valid possibility.

Constraints and evolvability

A constraint is any mechanism that may limit or bias the response to selection (see Wagner 1986; Arnold 1992; Houle, 2001 for review). We make a distinction between variational and selective constraints. Variational constraints are due to limitation and biases in the variability of characters. Developmental constraints are sometimes also defined in this way (e.g. Maynard Smith et al. 1985), but we prefer the more general term, as character variability need not be a consequence of development, as in the case of many cellular or biochemical traits. Selective "constraints" derive from conflicting selection pressures, and are constraints only from the perspective of achieving specific adaptations, and not from the perspective of optimizing the fitness of the organism as a whole. The shape of the fitness landscape itself as a selective constraint has also been widely discussed (e.g. Fear & Price 1988; Kaufmann 1993; Arnold et al. 2001), but we will not consider this further here. As we will see, the distinction between selective and variational constraints is sometimes a matter of perspective.

The most common objection to constraint as an explanation of stasis is the notion that quantitative characters are very evolvable because they show ample amounts of standing additive genetic variance and new mutational variation. In the following we will show that this variation is not sufficient to ensure evolvability. Levels of additive genetic and mutational variation have traditionally not been measured in a way that is operationally linked to evolvability. Furthermore, it is the quality and not the quantity of variation that is important. Finally, the evolution of variability itself must also be taken into account.

Genetic constraints and short-term evolvability. In an important contribution, Bradshaw (1991) reviewed cases were a failure to adapt in the face of unambiguous evidence for selection is plausibly due to a lack of appropriate genetic variation. For example, although the evolution of heavy-metal tolerance in plants is a celebrated textbook example of rapid evolution, Bradshaw pointed out that there are many plant populations that have failed to adapt in this way, and he showed that this is linked to an absence of variation for metal tolerance in the candidate populations.

Still, most traits exhibiting stasis appear to be genetically variable. This leads us to a short-term version of the paradox of stasis. There are now several well-documented examples of traits under directional selection in the field, such as clutch size in birds, that show no evolutionary response despite demonstrable heritability (Price & Liou 1989; Cooke et al. 1990; Frank & Slatkin 1992; Merila et al. 2001). A number of plausible explanations can be evoked, including soft selection, confounding effects of condition, poor estimates of selection or genetic variance, GxE interactions, environmental deterioration, and a failure to account for selection throughout the life cycle. Despite this we believe that it is worthwhile to take a closer look at the evolvability of quantitative traits, and ask whether such cases of evolutionary failure may also be caused by a lack of useful genetic variation.

We have argued (Houle 1992; Hansen et al. 2002a) that the use of heritability, h^2, as a measure of evolvability is misleading. One reason is that there is a strong correlation between additive genetic and phenotypic variation, which means that the heritability is poorly correlated with additive genetic variation (see Figure 2a for an example). A second reason is that heritability is not independent of its corresponding measure of selection strength, the selection differential. Under directional selection, the selection differential, S, is proportional to the phenotypic variance, which also enters in the denominator of the heritability. Thus, if h^2 is high, we may well expect a

proportionally smaller S for the same fitness function. This makes heritability highly suspect as an *a priori* predictor of evolvability.

Thus, h^2 and S, the components of the breeder's equation ($R = h^2S$), do not constitute proper measures of evolvability and selection, as is often implicitly assumed. This suggests that the theoretically preferred separation should be based on the selection gradient, □, and the additive genetic variance, G, as in the Lande (1976, 1979; Lande & Arnold 1983) equation ($R = G$□). The selection gradient is a descriptor of the adaptive landscape and thus of the causal basis of selection. In our view, evolvability should be seen as the ability to respond to an externally imposed selection regime as represented by a fitness function.

Following Houle (1992), Hansen et al. (2003a) showed that I_A, the additive genetic variance scaled with the square of the trait mean, is an operational measure of evolvability, as it can be interpreted as expected percent change per generation per unit strength of directional selection. This holds for traits on a ratio scale, and requires the use of mean-scaled selection gradients, or fitness elasticities, as measures of selection strength (see van Tienderen 2000). Mean-scaled selection gradients have a natural unit as they measure the strength of selection on fitness itself as one. Thus, I_A is interpretable as the expected proportional response if selection was to be causally as strong as on fitness itself.

The poor correlation between I_A and h^2 (e.g. Houle 1992) means that judgments of evolvability based on the latter are irrelevant. Houle (1992) found that the conclusion that life-history traits were less evolvable than morphological traits (e. g. Gustafsson 1986; Mousseau & Roff 1987) was turned on its head if evolvability was measured by coefficients of additive genetic variation ($CV_A = \sqrt{I_A}$) instead of heritability. Hansen et al. (2003a) found that the evolvabilities (I_A) of floral traits in a *Dalechampia scandens* population were typically less than half a percent (i.e., response to selection as strong as on fitness itself would be less than half a percent per generation), a significant constraint on adaptation to potentially rapidly changing pollination regimes. This limitation is not apparent from h^2 alone.

Genetic constraints and long-term evolvability. Regardless of how variation is measured, it might appear that the constraint hypothesis is falsified for traits such as Dipteran wing shape or mammalian body temperature, which are conserved on higher taxonomic levels, as there arguably are detectable levels of additive genetic variation in each case (see Lynch 1994 for BT and references above for wings). At sufficiently long time scales, any level of replenishable

variation could be turned into monumental changes. On longer time scales it is variability, the capacity of traits to vary (Wagner & Altenberg 1996), rather than the standing level of variation that dictates evolvability. The requirement that new genetic variation is produced by mutation is fulfilled for at least some aspects of wing variation in *Drosophila* (Santiago et al. 1992; Houle, unpublished). More generally, studies of a variety of traits in many organisms indicate that a fair amount of mutational variability is generated each generation (Lynch 1988; Houle et al. 1996; Lynch et al 1999).

As with additive genetic variation, mutational variation can be measured on a variance-standardized scale (i.e. mutational heritability) or on a mean-standardized scale (Houle et al. 1996; Houle 1998). It is instructive to consider the level of evolvability in wing morphology that can be maintained through mutation. A mutation-accumulation study by Santiago et al. (1992) showed that the mutational heritability of wing traits was a typical 0.001; however, estimates of mean-standardized mutational variance for wing morphology was ~ 2×10^{-6} (0.0002%), an unusually low value (cmp. Houle et al. 1996). This means that long-term directional selection of similar strength as selection on fitness would only be able to change the traits by 0.0002% per generation once standing variation from the base population is exhausted. This suggests fairly strong genetic constraints on wing shape. Once again, mutational heritabilities give a misleading picture of potential evolvability.

It is thus debatable to what time scales genetic constraints may extend. It seems likely that a lack of mutational and standing genetic variation may sometimes limit evolution on ecological time scales, but it is also reasonably clear that stasis on a true geological time scale, comprising millions of generations, cannot be explained in this way. Doubling of wing trait value at the low rate of evolution calculated above would only take 150,000 generations, or perhaps 15,000 years, for a *Drosophila melanogaster* population. If stasis in macroevolution is caused by constraints, these have to render even the low level of evolvability revealed by I_A evolutionarily irrelevant. The crucial question to which we now turn is whether genetic variation can be translated into adaptive changes without compromising other aspects of organismal function.

Variational constraints due to functional architecture

The variational properties of a character are determined by its genotype-phenotype map; its underlying functional architecture. The functional architecture is the collection of pathways that

lead from the genes to the character (Houle 1991, 2001). Thus, a deeper understanding of constraint must come from an understanding of the functional architecture. This includes the relationships among different characters, and the way in which genes are combined in the mapping from genotype to phenotype (termed functional epistasis by Hansen & Wagner 2001). Below, we discuss in more detail two forms of constraints due to functional architecture. Pleiotropic constraints stem from interactions among characters; epistatic constraints stem from interactions among genes. In each case we evaluate their potential as explanations of stasis, and suggest ways to operationalize the concepts to make them empirically accessible.

Pleiotropic constraints & conditional evolvability. All too many assessments of character evolvability are done in isolation. In such situations it is customary to point out that the character in question is, or is likely to be, genetically variable. One of the assumptions of this line of argument is that the character has a property that Lewontin (1978) called *quasi-independence,* meaning that it is possible to change the character without unduly disturbing other aspects of the organism.

Such independence may be compromised by pleiotropy, where the genes that create variation in a focal character also create variation in other characters. If the normal state of affairs is that most characters are under stabilizing selection most of the time, the evolvability of the focal character will often be compromised by pleiotropy. This possibility may also be viewed as a potential cause of apparent stabilizing selection, as pointed out above. There is an urgent need to assess how severe such pleiotropic constraints may be, both for standing and mutational variation. To do this we need to devise measures of evolvability that can be used to assess how much of the variation is effectively available to produce particular adaptations.

The concept of conditional evolvability is such an attempt (Hansen et al. 2003b; Hansen 2003). Conditional evolvability is defined as the evolvability of a character *y* in the event that a set of constraining characters *x* is not allowed to change. Under the usual assumptions of the Lande equation, conditional evolvability is determined by the conditional genetic variance (i.e. the residual variance when the genetic value of *y* is regressed on the genetic value of *x*). It can be further shown that the conditional variance is approximately valid as a predictor of evolvability regardless of the strength of stabilizing selection on *x* (Hansen 2003). When directional selection is placed on the focal character, the constraining characters are displaced from their optima by a

small amount, and after a few generations, the strength of stabilizing selection only affects the size of this displacement and not the degree of constraint on the focal character.

The conditional evolvability is always relative to the particular set of constraining characters that are included in *x*. It may for example be used to assess the constraining effect of particular characters that are believed to be under stabilizing selection. In Fig. 2b we show how the evolvabilities of the floral traits in the *Dalechampia scandens* population mentioned above could be further reduced by conditioning on two functionally important floral characters that may be under stabilizing selection (Hansen et al. 2003b; see also Armbruster et al. this volume).

Conditional evolvability may also be assessed on some measure of background fitness. This will efficiently capture the organism-level pleiotropic constraints generated by a large number of characters that are not directly studied. Weber (1996) provided an interesting example of this sort of reasoning when he assessed the fitness consequences of his artificial-selection experiment on wind-tunnel flight performance in *Drosophila*. Selection over 100 generations resulted in a response of over 30 standard deviations in performance (see Weber's Fig. 2), but competitive fitness dropped by only 6%, although the confidence limits on this estimate are wide. Weber concludes that performance in *Drosophila* is relatively unconstrained by pleiotropy. It certainly seems unlikely that such a moderate fitness reduction provides an insurmountable constraint. As for body temperature, it is unfortunate that no one has yet taken up Williams' (1992) excellent suggestion of selecting on BT in mice and see what happens.

Conditioning on background fitness may be particularly useful in assessing the adaptive potential of mutational variation. Mutational variation is likely to hold a much higher fraction of useless variation than standing genetic variation, as it is not yet filtered by selection (Houle et al. 1996; Houle 1998). It is possible that a large majority of mutational variation is due to mutations in housekeeping genes and general regulatory genes that are expressed in a multitude of tissues and circumstances. This may appear as variability in individual characters, but is hardly good material for building adaptations. Galis (1999) suggests an intriguing example, where apparent genetic variability in the number of mammalian neck vertebrae is rendered useless by pleiotropic effects that greatly elevate cancer risk (see also Galis & Metz 2001; Galis et al. 2001 for further examples). Stern (2000) made the interesting suggestion that adaptively useful mutations are largely limited to regulatory elements that have temporally and spatially restricted effects, which may be less pleiotropically constrained than other mutations. At present the quality of mutational

variation is basically unknown. Assessing the amount of conditional character evolvability provided by mutation is essential to evaluate the pleiotropic-constraint hypothesis.

The idea of conditional evolvability is related to a variety of proposals concerning “internal” selection pressures (see Wagner & Schwenck 2000 for review). The logic of these hypotheses is that the complexity of development links the variational properties of characters to each other. Some key characters may then become so entrenched in development that they completely lose their variational freedom. Riedl (1977, 1978) proposed entrenchment, or ‘developmental burden’ as an explanation for deep invariant homologies such as the vertebra in vertebrates.

Mammalian BT is a plausible example of this sort of entrenchment. Since the function of every protein is likely to be affected by a change in body temperature, evolution of a different BT might require adaptation throughout the genome. This may equally well be viewed as a selective constraint, as there is no conceptual difference between adaptation to the internal and the external environment. It is, however, not clear whether this mechanism can generate an absolute constraint on the position of the optimum. Mammalian tissues are tolerant to minor variations in temperature, and are able to continue operation within a few degrees of the normal BT. In other words, the physiological or ecological constraints are not absolute. This means that the optimum should be at least somewhat negotiable. A large change in, say, external temperature should produce at least a small shift in the BT optimum. After this, the physiology should be able to adapt to the new optimum, making possible another shift in the direction of ecological adaptation, and so on. In general, continuous variation combined with plasticity may make adaptive optima inherently responsive to external change. Again, formal models need be developed to assess the potential for entrenchment of quantitative characters due to interdependence with numerous secondary factors that may be either ecological (external) or physiological (internal).

The idea of entrenchment is based on the assumption that complex interdependence among traits is a constraint. This assumption has its theoretical basis in Fisher's (1930, pp. 38-41) geometrical model of adaptation. Fisher's model was an attempt to formalize the notion that a random change in a complex apparatus is less likely to improve function than a random change in a relatively simple one. Fisher showed that the probability of a mutation being advantageous was a decreasing function of the dimensionality of the trait. This result, however, needs qualification. If we take the perspective of the evolvability of an individual trait, pleiotropic links to other characters may act as a constraint, but also serve to increase the mutational target size and

therefore the evolvability of the trait. Hansen (2003) investigated this trade-off between selective constraints and evolvability in formal models and showed that an intermediate level of pleiotropy was likely to result in the most independently evolvable characters.

A neglected approach to the detection of pleiotropic constraints is the estimation of the dimensionality of phenotypic variation (reviewed in Houle 2001; Steppan et al. 2002). Multivariate variation may exist for only a few kinds of change in form, and be absent in all others. In such a case, evolution could only occur along these few variable directions. Quantitatively it may be far more likely for evolution to proceed along relatively variable aspects of the phenotype, the 'genetic lines of least resistance', than less variable ones (Schluter 1996). Indeed, several studies have found that a large amount of multivariate genetic variation within and among species is concentrated in only a few aspects of form (Kirkpatrick and Lofsvold 1992; Björklund 1996; Schluter 2000, Chapter 9). For wing shape in *Drosophila melanogaster*, however, preliminary data surprisingly suggest that genetic variance in wing shape is distributed across essentially all aspects of form (Houle, in prep). This perhaps renders pleiotropic constraint less plausible as an explanation of stasis in wing shape.

Epistatic constraints and the evolution of evolvability. In an influential 19th century critique of Darwin, Fleeming Jenkin argued that although selection may bring a species out towards the limit of its natural variation, it would get stuck, as there would be no more useful variation in the appropriate direction (see Gould 2002, p142). Jenkin thus imagined a species as a fixed sphere of variation around some immutable essence. While Jenkin's reasoning is a reflection of a flawed typological species concept, the hypothesis that the range of possible variation is limited is still plausible. Indeed such limits are inherent in any mutation model based on a finite number of alleles with fixed effects (e.g. Zeng et al. 1989). Empirically, Mackay et al. (1995) found that the divergence of *Drosophila* in bristle number between lines accumulating spontaneous mutation reached a plateau after less than two hundred generations. This notion of limits can be contrasted with its opposite extreme, the additive model of population genetics. In the additive model, the distribution of the effects of new mutations on the phenotype is independent of the genetic background. Thus, when selection moves a character towards more extreme values, the variational properties of the genome stay the same. This implies that new variation can thus be produced ad infinitum, and the sphere of variation moves along with the population mean.

In between these extremes of fixed and unlimited ranges of potential variation lies the probably more realistic idea that the range of variation is itself evolvable. The genotype-phenotype map is a very complex function, where the correct functioning of large numbers of genes is necessary to achieve a particular phenotype. This implies that functional epistatic interactions among genes must exist. By definition, epistasis means that the effects of alleles at a focal locus will change with changes in the rest of the genotype, the genetic background. Any sort of evolutionary change in the genetic background may affect the variation expressed by variants at a particular locus (Hansen & Wagner 2001). Epistasis thus makes evolvability evolvable. Note however, that this evolution could either enhance or diminish the potential for adaptation. Thus another potential explanation for stasis is that epistatic interactions tend to restrict variation under selection.

If genes interact in such a way as to diminish each other's effects as the population is selected in a particular direction, then evolvability will be diminished when natural selection pushes the population in that direction. We call this negative epistasis. In contrast, if genes interact positively by mutually reinforcing each other, evolvability will be enhanced. If negative epistasis dominates in all directions in phenotype space, then a form of Jenkin's sphere may be generated, as the variability will be diminished when the population moves to more extreme character values. We call this an epistatic constraint. Note that a population subject to an epistatic constraint might possess ample genetic variance and mutational variability in each trait, but the additive and mutational variance will be scaled down or become biased as selection changes the character mean.

Formal equations describing these effects in the case of multilinear epistasis are given in Hansen & Wagner (2001). Epistasis may generally be described by a set of epistasis factors, *f*, that describes how a change in the genetic background changes the effect of any particular gene substitution. If □ is the effect of the gene substitution in some reference genotype, then *f*□ is the effect in an alternative genetic background. The epistasis factor is thus a function of the genetic background. If $f < 1$ with respect to a particular perturbation, then epistasis is negative, while $f > 1$ corresponds to positive epistasis. The average values of the epistasis factors with respect to particular genetic perturbances are therefore the parameters we need to predict the evolution of evolvability.

The hypothesis of epistatic constraints is thus operational, and can be tested by estimation of epistasis factors in real genetic architectures. We are currently preparing to do this

for fly wing characters. Estimation of epistasis factors requires the creation of different genetic backgrounds followed by a comparison of the effects of known genetic elements in the different backgrounds. To test the epistatic-constraint hypothesis, backgrounds with extreme character values need be compared to backgrounds with less extreme values. These backgrounds may be created through artificial selection, for example.

The question that emerges from the epistatic-constraint hypothesis is whether we should expect genetic architectures to exhibit negative epistasis. We may distinguish two sub-questions. First, are there system-theoretic reasons why functional architectures should produce particular patterns of epistasis? Second, does evolution lead to functional architectures with particular patterns of epistasis? Despite the novelty of these questions, there are some relevant observations.

With regard to the first question, there are certainly aspects of functional architectures that lead to particular patterns of interaction among components (e.g. Kauffman 1993). For example, if there is widespread redundancy in the genotype-phenotype map, we may expect negative epistasis, as there are many subsystems that can produce adaptively relevant variation but when change is achieved through one of these, the variability of the others becomes adaptively irrelevant. The theoretical study of epistasis in model architectures, as exemplified by Szathmary's (1993) study of epistasis in metabolic pathways or Gibson's (1996) study of epistasis in gene regulation, may prove illuminating here.

The evolution of functional architectures has usually been addressed in terms of the evolution of canalization or robustness. The phenomenon of genetic assimilation, first demonstrated in fly wing morphology (Waddington 1953), shows that increased evolvability can evolve when directional selection is applied in a novel environment. Indeed, when we conclude that the wild type is genetically canalized, we are saying that the epistasis factors are positive with respect to perturbations of the wild-type genotype. The epistatic-constraint hypothesis thus runs into the difficulty of positing a (relatively) decanalized wild type, so that there is room to diminish the variability when selection perturbs the population. Evidence for canalization of the wild type has come from a series of empirical studies demonstrating that severe mutational or environmental disturbances can release "hidden" genetic variation (reviewed in Scharloo 1991; Moreno 1994; Gibson & Wagner 2000; Rutherford 2000). The changes in genetic background that produce such decanalization are usually extreme, however, and likely to be associated with severe deleterious side effects, making them less relevant to microevolutionary changes.

Similarly, the effects detected in such studies are themselves very large, leaving it unclear whether alleles of small effect are also unmasked by decanalizing perturbations.

Since Waddington it has been thought that stabilizing selection may lead to the evolution of canalization. This prediction is supported by some models (Wagner et al. 1997; Rice 1998), but the effects are weak. Hermisson et al. (2003) have further shown that these results may not hold with more realistic representations of multi-locus dynamics. Although there is selection for canalization at individual loci, the interactions among loci means that many loci will still become decanalized. Loci with the highest mutation rates are usually canalized, but the net phenotypic effect depends on how genes interact and stabilizing selection may well produce a somewhat decanalized phenotype with high mutational variability.

The epistatic-constraint hypothesis is thus at least tenable, but needs to be further evaluated with modeling and empirical studies of functional epistasis within the range of "normal" variation. It is particularly important to consider the effects of functional epistatic architecture on the evolution of pleiotropic patterns; i. e., on the evolution of pleiotropic constraints. It has been proposed that the functional architecture may evolve to resemble the adaptive landscape, either through correlated stabilizing selection (Olson & Miller 1958; Cheverud 1984, 2001), or through correlated shifting directional selection (Wagner 1996). The mechanisms behind these hypotheses need more study and this will necessarily involve the study of patterns of epistasis.

Conclusions

Character conservation remains an important challenge for evolutionary theorists. None of the many hypotheses we have considered provides a completely satisfactory explanation for stasis in our two examples: fly wing shape and mammalian BT. Stabilizing selection undoubtedly contributes to stasis, but alone is an insufficient explanation for it. Selective explanations of stasis will need to focus on the dynamics of adaptive landscapes (Simpson 1944; Arnold et al. 2001), and it will be necessary to further develop and test ideas for how adaptive optima themselves can be stabilized. Currently, variational constraints can neither be generally rejected nor generally accepted as an explanation of stasis. The constraint hypothesis has suffered from a certain vagueness, and perhaps from its association with the untenable notion that evolution only happens in association with speciation. The recent interest in evolvability (e.g. Raff 1996; Wagner & Altenberg 1996; Gerhart & Kirschner 1997) provides a new perspective on constraints that may

prove helpful. As outlined above, the theory of evolvability is in the process of being operationalized on several levels, and may eventually provide the tools for assessing the prevalence and power of constraint in both micro- and macroevolution.

Acknowledgements. We thank Ashley Carter, Massimo Pigliucci, and Katherine Preston for helpful comments on the manuscript.

Literature Cited

Ancel, L.W. 2000. Undermining the Baldwin expediting effect: Does phenotypic plasticity accelerate evolution? Theoretical Population Biology 58:307-319.

Arnold, S. J. 1992. Constraints on phenotypic evolution. Am. Nat. 140:S85-S107.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112/113: 9-32.

Baldwin, J.M. 1896. A new factor in evolution. Am. Nat. 30:354-451.

Bell, M. A. 1989. Stickleback fishes: Bridging the gap between population biology and paleobiology. TREE 3: 320-325.

Björklund, M. 1996. The importance of evolutionary constraints in ecological time scales. Evolutionary Ecology 10: 423-431.

Bradshaw, A.D. 1991. Genostasis and the limits to evolution. Phil. Trans. R. Soc. Lond. B 333: 289-305.

Cavicchi, S., G. Giorgi, and C. Pezzoli. 1981. Effects of environmental temperature on developmental patterns in *Drosophila melanogaster* (Meig). Boll. Zool. 48:227-231.

Cavicchi, S., D. Guerra, G. Giorgi, and C. Pezzoli. 1985. Temperature-related divergence in experimental populations of *Drosophila melangaster* I. Genetic and developmental basis of wing size and shape variation. Genetics 109:665-689.

Charlesworth, B., R. Lande, and M. Slatkin. 1982. A neodarwinian commentary on macroevolution. Evolution 36: 474-498

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol. 110: 155-171.

Cheverud, J. M. 2001. The genetic architecture of pleiotropic relations and differential epistasis. Pp 411-433 In Wagner, GP (ed), The character concept in evolutionary biology. Academic Press, San Diego.

Cooke, F., P. D. Taylor, C. M. Francis, and R. F. Rockwell. 1990. Directional selection and clutch size in birds. Am. Nat. 136:261-267.

Coyne, J. A., and E. Beecham. 1987. Heritability of two morphological characters within and among natural populations of *Drosophila melanogaster*. Genetics 117:727-737.

Eldredge, N. 1999. The pattern of evolution. Freeman, New York.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. *in* Schopf, ed. Models in paleobiology. Freeman.

Endler, J. A. 1986. Natural Selection in the Wild. Monographs in population biology 21.

Fear, K. K., and T. Price. 1998. The adaptive surface in ecology. Oikos 82: 440-448.

Fisher, R. A. 1930. The Genetical Theory of Natural Selection. Oxford University press. A complete variorum edition. 1999.

Frank, S. A., and M. Slatkin. 1992. Fisher's Fundamental Theorem of Natural Selection. TREE 7: 92-95.

Galis, F. 1999. Why do almost all mammals have seven cervical vertebrae? Developmental constraints, Hox genes, and cancer. J. Exper. Zool. (MDE) 285: 19-26.

Galis, F., and J. A. J. Metz. 2001. Testing the vulnerability of the phylotypic stage: on modularity and evolutionary conservation. J. Exper. Zool. (MDE) 291: 195- 204

Galis, F., J. M. van Alphen, and J. A. J. Metz. 2001. Why five fingers? Evolutionary constraint on digit numbers. TREE 16: 637-646.

Galpern, P. 2000. The use of common principal component analysis in studies of phenotypic evolution: an example from the Drosophilidae. M.S. Thesis. Department of Zoology, University of Toronto, Toronto, Canada.

Gerhart, J., and M. Kirschner. 1997. Cells, embryos and evolution: towards a cellular and developmental understanding of phenotypic variation and evolutionary adaptability. Blackwell

Gibson, G. 1996. Epistasis and pleiotropy as natural properties of transcriptional regulation. Theor. Pop. Biol. 49: 58-89.

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

Gilchrist, G. W., R. B. Huey, and L. Serra. 2001. Rapid evolution of wing size clines in *Drosophila subobscura*. Genetica 112/113: 273-286.

Gilchrist, A. S., R. B. R. Azevedo, L. Partridge, and P. O'Higgins. 2000. Adaptation and constraint in the evolution of *Drosophila melanogaster*. Evolution and Development 2:114-124.

Gould, S. J. 2002. The structure of evolutionary theory. Belknap Press, Cambridge MA.

Gould, S. J., and N. Eldredge. 1993. Punctuated equilibrium comes of age. Nature 366: 223-227.

Gustafsson, L. 1986. Lifetime reproductive success and heritability: empirical support for Fisher's fundamental theorem. Am. Natur. 158: 761-764.
Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. Evolution 51: 1341-1351.
Hansen, T. F. 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. Biosystems, 69: 83-94..
Hansen, T. F., Armbruster, W. S., Carlson, M. L. & Pelabon, C. 2003b. Evolvability and genetic constraint in *Dalechampia* blossoms: Genetic correlations and conditional evolvability. J. Exper. Zool. 296B: 23-39.
Hansen, T. F., Pelabon, C, Armbruster, W. S, Carlson, M. L. 2003a. Evolvability and constraint in *Dalechampia* blossoms: Components of variance and measures of evolvability. J. Evol. Biol. 16: 754-765.
Hansen, T. F., Wagner, G. P. 2001. Modeling genetic architecture: A multilinear model of gene interaction. Theoretical Population Biology: 59: 61-86.
Hendry, A. P., and M. T. Kinnison. 1999. The pace of modern life: measuring rates of contemporary microevolution. Evolution 53: 1637-1653
Hermisson, J., Hansen, T. F. & Wagner, G. P. 2003. Epistasis in polygenic traits and the evolution of genetic architecture. Am. Nat. 161: 708-734.
Hoffman, A. 1989. Arguments on evolution: A paleontologist's perspective. Oxford University press.
Holt, R. D. 1996. Demographic constraints in evolution: towards unifying the evolutionary theories of senesence and niche conservativism. Evol. Ecol. 10:1-11.
Holt, R. D. and M. S. Gaines. 1992. Analysis of adaptation in heterogeneous landscapes: Implications for the evolution of fundamental niches. Evol. Ecol. 6: 433-447.
Houle, D. 1991. Genetic covariance of fitness correlates: what genetic correlations are made of and why it matters. Evolution 45:630-648.
Houle, D. 1992. Comparing evolvability and variability of quantitative traits. Genetics 130: 195-204.
Houle, D. 1998. How should we explain variation in the genetic variance of traits? Genetica 102/103: 241-253.
Houle, D. 2001. Characters as the units of evolutionary change. In G. P. Wagner, ed. The character concept in evolutionary biology. Pp. 109-140. Academic Press. San Diego.
Houle, D., K. A. Hughes, D. K. Hoffmaster, J. Ihara, S. Assimacopoulos, D. Canada, and B. Charlesworth. 1994. The effects of spontaneous mutation on quantitative traits. I. Variances and covariances of life history traits. Genetics 138: 773-785
Houle, D., B. Morikawa, and M. Lynch. 1996. Comparing mutational variabilities. Genetics 143: 1467-1483.
Huey, R. B., G. W. Gilchrist, M. L. Carlson, D. Berrigan, and L. Serra. 2000. Rapid evolution of a geographic cline in size in an introduced fly. Science 287: 308-309.
Kauffman, S. A. 1993. The Origins of Order. Self-Organization and Selection in Evolution. Oxford Univ. press.
Kawecki, T. J. 1995. Demography of source-sink populations and the evolution of ecological niches. Evol. Ecol. 9: 38-44.
Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J.M., Berrigan, D., Vignnieri, S.N., Hill, C. E., Hoang, A., Gibert, P., Beerli, P. 2001. The strength of phenotypic selection in natural populations. Am. Nat. 157: 245-261
Kinnison, M. T., and A. P. Hendry. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. Genetica 112/113: 145-164.
Kirkpatrick, M., and D. Lofsvold. 1992. Measuring selection and constraint in the evolution of growth. Evolution 46:954-971.
Lande, R. 1976. The maintenance of genetic variability by mutation in a polygenic character with linked loci. Genetical Research 26:221-235.
Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain:body size allometry. Evolution 33: 402-416.
Lande, R. 1985. Expected time for random genetic drift of a population between stable phenotypic states. Proc. Natl. Acad. Sci. 82: 7641-7645.
Lande, R. 1986. The dynamics of peak shifts and the pattern of morphological evolution. Paleobiology, 12: 343-354.
Lande, R, and Arnold. S. J. 1983. The measurement of selection on correlated characters. Evolution 37:1210-1226.
Lewontin, RC. 1978. Adaptation. Sci. Am. 239: 212-231.
Lieberman, B. S., and S. Dudgeon. 1996. An evaluation of stabilizing selection as a mechanism for stasis. Paleaeogeography, Paleaeoclimatology, Palaeoecology 127: 229-238.

Lynch C. B. 1994. Evolutionary inferences from genetic analyses of cold adaptation in laboratory and wild populations of the house mouse. Pp. 278-304 *in* C. R. B. Boake, ed. Quantitative genetic studies of behavioral evolution. The university of Chicago press, Chigago.
Lynch, M. 1988. The rate of polygenic mutation. Genet. Res. Camb. 51: 137-148.
Lynch, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. Am. Nat. 136: 727-741.
Lynch, M. 1993. Neutral models of phenotypic evolution. Pp. 86-108 *in* L. Real, ed. Ecological genetics. Princeton Univ. Press, Princeton NJ.
Lynch, M., and W. G. Hill. 1986. Phenotypic evolution by neutral mutation. Evolution 40: 915-935.
Lynch, M., Blanchard, J., Houle, D., Kibota, T., Schultz, S., Vassilieva, L., Willis, J. 1999. Perspective: Spontaneous deleterious mutation. Evolution 53: 645-663
McAlpine, J. F. 1989. Phylogeny and classification of the Muscomorpha. Pages 1397-1518 *in* J. F. McAlpine and D. M. Wood, eds. Handbook of Nearctic Diptera, vol. 3. Research Branch of Agriculture Canada, Monograph 32, Ottawa.
.
Mackay, T. F. C., R. F. Lyman, and W. G. Hill. 1995. Polygenic mutation in *Drosophila melanogaster*: non-linear divergence among unselected strains. Genetics 139:849-859.
Maynard Smith, J. 1983. The genetics of stasis and punctuation. Ann. Rev. Genet. 17: 11-25.
Maynard Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup, and L. Wolpert. 1985. Developmental constraint and evolution. Quart. Rev. Biol., 60: 265-287.
Merila, J., B. C. Sheldon, and L. E. B. Kruuk. 2001. Explaining stasis: microevolutionary studies in natural populations. Genetica 112/113:199-222.
Mitchell, W. A., and T. J. Valone. 1990. The optimization research program: studying adaptations by their function. Quart. Rev. Biol. 65: 43-52.
Moreno, G. 1994. Genetic architecture, genetic behavior, and character evolution. Ann. Rev. Ecol. Syst. 25: 31-45.
Mousseau, T. A., and D. A. Roff. 1987. Natural Selection and the heritability of fitness components. Heredity 59: 181-197.
Olson, E. C. and Miller, R. L. 1958. Morphological integration. Chicago Univ. Press. Chicago.
Parker, G. A., and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. Nature 348: 27-33.
Powell, J. R. 1997. Progress and Prospects in Evolutionary Biology: the *Drosophila* Model. Oxford, New York.
Price T., and L. Liou. 1989. Selection on clutch size in birds. Am. Natur. 134: 950-959.
Raff, R. A. 1996. The shape of life: genes, development, and the evolution of animal form. Univ. Chicago press.
Riedl, R. J. 1977. A systems-analytical approach to macro-evolutionary phenomena. Quart. Rev. Biol. 52: 351-370.
Riedl, R. J. 1978. Order in living organisms: a systems analysis of evolution. Wiley, 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.
Robertson, F. W. 1959. Studies in quantitative inheritance. XII. Cell size and number in relation to genetic and environmental variation of body size in *Drosophila*. Genetics 44:869-896.
Robertson, F. W. 1962. Changing the relative size of the body parts of *Drosophila* by selection. Genetical Research 3:169-180.
Robertson, F. W., and E. C. R. Reeve. 1952. Studies in quantitative inheritance. I. The effects of selection for wing and thorax length in *Drosophila melanogaster*. Journal of Genetics 50:416-448.
Robinson, B.W., and R. Dukas. 1999. The influence of phenotypic modifications on evolution: The Baldwin effect and modern perspectives", Oikos 85: 582-89.
Rutherford, S. L. 2000. From genotype to phenotype: buffering mechanisms and the storage of genetic information. BioEssays 22: 1095-1105.
Santiago, E., J. Albornoz, A. Dominguez, M. A. Toro, and C. López-Fanjul. 1992. The Distribution of Spontaneous Mutations on Quantitative Traits and Fitness in *Drosophila melanogaster*. Genetics 132: 771-781.
Scharloo, W. 1991. Canalization: Genetic and developmental aspects. Ann. Rev. Ecol. Syst. 22: 65-93
Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50: 1766-1774.
Schluter, D. 2000. The ecology of adaptive radiation. Oxford university press, Oxford.
Sheldon, P. R. 1996. Plus ca change - a model for stasis and evolution in different environments. Paleaeogeography, Paleaeoclimatology, Palaeoecology 127: 209-227.

Simpson, G. G. 1944. Tempo and mode in evolution. Colombia Univ. Press.
Stenseth, N. C., and J. Maynard Smith. 1984. Coevolution in ecosystems: red queen evolution or stasis? Evolution, 38: 870-880.
Steppan, S. J., P. C. Phillips, and D. Houle. 2002. Comparative quantitative genetics: evolution of the G matrix. TREE 17: 320-327.
Stern, D. 2000. Perspective: Evolutionary developmental biology and the problem of variation. Evolution 54: 1079-1091.
Szathmáry, E. 1993. Do deleterious mutations act synergistically? Metabolic control theory provides a partial answer. Genetics 133: 127-132
Travis, J. 1989. The role of optimizing selection in natural populations. Ann. Rev. Ecol. Syst. 20: 279-296.
van Tienderen, P. M. 2000. Elasticities and the link between demographic and evolutionary dynamics. Ecology 81: 666-679.
Van Valen, L. M. 1982. Integration of species, stasis and biogeography. Evolutionary theory 6: 99-112
Venables, W.N., and B.D. Ripley. 1994. Modern Applied Statistics with S-Plus. Springer-Verlag, New York.
Waddington, C. H. 1953. Genetic assimilation of an acquired character. Evolution 7: 118-126
Wagner, G. P. 1986. The systems approach: An interface between development and population genetic aspects of evolution, *in* D. M. Raup and D. Jablonski, eds. Patterns and processes in the history of life. Pp 149-165. Springer.
Wagner, G. P. 1996. Homologues, natural kinds and the evolution of modularity. Am. Zool. 36: 36-43.
Wagner, G. P, Altenberg. L.1996. Complex adaptations and evolution of evolvability. Evolution 50: 967-976.
Wagner, G. P., Booth, G, Bagheri-Chaichian, H. 1997. A population genetic theory of canalization. Evolution 51: 329-347.
Wake, D. B., G. Roth, and M. Wake. 1983. On the problem of stasis in organismal evolution. J. Theor. Biol. 101: 211-224
Weber, K. E. 1990. Selection on wing allometry in *Drosophila melanogaster*. Genetics 126:975-989.
Weber, K. E. 1992. How small are the smallest selectable domains of form? Genetics 130:345-353.
Weber, K. E. 1996. Large genetic changes at small fitness cost in large populations of *Drosophila melanogaster* selected for wind tunnel flight: rethinking fitness surfaces. Genetics 144:205-213.
Weber, K., R. Eisman, L. Morey, A. Patty, J. Sparks, M. Tausek, and Z.-B. Zeng. 1999. An analysis of polygenes affecting wing shape on chromosome 3 in *Drosophila melanogaster*. Genetics 153:773-786.
Williams, G. C. 1992. Natural Selection: Domains, Levels, and Challenges. Oxford Univ. press. Oxford.
Williamson P. G. 1987. Selection or constraint? A proposal on the mechanism for stasis. In Campbell, M.F., ed. Rates of evolution. Allen & Unwin.
Zeng, Z.-B., H. Tachida, and C. C. Cockerham. 1989. Effects of mutation on selection limits in finite populations with multiple alleles. Genetics 122: 977-984.

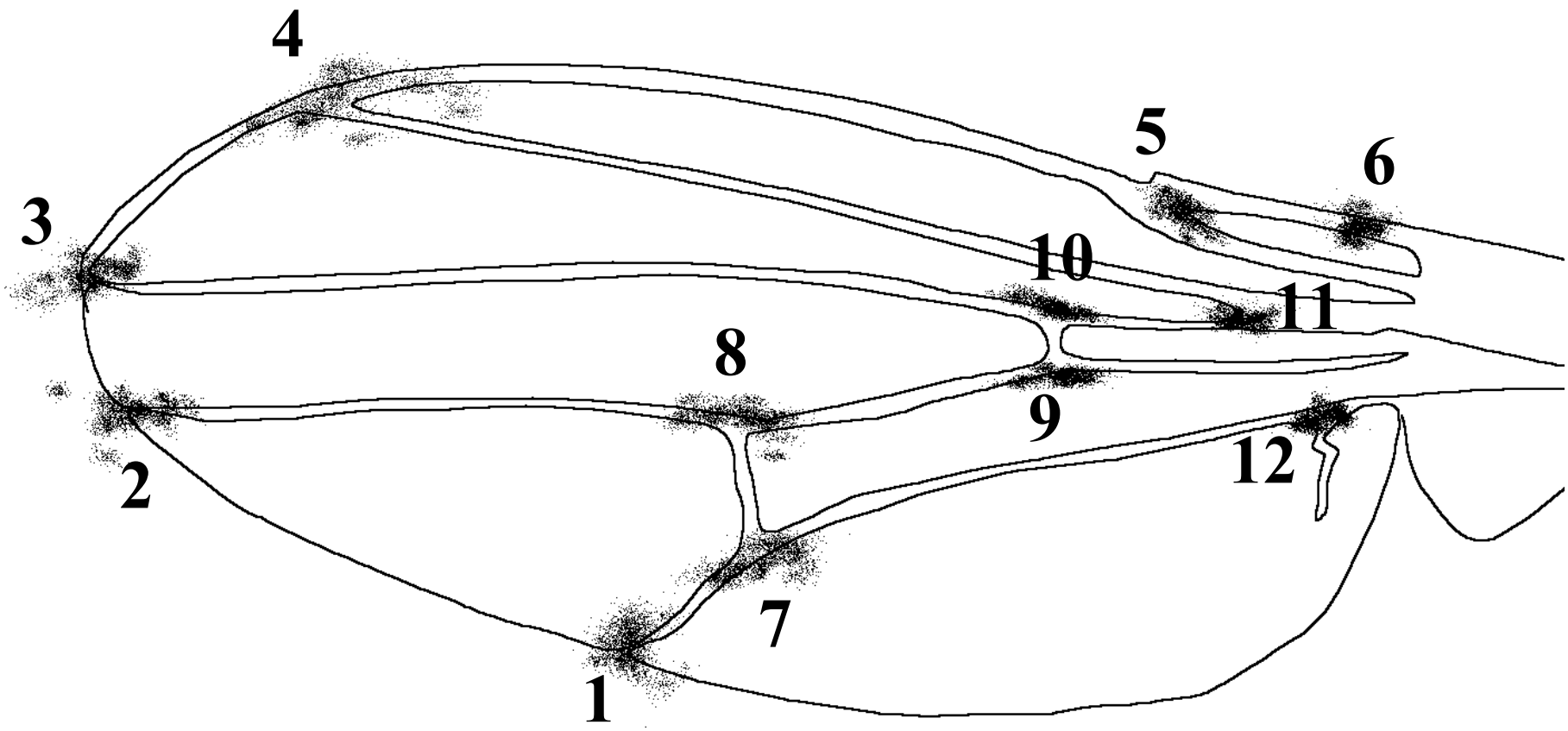
Table 1: Discriminant function analysis of Drosophilid wing shape data. The discriminant function was estimated based on 2/3 of the data, and tested on the remaining 1/3 using S-Plus routine lda (Venables and Ripley, 1994 pp 315-316).

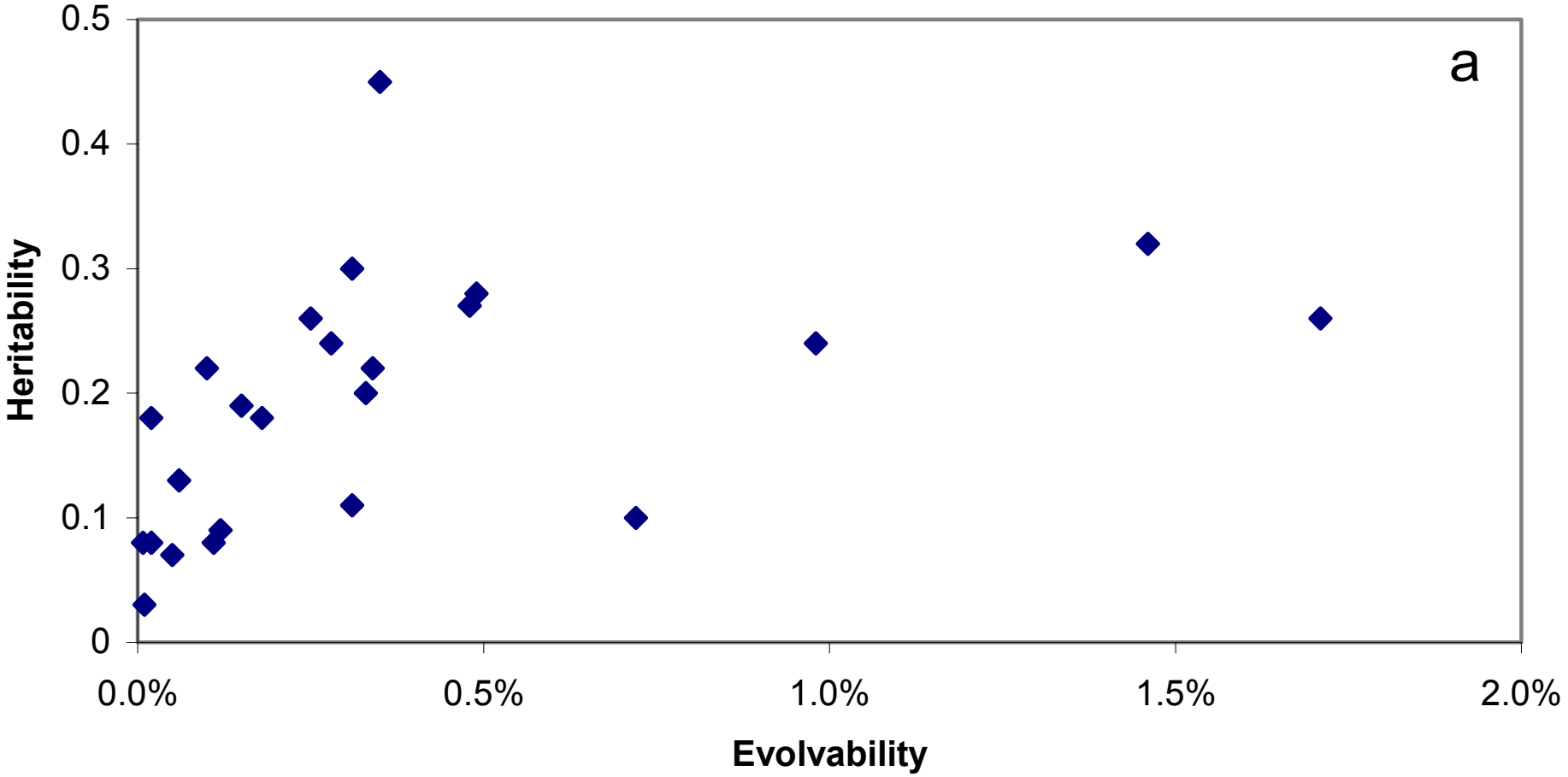
	Number of Specimens			
Species	Trained	Tested	Classification errors	Error rate
algonquin	43	21	1	0.05
athabasca	52	26	5	0.19
hydeii	119	61	13	0.21
melanogaster	128	65	3	0.05
simulans	78	40	9	0.22
18 other species	1177	592	0	0.00
Total	1597	805	31	0.036

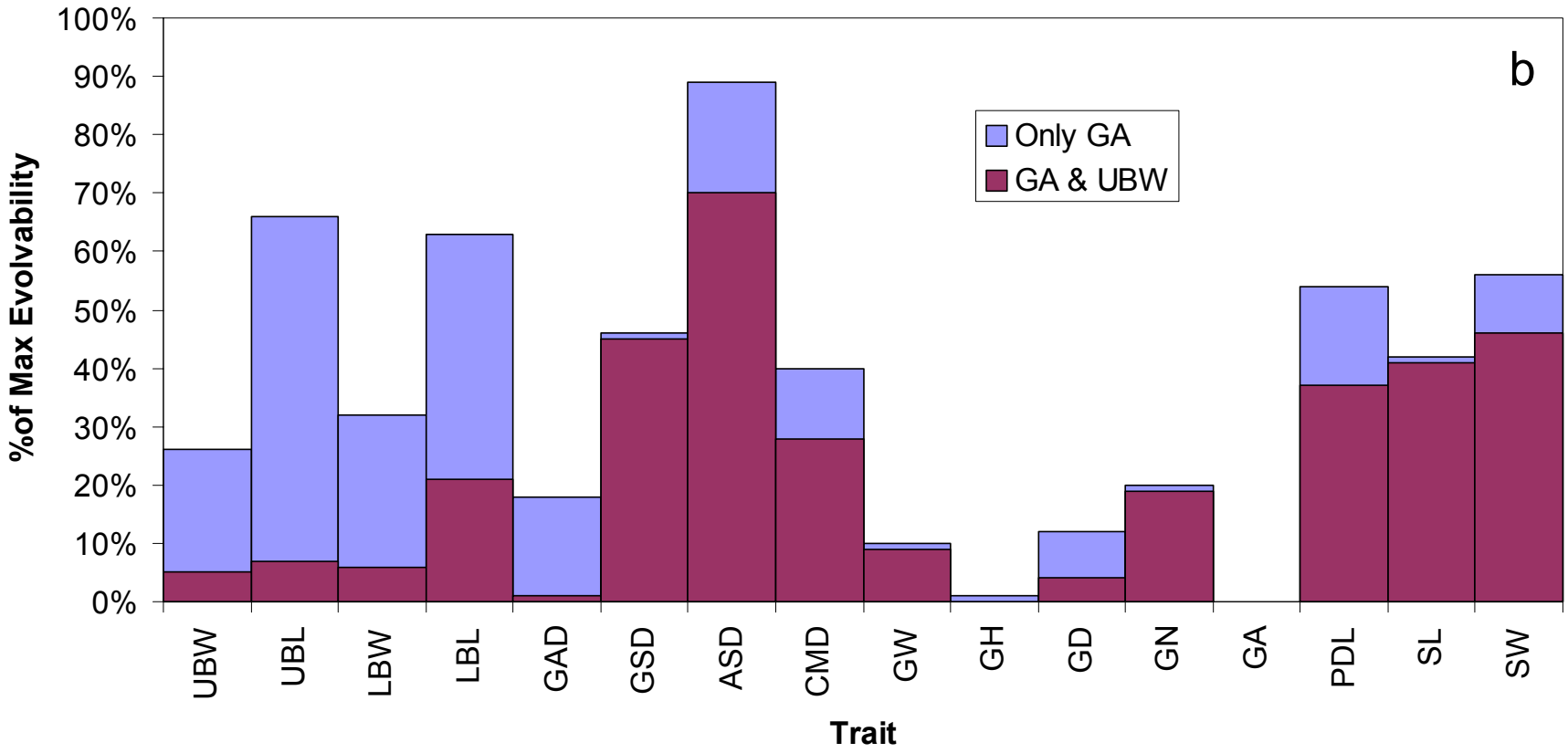
Captions to Figures

Figure 1. The wing vein landmark coordinates of Drosophilid flies when centroid size is removed and the data are aligned to a consensus configuration (Galpern 2000). The data consist of 2774 individuals of 23 species: 18 *Drosophila* species (*busckii, nebulosa, willistoni, saltans, sturtevanti, melanogaster, simulans, algonquin, athabasca, immigrans, sulfurigaster, falleni, guttifera, micromelanica, robusta, americana, virilis, hydeii*, and *repleta*), *Scaptodrosophila lebanonensis* and *S. stonei*, *Hirtodrosophila pictiventris*, and *Chymomyza procnemis*.

Figure 2. a) Relationship between heritabilities and "evolvabilities", measured as I_A, in a set of floral traits from a population of *Dalechampia scandens*. An "evolvability" of 1% means that the response to directional selection of the same strength as on fitness itself would be 1% per generation. Generally the two measures are unrelated, with the exception of a few shape characters in the lower left part of the plot that were practically devoid of genetic variation. b) Effects of conditioning evolvabilities on two functionally important traits (GA is size of a gland that produce resins as a pollinator reward, and UBW is the size of an involucral bract with an important protective function). The graph shows reduction in evolvability due to conditioning on either GA alone, or on both traits together. From Hansen et al. (2003a,b; see also Armbruster et al. this volume).


4
5
6
3
10
11
8
9
2
12
7
1


a
Heritability
0.5
0.4
0.3
0.2
0.1
0
0.0%
0.5%
1.0%
1.5%
2.0%
Evolvability


b
Only GA
GA & UBW
%of Max Evolvability
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
UBW
UBL
LBW
LBL
GAD
GSD
ASD
CMD
GW
GH
GD
GN
GA
PDL
SL
SW
Trait