

Variation and selection — what are we measuring?

Mats Björklund

Department of Animal Ecology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE-752 36 Uppsala, Sweden (e-mail: mats.bjorklund@ebc.uu.se)

Received 10 July 2003, revised version received 14 Aug. 2003, accepted 14 Aug. 2003

Björklund, M. 2003: Variation and selection — what are we measuring? — *Ann. Zool. Fennici* 40: 387–394.

Variation is a necessity for evolution and consequently patterns of variation have been analysed extensively. This analysis can be divided into two parts, a predictive part, where variation is analysed in terms of possible future evolutionary change (adaptive significance), and a narrative, which aims to understand the evolutionary history of current variation. Most often variation in morphology can be found in terms of overall size, with minor variation in shape. This suggests either that selection on size has been historically unimportant (selectively neutral), or that stochastic events override causal relations between size and fitness. The pattern of current multivariate variation suggests that selection on trait combinations has been far more important than selection on size, and should therefore be the focus of more intensive studies.

Introduction

One of the central tenets of evolutionary biology is the availability of phenotypic variation for selection to act on, and genetic variation underpinning the phenotypic variation. Without variation in fitness among individuals and a correlation between phenotypic traits and fitness evolution cannot occur. This was recognized already by Darwin (1859) who clearly identified variation as the key component for his theory of evolution (Gould 2002). In Darwin's view, variation was assumed to be copious, small in extent and undirected (Gould 2002). This means that variation is assumed to be unbounded so that at any given time we can expect that there is variation in any given direction, or at least within a reasonable time scale, that evolution is not constrained by the lack of variation.

One of the main tasks for present-day evolutionary biologists is to describe patterns of

genetic and phenotypic variation and from them try to understand the evolutionary implications. For example, ecologists might ask why mating systems are variable among populations in some species and not in others, or what causes variation in reproductive success among individuals and what are the evolutionary consequences of this variation? For population geneticists the description of genetic variation among and within populations is always a starting point for further analysis of the reasons and implications of the particular pattern of genetic variation (Wright 1978). For ecologists, questions such as why some traits vary largely among species while others do not are common, and, macro-evolutionary questions concerning variance in species number among genera and families abound (Stanley 1979). The starting point here is always a description of variation among individuals, populations, and species or at any level of interest. From there two questions arise: first,

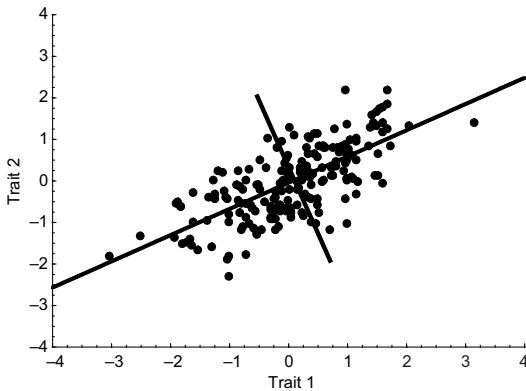


Fig. 1. Covariance between two morphological traits. The lines show the major and minor axes, respectively.

why do we have this pattern of variation and not any other, and second, given the variation, what evolutionary paths will be taken next? In other words, the questions in focus are either narrative or predictive (O'Hara 1988).

In this paper, I will discuss the analysis of variation and selection from both the predictive and the narrative view, and argue that these two views are sometimes confounded, which can have important implications for the way we view variation. The arguments are slightly twisted, on purpose, to stimulate discussion around these issues.

Patterns of variation

One of the apparent results when analysing patterns of univariate morphological variation is that there is (almost) always abundant phenotypic variation for each trait analysed, and if appropriate sample sizes are available, most often these traits have considerable additive genetic variation as well. Strikingly, often around 50% of the phenotypic variation is accounted for by an underlying genetic variation, regardless of taxa analysed (Mousseau & Roff 1987). This suggests that for a predictive analysis of the evolution of single traits the lack of variation should not be a problem, thus supporting the assumption from Darwin. This has also been used very frequently as an argument that in given time almost anything can evolve, and also as an argument for the view that similarities between species are

a result of convergence rather than a result of shared history (review by Harvey & Pagel 1991). If the assumption of sufficient variation is true the predictive questions are quite easy to answer since if we know the patterns of selection acting on the population we also know the evolutionary trajectory of the population. The narrative part is more difficult, and I will return to that question later.

The picture obtained from multivariate analyses of morphological variation is often strikingly different from the univariate analyses in one important aspect. Although variation is found in each trait separately, the level of covariation reveals another picture, namely that variation is limited and directed rather than unlimited and undirected as assumed by Darwin. This can easily be seen in Fig. 1 where two traits are plotted against each other. It is obvious that both traits exhibit considerable variation, but they covary to a fairly large extent resulting in a total pattern of variation that is limited in dimensionality. Almost all variation is in terms of the major axis describing the covariation between these two traits and the residual variation around the minor axis is clearly smaller than along the major axis. This results in large areas in the plot where there are no individuals to be found ('white areas'), meaning simply that there are no individuals found in the population with these particular combinations of traits. Thus, variation is constrained to these two dimensions and is certainly not undirected. The predictive consequences of this pattern have been the focus of much theoretical work and are well known: if the pattern of covariation is strong, almost any pattern of selection will result in a change in the direction of the major axis, and not in the direction of the selection vector (Zeng 1988, Björklund 1996, Schluter 1996). It is straightforward to show that the response to selection in one direction in this kind of situation will often lead to a response in the opposite direction, a clearly maladaptive response. This means that the Darwinian view based on univariate patterns of variation that selection can push the population in any direction has to be modified.

It is worth using some space on why we get this surprising result. Selection acts on variation and is more efficient the more variation there is,

since there is ‘more to choose between’. This means that selection in the direction of the major axis will be very efficient (actually more so than for each trait separately; Björklund 1994a), while selection in the direction of the minor axis will be less efficient since there is much less variation in this direction: most individuals have the same trait combination but differ mainly in size (major axis) (Björklund 1984b). Selection for other trait combinations is hampered by the lack of sufficient variation. Since most of the variation in the line of selection is in terms of size, the response will also be in terms of size. This pattern has been called ‘evolution along the lines of least resistance’ (Schluter 1996), which perhaps is a bit of an unfortunate term since it does not recognise that the cause of the pattern of evolution is the pattern of variation in the population. Rather than being the easy way out, as implied by the term ‘least resistance’ it might be better to focus on the fact that the resulting response to selection is a necessity given the pattern of variation.

This pattern becomes even more apparent when we add traits to the analysis. Variation for certain trait combination decreases, and it is not uncommon to find situations where there are trait combinations with no measurable variation at all (last principal component(s) equal zero). This means that all individuals in the population are identical with regard to this trait combination, and hence no selection can occur. The latter case has been termed ‘evolutionary forbidden trajectories’ by Kirkpatrick and Lofsvold (1992), in contrast to ‘evolutionary permitted trajectories’ where variation actually exists. This finding is in direct opposition to Darwin’s assumption of unlimited variation. For the predictive question this pattern is not a problem — no variation, no selection, no evolution. There are good reasons to take this finding seriously since almost always whenever a multivariate analysis is made this kind of pattern is found. This suggests that in natural populations variation is commonly limited in amount and in directionality, which in turn suggests that evolution is limited and constrained as well. Even though this is in obvious contrast to Darwin’s assumption, it is not a refutation of the basic idea of variation driven evolution, but an important modification to it.

An issue that has received considerable attention in this context is the stability of the genetic variance–covariance matrix over time and environments. If the covariances are strong and persistent over time then evolutionary pathways are highly predictable from the knowledge of the patterns of genetic covariation (Björklund 1996, Schluter 1996). If this were true then we would see highly constrained patterns of phenotypic variation among populations within a species, or among species within a genus over time. In fact, this is true for many taxa (reviews in Björklund 1996, Schluter 1996). This is consistent with a persistent bias in the variance, but not evidence for it as this can be a result of other processes as well. Strictly speaking, estimates of genetic variances and covariances are particular to the population and environment they are measured in, and cannot be extrapolated to other populations and environments (Falconer & Mackay 1996). The few empirical tests of stability of the genetic variance–covariance matrix over time and environments for the same population have so far given mixed results (Roff 2000, Stepan *et al.* 2002). This implies that for the predictive question, we are unable to make firm conclusions about the future response.

The narrative view

So far I have only discussed the predictive question, i.e. given the pattern of variation we have today, what would we expect to see tomorrow? Of importance is also the narrative question, i.e. what processes in the past have resulted in the patterns of variation we see today? In other words, why do we have a particular pattern of variation and covariation and not any other pattern? Why are some traits more variable than others? Why are some trait combinations more variable than others? The obvious answer to most evolutionary biologists is selection. The pattern of variation observed is a result of selection in the past. This simple conclusion has led many to try to understand what kind of selection has acted in the past given the pattern today (Lande 1979, Grant 1986, Merilä *et al.* 1994). As in all studies about history we cannot know exactly what has happened but we can get a picture of

which patterns of selection are more likely and which are less so. Naturally, these kinds of studies make rather strong assumptions concerning the stability of genetic variances and covariances over time, and are in fact heavily dependent on these assumptions. Even minor changes in the patterns of covariation can lead to major differences in the estimates of past selection. Still, this is the best way to understand past selection.

Suppose we have a pattern of strong covariation among traits in a population (this refers both to the genetic and phenotypic level). Most people would attribute this to selection at any time in the past, not necessarily in the most recent past. But there are alternative explanations, and one alternative that I am sure most people would find disappointing, and perhaps even boring, concerns the way we define traits. Most people dealing with morphology define traits based on what we as humans identify as more or less separate parts, and the criterion for defining a trait is usually that it can be easily seen and measured with low rates of error. It is important here to remember that these definitions are constructs of a human eye and brain and may or may not refer to biologically relevant properties of the organism in question.

The question of what is a character has been the focus of intensive debate among systematists and evolutionary oriented developmental biologists, since the concept of homology is so important in this context (Hall 1994). If phylogenies are to be erected on the basis of shared homologies, then putative homologies need to be defined and identified. Without going into details on this issue, it is apparent that the concept of a homology is not as easy to define as one would hope for, and what actually constitutes a character is not obvious at all (Hall 1994). This has relevance here since a pattern of strong covariation can arise if we measure the same trait but from different angles, i.e. what we define as separate traits are really different aspects of the same trait. Anything but strong covariation is in this case highly surprising. This has not escaped attention and, for example, the use of principal components analysis (PCA) has been suggested to identify the 'real' traits. A PCA can indeed be important to identify redundant 'traits', i.e. traits that actually only are linear combinations

of other traits. However, the PCA has a strong disadvantage in this context since the principal components (eigenvectors) are defined as being orthogonal to each other. This means that two components are uncorrelated with each other by definition, and if these components are viewed as traits then these traits are, again by definition, uncorrelated. This obscures true character correlations, i.e. if we can identify the 'real' characters these may or may not be correlated for any biological reason, but the PCA-based traits will automatically be treated as independent. I am not aware of any study that deals with this question empirically, and the problem might not be large, but it is well worth keeping in mind as a possible source of error.

Let us put this distressing possibility aside and return to explanations based on evolutionary processes such as selection. An obvious explanation for the apparent lack of variation in a trait or trait combination is stabilising selection. Stabilising selection reduces variation and if this selection is persistent over a long time, variation disappears from the population. Viewed on a broader scale, functionally important traits, or trait combinations, can be expected to display little variation, while as traits become more selectively neutral they show more and more variation since all sorts of trait values are then 'allowed' by selection, i.e. there is very little or no relation between the trait value and fitness. Returning to Fig. 1, there is considerable variation along the major axis which corresponds to size (positive covariance — individuals larger than the mean in one trait are expected to be larger than the mean also in the other trait, and vice versa), while there is much less variation along the minor axis. Now, if traits with considerable variation have little importance for fitness, this means that the amount of variation in size we see in Fig. 1 is a result of relaxed selection for size. The considerably smaller variation along the minor axis is then a result of stronger stabilising selection. In more colloquial terms, it does not matter which size you have, as long as you have the right proportions. Wrong trait proportions are penalized by selection, while size is not, resulting in a pattern of large amount of variation in terms of size, and less so in terms of the combination of the traits. Extending to many

more traits and cases where there is no variation suggests that there are combinations that are so fundamentally important that any deviation is strongly selected against.

This finding has an important consequence for students of selection in natural populations. Usually traits that are chosen for study show variation in the population in question, for example size, and the variation is analysed in relation to some measure of fitness. However, since we can expect a negative relationship between the amount of variation in a trait and its importance for fitness, we would expect to find only a weak relationship between, for example, size and the fitness measure. This is also true in many cases: although significant relationships are found, selective gradients are in most cases quite low and in many cases non-significant (Kingsolver *et al.* 2001). Non-significant selection gradients can be a result of insufficient sample sizes, but this is a problem for low values; as the true gradients are getting larger, sample size becomes less and less a problem. This in turn suggests that what most scientists are measuring are highly variable traits that are not expected to be under strong selection anyway. Note that this holds for both directional and quadratic (stabilising and disruptive) selection. Significant selection can be found but in many cases it is low or not different from zero (Kingsolver *et al.* 2001).

The most important traits, the ones under strong selection (directional or stabilising), are those with low variance. This includes not only univariate traits, but also most often trait combinations because, as we have seen above, these are the ones that lack variation most often. Studies of selection on character correlations are surprisingly small (Björklund & Senar 2001). To some extent this can be explained by the fact that fairly large sample sizes are usually needed for a robust analysis of selection (most often more than 100), but this cannot be the whole story since there are a sufficient number of studies published that actually have large sample sizes (> 100 in Kingsolver *et al.* 2001). Another reason might also be the intractability of the statistics of selection on character correlations, but today with fast computers and statistical software that can within a few seconds do all the analyses one wants to do this is not a good explanation. How-

ever, computer limitations could certainly have been a major problem only a decade ago.

Another explanation might reside in our perception of variance. Variance is a fundamental prerequisite for evolution, and the more variance there is the faster the evolutionary rate (Fisher's Fundamental Theorem). However, this is the predictive view: a population with large variance will respond better to selection than a population with smaller variance given appropriate selection. Measuring the amount of variation in a population, therefore, one measures the evolutionary potential of the population. The reason for the large variance (the narrative view) in a certain population cannot be known from the knowledge of the amount of variance alone. However, a large variance suggests that selection in the past has not been strong. To some extent the focus on traits with a large variance can be a result of a confusion of the predictive and the narrative view. The very fact that variance is necessary for future possible evolution is not the same as to say it is important today.

There is also a more technical issue in this question. Our statistical tools for analysing selection are variance-based for many good reasons. But how do we analyse selection on things that do not exist? In some cases manipulative experiments are possible to create phenotypes not observed in the population, and measure selection on these phenotypes. In other cases, like bill proportions in birds, this is certainly not easy, and when it comes to more complicated patterns manipulation is in most cases simply not possible.

Is size selectively neutral?

The discussion above suggests that highly variable traits such as size have a history of relaxed selection, while proportions between traits have a history of stronger stabilising selection as measured by the low variance. The large number of studies showing rather moderate or nonsignificant selection for size supports this. At this point I am sure many readers will object, arguing that size has indeed been shown to be important in many studies where direct selection has not been measured but other measures of importance

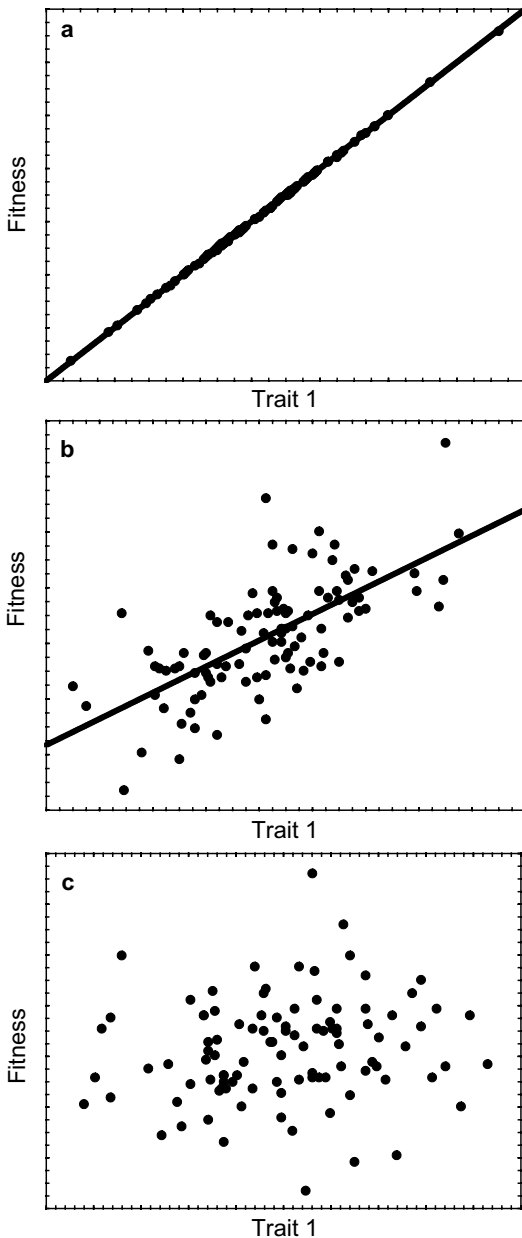


Fig 2. The relationship between fitness and size, $\sigma^2(P)$, in relation to different levels of stochastic phenotype-independent variation in fitness, $\sigma^2(E)$. — **a:** $\sigma^2(E) = 0$, — **b:** $\sigma^2(E) = \sigma^2(P)$; — **c:** $\sigma^2(E) \gg \sigma^2(P)$.

(like probability of winning contests and mating success). Still whenever selection is properly measured, i.e. the relationship between size and a measure of fitness such as survival and reproductive output, the actual level of selection is often quite low. This strongly suggests that size

might not be important over evolutionary time and this can account for the large variance in size usually found in natural populations.

One alternative explanation can be the following. The variance in fitness, say survival (S), can be decomposed into two parts: one part describing the actual relationship between survival and the phenotype (P), and the other part the variance in survival due to phenotype-independent events (E) or:

$$\sigma^2(S) = \sigma^2(P) + \sigma^2(E) + 2 \text{cov}[P, s]$$

Here E can stand for any environmental cause of death that is unrelated to the phenotype. $\sigma^2(E)$ means that to some extent individuals die due to bad luck even though they might have a phenotype that makes them more likely to survive than other phenotypes. The first term is deterministic, while the second term is stochastic. The third term, the covariance, describes the extent to which certain phenotypes associate themselves with situations that are more likely to entail bad luck. For example, if some phenotypes are more likely to engage in risky behaviour then they are more likely to die due to accidents than others, i.e. bad luck is not randomly distributed among phenotypes. For the sake of the following argument I will assume that this covariance is zero.

The possibility to find a statistically significant relation between survival and the phenotype then according to the relationship above is determined by the relation between $\sigma^2(E)$ and $\sigma^2(P)$. For most sample sizes, $\sigma^2(P)$ needs to be larger, or much larger, than $\sigma^2(E)$. This suggests that there might be a relationship between e.g. size and survival, but that this is clouded by the large environmental variance. This is incorporated in the standard regression analysis of selection (Lande-Arnold) since the test of significance is indeed the variance accounted for by the regression $\sigma^2(P)$ in relation to the residual variance $\sigma^2(E)$. A non-significant result does not actually mean that there is no relationship, only that $\sigma^2(E)$ was not small enough in relation to $\sigma^2(P)$ given the sample size. Figure 2 gives an example. Here fitness is strongly related to the phenotype (slope = 1.0), and if there is no environmental variance this is exactly what

would be estimated from the data. However, as the environmental variance increases, the estimated slope decreases and if $\sigma^2(E)$ is large enough we get a slope of zero, and conclude that there is no relation between the phenotypic trait measured and fitness. However, the size of the trait might still be important, but this is hidden by the stronger environmental component. It is important to remember that this is not only an estimation problem, but also relates to selection, since the response to selection is a function of the covariance between fitness and the trait (The Second Fundamental Theorem).

An example of how this can work can be taken from a small passerine species, such as the great tit. In many small passerines there is a fitness advantage of breeding early, and if there is a correlation between size and start of breeding, then there is a relationship between size and fitness, which indeed has been shown repeatedly. However, at northern latitudes weather conditions can be highly capricious and instances of temperatures below zero at the time when young are most vulnerable are not uncommon. If this occurs when the first breeders have young then these young will die, while broods still being incubated are not affected to the same degree. This will result in a fitness advantage for the smallest late breeding individuals, not at all as a result of an advantage of being small, but by purely stochastic reasons. This example is somewhat extreme (but real, pers. obs.). In other cases, the stochastic effects are more subtle but nevertheless important. This view is strengthened by the recent analyses of selection in natural populations (Merilä *et al.* 2001). In other words, what causes some individuals to survive and some to die is to a very large extent due to chance.

This view can also reconcile the contrasting findings that on the one hand size can be found to be important in experiments where the environmental component is greatly reduced, and a weak overall relation to fitness as measured by selection analyses since the environmental component is so much larger. If most of the variation in fitness is a result of environmental causes then selection is expected to be ineffective, and the result is a large phenotypic variance. The problem is that from a selection analysis alone it is not possible to discern the two possible

explanations; no real relationship between fitness and size or a large environmental variance in fitness.

Conclusions

Evolutionary biologists are interested in variance at all possible levels. This interest can be roughly decomposed into two parts, one narrative, or historical, where the questions about the origin and cause of present day variation is analysed, and a predictive where the variation is analysed in order to understand current functional relationships for predicting future evolutionary change. A common pattern when analysing morphological variation is that individuals vary quite extensively in terms of size, but less so in various shape components and there are cases where no variation at all could be found in some shapes. If we acknowledge that selection generally acts to reduce variance, then we can conclude that a 'trait' such as size is quite often not under very strong selection, whereas proportions quite commonly are. From this viewpoint it is ironical that so many analyses are concerned mainly with selection on size rather than selection of proportions (selection on correlations). The difference in variance alone should spur an interest, albeit this is uncommon today. A low relation between size and fitness does not necessarily mean that there is no relationship, since size has been shown to be important in many different contexts, but only that the environmentally induced variation in fitness is so large, so with most realistic sample sizes we are unable to find a relationship.

Acknowledgements

I thank Jon Stone for comments on the ms. This work was supported by grants from the Swedish Science Council.

References

- Björklund, M. 1994a: Species selection for organismal integration. — *J. Theor. Biol.* 171: 427–430.
- Björklund, M. 1994b: Processes generating macroevolutionary patterns of morphological variation in birds — a

- simulation study. — *J. Evol. Biol.* 7: 727–742.
- Björklund, M. 1996: The importance of evolutionary constraints in ecological time scales. — *Evol. Ecol.* 10: 423–431.
- Björklund, M. & Senar, J. C. 2001: Sex differences in survival selection in the serin, *Serinus serinus*. — *J. Evol. Biol.* 14: 841–849.
- Darwin, C. 1859: *The origin of species*. — Murray, London.
- Falconer, D. S. & Mackay, T. F. C. 1996: *Introduction to quantitative genetics*. — Longman, Burnt Mill.
- Gould, S. J. 2002: *The structure of evolutionary theory*. — Belknap Press, Cambridge, MA.
- Grant, P. R. 1986: *Ecology and evolution of Darwin's finches*. — Princeton University Press, Princeton, NJ.
- Hall, B. K. 1994: *Homology: the hierarchical basis of comparative biology*. — Academic Press, New York.
- Harvey, P. H. & Pagel, M. D. 1991: *The comparative method in evolutionary biology*. — Oxford Univ. Press, Oxford.
- Kingsolver, J. G., Hoekstra, H. E., Hoekstra, J. M., Berrigan, D., Vignieri, S. N., Hoang, A., Gibert, P. & Beerli, P. 2001: The strength of phenotypic selection in natural populations. — *Am. Nat.* 157: 245–261.
- Kirkpatrick, M. & Lofsvold, D. 1992: Measuring selection and constraint in evolution of growth. — *Evolution* 46: 954–971.
- Lande, R. 1979: Quantitative genetic analysis of multivariate evolution, as applied to brain: body size allometry. — *Evolution* 33: 402–416.
- Merilä, J., Björklund, M. & Gustafsson, L. 1994: Evolution of morphological differences with moderate genetic correlations among traits as exemplified by two flycatcher species (*Ficedula*; *Muscicapidae*). — *Biol. J. Linn. Soc. Lond.* 52: 19–30.
- Merilä, J., Kruuk, L. E. B. & Sheldon, B. C. 2001: Cryptic evolution in a wild bird population. — *Nature* 412: 76–79.
- Mousseau, T. A. & Roff, D. A. 1987: Natural selection and the heritability of fitness components. — *Heredity* 59: 181–197.
- O'Hara, R. J. 1988: Homage to Clio, or, toward an historical philosophy for evolutionary biology. — *Syst. Zool.* 37: 142–155.
- Roff, D. 2000: The evolution of the G matrix: selection or drift. — *Heredity* 84: 135–142.
- Schluter, D. 1996: Adaptive radiation along genetic lines of least resistance. — *Evolution* 50: 1766–1774.
- Stanley, S. M. 1979: *Macroevolution: Pattern and process*. — W.H. Freeman, San Francisco.
- Steppan, S. J., Phillips, P. C. & Houle, D. 2002: Comparative quantitative genetics: evolution of the G matrix. — *Trends Ecol. Evol.* 17: 320–327.
- Wright, S. 1978: *Evolution and the genetics of populations. 4. Variability within and among natural populations*. — Chicago University Press, Chicago.
- Zeng, Z.-B. 1988: Long-term correlated response, interpopulation covariance, and interspecific allometry. — *Evolution* 42: 363–374.