REVIEW

Phenotypic integration: studying the ecology and evolution of complex phenotypes

Massimo Pigliucci

Departments of Botany and of Ecology & Evolutionary Biology, University of Tennessee, Knoxville, TN 37996-1100, USA E-mail: pigliucci@utk.edu; URL www.genotype-environment.org

Abstract

Phenotypic integration refers to the study of complex patterns of covariation among functionally related traits in a given organism. It has been investigated throughout the 20th century, but has only recently risen to the forefront of evolutionary ecological research. In this essay, I identify the reasons for this late flourishing of studies on integration, and discuss some of the major areas of current endeavour: the interplay of adaptation and constraints, the genetic and molecular bases of integration, the role of phenotypic plasticity, macroevolutionary studies of integration, and statistical and conceptual issues in the study of the evolution of complex phenotypes. I then conclude with a brief discussion of what I see as the major future directions of research on phenotypic integration and how they relate to our more general quest for the understanding of phenotypic evolution within the neo-Darwinian framework. I suggest that studying integration provides a particularly stimulating and truly interdisciplinary convergence of researchers from fields as disparate as molecular genetics, developmental biology, evolutionary ecology, palaeontology and even philosophy of science.

INTRODUCTION: BACK TO INTEGRATION

The word 'integration,' with reference to complex aspects of an organism's phenotype, has been around in ecology and evolutionary biology for quite some time. Olson & Miller (1958) wrote a whole book about it just a few years after the discovery of the structure of DNA, while their botanical colleagues, Clausen and collaborators (Clausen & Hiesey 1960) were carrying out a multiple decade-long research programme aimed at studying what they referred to as 'character coherence'. In the mean time, Berg (1960) proposed specific hypotheses concerning the ecological circumstances under which higher or lower degrees of integration among characters should be favoured, hypotheses that are used until now to guide empirical research (e.g. Armbruster *et al.* 1999).

Yet, studies of phenotypic integration have been in the background of evolutionary ecological research throughout the second half of the 20th century, partly because of the unparalleled explosion of molecular techniques that has focused attention elsewhere and partly because long-standing conceptual and analytical problems have remained largely unaddressed for a great part of that time. Among the problems faced by researchers interested in integration were rather basic obstacles, such as the absence of a coherent and practical definition of 'integration' itself, the vague conceptual framework in which to fit empirical and theoretical studies of integration, and the formidable challenges posed by multivariate statistical analyses of the relevant quantities.

Things have begun to change during the last decade, and the field is now rapidly maturing to the point that a book attempting to highlight its major advances and challenges could be conceived (Pigliucci & Preston in press). Perhaps the beginning of the renewal of interest in integration came from the realization, in the mid-1980s (Schlichting 1986, 1989a), of its close ties with another long-neglected biological phenomenon that was beginning to experience a renaissance, phenotypic plasticity (Pigliucci 2001). Soon afterwards, the so-called 'new morphometrics', an ensemble of sophisticated multivariate statistical techniques for the analysis of shape changes, came of age (Rohlf & Marcus 1993; Marcus *et al.* 1996), finally making it possible to rigorously quantify an approach intuitively portrayed by

Thompson (1917) at the beginning of the 20th century through his famous 'grid diagrams' of shape change.

Another important set of advances started in the mid-1990s, when Gunter Wagner and his collaborators tackled the difficult task of providing a coherent evolutionary framework for studying integration. They first discussed the role of modularity in the evolution of body plans, distinguishing between the phenomena of integration (increased genetic and functional relationship among traits) and parcellation (i.e. a decrease in integration via the decoupling of formerly related traits) (Wagner 1995; Wagner & Altenberg 1996). They then provided an overarching conceptual treatment of integration as it relates to adaptation and constraints over evolutionary time scales (Wagner & Schwenk 2000), and explored the very concept of 'character' in evolutionary biology (Wagner 2001).

In the following sections, current conceptual and empirical issues surrounding the study of phenotypic integration will be discussed. The goal is not only to provide a reference point for where we are now and where we are presumably going in the near future, but most particularly to stimulate colleagues and graduate students to enter the field and take on its many fascinating challenges. This is an area in which ecology, evolutionary biology, developmental biology, genetics and molecular biology all contribute to our deeper understanding of why organisms are the way they are. Such an understanding has been the goal and the unifying theme of biological research ever since Darwin (1859). It is still have a long way to go in order to complete the puzzle.

INTEGRATION, ADAPTATION AND CONSTRAINTS

Let me start by attempting to provide a definition of integration. Precise definitions of complex ideas often hinder, rather than help, scientific research – as it is, for example clear in the case of the never-ending saga of species concepts. Furthermore, there are positive philosophical reasons for thinking of things such as 'species' or 'integration' which Wittgenstein (1953/1973) referred to as 'family resemblance concepts,' i.e. as ideas that by their intrinsic nature do not have an essential definition, but are rather best understood as identified by a complex web of properties. Any one of these properties may play a major role or be left entirely out of any specific instantiation of the concept. Accordingly, I think that the most reasonable definition of 'phenotypic integration' is a rather generic one, such as the pattern of functional, developmental and/or genetic correlation (however measured) among different traits in a given organism. Perhaps more importantly, I think it is crucial to realize that it probably makes no sense to think of an organism as 'more or less integrated' than another. Integration is measured at the level of functional

groups of characters, and some groups may be more or less integrated within the same organism. Calculating an average, organism-wide, measure of integration probably makes as much sense as calculating an average degree of phenotypic plasticity or heritability across many traits and environments, which is to say, not much.

The next question that naturally arises, then, is: should we think of phenotypic integration as an adaptation or as a constraint? Most early authors studying integration, from Olson and Miller to Berg and to Clausen et al., clearly thought in adaptationist terms. More recently, many authors have been investigating patterns of genetic variance–covariances among characters as an indication of constraints on future evolution (e.g. Antonovics 1976; Cheverud 1988). The current thinking seems to strike a necessary compromise between the two views: upcoming essays by Schwenk & Wagner (in press) and by myself (Pigliucci in press) independently argue that we need to think of adaptation and constraints as players engaged in a continuous dialectic throughout evolutionary history. On the one hand, patterns of phenotypic integration can surely be modified by natural selection in order to improve adaptation to the external environment and/or maintain the coherence of the internal developmental system. On the other hand, obviously whatever functional, developmental or genetic relationships among traits are now present in an organism will limit and channel at least its short and mid-term future evolutionary trajectory (Jernigan et al. 1994).

Perhaps surprisingly, philosophers of science have arrived at a similar conclusion by way of a completely different, but intellectually intriguing, route. For example, Matthen & Ariew (2002) have argued that what we call developmental or genetic 'constraints' are actually the necessary substrate for natural selection to work (and hence for adaptation to occur). Think of it this way: the rule of Mendelian inheritance may be thought of as 'constraining' the occurrence of certain phenotypes in a population. Yet, if inheritance were, say, blended (as Darwin thought), natural selection would soon run out of fuel and could not play the role of constructive causative agent that the modern theory of evolution asserts it plays. An interesting consequence of Matthen & Ariew's views is that the infamous debate between 'adaptationists' and 'spandrelists' (Pigliucci & Kaplan 2000) is seen to be based on a largely mistaken conception of evolution: generally speaking, as the genetic-developmental milieu is necessary for natural selection to operate, it literally does not make sense to ask, as a pan-adaptationist might, 'what would selection do in the absence of constraints?' The question is a categorical mistake, rather like asking what is the sound of the colour red. (However, using the absence of specific constraints as a null hypothesis adopted to investigate the evolution of particular characters is in fact appropriate and even necessary.)

A related question that research on the evolution of phenotypic integration may begin to answer is the one posed by the frequent observation of evolutionary stasis. As Gould (Gould & Eldredge 1993; Gould 2002) repeatedly observed, perhaps the most difficult feature of the fossil record to explain is not the more or less gradual or punctuated pattern of morphological change, but the long periods of stasis, reflected in the common observation that there are easily recognizable morphological gaps among most extant taxa. In a essay to be published, Hansen & Houle (in press) make a strong argument for the rejection of the classical neo-Darwinian explanation of stasis: millions of years of strong stabilizing selection on individual traits in the face of an unchanging external environment is simply not a sensible way to think about the problem. Hansen and Houle instead argue that it is the quality, not the quantity, of variation that is crucial to understanding phenotypic evolution. By quality they mean the patterns of phenotypic integration defined by the underlying pleiotropic and epistatic interactions among genes. Too often we have focused on one trait at a time, which has made it difficult to see why that particular trait does not change over evolutionary scales, often despite the measurable presence of genetic variation for the trait itself. It is only when we examine the complex web of co-variation of that trait with many others to which it is genetically and functionally linked that we gain a clearer picture of why, say, the shape of Drosophila wings has remained unchanged for 50 million years (Hansen & Houle in press).

Many questions about phenotypic integration as both constraint and adaptation remain, of course, open. We still do not have many examples of sophisticated biological hypotheses concerning the functionality of specific patterns of integration, and in this the field has not progressed much from the time of Berg's pioneering studies. However, we do know where to start: the most promising published examples of specific evolutionary hypotheses about phenotypic integration are found in instances in which well-understood functional modules can be quantified and compared among species. For instance, studies of the evolution of cranial morphology in monkeys (Marroig & Cheverud 2002), or of skull and dental morphology in the deer mouse (Myers *et al.* 1996) are conceptually rich because we have a good understanding of the function and development of these structures, information that leads to specific hypotheses that are much more intellectually satisfying than more generic exploratory analyses. A similar case for plants can be made with research on the relationship between floral integration and pollinator pressures (Herrera *et al.* 2002) – essentially the question of interest in Berg's classical studies.

Another general area awaiting major developments stems from our discussion of constraints and selection as continuously interacting phenomena locked in a never-ending dialectics. How are we to study such interactions given that most of our conceptual and analytical tools (such as, for example, analyses of variance) are designed with the quantification of 'major effects' as the overarching goal? Similarly, it seems clear that a clever use of the phylogenetic comparative method is paramount to making progress in the study of integration (Ackerly & Donoghue 1998; Schwenk & Wagner 2001), but even there, most thinking so far has been along the lines of neatly separating 'phylogenetic' from 'ecological' effects, with little consideration for the likelihood that the two may in fact turn out to often be inextricably intertwined (see Westoby *et al.* 1995).

If the ability to formulate sophisticated testable hypotheses concerning the evolution of phenotypic integration is still an urgent problem, so is the study of its genetic and molecular bases, although new techniques and ways of thinking have certainly made remarkable inroads in recent years. It is to this field that I now turn my attention.

THE GENETICS AND MOLECULAR BIOLOGY OF PHENOTYPIC INTEGRATION

For both empirical and conceptual reasons, I am becoming increasingly sceptical of the idea that general biological phenomena such as phenotypic plasticity or integration can be thought of as having a particular 'genetic basis' (e.g. Schlichting & Pigliucci 1993; Via 1993). Surely, specific instances of plasticity, integration, or whatever else one is interested in, do have a (often complex and difficult to untangle) genetic basis. By this I mean that certain gene products – if altered – will in turn modify the observable plasticity, integration, etc. However, as it is extremely difficult to make a convincing case that specific genes are 'for' a given, relatively well-defined, trait (Kaplan & Pigliucci 2001), it is a fortiori much more difficult to envisage a coherent genetic underpinning of heterogeneous phenomena such as those that is discussed here.

My prediction, therefore, is that the genetic/molecular bases of phenotypic integration will turn out to be of the same general kind as those of phenotypic plasticity, allometric relationships, or even quantitative individual traits. Quantitative trait loci (QTL) studies (Lynch & Walsh 1998; Phillips 1999) will undoubtedly uncover a varied number of genomic regions involved in patterns of integration in any given species, with some regions corresponding roughly to candidate genes already known through molecular/mutagenic screening. Several of these genomic regions will have complex pleiotropic effects and epistatic interactions, and the magnitude of their effects will vary from explaining a moderate amount of phenotypic variance to the lowest level of detectability allowed by the particular experimental design employed. Indeed, we already

have examples of all the above (Juenger *et al.* 2000; Mezey *et al.* 2000).

More detailed molecular analyses will also confirm the general picture that is emerging from the massive effort to study the molecular mechanisms underlying many other aspects of an organism's phenotype. In particular, we will find that some of the variation among individuals, populations and species is traceable to a certain number of regulatory elements (generically defined as any gene producing a product whose function is to turn on or off the action of other genes), or to the regulatory regions upstream of genes known to play important roles in development. Again, we already have several examples at hand (e.g. Nijhout 1994; Schmidt *et al.* 1994; Carroll *et al.* 1995; Purugganan & Suddith 1999). Indeed, this general outcome was clearly predicted as early as the mid-1970s by Jacob (1977), when he proposed not only that evolution is not an optimizing process, but that it proceeds by a (rather messy) re-assembly of whatever components are already available, similar to the work of a 'bricoleur' (do-it-yourself). In other words, the products of evolution clearly show the traces of haphazard historical events. Unfortunately, it is this very historicity of the outcomes of evolution that strictly limits the sort of generalizations we can obtain from empirical studies.

Phenotypic variance that cannot be explained with QTL and candidate loci will be attributed to a variable number of 'modifiers', a generic label attached to any gene whose product contributes slightly to the observable variation at the phenotypic level (e.g. Templeton *et al.* 1993; Cohan *et al.* 1994; Pooni & Treharne 1994). All of this, of course, is not to say that the study of the genetics and molecular biology of phenotypic integration is not of interest, but that the issues raised by it, and the generalizations likely to be gained, are not peculiar to the phenomenon of integration.

TWIN COMPLEXITY: PHENOTYPIC PLASTICITY AND INTEGRATION

When we come to consider the related issue of phenotypic plasticity, there is an important conceptual distinction to make, the one that has been clearly outlined by Schlichting (1986). On the one hand, we can think of the phenotypic plasticity of whatever measure of integration we are using. For example, correlations or covariances between traits can be altered by certain environmental conditions (e.g. Marshall *et al.* 1986; Schlichting 1989b; Stearns *et al.* 1991), i.e. they may be plastic. On the other hand, the plasticities of different traits may themselves be integrated, meaning that they can be correlated to each other. This is the far less-studied phenomenon of plasticity integration (Schlichting 1989a).

One can learn a lot from having studied phenotypic plasticity when beginning to do research on integration. In particular, the major challenges and the best strategies, I believe, are the same in the two cases. For example, as anybody who studies plasticity soon realizes, the major obstacle is simply a matter of logistics: one has to investigate the reaction norms of several genotypes from different populations (possibly representative of different species); do so for a large number of traits, and expose the organisms to a series of levels of multiple environmental factors. If one wishes to maintain some statistical power, the sheer number of individuals to use in each study easily reaches thousands, which is seldom practical for anything but a few model organisms (and even then, at great expenditure of time and resources).

Similarly in the case of integration, a researcher faces the same sort of rapidly multiplying number of individuals to measure. Even if one limits a study to one 'standard' environment (whatever that means, biologically), one would probably want to follow the change in integration patterns through development, with each point along an ontogenetic trajectory taking the place of an environmental factor in a typical plasticity study. We can only then imagine what sort of difficulties face a student of both plasticity and integration, either in the sense of the reaction norms of measures of integration, or in the more seldom addressed one of plasticity integration. I most certainly do not want to discourage the reader from pursuing this line of research, but this sort of considerations accounts for the limited output of studies of plasticity and integration in the same organism.

It seems to me that the best way around the logistical problems, in this as in any other field of science, is to abandon brute force approaches (i.e. larger and larger experimental set ups yielding barely significant statistical results), and concentrate instead on more specific questions that can then be dissected in a more powerful manner, sort of an evolutionary detective approach (Hilborn & Mangel 1997). Again, we can learn from the history of research on phenotypic plasticity. The most informative studies of plasticity I have come across are not the all too-common kind that simply describes a bunch of reaction norms for yet another organism subjected to yet another environmental gradient. Rather, they are those that set out to test specific functional hypotheses about particular kinds of plasticity, usually those that are somewhat understood either ecologically or developmentally. Examples include research on shade avoidance in plants (e.g. Dorn *et al.* 2000; Weinig 2000; Donohue *et al.* 2001; Gilbert *et al.* 2001; Callahan & Pigliucci 2002), the evolution and developmental biology of polyphenisms of the 'eye' on butterflies' wings (e.g. Windig 1994; Kingsolver 1995; Roskam & Brakefield 1999), or metamorphosis in amphibians (e.g. Newman 1992; Denver 1997; Denver *et al.* 1998), to name but a few. In other words, it is the use of detailed hypotheses that allows one to

make progress in ecology and evolutionary biology, too often plagued by an inflation of generic results that costs us much in terms of human and financial resources (Pigliucci 2002). This is never so evident as in cases such as research on plasticity and/or integration, where the logistics are so often strictly limiting our ability to design experiments or conduct field observations.

A NOTE ON MACROEVOLUTION

Although this review concerns evolutionary ecology, it is worth noting that an interesting aspect of the study of phenotypic integration concerns macroevolutionary questions, such as the stability of patterns of integration of facial characters during the evolution of primates, and hominids in particular (Ackermann & Cheverud in press), or the study of integration of cranial traits in tamarins (Cheverud 1995).

There are two approaches to macroevolutionary studies of phenotypic integration: one that examines the fossil record (as in the case of the studies referenced above) and the other that uses the phylogenetic comparative method (e.g. Ackerly & Donoghue 1998; Schwenk & Wagner 2001). There are distinct advantages and disadvantages to the two approaches, which I will briefly discuss. The palaeontological record is notoriously incomplete, of course, and – perhaps more importantly – it is simply not possible to adequately account for important phenomena affecting integration, such as phenotypic plasticity and degrees of within- and across-population genetic variation. These problems are much less significant in the case of phylogenetic comparative studies of extant taxa. Alternatively, the reconstruction of past patterns of covariation among characters based on currently living taxa and their inferred phylogenetic relationships (e.g. Steppan 1997) is at best a delicate matter involving quite a few assumptions and educated guesses, not the least of which concern both the accuracy of the phylogenetic hypothesis used and the assumptions concerning the prevalent mode of character evolution during periods of millions of years.

In these cases, perhaps even more so than in the instance of research on the plasticity of integration patterns discussed above, again being clever with hypothesis testing allows one to make more of an inroad than a brute force approach. The latter would be useless with most of the fossil record, given the scarcity of material, while in the case of the phylogenetic comparative method one would soon run into the same issues of logistics I already discussed with the comparative method, given the high number of extant taxa which one can study.

Indeed, the papers I cited in this section are remarkable for their use of multiple lines of evidence and detailed reasoning in terms of the functional morphology of the investigated patterns of integration. Particularly in quantita-

tive historical sciences such as evolutionary biology, the so-called 'method of multiple hypotheses' (Chamberlain 1897) and manifold approaches to enquiry are most likely to bring about a 'consilience' of the evidence (Whewell 1840), the convergence of the available pieces of the puzzle towards the outlining of the same underlying picture. The study of phenotypic plasticity, integration and macroevolution are particularly fertile grounds for the testing of these ideas from philosophy of science.

FROM DATA ANALYSIS TO THE THEORY OF COMPLEX EVOLUTION

The last broad category of issues concerning the study of phenotypic integration that I wish to briefly address here concerns the twin problems of how to analyse the complex data sets typical of integration studies and, consequently, of how to frame such results in a way that advances our theoretical understanding of the issues. I obviously have no space here for even a superficial treatment of the statistical methods involved, from the already mentioned suite of 'new morphometrics' approaches (Rohlf & Marcus 1993; Marcus *et al.* 1996) to the recent revival of common principal components and the associated Flury hierarchy of hypotheses of matrix similarity (Flury 1988; Phillips & Arnold 1999; Houle *et al.* 2002; Steppan *et al.* 2002), to even more recent alternatives for studying the causal factors influencing differences in correlation or covariance matrices (Roff 2002).

The basic issues involved concern the problem of adapting or expanding the wealth of available bivariate and multivariate statistical tools to the challenges posed by phenotypic integration data sets. These typically consist of the measurement of a large number of inter-related characters, relationships about which we often have specific *a priori* hypotheses that need to be incorporated in the analysis (Rudge 1998). Furthermore, the best data sets are, as indicated above, representative of phylogenetically related taxa, which brings in all the additional issues of phylogenetic reconstruction (e.g. Gittleman & Luh 1994; Huelsenbeck *et al.* 2000; Emerson *et al.* 2001), and of their use to study the evolution of quantitative traits (e.g. Martins 2000; Rohlf 2001; Martins *et al.* 2002).

The general point to be made is that the field as a whole has been held back for decades by the lack of proper statistical and analytical tools, much in the way its advance used to be limited by the lack of a clear conceptual framework as I discussed at the beginning of this essay. I think it is a testimonial to the increased perceived importance of understanding the evolution of complex phenotypes that so many researchers have finally decided to tackle the complexities inherent in developing a host of new analytical tools. Judging from the most recent articles and from the number of new papers and essays that are about to

come out, I doubt that these debates are in any way settled. However, far from being a problem, the current fluid stage of thinking about these matters is a sign of vitality in a field that has remained largely in the background for almost a century and is finally coming of age.

Things are still much more nebulous when it comes to incorporating the conceptual and analytical advances that have characterized the study of phenotypic integration into the more general framework of the neo-Darwinian theory of evolution. Attempts are surely being made (Schwenk & Wagner in press; Wolf *et al.* in press), but they are clearly only the initial forays into what promises to be a stimulating intellectual enterprise over the next decade or two. Let me conclude by outlining what I think are the major challenges propelling the expansion of current thinking on phenotypic evolution to comprise integration:

- Research on the relative advantages of different analytical methods to analyse integration data sets needs to continue to fuel the development of yet newer methods. It will not be necessary to settle on one 'best' approach, in the same way in which many other problems in quantitative biology can be tackled using more than one statistical tool. However, we do need a better understanding of the robustness of our approaches and of the limits they pose on our ability to infer causal connections among interesting biological phenomena (Shipley 2000).
- We need new theoretical and empirical tools that allow us to operationalize the insight that natural selection, and the genetic-developmental milieu in which it occurs, are both necessary players in a continuous dialectical relationship, and cannot simply be thought of as dichotomous choices ('constraints' vs. selection).
- The metaphor of evolution on 'adaptive landscapes', which has recently been fundamentally criticized in its classic form for the misleading imagery it evokes when one considers unrealistically low-dimensional landscapes (Gavrilets 1997; Gavrilets *et al.* 1998), needs to be largely revised to be useful when it comes to the high-dimensionality of research problems in phenotypic integration (Wolf *et al.* in press).
- Research on integration (and, I would argue, more generally in evolutionary ecology) will benefit greatly from a more nuanced philosophy of science, one that emphasizes competition among different complex models as opposed to more standard, sometimes simplistic, null hypotheses. Similarly, we need to explore methods of investigation that aim at converging on the best inference from different, partial, angles of attack to a particular problem (consilience of evidence), instead of continuing with the practice of standard falsificationism (Popper 1968) or 'strong inference' (Platt 1964) more typical of the 'hard' sciences.

- Finally, we need to assess to what extent the view of phenotypic evolution that emerges from integration studies (e.g. the remarkable conservation of molecular pathways underlying the same structures in phylogenetically widely separated groups: Treisman 1999) fits well with the standard neo-Darwinian paradigm, or perhaps prompts us to an expansion of such paradigm (Gould 2002) to include new insights from the field of the evolution of development (Oyama *et al.* 2001).

This is a tall order indeed, and it will require the joint efforts of researchers interested in developmental biology, molecular genetics, biostatistics, evolutionary ecology and even philosophy of science. A repeatedly invoked (Schlichting & Pigliucci 1998) and truly satisfactory integration of evolutionary and developmental biology will be the ultimate prize for our efforts.

ACKNOWLEDGMENTS

REFERENCES

Ackerly, D.D. & Donoghue, M.J. (1998). Leaf size, sapling allometry, and Corner's rules: phylogeny and correlated evolution in maples (Acer). *Am. Nat.*, 152, 767–791.

Ackermann R.R. & Cheverud J.M. (in press). Morphological integration in primate evolution. In: *The Evolutionary Biology of Complex Phenotypes* (eds. Pigliucci, M. & Preston, K.). Oxford University Press, Oxford.

Antonovics, J. (1976). The nature of limits to natural selection. *Ann. Missouri Botan. Gardens*, 63, 224–247.

Armbruster, W.S., Stilio, V.S.D., Tuxill, J.D., Flores, T.C. & Runk, J.L.V. (1999). Covariance and decoupling of floral and vegetative traits in nine Neotropical plants: a re-evaluation of Berg's correlation-pleiades concept. *Am. J. Botany*, 86, 39–55.

Berg, R.L. (1960). The ecological significance of correlation pleiades. *Evolution*, 14, 171–180.

Callahan, H. & Pigliucci, M. (2002). Shade-induced plasticity and its ecological significance in wild populations of Arabidopsis thaliana. *Ecology*, 83, 1965–1980.

Carroll S.B., Weatherbee S.D. & Langeland J.A. (1995) Homeotic genes and the regulation and evolution of insect wing number. *Nature*, 375, 58–61.

Chamberlain, T.C. (1897). The method of multiple working hypotheses. *Science*, 15, 92–96.

Cheverud, J.M. (1988). The evolution of genetic correlation and developmental constraints. In: *Population genetics and evolution* (ed. Jong G.D.). Springer-Verlag, Berlin,pp. 94–101.

Cheverud, J.M. (1995). Morphological integration in the saddle-back tamarin (Saguinus fuscicollins) cranium. *Am. Nat.*, 145, 63–89.

Clausen, J. & Hiesey, W.M. (1960). The balance between coherence and variation in evolution. *Proc. Natl Acad Sci U S A*, 46, 494–506.

Cohan, F.M., King, E.C. & Zawadski, P. (1994). Amelioration of the deleterious pleiotropic effects of an adaptive mutation in Bacillus subtilis. *Evolution*, 48, 81–95.

Darwin, C. (1859). *The Origin of Species by Means of Natural Selection: or, the Preservation of Favored Races in the Struggle for Life*. A.L. Burt (1910), New York.

Denver, R.J. (1997). Proximate mechanisms of phenotypic plasticity in amphibian metamorphosis. *Am. Zool.*, 37, 172–184.

Denver, R.J., Mirhadi, N. & Phillips, M. (1998). Adaptive plasticity in amphibian metamorphosis: response of Scaphiopus hammondii tadpoles to habitat desiccation. *Ecology*, 79, 1859–1872.

Donohue, K., Messiqua, D., Pyle, E.H., Heschel, M.S. & Schmitt, J. (2001). Evidence of adaptive divergence in plasticity: density- and site-dependent selection on shade-avoidance responses in Impatiens capensis. *Evolution*, 55, 1956–1968.

Dorn, L.A., Pyle, E.H. & Schmitt, J. (2000). Plasticity to light cues and resources in Arabidopsis thaliana: testing for adaptive value and costs. *Evolution*, 54, 1982–1994.

Emerson, B.C., Ibrahim, K.M. & Hewitt, G.M. (2001). Selection of evolutionary models for phylogenetic hypothesis testing using parametric methods. *J. Evol. Biol*, 14, 620–631.

Flury, B. (1988). *Common Principal Components and Related Multivariate Models*. Wiley, New York.

Gavrilets, S. (1997). Evolution and speciation on holey adaptive landscapes. *Trends Ecol. Evol.*, 12, 307–312.

Gavrilets, S., Li, H. & Vose, M.D. (1998). Rapid parapatric speciation on holey adaptive landscapes. *Proc. R. Soc. Lond.*, B265, 1483–1489.

Gilbert, R., Jarvis, P.G. & Smith, H. (2001). Proximity signal and shade avoidance differences between early and late successional trees. *Nature*, 411, 792–795.

Gittleman, J.L. & Luh, H.-K. (1994). Phylogeny, evolutionary models and comparative methods: a simulation study. In: *Phylogenetics and ecology*.The Linnean Society, London, pp. 103–122.

Gould, S.J. (2002). *The Structure of Evolutionary Theory*. Harvard University Press, Cambridge, MA.

Gould, S.J. & Eldredge, N. (1993). Punctuated equilibrium come of age. *Nature*, 366, 223–227.

Hansen, T.F. & Houle, D. (In press). Evolvability, stabilizing selection, and the problem of stasis. In: *The Evolutionary Biology of Complex Phenotypes* (eds Pigliucci, M. & Preston, K.). Oxford University Press, Oxford.

Herrera, C.M., Cerda, X., Garcia, M.B., Guitian, J., Medrano, M., Rey, P.J. & Sanchez-Lafuente, A.M. (2002). Floral integration, phenotypic covariance structure and pollinator variation in bumblebee-pollinated Helleborus foetidus. *J. Evol. Biol.*, 15, 108–121.

Hilborn, R. & Mangel, M. (1997). *The Ecological Detective: Confronting Models with Data*. Princeton University Press, Princeton, NJ.

Houle, D., Mezey, J. & Galpern, P. (2002). Interpretation of the results of common principal components analyses. *Evolution*, 56, 433–440.

Huelsenbeck, J.P., Rannala, B. & Masly, J.P. (2000). Accomodating phylogenetic uncertainty in evolutionary studies. *Science*, 288, 2349–2350.

Jacob, F. (1977). Evolution and tinkering. *Science*, 196, 1161–1166.

Jernigan, R.W., Culver, D.C. & Fong, D.W. (1994). The dual role of selection and evolutionary history as reflected in genetic correlations. *Evolution*, 48, 587–596.

Juenger T., Purugganan M. & Mackay T.F.C. (2000) Quantitative trait loci for floral morphology in Arabidopsis thaliana. *Genetics*, 156, 1379–1392.

Kaplan J. & Pigliucci M. (2001) Genes 'for' phenotypes: a modern history view. *Biology and Philosophy*, 16, 189–213.

Kingsolver J.G. (1995) Fitness consequences of seasonal polyphenism in western white butterflies. *Evolution*, 49, 942–954.

Lynch M. & Walsh B. (1998) *Genetics and analysis of quantitative traits*. Sinauer, Sunderland, MA.

Marcus L.F., Corti M., Loy A., Naylor G.J.P. & Slice D.E. (1996) *Advances in morphometrics*. Plenum Press, New York, NY.

Marroig G. & Cheverud J.M. (2002) A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. *Evolution*, 55, 2576–2600.

Marshall D.L., Levin D.A. & Fowler N.L. (1986) Plasticity of yield components in response to stress in Sesbania macrocarpa and Sesbania vesicaria (Leguminosae). *American Naturalist*, 127, 508–521.

Martins E.P. (2000) Adaptation and the comparative method. *Trends in Ecology & Evolution*, 15, 296–299.

Martins E.P., Diniz-Filho J.A.F. & Hoursworth E.A. (2002) Adaptive constraints and the phylogenetic comparative method: a computer simulation test. *Evolution*, 56, 1–13.

Matthen M. & Ariew A. (2002) Two ways of thinking about fitness and natural selection. *Journal of Philosophy*, 49, 55–83.

Mezey J.G., Cheverud J.M. & Wagner G.P. (2000) Is the genotype-phenotype map modular? A statistical approach using mouse Quantitative Trait Loci data. *Genetics*, 156, 305–311.

Myers P., Lundrigan B.L., Gillespie B.W. & Zelditch M.L. (1996) Phenotypic plasticity in skull and dental morphology in the prairie deer mouse (Peromyscus maniculatus bairdii). *Journal of Morphology*, 229, 229–237.

Newman R.A. (1992) Adaptive plasticity in amphibian metamorphosis. *BioScience*, 42, 671–678.

Nijhout H.F. (1994) Developmental perspectives on evolution of butterfly mimicry. *BioScience*, 44, 148–156.

Olson E.C. & Miller R.L. (1958) *Morphological Integration*. University of Chicago Press, Chicago.

Oyama S., Griffiths P.E. & Gray R.D. (2001) *Cycles of Contingency: Developmental Systems Theory and Evolution*. MIT Press, Cambridge, MA.

Phillips P.C. (1999) From complex traits to complex alleles. *Trends in Genetics*, 15, 6–8.

Phillips P.C. & Arnold S.J. (1999) Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution*, 53, 1506–1515.

Pigliucci M. (2001) *Phenotypic Plasticity: Beyond Nature and Nurture*. Johns Hopkins University Press, Baltimore, MD.

Pigliucci M. (2002) Are ecology and evolutionary biology 'soft' sciences? *Annales Zoologici Finnici*, 39, 87–98.

Pigliucci M. (in press) Studying mutational effects on G matrices. In: *The Evolutionary Biology of Complex Phenotypes* (ed. Pigliucci M). Oxford University Press, Oxford.

Pigliucci M. & Kaplan J. (2000) The fall and rise of Dr. Pangloss: adaptationism and the Spandrels paper 20 years later. *Trends in Ecology & Evolution*, 15, 66–70.

Pigliucci M. & Preston K. (in press) *The Evolutionary Biology of Complex Phenotypes*. Oxford University Press, Oxford.

Platt J.R. (1964) Strong inference. *Science*, 146, 347–353.

Pooni H.S. & Treharne A.J. (1994) The role of epistasis and background genotype in the expression of heterosis. *Heredity*, 72, 628–635.

Popper K.R. (1968) *Conjectures and refutations: the growth of scientific knowledge*. Harper & Row, New York, NY.

Purugganan M.D. & Suddith J.I. (1999) Molecular population genetics of floral homeotic loci: departures from the equilibrium-neutral model at the APETALA3 and PISTILLATA genes of Arabidopsis thaliana. *Genetics*, 151, 839–848.

Roff D. (2002) Comparing G matrices: a MANOVA approach. *Evolution*, 56, 1286–1291.

Rohlf F.J. (2001) Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution*, 55, 2143–2160.

Rohlf F.J. & Marcus L.F. (1993) A revolution in morphometrics. *Trends in Ecology and Evolution*, 8, 129–132.

Roskam J.C. & Brakefield P.M. (1999) Seasonal polyphenism in Bicyclus (Lepidoptera: Satyridae) butterflies: different climates need different cues. *Biological Journal of the Linnean Society*, 66, 345–356.

Rudge D.W. (1998) A Bayesian analysis of strategies in evolutionary biology. *Perspectives on Science*, 6, 341–360.

Schlichting C.D. (1986) The evolution of phenotypic plasticity in plants. *Annual Review of Ecology and Systematics*, 17, 667–693.

Schlichting C.D. (1989a) Phenotypic integration and environmental change. *BioScience*, 39, 460–464.

Schlichting C.D. (1989b) Phenotypic plasticity in Phlox. II. Plasticity of character correlations. *Oecologia*, 78, 496–501.

Schlichting C.D. & Pigliucci M. (1993) Evolution of phenotypic plasticity via regulatory genes. *The American Naturalist*, 142, 366–370.

Schlichting C.D. & Pigliucci M. (1998) *Phenotypic Evolution, A Reaction Norm Perspective*. Sinauer, Sunderland, MA.

Schmidt D.L., Jong A.J.d. & Vries S.C.d. (1994) Signal molecules involved in plant embryogenesis. *Plant Molecular Biology*, 26, 1305–1313.

Schwenk K. & Wagner G.P. (2001) Function and the evolution of phenotypic stability: connecting pattern to process. *American Zoologist*, 41, 552–563.

Schwenk K. & Wagner G.P. (in press) The relativism of constraints on phenotypic evolution. In: *The Evolutionary Biology of Complex Phenotypes* (eds. Pigliucci M & Preston K). Oxford University Press, Oxford.

Shipley B. (2000) *Cause and Correlation in Biology: a User's Guide to Path Analysis, Structural Equations and Causal Inference*. Cambridge University Press, Cambridge, England.

Stearns S., de-Jong G. & Newman B. (1991) The effects of phenotypic plasticity on genetic correlations. *Trends in Ecology and Evolution*, 6, 122–126.

Steppan S.J. (1997) Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. *Evolution*, 51, 587–594.

Steppan S.J., Phillips P.C. & Houle D. (2002) Comparative quantitative genetics: evolution of the G matrix. *Trends in Ecology & Evolution*, 17, 320–327.

Templeton A.R., Hollocher H. & Johnston J.S. (1993) The molecular through ecological genetics of abnormal abdomen in Drosophila mercatorum. V. Female phenotypic expression on natural genetic backgrounds and in natural environments. *Genetics*, 134, 475–485.

Thompson D. (1917) *On growth and form*. Cambridge University Press, Cambridge.

Treisman J.E. (1999) A conserved blueprint for the eye? *BioEssays*, 21, 843–850.

Via S. (1993) Adaptive phenotypic plasticity: target or by-product of selection in a variable environment? *The American Naturalist*, 142, 352–365.

Wagner G. (2001) *The Character Concept in Evolutionary Biology*. Academic Press, San Diego, CA.

Wagner G.P. (1995) Adaptation and the modular design of organisms. In: *Advances in Artificial Life* (eds. Moran F, Moreno A, Merelo JJ & Chacon P), pp. 317–328. Springer, Berlin.

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

Wagner G.P. & Schwenk K. (2000) Evolutionarily stable configurations: functional integration and the evolution of phenotypic stability. *Evolutionary Biology*, 31, 155–217.

Weinig C. (2000) Limits to adaptive plasticity: temperature and photoperiod influence shade-avoidance responses. *American Journal of Botany*, 87, 1660–1668.

Westoby M., Leishman M.R. & Lord J.M. (1995) On misinterpreting the 'phylogenetic correction'. *Journal of Ecology*, 83, 531–534.

Whewell W. (1840) *The philosophy of the inductive sciences, founded upon their history*. J.W. Parker, London.

Windig J.J. (1994) Reaction norms and the genetic basis of phenotypic plasticity in the wing pattern of the butterfly Bicyclus anynana. *Journal of Evolutionary Biology*, 7, 665–695.

Wittgenstein L. (1953/1973) *Philosophical Investigations*. Macmillan, New York, NY.

Wolf J.B., Allen C.E. & Frankino W.A. (in press) Multivariate phenotypic evolution in developmental hyperspace. In: *The Evolutionary Biology of Complex Phenotypes* (eds. Pigliucci M & Preston K). Oxford University Press, Oxford.