

Landscapes, Surfaces, and Morphospaces: What Are They Good For?

Massimo Pigliucci

3.1 Introduction

Few metaphors in biology are more enduring than the idea of Adaptive Landscapes, originally proposed by Sewall Wright (1932) as a way to visually present to an audience of typically non-mathematically savvy biologists his ideas about the relative role of natural selection and genetic drift in the course of evolution. The metaphor, however, was born troubled, not the least reason for which is the fact that Wright presented different diagrams in his original paper that simply cannot refer to the same concept and are therefore hard to reconcile with each other (Pigliucci 2008). For instance, in some usages, the landscape's non-fitness axes represent combinations of individual genotypes (which cannot sensibly be aligned on a linear axis, and accordingly were drawn by Wright as polyhedrons of increasing dimensionality). In other usages, however, the points on the diagram represent allele or genotypic *frequencies*, and so are actually populations, not individuals (and these can indeed be coherently represented along continuous axes).

Things got even more confusing after the landscape metaphor began to play an extended role within the Modern Synthesis in evolutionary biology and was appropriated by G. G. Simpson (1944) to further his project of reconciling macro- and microevolution, i.e. to reduce paleontology to population genetics (some may object to the characterization of this program as a reductive one, but if the questions raised by one discipline can be reframed within the conceptual framework of another, that

is precisely what in philosophy of science is meant by reduction; see Brigandt 2008). This time the non-fitness axes of the landscape were phenotypic traits, not genetic measures at all. Even more recently, Lande and Arnold (1983) proposed a mathematical formalism aimed at estimating actual (as opposed to Simpson's hypothetical) fitness surfaces, making use of standard multiple regression analyses. But while Simpson was talking about macroevolutionary change involving speciation, Lande and Arnold were concerned with microevolutionary analyses within individual populations of a single species.

In principle, it is relatively easy to see how one can go from individual-genotype landscapes to genotypic-frequency landscapes (the two Wright versions of the metaphor). However, the (implied) further transition from either of these to phenotypes (either in the Lande–Arnold or in the Simpson version) is anything but straightforward because of the notorious complexity and non-linearity of the so called genotype–phenotype mapping function (Alberch 1991; Pigliucci 2010). This is a serious issue if—as I assume is the case—we wish to use the landscape metaphor as a unified key to an integrated treatment of genotypic and phenotypic evolution (as well as of micro- and macroevolution). Without such unification evolutionary biology would be left in the awkward position of having two separate theories, one about genetic change, the other about phenotypic change, and no bridge principles to connect them.

One more complication has arisen in more recent years, this one concerning Wright-style fitness landscapes. Work by Gavrillets (2003; see also Chapter 17

this volume) and collaborators, made possible by the availability of computing power exceeding by several orders of magnitudes what was achievable throughout the twentieth century, has explored the features of truly highly dimensional landscapes—as opposed to the standard two- or three-dimensional ones explicitly considered by Wright and by most previous authors. As it turns out, evolution on so-called “holey” landscapes is characterized by qualitatively different dynamics from those suggested by the standard low-dimensional version of the metaphor—a conclusion that has led some authors to suggest abandoning the metaphor altogether, in favor of embracing directly the results of formal modeling (Kaplan 2008; though see Plutynski 2008 and Chapter 2 for a somewhat different take).

In this essay I wish to discuss the implications of four versions of the metaphor, often referred to as fitness landscapes, Adaptive Landscapes, fitness surfaces, and morphospaces. I will argue that moving from one to any of the others is significantly more difficult than might at first be surmised, and that work with morphospaces has been unduly neglected by the research community.

3.2 Four types of landscapes

As I mentioned earlier, there are several versions of the “landscape” metaphor that have proliferated in the literature since Wright’s original paper. Indeed, Wright himself was referring—in that very same paper—to at least two conceptions of landscapes, one with individual genotypes as the points on the (necessarily non-continuous) landscape, the other with populations identified by gene or genotype frequencies (in a continuous space). I will ignore the distinction between the two types of Wright landscapes for two reasons: first, they are connected in a conceptually straightforward manner, since it is obviously possible to go from individual genotypes to populations of genotypes without any theoretical difficulty. Second, most of the post-Wright literature, including the presentation of the metaphor in textbooks and the more recent work on “holey” landscapes, is framed in terms of gene/genotype frequencies, not individual genotypes. This is understandable within the broader

context of the Modern Synthesis (Mayr and Provine 1980) as a theory essentially rooted in (though certainly not limited to) population genetics, where evolution is often simply defined as change in allelic frequencies (Futuyma 2006).

Despite some confusion due to the often interchangeable use of terms like “fitness” and “Adaptive” Landscape, I will use and build here on the more rigorous terminology followed by authors such as McGhee (2007) and distinguish the following four types of landscapes:

Fitness landscapes. These are the sort of entities originally introduced by Wright and studied in a high-dimensional context by Gavrillets and collaborators. The non-fitness dimensions are measures of genotypic diversity. The points on the landscape are population means, and the mathematical approach is rooted in population genetics.

Adaptive Landscapes. These are the non-straightforward “generalizations” of fitness landscapes introduced by Simpson, where the non-fitness dimensions now are phenotypic traits. The points on the landscape are populations speciating in response to ecological pressures or even above-species level lineages (i.e. this is about macroevolution).

Fitness surfaces. These are the Lande–Arnold type of landscapes, where phenotypic traits are plotted against a surrogate measure of fitness. They are statistical estimates used in quantitative genetic modeling, and the points on the landscape can be either individuals within a population or population means, in both cases belonging to a single species (i.e. this is about microevolution).

Morphospaces. These were arguably first introduced by Raup (1966), and differ dramatically from the other types for two reasons: (a) they do not have a fitness axis; (b) their dimensions, while representing phenotypic (“morphological”) traits, are generated via a priori geometrical or mathematical models, i.e. they are not the result of observational measurements. They typically refer to across-species (macroevolutionary) differences, though they can be used for within-species work.

Let us take a look in a bit more detail at each type of landscape, to familiarize ourselves with their respective similarities and differences. Throughout

I will use both the terminology just summarized and the names of the relevant authors, to reduce confusion at the cost of some redundancy. The first thing we notice from even a cursory examination of the literature is that there are few actual biological examples of fitness landscapes (Wright-style) or Adaptive Landscapes (Simpson-style) available, while there is a good number of well understood examples of morphospaces (Raup-style) and particularly of adaptive surfaces (Lande–Arnold style). These differences are highly significant for my discussion of the metaphor.

Beginning with Wright-type, fitness landscapes, many of the examples in the literature are entirely conceptual, i.e. presented by various authors only for heuristic purposes. It should be obvious why it is so: to actually draw a “real” fitness landscape we need a reasonably complete description of the genotype \rightarrow fitness mapping function, i.e. we need data about the fitness value of each relevant combination of genotypes that we happen to be interested in. For most purposes this is next to impossible. Dobzhansky (1970, p. 25) famously put it this way:

Suppose there are only 1000 kinds of genes in the world, each gene existing in 10 different variants or alleles. Both figures are patent underestimates. Even so, the number of gametes with different combinations of genes potentially possible with these alleles would be 10^{1000} . This is fantastic, since the number of subatomic particles in the universe is estimated as a mere 10^{78} .

Indeed, things are complicated even further by the fact that the genotype \rightarrow fitness function can be thought of as the combination of two subfunctions: genotype \rightarrow phenotype and phenotype \rightarrow fitness. The second function requires an understanding of fitness (Lande–Arnold) surfaces, which can then be translated into fitness landscapes through the first function. I know of very few instances in which anything like that has even been attempted, given the so far (and possibly in principle) enormous difficulties, both empirical and computational.

The major class of exceptions to the paucity of actual fitness (Wright) landscapes is in itself highly illuminating of both the potential and limitations of the metaphor: studies of the evolution of RNA and protein structures. Consider for instance the

work of Cowperthwaite and Meyers (2007) on 30-nucleotide long binary RNA molecules. This is arguably one of the simpler models of genotype–phenotype–fitness relations, and it is computationally tractable and empirically approachable. Still, 30-nucleotide binary molecules correspond to a bewildering one billion unique sequences! These in turn generate “only” about 220,000 unique folding shapes in a G/U landscape and a “mere” 1000 shapes in the A/U landscape, both of which these authors have tackled. In other words, a genotypic space of a billion sequences corresponds (obviously in a many-to-many manner) to a phenotypic space of thousands to hundreds of thousands of possibilities, each characterized by its own (environment-dependent, of course) fitness value.

Things get even more complex when we move from RNA to proteins: Wroe et al. (2007) have explored evolution in protein structure-defined phenotypic space, focusing in particular on so-called “promiscuity,” the ability of a given protein to perform two functions because of the alternation between multiple thermodynamically stable configurations. Proteins are, of course, much more complicated biochemical objects than RNA molecules. They are made of more types of building blocks, and their three-dimensional structure is more difficult to predict from their linear sequence. Still, studies like those of Wroe et al. show some interesting similarities between RNA and protein landscapes, perhaps one of the most significant being that—as predicted by the “holey” landscape models studied by Gavrillets—large portions of phenotypic hyperspace are actually neutral in terms of the fitness of the forms that define that space, which means that much of the time evolution is a matter of sliding around a fitness landscape via genetic drift, with occasional “punctuations” of selective episodes.

Let us turn now to Adaptive (Simpson) Landscapes. The original idea was to use the metaphor to show how macroevolutionary events like ecological speciation can be understood in terms of the then nascent theory of population genetics, developed to directly address microevolutionary processes. Simpson, for instance, presented a (hypothetical) analysis of speciation in the *Equidae* (horse) lineage from the Cenozoic (Simpson 1953). The sub-

family (i.e. significantly above species level) *Hyracotheriinae* is represented as occupying an adaptive peak characterized by their browsing-suited teeth morphology during the Eocene. Simpson also imagined a contemporary, empty, adaptive peak for species whose teeth are suitable for grazing. The diagram then shows how gradually, during the Oligocene, the subfamily *Anchitheriinae* evolved anagenetically (i.e. by replacement) from their *Hyracotheriinae* ancestors. The new subfamily, as a result of evolving greater body size, also acquired a teeth morphology that corresponded to a gradual movement of the peak itself closer to the empty grazing zone. Finally, in the late Miocene, the *Anchitheriinae* gave origin, cladistically (i.e. by splitting) to two lineages, one of which definitely occupied the grazing peak and gave rise to the *Equinae*. Simpson's reconstruction of the events is, of course, compatible with what we knew then of the phylogeny and functional ecology of the horse family, but the landscape's contour and movements are by necessity entirely hypothetical. This is a truly heuristic device to make a conceptual point, not the result of any quantitative analysis of what actually happened.

The situation is quite different for fitness (Lande–Arnold) surfaces, which are supposed to be closely related to Simpson's landscapes, but in fact work significantly differently in a variety of respects. For one thing, the literature is full of Lande–Arnold type selection analyses, because they are relatively straightforward to carry out empirically and in terms of statistical treatment. As is well known, phenotypic selection studies of this type are based on the quantification of selection vectors by means of statistical regression of a number of measured traits on a given fitness proxy, appropriate or available for the particular organism under study. There are, however, several well-known problems of implementation (Mitchell-Olds and Shaw 1987). Among other things, surveys of Lande–Arnold type studies show that many actual estimates of selection coefficients are unreliable because they are based on too small sample sizes (Kingsolver et al. 2001; Siepielski et al. 2009). Moreover, they tend to have very low replication either spatially (from one location to the other) or temporally (from one season to another), thereby undermining any claim to the generality or reproducibility of the results

(unlike the much more experimentally confined cases of RNA and protein structures mentioned earlier). Most crucially, of course, there is essentially no connection between fitness surfaces and either fitness (Wright) landscapes or adaptive (Simpson) landscapes. The reason for the former lack of connection is that we do not have any idea of the genotype \rightarrow phenotype mapping function underlying most traits studied in selection analyses, so that we cannot articulate any transition at all from Wright landscapes to Lande–Arnold surfaces. In terms of a bridge from Lande–Arnold surfaces to Simpson landscapes, this would seem to be easier because they are both expressed in terms of phenotypic measurements versus fitness estimates. But there is where the similarity ends: Lande–Arnold selection coefficients cannot be compared across species, and they do not incorporate any of the functional ecological analyses of the type envisioned by Simpson. Indeed, a major problem with the literature on selection coefficients is that they simply ignore functional ecology altogether: we can measure selection vectors, but we usually have no idea of what causes them, and no such idea can come from multiple regression analyses (this is simply a case of the general truth that correlation does not imply causation, let alone a specific type of causation). All of this means that—despite claims to the contrary (Arnold et al. 2001; Estes and Arnold 2007; see also Chapter 13) adaptive surfaces are *not* a well worked out bridge between micro- and macroevolution (see Kaplan 2009 for a detailed explanation based on the specific claims made by Estes and Arnold 2007).

The situation is again different when we move to morphospaces. Arguably the first example of these was proposed by Raup (1966), who generated a theoretical space of all possible forms of bivalved shells based on a simple growth equation, with the parameters of the equation defining the axes of the morphospace. As I have already noted, there is no fitness axis in morphospaces, though as we shall see later, fitness/adaptive considerations do enter into how morphospaces are *used*. Moreover, morphospaces do not depend on any actual measurement at all: they are not constructed empirically, by measuring gene frequencies or phenotypic traits, they are drawn from a priori—geometrical

or mathematical—considerations of what generates biological forms. As documented in detail by McGhee (2007), morphospaces have been generated for a variety of organisms, from invertebrates to plants, and they have been put to use by comparing actually existing forms with theoretically possible ones that are either extinct or that for some reason have never evolved. Which brings us to the question of what biologists actually do with the various types of landscapes.

3.3 What are landscapes for?

When it comes to asking what the metaphor of landscapes in biology is for we need to begin by distinguishing between the visual metaphor, which is necessarily low-dimensional, and the general *idea* that evolution takes place in some sort of hyperdimensional space. Remember that Wright introduced the metaphor because his advisor suggested that a biological audience at a conference would be more receptive toward diagrams than toward a series of equations. But of course the diagrams are simply not necessary for the equations to do their work. More to the point, the recent papers by Gavrillets and his collaborators, mentioned previously, have shown in a rather dramatic fashion that the original (mathematical) models were too simple and that the accompanying visual metaphor is therefore not just incomplete, but highly misleading. Gavrillets keeps talking about landscapes of sorts, but nothing hinges on the choice of that particular metaphor as far as the results of his calculations are concerned—indeed, arguably we should be using the more imagery-neutral concept of hyperdimensional spaces, so not to deceive ourselves into conjuring up “peaks” and “valleys” that do not actually exist.

In a very important sense Wright’s metaphor of what we have been calling fitness landscapes was meant to have purely heuristic value, to aid biologists to think in general terms about how evolution takes place, not to actually provide a rigorous analysis of or predictions about the evolutionary process (it was for the math to do that work). Seen from this perspective, fitness landscapes have been problematic for decades, generating research aimed at solving problems—like the “peak shift”

one (Whitlock et al. 1995)—that do not actually exist as formulated, or that at the very least take a dramatically different form, in more realistic hyperdimensional “landscapes.” Even when (relatively) low-dimensional scenarios actually apply, as in the cases of RNA and protein functions briefly discussed earlier, the work is done by intensive computer modeling, not by the metaphor, visual or otherwise.

The peak shift problem started captivating researchers’ imagination soon after Wright’s original paper, and consists in explaining how natural selection could move a population off a local fitness peak. The landscape metaphor seems to make it obvious that there is a problem, and that it is a significant one, because of course natural selection could not force a population down a peak to cross an adaptive valley in order to then climb up a nearby (presumably higher) peak. That would amount to thinking of natural selection as a teleological process, something that would gratify creationists of all stripes, but is clearly not a viable solution within the naturalistic framework of science (that said, there are situations described in population genetics theory where natural selection does not always increase fitness, as in several scenarios involving frequency dependent selection (Hartl and Clark 2006)).

Wright famously proposed his shifting balance theory of interdemetic selection as an alternative mechanism to explain peak shifts (Wright 1982): genetic drift would move small populations off-peak, and natural selection would then push some of them up a different peak. This particular solution to the problem has proven theoretically unlikely, and it is of course very difficult to test empirically (Coyne et al. 1997). Several other answers to the peak shift problem have been explored, including the (rather obvious) observation that peaks are not stable in time (i.e. they themselves move), or that phenotypic plasticity and learning may help populations make a local “jump” from one peak to another. It is also interesting to observe that of course there is no reason to think that natural selection *has* to provide a way to shift between peaks, since it is a satisficing, not an optimizing, process. Getting stuck on a local peak and eventually going extinct is the fate shared by an overwhelming

majority of populations and species (van Valen 1973). The current status of this particular problem, as I said, is that Gavrilets's work has shown that there simply aren't any such things as peaks and valleys in hyperdimensional genotypic spaces, but rather large areas of quasi-neutrality (where therefore there is ample room for drift, in that sense at least indirectly vindicating Wright's original intuition of the importance of stochastic events), punctuated by occasional fitness holes and "multidimensional bypasses" (Gavrilets 1997), i.e. connections between distant areas of the hyperspace that can be exploited by natural selection to "jump," though the term now means nothing like what the classical literature on peak shift refers to.

Similar considerations to the ones made in the case of fitness landscapes apply to Adaptive Landscapes, the phenotype-based version of the metaphor introduced by Simpson—albeit with some caveats. Again, the few actual visual examples of such landscapes to be found in the literature have heuristic value only, though at least they are potentially less misleading than Wright-type fitness landscapes simply because low dimensionality is a more realistic situation when we are considering specific aspects of the phenotype. (The phenotype *tout court* of course is a high-dimensionality object, but biologists are rarely interested in that sort of phenotypic analyses, focusing instead either on a small sample of characters, or on a particular aspect of the phenotype—such as skull shape or leaf traits—that can be studied via a small number of variables.) As we have seen, Simpson's classic example concerned the evolution of equids, the horse family, and the corresponding landscape first appeared in his *The Major Features of Evolution* (Simpson 1953). That diagram makes for a great story, which could even be true (and is certainly consistent with the paleontological data), but the whole point of the diagram is to capture the reader's imagination, not to present empirical data or provide testable hypotheses about the observed morphological shifts and alleged ecological context.

Things are somewhat different for (Lande–Arnold; see Chapters 7 and 9) fitness surfaces, because studies quantifying selection coefficients in natural populations are common and can fairly be thought of as statistical analyses of multidimen-

sional fitness (but strictly speaking not adaptive) surfaces. Tellingly, though, what does the work is not the occasional graph of a partial surface (or its often hard to interpret multivariate rendition) but, instead, a tabular output from Lande–Arnold style multiple regression analyses. What understanding we do have of fitness surfaces comes from the actual statistics and our ability to make sense of them (and of their limitations (Mitchell-Olds and Shaw 1987; Pigliucci 2006)), not from visualizations of biologically interpretable surfaces.

The situation is significantly better, I suggest, in the case of the fourth type of landscape: Raup-style morphospaces. McGhee (2007) discusses several fascinating examples, but I will focus here on work done by Raup himself, with crucial follow-up by one of his graduate students, John Chamberlain. It is a study of potential ammonoid forms that puts the actual (i.e. not just heuristic) usefulness of morphospaces in stark contrast with the cases of fitness and Adaptive Landscapes/surfaces discussed so far.

Raup (1967) explored a mathematical-geometrical space of ammonoid forms defined by two variables: W , the rate of expansion of the whorl of the shell; and D , the distance between the aperture of the shell and the coiling axis. As McGhee shows in his detailed discussion of this example, Raup arrived at two simple equations that can be used to generate pretty much any shell morphology that could potentially count as "ammonoid-like," including shells that—as far as we know—have never actually evolved in any ammonoid lineage. Raup then moved from theory to empirical data by plotting the frequency distribution of 405 actual ammonoid species in W/D space and immediately discovered two interesting things: first, the distribution had an obvious peak around $0.3 < D < 0.4$ and $W \sim 2$. Remember that this kind of peak is not a direct measure of fitness or adaptation, it is simply a reflection of the actual occurrence of certain forms rather than others. Second, the entire distribution of ammonoid forms was bounded by the $W = 1/D$ hyperbola, meaning that few if any species crossed that boundary on the morphospace. The reason for this was immediately obvious: the $1/D$ line represents the limit in morphospace where whorls still overlap with one another. This

means that for some reason very few ammonites ever evolved shells in which the whorls did not touch or overlap.

Raup's initial findings were intriguing, but they were lacking a sustained functional analysis that would account for the actual distribution of forms in W/D space. Why one peak, and why located around those particular coordinates? Here is where things become interesting and the morphospace metaphor delivers much more than just heuristic value. John Chamberlain, a student of Raup, carried out experimental work to estimate the drag coefficient of the different types of ammonoid shells. His first result (Chamberlain 1976) clarified why most actual species of ammonoids are found below the $W = 1/D$ hyperbola: shells with whorl overlap have a significantly lower drag coefficient, resulting in more efficiently swimming animals.

However, Chamberlain also found something more intriguing: the experimental data suggested that there should be *two* regions of the W/D morphospace corresponding to shells with maximum swimming efficiency, while Raup's original frequency morphospace detected only one peak. It seemed that for some reason natural selection found one peak, but not the other (Fig. 3.1). Four decades had to pass from Raup's paper for the mystery of the second peak to be cleared up: the addition of 597 new species of ammonoids to the original database showed that indeed the second peak had also been occupied (Fig. 3.2)! Notice that this is a rather spectacular case of confirmed prediction in evolutionary biology, not exactly a common occurrence, particularly in paleontology.

Let me briefly go over a second example—from the same line of inquiry—of how practically (as opposed to simply heuristically) useful morphospaces can be. Fig. 3.3 again shows a W/D space, this time occupied by two different groups of animals with similar morphology and ecology. The top diagram plots the frequency distributions of Cretaceous ammonoids (the large area on the right of the figure) and of nautilids of the same period (the narrow area on the top-left section of the morphospace). Ammonoids, but not nautilids, went extinct at the end of the Cretaceous, a fact reflected by the lower graph, plotting the distribution of Cenozoic nautilids in the same morphospace. Two

things need to be noted about this second graph: first, the nautilids shifted their major peak in a position previously occupied by ammonoids. This may represent a nice example of competitive exclusion that got released by extinction. Second, why have the nautilids—which are structurally and developmentally similar to ammonoids—not expanded to occupy the full morphospace left empty by the demise of the ammonoids? We do not currently have a satisfactory answer to that question. It seems unlikely that natural selection hasn't had the time to explore the empty morphospace (after all, the ammonoids went extinct 65 million years ago). Given the similarity in architecture and development between the two types of organisms, it also seems unlikely—though certainly not impossible—that a developmental constraint played a role, and we know that the empty space can be colonized, since that's where the extinct ammonoids were to begin with. Finally, considering that there doesn't seem to be competition by any other ammonoid-like group, we are left with the possibility of some sort of genetic constraint, a hypothesis that is however difficult to test given the current status of nautilid genomics.

3.4 What to do with the landscapes metaphor(s)

It is time to put together our thoughts about what the various landscape metaphors are supposed to accomplish on one hand, and what they have so far actually accomplished on the other hand, so to arrive at some conclusion concerning whether any of the four variants of the metaphor examined in this essay is actually useful to the biological scientific community.

Wright-style fitness landscapes are supposed to help biologists think through how the genetic makeup of populations changes in response to evolutionary mechanisms, chiefly natural selection (one can also visualize drift, but not the other classical mechanisms, such as mutation, assortative mating, and migration). The metaphor was confused from the beginning, sometimes referring to individual genotypes and at other times to populations; it has historically been used in the low-dimensional version (typically with two "genetic" dimensions

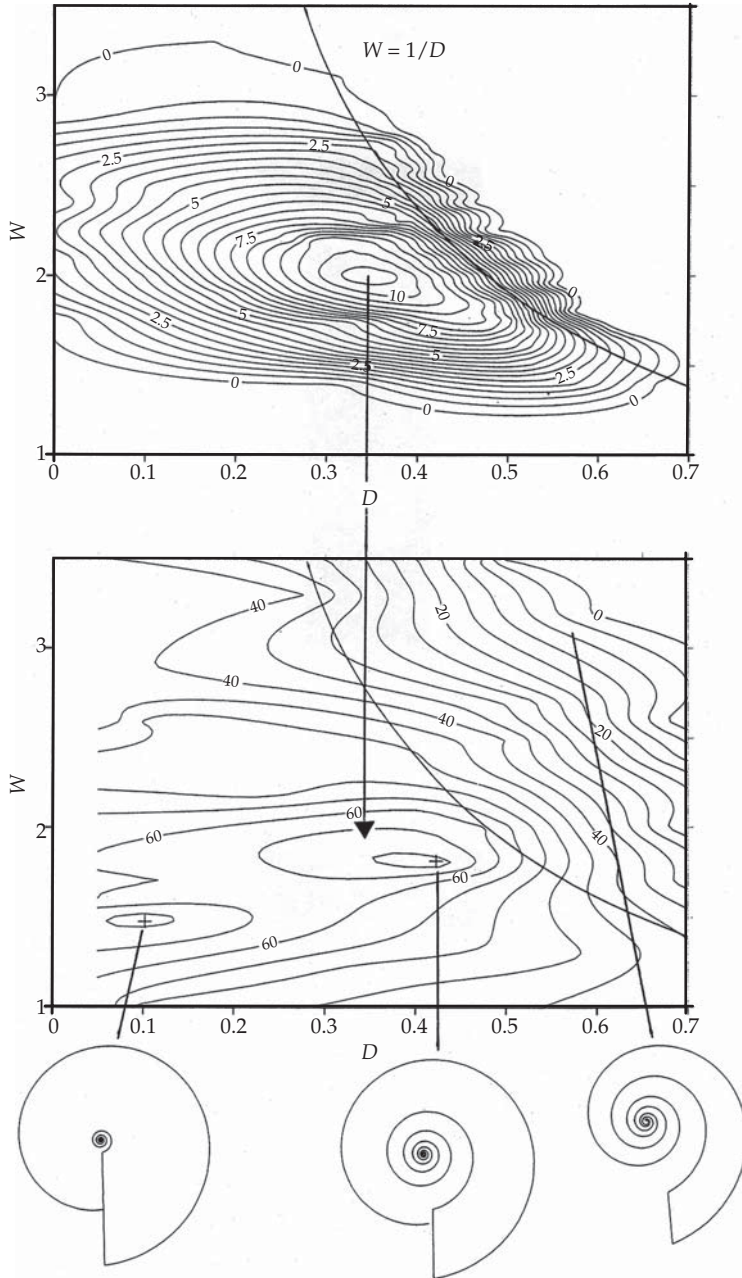


Figure 3.1 A frequency plot and functional analysis of the W/D morphospace of ammonoid shells, following work by Raup and Chamberlain. The upper graph is based on Raup's (1967) original paper, and shows that of the 405 species of ammonoids known at that time all of them could be found below the $W = 1/D$ hyperbole—in agreement with the fact that shells found within that parameter space have higher swimming efficiency. Notice the one frequency peak around the $0.3 < D < 0.4$ and $W \sim 2$ coordinates. The lower graph is based on Chamberlain's (1981) data, and shows two adaptive peaks in terms of swimming efficiency. One peak corresponds to the actual ammonoid peak shown in the top graph, but the second one seemed to indicate that natural selection had somehow "missed" a second W/D combination that maximizes swimming efficiency. (Graphs from McGhee (2007), reproduced with permission from Cambridge University Press.)

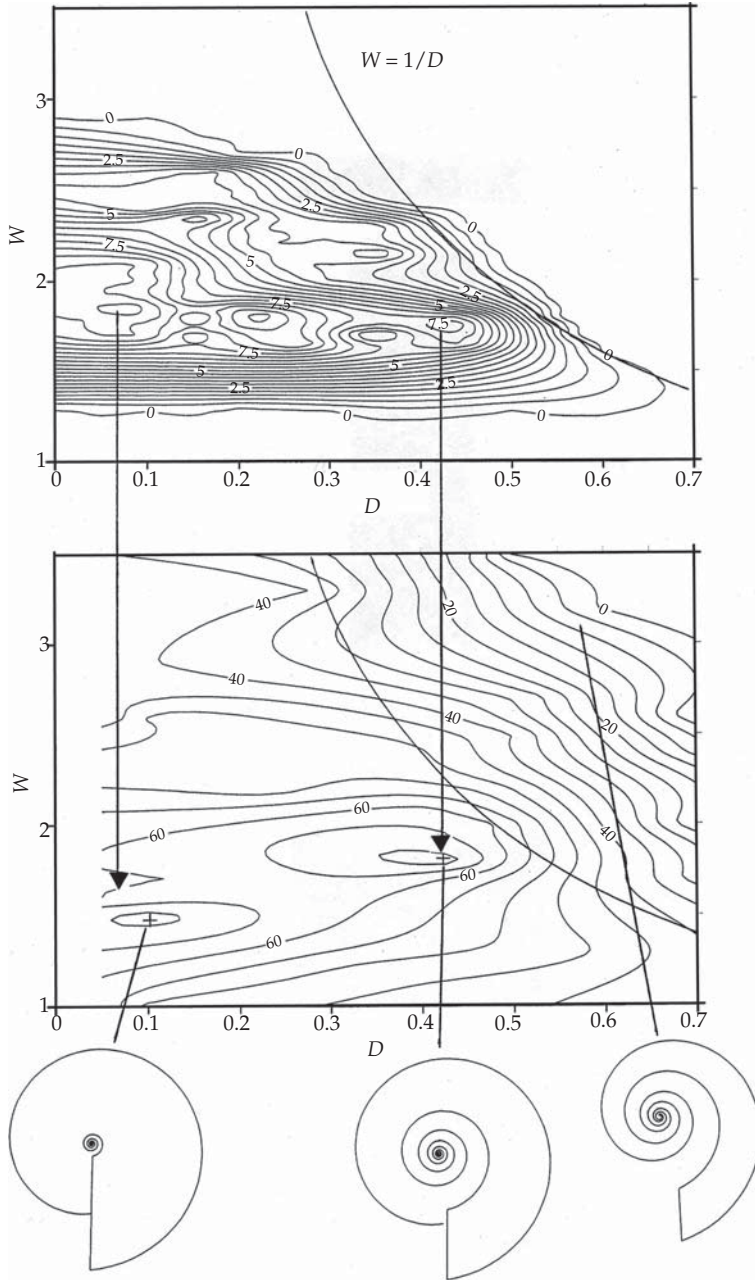


Figure 3.2 The same W/D morphospace originally studied by Raup and Chamberlain (Fig. 3.1), now augmented with an additional 597 newly discovered species of ammonoids. The frequency distribution of actual forms (top), showing two peaks, now nicely corresponds with the two predicted adaptive peaks based on drag coefficients (or its reverse, swimming efficiency, bottom). (Graphs from McGhee (2007), reproduced with permission from Cambridge University Press.)

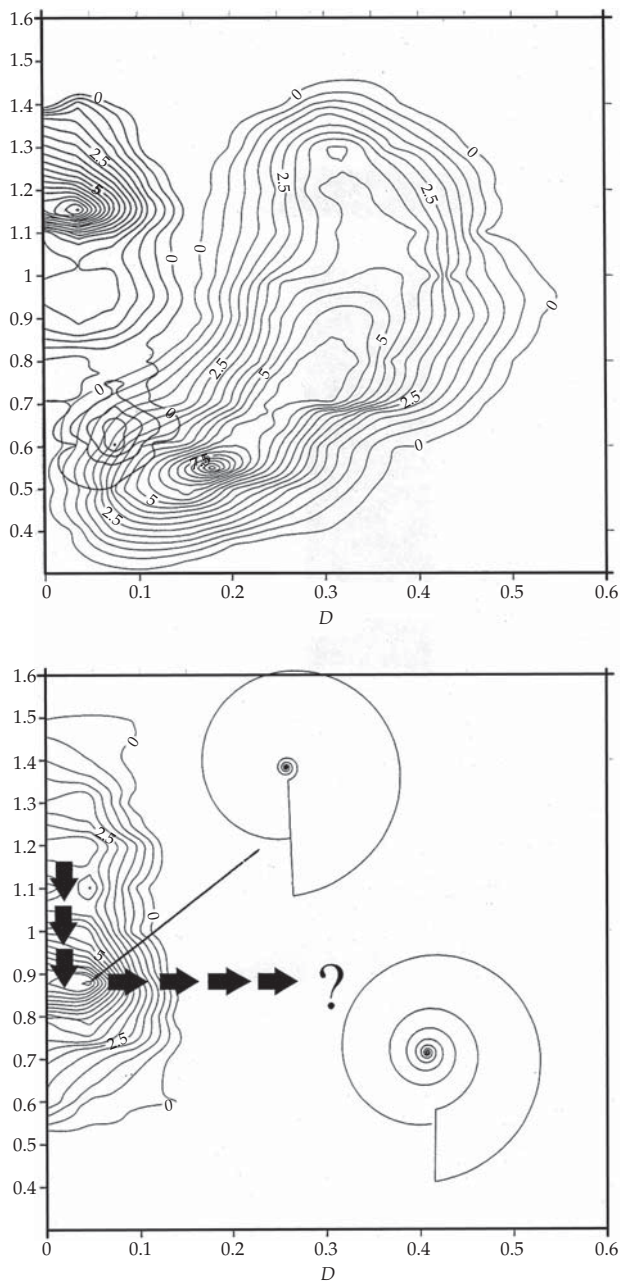


Figure 3.3 Two more morphospaces defined by W/D : the top graph shows the frequency distribution of Cretaceous ammonoids (large area on the center-right) and nautilids (smaller area on the upper left). The bottom graph shows what happened after the extinction of the ammonoids: the nautilids moved their main peak down, but failed to expand toward the right of the morphospace. Possible reasons are discussed in the main text. (Graphs from McGhee (2007), reproduced with permission from Cambridge University Press.)

and one fitness dimension), while recent work clearly shows that whatever intuitions one derives from low-dimensionality fitness landscapes they are likely to prove profoundly misleading. Indeed, an argument can be made that entire research programs, such as the search for mechanisms causing “peak shifts,” have been informed by a faulty assumption, since more realistic hyperdimensional genotypic spaces simply do not have anything that resembles peaks and valleys. It seems like the rational thing to do in this case would in fact be to follow Kaplan’s (2008) advice, abandon the metaphor altogether and simply embrace directly the results of formal modeling—as both the cases of Gavrilets’ “holey” spaces and the research on the evolution of RNA and protein function elegantly illustrate. Wright may have needed to soften his math with pictures in the 1930s, but surely modern biologists ought to be able to take on the full force of the mathematical theory of evolution.

Simpson-style Adaptive Landscapes also aimed from the beginning at a heuristic value, as demonstrated by the fact that there are few examples in the literature that are not hypothetical. However, it is arguable that these landscapes are in fact less misleading than fitness landscapes, since their low dimensionality reflects the real fact that often biologists are interested in selected aspects of an organism’s phenotype, and rarely consider simultaneously hundreds or thousands of characters (again, unlike the genetic scenarios). Still, the point of Adaptive Landscapes is to study adaptation and its macroevolutionary consequences in terms of speciation and lineage divergence, which means that far more than Simpson-style suggestive pictures are necessary. Serious studies of Adaptive Landscapes need to integrate historical records (via paleontology and/or cladistics), functional ecology, as well as morphology—a splendid area for fertile interdisciplinary work, much of which remains to be done.

Lande–Arnold style multiple regression analyses of natural selection—and the graphical rendition of the underlying fitness surfaces—has been well underway, and arguably represents the bulk of the empirical work inspired by landscape metaphors. Yet Lande–Arnold selection analyses themselves are characterized by many well-known issues, such

as problems with multicollinearity of different traits, the problem of the effect of “missing” (i.e. unmeasured) traits, the assumption of linearity of the statistical models, the dearth of spatially and temporally replicated studies with sufficient sample sizes, etc. (Mitchell-Olds and Shaw 1987; Kingsolver et al. 2001; Pigliucci 2006; Siepielski et al. 2009). None of this affects the conclusion that fitness surfaces can be rigorously quantified, though their visualization must be left largely to ineffective multivariate compound variables.

Where both Simpson-type Adaptive Landscapes and Lande–Arnold type fitness surfaces do not deliver is when we assume that they are phenotypic versions of Wright-type landscapes, as Simpson himself surely did. This simply cannot be the case because of the high complexity and non-linearity of the genotype → phenotype mapping function, as argued elsewhere (Alberch 1991; Pigliucci 2010). This is a problem insofar as we are interested in an evolutionary theory that provides us not just with an account of genetic change (as given by population and quantitative genetics), but also with accounts of phenotypic change and of how, precisely, the two are connected.

We finally get to Raup-style morphospaces, which I think are not just heuristically useful (to visualize the range of possible organismal forms within particular aspects of the phenotype), but actually have a nice if small record of generating new understanding, as well as testable hypotheses, in biological research, despite still being somewhat of a backwater topic in need of further attention. As philosopher of science James Maclaurin (2003) aptly put it: “Theoretical morphology might allow us to sort life into the actual, the non-actual and the impossible,” a research agenda splendidly illustrated by the examples of ammonoids and nautilids shell shapes discussed above and that can be traced back to general attempts at theorizing about biological form that even predate an explicitly evolutionary approach (Thompson 1917). However, a survey of the literature on morphospaces (McGhee 2007) shows that most of the available examples are restricted to a small range of animal taxa and limited aspects of their phenotype (with some exceptions concerning plants (Niklas 2004)). This could

be due to the fact that comparatively few scientists (mostly drawn from paleontology) have even thought about organizing their research using the framework of morphospaces, because until recently the study of phenotypic evolution had taken a backseat in biology (Schlichting and Pigliucci 1998), or because there are actual conceptual issues to be dealt with that may limit the general applicability of morphospaces across living organisms and types of characters. Only further research will be able to tell.

References

- Alberch, P. (1991). From genes to phenotypes: dynamical systems and evolvability. *Genetica*, 84, 5–11.
- Arnold, S.J., Pfrender, M.E., and Jones, A.G. (2001). The Adaptive Landscape as a conceptual bridge between micro- and macroevolution. *Genetics*, 112–113, 9–32.
- Brigandt, I. (2008). Reductionism in biology. Stanford Encyclopedia of Philosophy, <http://plato.stanford.edu/entries/reduction-biology/>
- Chamberlain, J.A. (1976). Flow patterns and drag coefficients of cephalopod shells. *Paleontology*, 19, 539–563.
- Chamberlain, J.A. (1981). Hydromechanical design of fossil cephalopods. Systematics Association, Special Volume 18, 289–336.
- Cowperthwaite, M.C. and Meyers, L.A. (2007). How mutational networks shape evolution: lessons from RNA models. *Annual Review of Ecology Evolution and Systematics*, 38, 203–230.
- Coyne, J.A., Barton, N.H., and Turelli, M. (1997). A critique of Sewall Wright's shifting balance theory of evolution. *Evolution*, 51, 643–671.
- Dobzhansky, T. (1970). *Genetics of the Evolutionary Process*. Columbia University Press, New York.
- Estes, S. and Arnold, S.J. (2007). Resolving the paradox of stasis: models with stabilizing selection explain evolutionary divergence on all timescales. *American Naturalist*, 169, 227–244.
- Futuyma, D. (2006). *Evolutionary Biology*. Sinauer, Sunderland, MA.
- Gavrilets, S. (1997). Evolution and speciation on holey adaptive landscapes. *Trends in Ecology & Evolution*, 12, 307–312.
- Gavrilets, S. (2003). Evolution and speciation in a hyperspace: the roles of neutrality, selection, mutation, and random drift. In J.P. Crutchfield and P. Schuster (eds.) *Evolutionary Dynamics: Exploring the Interplay of Selection, Accident, Neutrality, and Function*. Oxford University Press, Oxford, pp. 135–162.
- Hartl, D.L. and Clark, A.G. (2006). *Principles of Population Genetics*. Sinauer, Sunderland, MA.
- Kaplan, J. (2008). The end of the Adaptive Landscape metaphor? *Biology and Philosophy*, 23, 625–638.
- Kaplan, J. (2009). The paradox of stasis and the nature of explanations in evolutionary biology. *Philosophy of Science*, 76, 797–808.
- Kingsolver J.G., Hoekstra, H.E., Hoekstra, J.M., Berrigan, D., Vignieri, S.N., Hill, C.E., et al. (2001). The strength of phenotypic selection in natural populations. *American Naturalist*, 157, 245–261.
- Lande, R. and Arnold, S.J. (1983). The measurement of selection on correlated characters. *Evolution*, 37, 1210–1226.
- Maclaurin, J. (2003). The good, the bad and the impossible. *Biology and Philosophy*, 18, 463–476.
- Mayr, E. and Provine, W. (1980). *The Evolutionary Synthesis: Perspectives on the Unification of Biology*. Harvard University Press, Cambridge, MA.
- McGhee, G.R. (2007). *The Geometry of Evolution: Adaptive Landscapes and Theoretical Morphospaces*. Cambridge University Press, Cambridge.
- Mitchell-Olds, T. and Shaw, R.G. (1987). Regression analyses of natural selection: statistical inference and biological interpretation. *Evolution*, 41, 1149–1161.
- Niklas, K.J. (2004). Computer models of early land plant evolution. *Annual Review of Earth and Planetary Sciences*, 32, 47–66.
- Pigliucci, M. (2006). Genetic variance covariance matrices: a critique of the evolutionary quantitative genetics research program. *Biology and Philosophy*, 21, 1–23.
- Pigliucci, M. (2008). Sewall Wright's adaptive landscapes: 1932 vs. 1988. *Biology and Philosophy*, 23, 591–603.
- Pigliucci, M. (2010). Genotype-phenotype mapping and the end of the “genes as blueprint” metaphor. *Philosophical Transactions of the Royal Society*, B365, 557–566.
- Plutynski, A. (2008). The rise and fall of the Adaptive Landscape? *Biology and Philosophy*, 23, 605–623.
- Raup, D.M. (1966). Geometric analysis of shell coiling: general problems. *Journal of Paleontology*, 40, 1178–1190.
- Raup, D.M. (1967). Geometric analysis of shell coiling: coiling in ammonoids. *Journal of Paleontology*, 41, 43–65.
- Schlichting, C.D. and Pigliucci, M. (1998). Phenotypic Evolution: A Reaction Norm Perspective. Sinauer, Sunderland, MA.

- Siepielski, A.M., DiBattista, J.D., and Carlson, S.M. (2009). It's about time: the temporal dynamics of phenotypic selection in the wild. *Ecology Letters*, 12, 1261–1276.
- Simpson, G.G. (1944). *Tempo and Mode in Evolution*. Columbia University Press, New York.
- Simpson, G.G. (1953). *The Major Features of Evolution*. Columbia University Press, New York.
- Thompson, d'A. W. (1917). *On Growth and Form*. Cambridge University Press, Cambridge.
- van Valen, L. (1973). A new evolutionary law. *Evolutionary Theory*, 1, 1–30.
- Whitlock, M.C., Phillips, P.C., Moore, F.B.G., and Tonsor, J. (1995). Multiple fitness peaks and epistasis. *Annual Review of Ecology and Systematics*, 26, 601–629.
- Wright, S. (1931). Evolution in Mendelian populations. *Genetics*, 16, 97–159.
- Wright, S. (1982). The shifting balance theory and macroevolution. *Annual Review of Genetics*, 16, 1–20.
- Wroe, R., Chan, H.S., and Bornberg-Bauer, E. (2007). A structural model of latent evolutionary potentials underlying neutral networks in proteins. *HFSP Journal*, 1, 79–87.