# 5
The Topology of the Possible

Walter Fontana

The problem of change

The classical framework of evolutionary change is based on two aspects of biological organization: phenotype and genotype. The notion of phenotype refers to the physical, organizational and behavioral expression of an organism during its lifetime. It emphasizes the systemic nature of biological organizations. The notion of genotype refers to a heritable repository of information that instructs the production of molecules whose interactions, in conjunction with the environment, generate and maintain the phenotype.

In its simplest manifestation, evolution is driven by the selection of phenotypes, which causes the amplification of their underlying genotypes, and the production of novel phenotypes through genetic mutations (Figure 5.1, top). These two factors of evolutionary change are the focal points of distinct research agendas. Selection is a dynamical phenomenon that arises spontaneously in constrained populations of autocatalytically reproducing entities. The research agenda focused on selection aims at characterizing the conditions under which a phenotypic innovation can, once generated, invade an existing population. The classical fields of inquiry concerned with selection are population genetics and ecology. The main variables are frequencies of genes or species representatives whose change is typically described by a nonlinear dynamical system. This description is not concerned with a mechanistic understanding of how change occurs in the first place nor with how the internal architecture of biological entities arises and how it influences their capacity to vary. Rather, evolving entities are treated as black-boxes whose variation or 'innovation' is generated by some stochastic process with simple characteristics in an isotropic phenotype space, typically motivated by the need for mathematical tractability.

The question of how a phenotypic innovation arises in the first place has, so far, been the concern of a different research track. The heritable modification of a phenotype usually does not involve a direct intervention at the phenotypic level, but proceeds indirectly through change at the genetic level
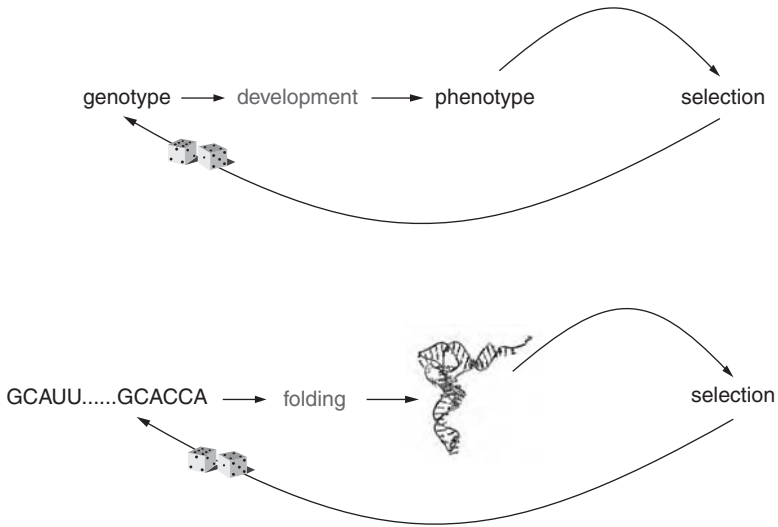
Figure 5.1 The folding of RNA sequences into shapes as a proxy of a genotype–phenotype map. Mutations occur at the genetic level. Their consequences at the phenotypic level are mediated by development, the suite of processes by which phenotype is constructed from genotype. RNA folding is a transparent and tractable model that captures this indirection of innovation within a single molecule. The RNA folding map is characterized by a number of remarkable statistical regularities with profound evolutionary consequences. These regularities may generalize to more complex forms of development.

(Figure 5.1). This forces the processes that link genotype to phenotype into the picture. In biology, these processes are known as development. Evolutionary trajectories, or histories, depend on development, because development mediates the phenotypic effects of genetic mutations and thus the accessibility of one phenotype from another. Aside from constraining and promoting variation, the mechanisms of development are themselves subject to evolution, creating a feedback between evolution and development.

In the early days of the neo-Darwinian school of evolutionary thought, insufficient knowledge of developmental processes justified ignoring the relationship between genotype and phenotype. This pragmatic approach produced a powerful conceptual and formal framework that has been exported – with mixed success – to fields outside biology, particularly 'evolutionary' flavors of economics and the social sciences. The problem remains that important phenomena of phenotypic evolution do not result naturally from a purely adaptationist framework that assumes phenotypes to be completely malleable by natural selection. These phenomena comprise the punctuated mode (the sudden nature) of evolutionary change (Eldredge and Gould, 1972), constraints to variation (Maynard-Smith *et al.*, 1985), the origin of

novelties (Müller and Wagner, 1991), directionality in evolution, path-dependency ('frozen accidents') in organismal structure, and phenotypic stability over evolutionary time, such as homology (for a discussion see Gould, 2002).

The origin of the problem is that the arrival of a new phenotype must necessarily precede its survival and spread in a population through selection. While selection is clearly an important driving force of evolution, the dynamics of selection doesn't teach us much about how evolutionary innovations arise in the first place. To say that a mutation is advantageous, means that it generated a phenotype favored by selection, but reveals nothing about why or how that mutation could innovate the phenotype. A model of the genotype–phenotype relation (development) is needed to illuminate how genetic change maps into phenotypic change. In this chapter, I informally discuss a simple molecular instance of such a relation that is based on the shapes of RNA sequences (Fontana and Schuster, 1998a,b; Stadler *et al.*, 2001), see Figure 5.1, bottom. RNA is an extreme case, because genotype (sequence) and phenotype (folded sequence or shape) are different aspects of the same molecule, as illustrated in Figure 5.2. The sequence of an RNA molecule functions as a genotype, since an RNA sequence can be directly copied (or replicated) by suitable enzymes. It is the sequence that is copied, not the shape of the molecule. An RNA sequence acquires a shape through a process in which the sequence folds into a spatial structure guided by many competing interactions between the chemical building blocks (the 'letters of the RNA alphabet') that constitute the sequence. Like a phenotype, the shape conveys biochemical behavior to an RNA molecule and is therefore a target of selection. The mapping from RNA sequences to shapes constitutes the perhaps simplest biophysically grounded example of a mapping from genotypes to phenotypes that is tractable both theoretically and experimentally.

The RNA model allows to study an important question: given a particular 'physics' (that is, a system of causation, which I shall simply call a 'mechanism,' ignoring post-modern worries in other disciplines) that generates phenotype from genotype, what are the statistical characteristics of the relation between genotype and phenotype and how do they affect the dynamics of evolutionary change? In other words, what can we say statistically about how phenotype changes with genotype? It is important to keep in mind the limitations of the model. While RNA is an instance of a genotype–phenotype relation, it is not a model of organismal development. First, the regulatory networks of gene expression and signal transduction that coordinate the production of phenotype (for an overview see Carroll *et al.*, 2001) have no analogue in RNA folding. Second, the molecular processes underlying organismal development are themselves genetically influenced and thus subject to evolution. In fact, major innovations in the history of life are associated with the emergence of new developmental mechanisms (Kirschner and Gerhart, 1998). In contrast, the folding of an RNA sequence into a shape is essentially governed by physico-chemical principles that lie
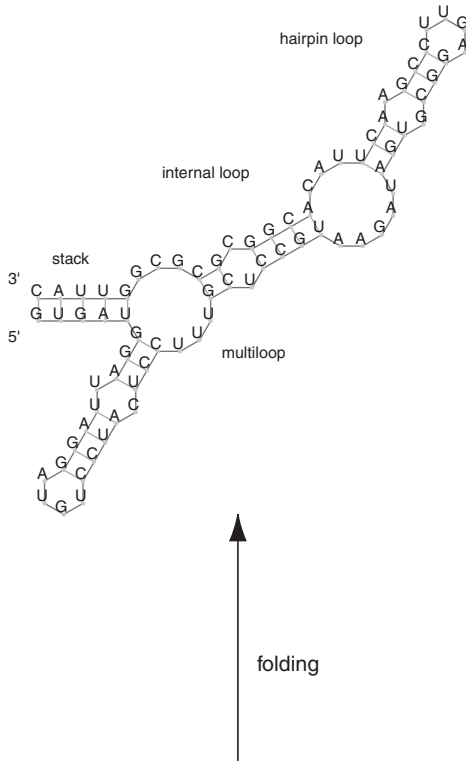
Figure 5.2   RNA shape*

* At the level of resolution considered here, RNA shape is a graph of contacts between building blocks (nucleotides) at positions $i = 1, \ldots, n$ along the sequence. Position 1 is the 5′-end. The graph has two types of edges: the backbone connecting nucleotide i with nucleotide $i + 1$ and hydrogen bonded base pairings between positions.

beyond the control of the sequence. This said, I believe that the RNA model is relevant, because it offers *perspectives* that are potentially generalizable to other, far more complex, situations.

RNA shape as a systemic phenotype

Figure 5.2 sketches all one needs to know about RNA folding for the present discussion. At one level, we are given a sequence of fixed length over an alphabet of four letters: A, U, G, and C. The letters represent certain molecular building blocks. RNA is chemically a very close relative of DNA. As in DNA, specific pairs of building blocks interact with one another preferentially: A pairs with

U, G pairs with C, and G also pairs with U. We call building blocks that pair with one another 'complementary'. The difference to DNA is that RNA occurs single-stranded. Rather than pairing up with a second complementary strand, as in the DNA double-helix, an RNA sequence folds up on itself by matching complementary segments along the sequence, as shown in Figure 5.2. The notion of structure depicted in Figure 5.2 simply consists in the pattern of pairings between positions along the sequence. The structure formation is driven by changes in (free) energy upon pairing. Stretches of pairs (the 'ladders' or stacks in Figure 5.2) stabilize a structure, while loops destabilize it. Notice that the pairing of two segments necessarily creates a loop. In this abstraction, pairings between positions located in different loops are not allowed. Consequently, the formation of a paired stretch (which is energetically 'good') generates not only a loop (which is energetically 'bad'), but it prevents the positions within that loop from pairing with any positions outside of it. A huge number of structures are possible for any given sequence. Yet, one or a few of these structures balance the trade-offs in an energetically optimal fashion. I shall call the energetically optimal fold of a sequence simply its 'shape'. This shape can be computed by clever and fast algorithms (Waterman and Smith, 1978; Nussinov and Jacobson, 1980; Zuker and Stiegler, 1981; Zuker and Sankoff, 1984) based on empirically measured energy parameters (Turner *et al.*, 1988; Jaeger *et al.*, 1989; He *et al.*, 1991; Walter *et al.*, 1994; Mathews *et al.*, 1999).

The actual shape of a sequence is a three-dimensional structure (see, for example, the bottom of Figure 5.1). Today, this three-dimensional structure cannot be computed directly from the sequence. Although the shape shown in Figure 5.2 is a very crude model, it is not a fiction either. The pairings in an actual structure can be established empirically and it turns out that the three-dimensional structure contains many of the pairings predicted by this more abstract notion of shape.

The pairing rules and their energetics establish a map (or a process) that assigns a shape to each sequence. The relevant point for the present discussion is that a shape cannot be changed directly, but requires modification of the sequence which is then interpreted by the folding process to generate the new shape. This situation mirrors the indirection in transforming phenotypes through genetic mutations (Figure 5.1).

The number of possible sequences of small to moderate length is unimaginably huge. For example, there are 4^{100} (that is, 10^{60}) sequences of length 100 – which is not very long for biological molecules. Our goal can only be to characterize the overall *statistical* regularities of the mapping from sequences to shapes.

Shape in RNA (as for bio-molecules in general) is a systemic property. Changing the 'letter' at a given sequence position alters many pairing possibilities throughout the sequence, possibly tipping the optimal pairing pattern. The effect at the shape level of changing a sequence position is

strongly dependent on the sequence context. In a four-letter alphabet, there are three possible substitutions, or mutations, at each sequence position. I call a position *neutral*, if it allows for at least one mutation that does not alter the shape of the sequence. Accordingly, such a mutation is a neutral mutation. Although a neutral mutation leaves the shape of a sequence unchanged, it has two important contextual effects illustrated in Figure 5.3. The shape shown at the top left of Figure 5.3 remains the same if C is substituted by G at the position labeled *x*. Yet, whether *x* is C or G determines the shape obtained when mutating the position *y* from G to A (Figure 5.3, lower half). This illustrates that the effect of a mutation at position *y* depends on the building block at position *x*. In biology, this dependency is called *epistasis*. I'd like to draw attention to another, more subtle, effect: although the replacement of C by G at *x* does not alter the shape, it alters the number and location of neutral positions (Figure 5.3, top right). Thus, while a neutral mutation does not change the phenotype, it affects the potential for change.
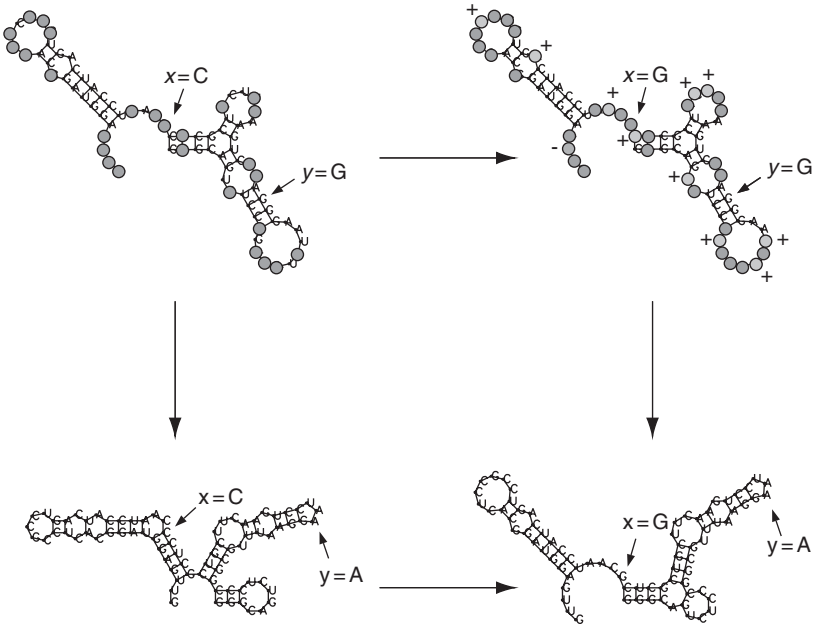

Figure 5.3   Epistasis*

* Bullets indicate a neutral position. In the top left sequence, position *x* is neutral because the substitution of G for C preserves the shape, as shown in the top right sequence. Yet, neutral *positions* do change as a consequence of a neutral mutation. The dark and open bullets in the top right sequence indicate positions that have gained or lost neutrality, respectively. The lower part illustrates the context sensitivity of mutational effects. The neutral mutation from C to G at *x* affects the consequences of swapping A for G at position *y*.

The tendency of a sequence to adopt a different shape upon mutation is called variability and is a prerequisite for its capacity to evolve in response to selective pressures (evolvability). Variability underlies evolvability. Figure 5.3 illustrates that variability (quantified as the number of non-neutral positions) is sequence dependent. Variability can therefore evolve (Wagner, 1996; Wagner and Altenberg, 1996). Note that the change in variability occurs here without the mechanisms of folding (development) themselves changing. This sequence-dependent change of variability is related to Waddington's concept of canalization (Waddington, 1942).

Neutral networks

I focus next on one statistical property of the RNA folding map that I believe to be a 'paradigm of change' of metaphorical import beyond RNA.

In the following it is useful to think of all possible sequences of a given length as related to one another by a notion of distance. The distance between two sequences is the smallest number of mutations required to convert one sequence into the other. The direct neighbors of a sequence are all sequences one mutation away. For example, a sequence of length 100 has 300 neighbors. This results in a very high-dimensional metric space. Despite its abstract nature, this so-called sequence space (Eigen, 1971), see Figure 5.4, is real in the sense that mutations are chemical events interconverting sequences with a certain probability. The probability of going from one
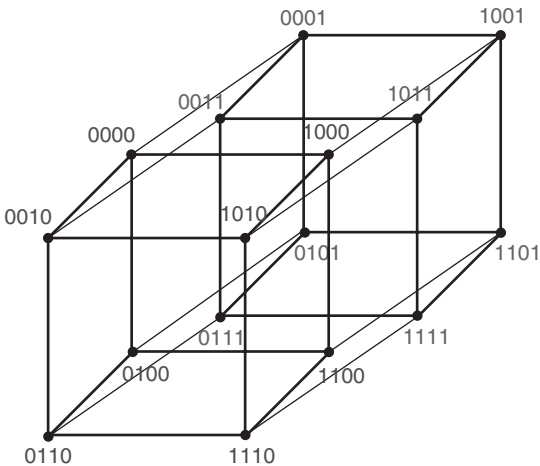

Figure 5.4   Sequence space for sequences of length 4 over the binary alphabet {0,1}*

* The connections correspond to single mutations that change only one position. The picture hints at how the four-dimensional sequence space (a hypercube) is constructed from the three-dimensional space (a cube).

sequence to another in a single event depends on their distance (the number of edges on the smallest path connecting them). The farther apart the sequences, the lower the probability for a direct jump.

Equipped with the notion of a sequence space, consider the sequence and its shape depicted at the top left of Figure 5.3. The positions marked by grey bullets are neutral. At least one mutation at each neutral position leads to a neighboring sequence with the same shape – a neutral neighbor. A given sequence has typically a significant fraction of neutral neighbors at a distance of one or two mutations. The same observation holds for each of these neutral neighbors. (For example, the sequence shown at the top right of Figure 5.3 also has several neutral positions, as indicated by grey bullets.) In this way, we can jump in steps of one or two mutations from sequence to sequence, mapping out a vast network in sequence space on which the phenotype is preserved. For this extensive, mutationally connected, network of sequences, we coined the term *neutral network* (Schuster *et al*., 1994), Figure 5.5A.

Robustness enables change

The possibility of changing the genotype while preserving the phenotype is a manifestation of a certain degree of phenotypic robustness toward genetic mutations. Yet, at the same time, it is a key factor underlying evolvability, the capacity of a system to evolve. It seems paradoxical, but robustness enables change. To understand this, imagine (with the help of Figure 5.5A) a population with phenotype 'star' in an evolutionary situation where phenotype 'triangle' would be advantageous or desirable. Phenotype 'triangle', however, may not be accessible in the vicinity of the population's current location in genotype space, say, somewhere in the northern portion of the 'star'-network of Figure 5.5A. In the popular image of a rugged fitness landscape, the population would be stuck at a local peak, forever waiting for an exceedingly unlikely event to deliver the right combination of several mutations to yield phenotype 'triangle'. Yet, if phenotype 'star' has an associated neutral network in genotype space, the population is not stuck, but can drift on that network into far away regions, vastly improving its chances of encountering the neutral network associated with phenotype 'triangle' (Huynen, 1996; Huynen *et al*., 1996; Fontana and Schuster, 1998ab; Nimwegen *et al*., 1999), see Figure 5.5A. Neutral networks therefore enable phenotypic innovation by permitting the accumulation of neutral (that is, phenotypically inconsequential) mutations. These mutations, however, alter the genetic context, enabling a subsequent mutation to become consequential.

Of course, there is no guidance on a neutral network, only drift. Yet, drift confined to the 'star'-network dramatically increases the chance of eventually encountering the 'triangle'-network. In the absence of neutral networks, an exceedingly unlikely direct jump from an isolated 'star'-location to some far-away 'triangle'-location would be required.
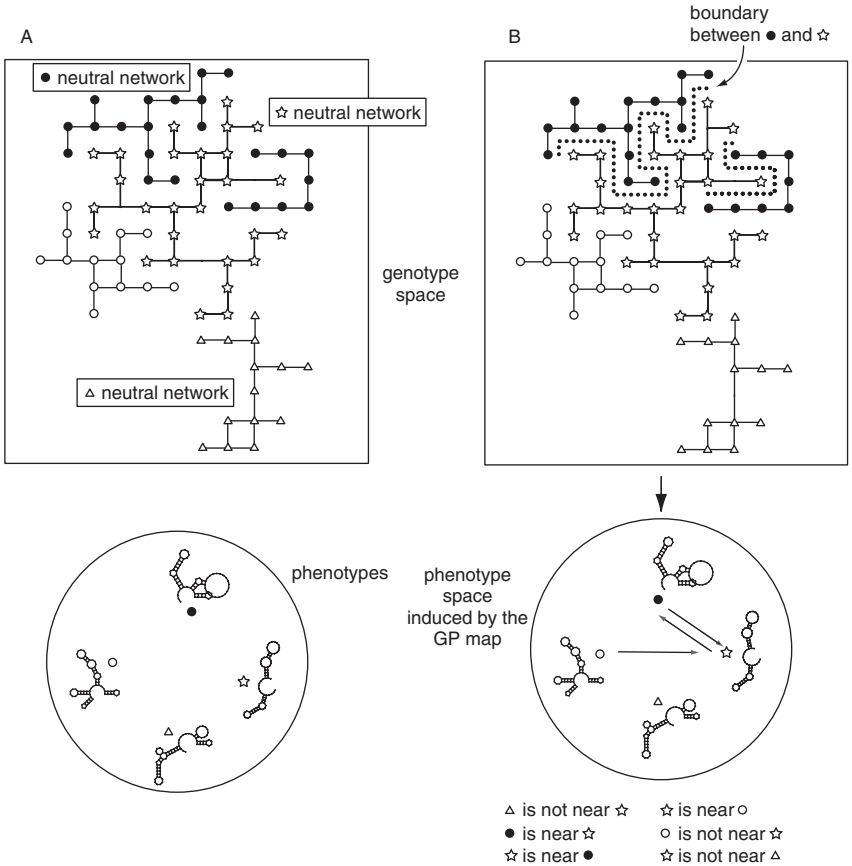
Figure 5.5   Neutral networks and shape space topology.
A: Schematic depiction of neutral networks in sequence space (upper part)*
B: Phenotypic neighborhood as fraction of shared boundary**

* The symbol of a sequence indicates its phenotype (lower part of the figure). A population located in the northern portion of the 'star' phenotype cannot access the 'triangle' phenotype, but it can diffuse on the 'star'-network until it encounters the 'triangle'-network.

** The 'star'-network is near the 'open circle'-network, because a random step out of 'open circle' has a high probability of yielding 'star'. The 'open circle'-network, however, is not near the 'star'-network, because a random step out of 'open circle' has a low probability of yielding 'star'. This effect can result from differently sized networks. In another case, the 'star' and 'triangle'-networks have similar sizes, but they border one another only rarely; 'star' and 'open circle' are not in each other's neighborhood. This is the case for shift transformations like the one shown in Figure 5.6A.

What appears to be a sudden and abrupt change at the phenotypic level becomes the result of several random neutral genetic changes. The existence of neutral paths in RNA sequence space was impressively demonstrated experimentally by Schultes and Bartel (2000). The importance of neutrality

for evolution was long recognized by the Japanese population geneticist Motoo Kimura (Kimura, 1968; Kimura, 1983). Population genetics, however, is concerned with the dynamics of gene frequencies. Neutrality is interesting in that context because it leads to a diffusion dynamics in the *frequency* of a neutral gene within the population. In the present context, I'm concerned with something quite different, namely the diffusion of a whole population over a neutral network in gene space. The next consequence of neutral networks that I wish to emphasize is the *organization of phenotype space*.

The topology of phenotype space

A space is formally a set of elements with a structure imposed by the relationships among the elements. A relation of distance gives rise to a metric space. For example, the distance between two RNA sequences is the number of positions in which they differ. For a metric space to be relevant, the distance must be defined in terms of naturally occurring (or technologically controllable) operations that interconvert elements. This is indeed the case with sequences, where the operations are mutations (or technologically targeted changes in the sequence). But what is a space of phenotypes, if phenotypes cannot be modified directly? Although any number of distance (or similarity) measures between phenotypes can be defined, a phenotype space so-constructed is of no help in understanding *evolutionary* histories, because it relates phenotypes in a manner that fails to take into account the indirection required to change them. Rather, what is needed is a criterion of *accessibility* of one phenotype from another by means of mutations in their underlying genetic representation. Such a notion of accessibility can then be used to define a concept of *neighborhood* that underlies the structure of phenotype space. Such a neighborhood does not rest on a notion distance (Fontana and Schuster, 1998b; Stadler *et al.*, 2001). This runs against common sense, which conceives 'neighborhood' in terms of 'small distance'. I shall not pursue this here, but a branch of mathematics known as point set topology clarifies the sense in which neighborhood is more basic than distance (see any text book on topology).

Recall that a neutral network is the set of all genotypes adopting a particular phenotype. In Figure 5.5B, I define a relation of accessibility between to (RNA) shapes, symbolized by 'star' and 'open circle', in terms of the adjacency of their corresponding neutral networks in genotype space (Fontana and Schuster, 1998a,b). The *boundary* of a neutral network consists of all sequences that are one mutation off the network. The intersection of the neutral network of 'open circle' with the boundary of the neutral network of 'star', relative to the total boundary of 'star's' network, is a measure of the probability that one step off a random point on the 'star'- network there is a sequence folding into 'open circle'. A rough schematic of such intersections is depicted in Figure 5.5B by a dashed line that traces the apposition of the 'star'-network and the 'filled circle'-network.

To fix ideas with a cartoon, think of Europe as sequence space and of a neutral network as a European state (a 'phenotype'), say, France. Now scan the boundary of France and measure the fraction of boundary it has in common with some other state, say Germany. That fraction measures the likelihood of ending up in Germany when making a random step out of France. Doing this for all states bordering France yields a distribution of relative border lengths. This distribution has an abstract average. Now define the neighbors of France as those states whose relative boundary with France is longer than that average. This construction is needed in actual genotype space because a neutral network is a very high-dimensional object that borders to a huge number of other neutral networks. (The example of geographic states fails miserably in this regard.) Most of these boundaries, however, are tiny and their states are not classified as 'neighbors' in this construction. (The actual distribution of relative boundary sizes is a generalized power-law (Fontana and Schuster, 1998b).) For important technical details the reader should consult Stadler *et al.*, 2001. Notice something funny, though. France is near Monaco, because a large fraction of Monaco's boundary is with France. Yet, Monaco is not near France, since France's boundary to Monaco is tiny. Returning to our original picture, a phenotype may be easily accessible from another phenotype by genetic mutations, but the reverse may not be true. A common sense manifestation of this is the observation that it can be very difficult to build structure (few paths), but very easy to destroy it (many paths). I will not provide RNA examples here, since their discussion requires too much terminology that is of little import to the target audience of this chapter. The important message is that accessibility is *not* a distance. A distance relation is always symmetric, while accessibility is not, as illustrated in Figure 5.5B. The RNA phenotype space is not a metric space, but a much weaker and less intuitive space (Stadler *et al.*, 2001).

Continuous and discontinuous change

One more step is needed to put together a full picture. Obviously, Austria is not near France, because Austria has a zero border to France. We can ask, however, whether there is a route from France to Austria such that at each step we always make a transition to a state in the current neighborhood – for these are the easy transitions. For example, we could step from France into Germany (easy) and from Germany into Austria (easy). Roughly speaking, mathematicians call a path continuous if at any given small step it proceeds along the neighborhood relations of the space the path is traversing; in other words, if you don't have to lift the pencil when drawing the path …

We can now ask an important question about the structure of phenotype space: Given any two phenotypes, is there always a continuous ('easy') path connecting them? That is, can we travel from one neutral network to another that is not a neighbor (in the boundary sense defined above), such that at each step (mutation) along the way we always transit to a neighboring network, until we reach our target network? In the RNA case, the answer is

no (Fontana and Schuster, 1998a,b). This means that certain shape transformations are irreducibly difficult to achieve, no matter which route we try. The formal construction I've sketched here, allows us to call such transformations discontinuous in the mathematical sense of the concept. When looked at in detail, these transformations are local changes in shape – nothing fancy – but they require the *simultaneous* rearrangement of several interactions (pairings) among building blocks, because any intermediate resulting from a serial rearrangement would be unstable or physically impossible (Figure 5.6A shows an example).
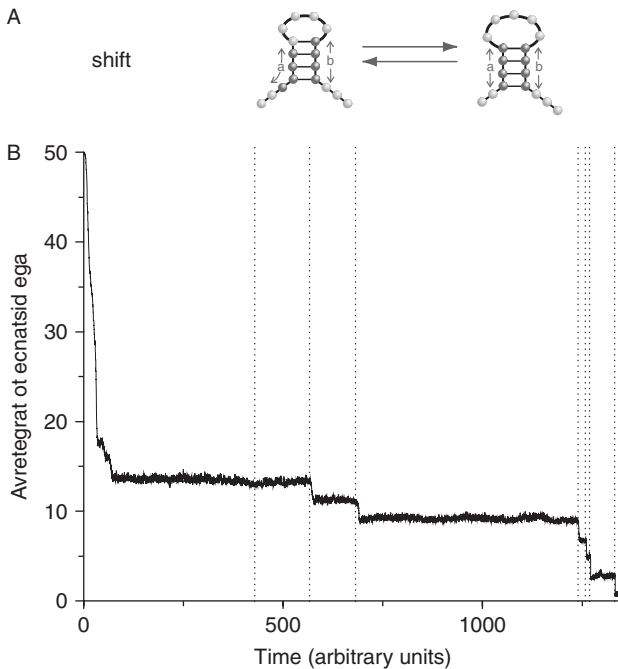

Figure 5.6 A: An example of a discontinuous shape transformation in RNA*
B: Punctuation in evolving RNA populations**

*In this case (a shift) one strand slides against the other in a paired segment. All pairings must slide in one single event, since any partial sliding would create bubbles destabilizing the intermediate. Triggering such a shift by a single mutation requires a special sequence context (Fontana and Schuster, 1998a,b).

**A population of RNA sequences evolves under selection for a specific target shape. The homing in on the target shape shows periods of stasis punctuated by sudden improvements. (Fitness is maximal when the distance to the target shape has become zero.) Yet, the phenotypic discontinuities (marker lines), as revealed by an *ex post* reconstruction of the evolutionary trajectory, are not always congruent with the fitness picture. In this example, the first two jumps in fitness turn out to be continuous in the phenotype topology discussed here and a crucial discontinuous transition is fitness neutral (first marker line).

A few observations deserve emphasis. Analyzing the mapping from genotypes to phenotypes in terms of neutral networks enables a mathematically rigorous definition of continuity (or discontinuity) in evolutionary trajectories. First, this notion of (dis)continuity cross-cuts (dis)similarity. Some transitions between similar shapes are discontinuous (e.g. Figure 5.6A) and some transitions between dissimilar shapes are continuous. Second, the notion of discontinuity defined here is *not* related to sudden jumps in fitness or the discrete nature of the notion of phenotype considered here. The categories of discontinuity are caused by the genotype–phenotype map. They remain the same regardless of the further evaluation of phenotypes in terms of fitness. (Of course, the particular shapes observed at discontinuous transitions in actual evolutionary trajectories will depend on the fitness map. Yet, the nature (type) of these transformations does not depend on the fitness map.) Third, the dynamical signature of this phenotype topology is punctuation (see Figure 5.6B). A population of replicating and mutating sequences under selection drifts on the neutral network of the currently best shape until it encounters a gateway to a network that conveys some advantage or is fitness-neutral. That encounter, however, is not under the control of selection, for selection cannot distinguish between neutral sequences. While similar to the phenomenon of punctuated equilibrium recognized by (Eldredge and Gould, 1972) in the fossil record of species evolution, punctuation as described here occurs in the *absence* of externalities (such as meteorite impact or abrupt climate change in the case of species). Here, punctuation reflects the variational properties of the developmental architecture.

Summary

Development matters

The processes that link genotype to phenotype typically generate a many-to-one mapping, that is, there are considerably fewer phenotypes than genotypes. Yet, this degeneracy also depends on the level of resolution at which we define phenotype. If, in the RNA case, we were to define shape in terms of atomic coordinates, there could be no redundancy; every sequence would have a unique shape and the topic of this paper is mute. I believe, however, that it is precisely the onset of robustness – that is, extended neutral networks – that identifies a meaningful level of resolution for phenotype. The notion of neutrality depends on an appropriate notion of equivalence, which is a notoriously difficult notion to pin down in biological, economic or social contexts.

Robustness enables change

The mathematical and statistical analysis of the mapping from RNA sequences to shapes has produced the concept of a neutral network – a

mutationally connected set of genotypes that map to the same phenotype. I believe that this concept holds for many genotype–phenotype mappings beyond the simple RNA case. Neutral networks are key to change. It is hard to understand how evolution could be successful at all, if developmental processes don't give rise to neutral networks. The popular image of a rugged evolutionary landscape, where populations get stuck at local optima, is certainly incomplete and perhaps even wrong. Populations may be pinned at the phenotypic level, but they constantly change at the genetic level, drifting on neutral networks, thereby dramatically increasing their chances for phenotypic innovation.

Variation is not innovation – the limits of selection

A neutral network has many neighbors (in the boundary sense). A transition from a given network to any of its neighbors corresponds to the acquisition of an 'easy' new phenotype. We have seen, however, that certain phenotypic changes *must* go through border-bottlenecks, involving a transition from one network to another that touches the former only rarely. I have distinguished (with some mathematical motivation) such phenotypic changes as being continuous or discontinuous. I'd like to belabor that point by suggesting that the continuous/discontinuous distinction reflects an intuitive distinction between variation that occurs readily (a better mouse trap) and innovation, that is, a new phenotype that is not readily achievable and that requires long periods of drift. The former variation is typically present in any reproducing population that drifts over a neutral network and explores its boundary by mutation. Such omni-accessible variation is therefore easily available to selection for adaptive responses. In the case of an innovation, however, selection can do nothing to guide a population over a neutral network in the direction of the rare boundary segment that marks an innovation, since selection cannot distinguish between genetic variants of the same phenotype. The occurrence of innovations is not under the control of selection, but, of course, their fate is. The economic literature makes a distinction between invention and innovation. Invention denotes the first occurrence of an idea, while innovation refers to its adoption in practice. From this perspective, what I call 'innovation' would be the economist's notion of invention, since my notion of innovation only reflects the constraints imposed by the genotype–phenotype map, independently of the fixation or rejection of an innovation through selection.

Beyond genotype and phenotype

I have described a paradigm of change in a rather specific molecular and biological setting. Underneath the surface, the notions of genotype and phenotype appear as convenient idealizations even in biology (Griesemer, 2000). They appear hardly meaningful in technological, economic and

social realms. I'd like to conclude, therefore, with a little translation for readers from other disciplines.

In social, economic and technological domains, we rely on systems consisting of many interacting heterogeneous components whose collective action gives rise to ordered system behavior. When we wish to change that behavior, we hit on a simple fact whose consequences I have described in this contribution: Behavior is not a thing, behavior is the property of a thing. It follows that the only way to change behavior is to change the thing that generates it. The change of a property is necessarily indirect.

As an example, consider a computer program. A computer program implements a function. A function is not something that can be touched and altered directly. To alter the function one must alter the program text. The mapping from program to function is what computer scientists call the semantics of the programming language in which the program is written. Substitute for computer program your favorite social, technological, or economic organization, for function the relevant qualitative behavior of that organization and for semantics the dynamical principles that govern the micro-interactions among the parts of the organization.

At a high level of abstraction, the structure of the situation may be represented by a mapping from systems to behaviors of those systems. An input to this mapping is the specification or description of a particular system configuration. The outcome of the mapping is the behavior or function of that system configuration. The mapping represents the dynamical principles that are characteristic for a particular class of systems – computer programs, molecules, electronic circuit diagrams, urban transportation systems, firms in a given industry. In other words, the mapping represents the rules or processes that are *constitutive* for a given class of systems. In social contexts, institutions would be a component of this mapping. The unfolding of this constitutive dynamics in the context of a particular system configuration yields the behavior associated with that configuration.

When we wish to change behaviors of systems, we often have a spatial metaphor in mind, such as going from 'here to there', where 'here' and 'there' are positions in the space of behaviors. But what exactly is the nature of this space? Who brought it to the party? It is a widespread fallacy to assume that the *space* of behaviors is there to begin with. This is a fallacy even when all possible behaviors are known in advance. How does this fallacy arise? When we are given a set of entities of any kind, we almost always can cook up a way of comparing two such entities, thereby producing a definition of similarity (or distance). A measure of similarity makes those entities hang together naturally in a familiar metric space. The fallacy is to believe that a so-constructed space has any significance. It isn't, because that measure of similarity is not based on available real-world operations, since we cannot act on behaviors directly. We only have configuration-editors, we don't have property-editors. Seen from this operational angle, that which

structures the space of behaviors is not the degree of similarity among behaviors but a rather different relation: operational *accessibility* of one behavior from another in terms of system-reconfigurations. This brings the mapping from system configurations to behaviors into the picture. The structure of behavior-space is induced by this mapping. It cannot exist independently of it.

Limits

In the previous sections I have described how this space-structure arises in the context of RNA molecules. With the above paragraphs in mind, substitute 'system' for 'RNA sequence' and 'behavior' for 'RNA shape'. It deserves emphasis, however, that the RNA model is a limiting (and highly idealized) case in which the mapping from sequences (systems) to shapes (behaviors) is exogenous to the system itself. In the game of Go, the rules are not determined by the board configuration; they are given independently. While this limiting case is useful to illustrate the induced character of behavior-space, the most intriguing cases are those in which the mapping itself is endogenously generated by the system. This is the case in all of organismal biology, where genes also code for products that implement development. The endogeneity of the mapping from systems to behaviors is clearly a feature of economic and social systems as well. Nothing prevents the rules of the game (and thus the game itself) from changing. It may be harder to change a rule than to change a game configuration, but this is a question of time scales, not of principle. In the case of an endogenous mapping, the structure of phenotype (behavior) space remains an induced one. However, the concept of space becomes significantly more subtle. Not only can a change in the mapping rearrange the accessibility structure between phenotypes, it can also bring into existence phenotypes that were not possible before. We still lack a formal grip on this radical form of change. In the mid sixties, C. H. Waddington vented his frustration with the mathematical apparatus of theoretical biology: '*The whole real guts of evolution – which is, how do you come to have horses and tigers, and things – is outside the mathematical theory*' (Quoted by (Gould, 2002, p. 584).) The challenge remains of bringing the guts of evolution within the scope of a mathematical theory – that is, of formalizing the emergence of new classes of biological objects.

Acknowledgements

References

Carroll, S. B., J. K. Grenier and S. D. Weatherfee (2001). *From DNA to diversity*. Malden, Massachusetts, Blackwell Science.

Eigen, M. (1971). 'Selforganization of matter and the evolution of biological macromolecules.' *Naturwissenschaften* 58: 465–523.

Eldredge, N. and S. J. Gould (1972). 'Punctuated equilibria: an alternative to phyletic gradualism.' *Models in Paleobiology*. T. J. M. Schopf. San Francisco, CA, Freeman, Cooper & Co.: 82–115.

Fontana, W. and P. Schuster (1998a). 'Continuity in evolution: on the nature of transitions.' *Science* 280: 1451–5.

Fontana, W. and P. Schuster (1998b). 'Shaping space: the possible and the attainable in {RNA} genotype–phenotype mapping.' *J. Theor. Biol.* 194: 491–515.

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

Griesemer, J. R. (2000). 'Reproduction and the reduction of genetics.' *The Concept of the Gene in Development and Evolution: Historical and Epistemological Perspectives*. P. Beurton, R. Falk and H.-J. Rheinberger (eds). Cambridge, MA, Cambridge University Press: 240–85.

He, L., R. Kierzek, Santa Lucia Jr, A. E. Walter and D. H. Turner (1991). 'Nearest-neighbour parameters for {G-U} Mismatches.' *Biochemistry* 30: 11124.

Huynen, M. A. (1996). 'Exploring phenotype space through neutral evolution.' *J. Mol. Evol.* 43: 165–9.

Huynen, M. A., P. F. Stadler and W. Fontana (1996). 'Smoothness within ruggedness: the role of neutrality in adaptation.' *Proc. Natl. Acad. Sci. USA* 93: 397–401.

Jaeger, J. A., D. H. Turner and M. Zuker (1989). 'Improved predictions of secondary structures for {RNA}.' *Proc. Natl. Acad. Sci. USA* 86: 7706–10.

Kimura, M. (1968). 'Evolutionary rate at the molecular level.' *Nature* 217: 624–6.

Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge, UK, Cambridge University Press.

Kirschner, M. and J. Gerhart (1998). 'Evolvability.' *Proc Natl Acad Sci USA* 95(15): 8420–7.

Mathews, D. H., J. Sabina et al. (1999). 'Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure.' *J. Mol. Biol.* 288: 911–40.

Maynard-Smith, J., R. Burian, S. Kauffman, P. Alberch, J. Campbell, B. Goodwin, R. Lande, D. Raup and L. Wolpert (1985). 'Developmental constraints and evolution.' *Quart. Rev. Biol.* 60: 265–87.

Müller, G. B. and G. P. Wagner (1991). 'Novelty in Evolution: Restructuring the Concept.' *Annu. Rev. Ecol. Syst.* 22: 229–56.

Nimwegen, E. van, J. P. Crutchfield and M. Huynen (1999). 'Neutral evolution of mutational robustness.' *Proc. Natl. Acad. Sci. USA* 96: 9716–20.

Nussinov, R. and A. B. Jacobson (1980). 'Fast algorithm for predicting the secondary structure of single-stranded RNA.' *Proc. Natl. Acad. Sci. USA* 77(11): 6309–13.

Schultes, E. A. and D. P. Bartel (2000). 'One sequence, two ribozymes: implications for the emergence of new ribozyme folds.' *Science* 289: 448–52.

Schuster, P., W. Fontana, P. F. Stadler and I. L. Hofacker (1994). 'From sequences to shapes and back: a case study in RNA secondary structures.' *Proc. Roy. Soc. (London) B* 255: 279–84.

Stadler, B. M. R., P. F. Stadler, G. Wagner and W. Fontana (2001). 'The topology of the possible: Formal spaces underlying patterns of evolutionary change.' *J. Theor. Biol.* 213: 241–74.

Turner, D. H., N. Sugimoto and S. M. Freier (1988). 'RNA structure prediction.' *Annual Review of Biophysics and Biophysical Chemistry* 17: 167–92.

Waddington, C. H. (1942). 'Canalization of development and the inheritance of acquired characters.' *Nature* 3811: 563–5.

Wagner, A. (1996). 'Does evolutionary plasticity evolve?' *Evolution* 50(3): 1008–23.

Wagner, G. P. and L. Altenberg (1996). 'Complex adaptations and the evolution of evolvability.' *Evolution* 50: 967–76.

Walter, A. E., D. H. Turner, J. Kim, M. Lyttle, P. Muller, D. Mathews and M. Zuker (1994). 'Coaxial stacking of helices enhances binding of oligoribonucleotides and improves prediction of RNA folding.' *Proc. Natl. Acad. Sci.* 91: 9218–22.

Waterman, M. S. and T. F. Smith (1978). 'RNA secondary structure: A complete mathematical analysis.' *Mathematical Biosciences* 42: 257–66.

Zuker, M. and D. Sankoff (1984). 'RNA secondary structures and their prediction.' *Bull. Math. Biol.* 46(4): 591–621.

Zuker, M. and P. Stiegler (1981). 'Optimal computer folding of larger RNA sequences using thermodynamics and auxiliary information.' *Nucleic Acids Research* 9: 133–48.