How does Complexity Arise in Evolution ?

Nature's recipes for mastering scarcity, abundance, and unpredictability.

BY PETER SCHUSTER

Information in biology has a quality that distinguishes it from information in chemistry and physics. It comes in encoded form and it is processed in a way that is closely related to information technology and computer science. Biological information is essentially stored in genotypes and transferred to future generation through inheritance, and less directly through epigenetic processes. Cellular metabolism is interpreted straighforwardly as information processing. Information is closely related to complexity: more complex things require more information to build and to operate. Any comprehensive understanding of biological phenomena requires an interpretation in evolutionary terms, as Theodosius Dobzhansky [4] pointed out in his famous phrase: "Nothing in biology makes sense except in the light of evolution". Understanding the complexity of biological systems is thus always incomplete if nothing is known about its origin.

It is commonplace to state that complexity has increased in the evolution of the biosphere. Human societies after all are indisputably more complex than animal societies. Animals, plants, and fungi, being multicellular organisms, are more complex than single cells. Eukaryotic cells are more complex than prokaryotic cells, and a cell is more complex than a bag of biopolymer molecules. At the same time the effective biological information, understood as genetic and epigenetic information, increases as the phenotypes become more complex in the above mentioned series from polynucleotide molecules to man. Information and complexity, however, do not seem to have gradually increased during the history of life on earth. Palaeontologists have discovered rather large and abrupt jumps in structural and functional complexity in the fossil record.

Irrespective of this evidence, increasing complexity in evolution is extremely hard to explain and very difficult to model, particularly since it apparently occurs on relatively short time scales. Progress in understanding the major complexity boosts is very slow; a model presented more than ten years ago [11] is hardly different from one that was recently proclaimed by Szathmáry and Maynard Smith [23, 33]. Their recent summary of the state of the art in understanding the major evolutionary transitions attempts to list the various biological data. Here we will be concerned with conceptual aspects of the origin of complexity.

Complexity has different aspects. For example, in [1] Brian Arthur measures it as ecological diversity, in the sense that an ecosystem with more species forming elaborated foodwebs and occupying sophisticated niches is more complex than an ecosystem with only few species. There is also the notion that complexity of construction makes simple things more complex in order to make them fulfil additional constraints (high performance, usability under extreme conditions, safety, etc.). The idea of internal complexity in the sense of (logical) depth [2] has also been proposed.

A major aspect of biological complextity is indeed related to logical depth but its roots can be traced more precisely. One might also call it **hierarchical complexity**, as it originates from integrating collections of objects of one level to entities of the next higher hierarchical level. Genes are combined and intergrated in chromosomes, several prokaryotic cells are united in a eukaryotic cell, cells are combined in multicellular organisms and organized in tissues, solitary individuals are united in societies, and there may have been other events of unification as well. Such unification events correspond to the large jumps in evolution. We shall come back to them later.

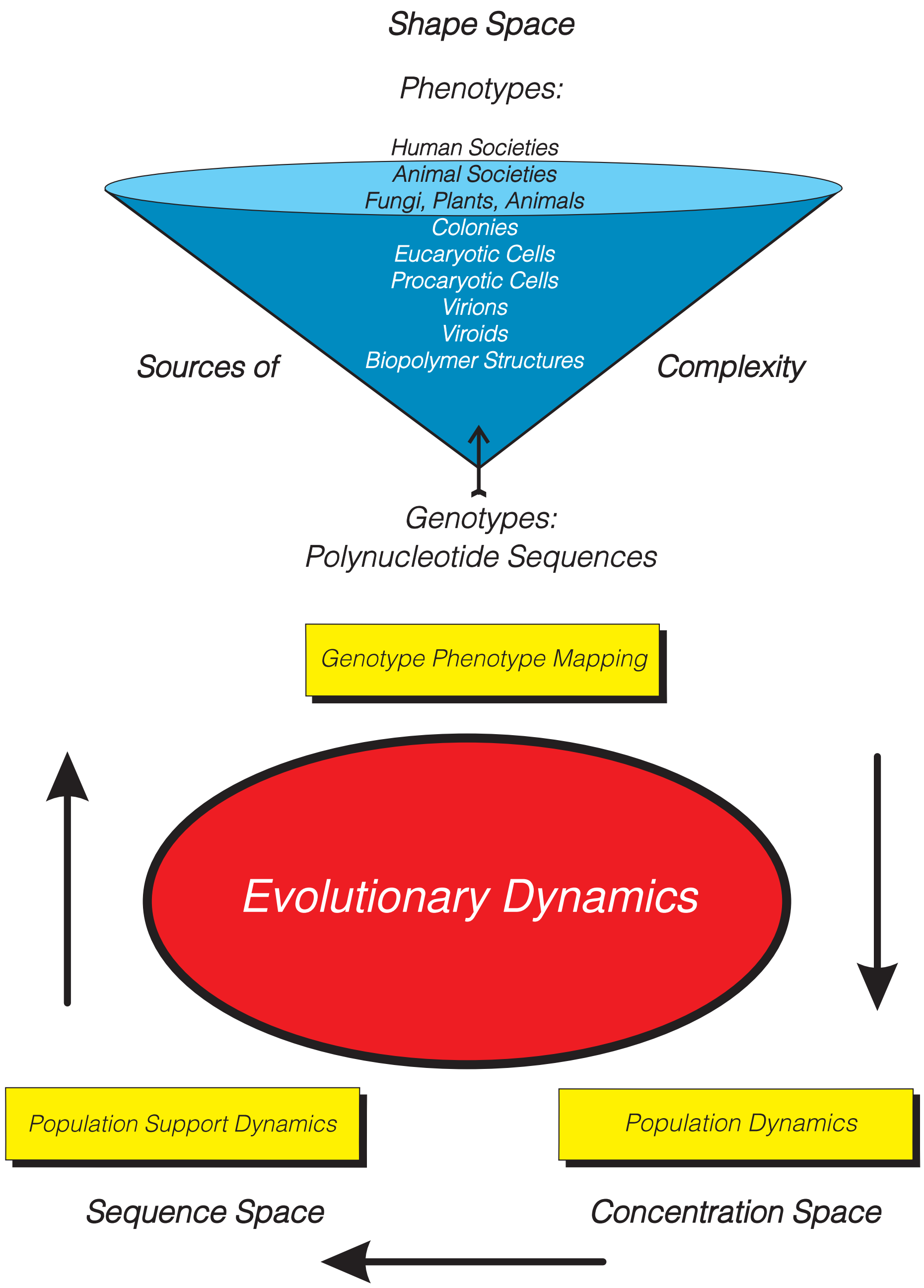

Figure 1: Evolutionary dynamics. The dynamics of evolution is partitioned into three simpler processes: (i) population dynamics, (ii) population support dynamics, and (iii) genotype-phenotype mapping. Increase in the complexity of life is confined to (iii): genotypes represented by RNA or DNA in genomes became longer and gave rise to more complex phenotypes during evolution. Major evolutionary transitions are characterized by radical innovations and large jumps in complexity.

EVOLUTIONARY DYNAMICS

Biological evolution is a highly sophisticated dynamical phenomenon, and its complexity is often confusing. It can be understood and modeled more easily if it is partitioned into three simpler processes [26], each of them highlighting one particular aspect of evolution (Figure 1). These three processes are **genotype-phenotype-mapping**, **population dynamics** and (population) **support dynamics**. They are properly described in three different abstract spaces: a space of population variables called ***concentration space***, a space of all (possible) genotypes called ***sequence space***, and a space of all possible phenotypes called ***shape space*** (In choosing this notion we refer to simple systems like molecules or viruses since various dynamical aspects including behavior are also part of the phenotypes of complex organisms).

Chemists and population geneticists are also familiar with concentration space. It is the space in which kinetics of chemical reaction or changes in populations take place. The variables count numbers of particles – for example molecules, virus particles, cells or organisms – and display their changes over time. Concentration space is restricted to the classes of genotypes actually present. When a mutant is produced, a new variable appears in concentration space; when a variant dies out, the corresponding variable disappears. The number of variables matches the number of currently existing genotype classes. The variables count the number of individuals of actually present genotype classes. The dynamics of real populations may be simple, like a monotonous approach towards a steady state. Optimization based on Darwin's principle of variation and "survival of the fittest", serves to illustrate such a process. It can be thought of as a walker climbing uphill on a (rugged) fitness landscape [21, 34]. Depending on the mechanism of reproduction and the nature of interactions between individuals, population dynamics may be as complex as spatio-temporal chaos. If one focusses on replicator dynamics [30] oscillations of concentrations [9], deterministic chaos [25], and spiral waves

[3] have all been reported. Is the population dynamics of human societies more complex than that of viruses? Certainly not. We need only recall chemistry: the Belousov-Zhabotinskii reaction exhibits all complex phenomena that are known from low-dimensional, nonlinear systems.

Sequence space is an impressive manifestation of the overwhelming diversity of possible genotypes. It is a result of the combinatorial building principle of nucleic acid molecules (Figure 2). Evolution in the entire time span from the origin of life to the present (or to the end of terrestrial life, as this would not make any recognizable difference) can explore only a negligibly small fraction of this hyperastronomically-large universe of potential biological information carriers. Considering this wealth of genetic coding possibilities, biological evolution is already a kind of open-ended process at the purely genetic level. Distances in sequence space are expressed in terms of elementary changes in genotypes during the course of evolution. These elementary changes are mutations and recombination events. Mutations come in three different types: point mutations or single base exchanges, insertions and deletions. In recombination two genotypes are split and joined in such a way that the new nucleic acid molecules contain one part of each parental molecule. The conventional representation of sequence space distance is the so-called Hamming distance, in which the distance between two genotypes is measured by the minimum number of point mutations needed to convert the genotypes into each other. Other genetic operations, such as recombination, bridge large distances in sequence space.

Support dynamics is about successful genetic operations, as it describes how populations change on their voyages through sequence space. Does support dynamics become more complex during evolution? To answer this question is somewhat trickier than in the case of population dynamics. Is the sequence space of multicellular organisms more complex than that of viruses? It is certainly larger.

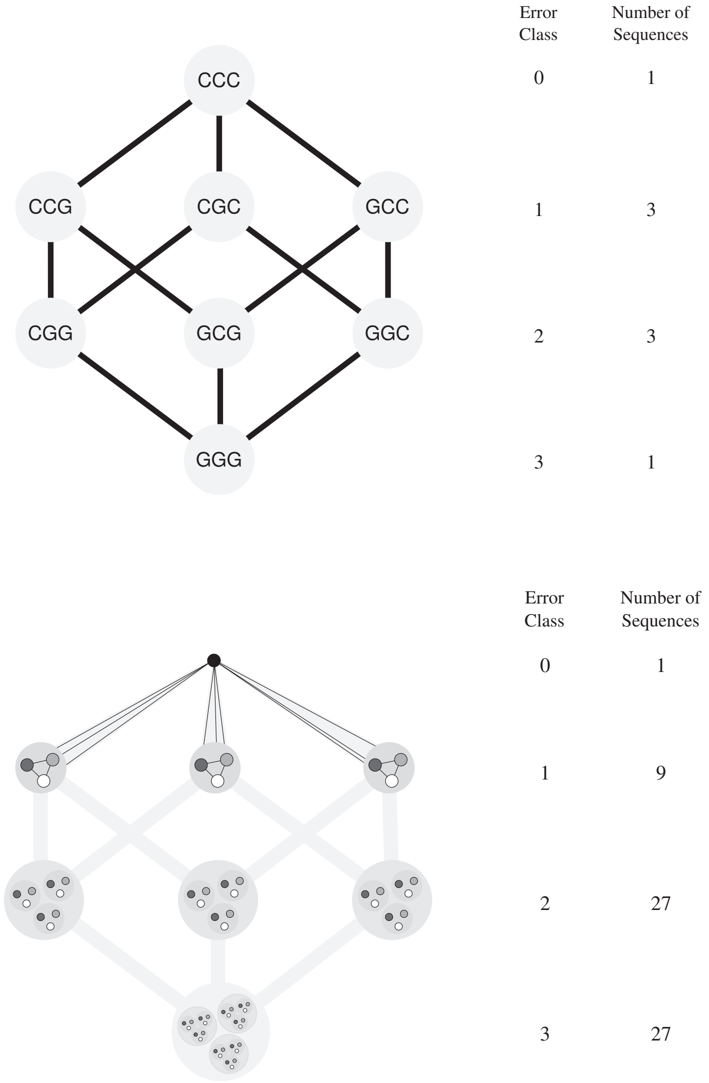

Figure 2: Sequence space. Sequences are represented by points in sequence space. Edges connect sequences of Hamming distance one, i.e., sequences that differ in a single position only. The index of the error class counts the Hamming distance from the reference sequence. The upper part shows the sequence space of **GC**-only sequences of chain length n=3. It has the form of a cube in ordinary three-dimensional space. In general, the sequence spaces of binary sequences are hypercubes of dimension n. The lower part contains the sequence space of natural (**AUGC**) sequences of chain length n=3. Note, that every base can be replaced by three other bases, for example, **A**$\rightarrow$**U**, **A**$\rightarrow$**G**, and **A**$\rightarrow$**C**. The different replacements are color coded in the sketch shown above. Edges are only shown for the error classes 0 and 1 (since the network of all connections is too sophisticated to be shown for the higher error classes).

Longer genomes, after all, have the combinatorial capacity for many more genotypes; thus, higher organisms explore a much smaller fraction of their sequence spaces. On the other hand, the areas actually populated are a negligible part of entire sequence space in all natural examples from viroids to man. The difference in support dynamics between prokaryotes and higher organisms thus boils down to haploidy versus diploidy and, in the latter case, to obligatory recombination. It appears correct to claim that support dynamics, just as population dynamics, has not become more complex during the course of evolution and one is doomed to failure if looking for the origin of complexity there.

Finally, shape space is the space of all phenotypes that are formed by processing genotypes in a given context. Like sequence space, it is a metric space. But the definition of a meaningful distance in shape space is an elaborate task requiring information on properties or functions of interest. Distances between RNA secondary structures have been defined that deal with formal interconversions of trees or other symbolic notations into each other [12, 17, 22]. A (trivial but evolutionarily highly relevant) distance is the difference in fitness between two phenotypes. Another appropriate measure of distance between two shapes might involve the average time it takes to reach a shape from the other one via an evolutionary process [18].

Figure 1 shows the existence of a cyclic relation between the three highlighted aspects of evolution. The genotype-phenotype mapping unfolds the properties of phenotypes, thus providing the parameters entering the kinetic equations of population dynamics. Population dynamics determine whether genotypes proliferate, stay at constant frequencies or decline and eventually go extinct. In this way population dynamics creates the input for support dynamics describing the development of populations at a higher (and coarse-grained) level; support dynamics provide a recording of genotypes that go extinct and register the newly-appearing genotypes. Thus, support dynamics does the accounting at the level of genotypes, and records the motion of entire populations. Support dynamics predicts the regions of sequence space into which populations migrate, and provides the input for the genotype-phenotype mapping thereby closing the cycle.

GENOTYPES AND PHENOTYPES

The view of evolutionary dynamics presented in Figure 1 allows us to focus on relations between genotypes and phenotypes in order to answer questions about the origin of natural complexity. Indeed, the objects and not the dynamics of evolution become more sophisticated in the course of the history of life. The way complexity increased becomes apparent when we recall the building plans and properties of living organisms. As mentioned earlier more complex things are made by integrating several (previously independent and already optimized) units into a larger entity by leaving unchanged as many details as possible. Hence, the smaller units retain some autonomy. For example, let us consider cellular metabolism.

Metabolism was invented in the early days of pre-prokaryotic life and was optimized so as to exploit a great variety of energy sources. The present day cells of highly developed organsisms make use of essentially the same metabolic machinery. Otherwise it would be impossible to express human genes in bacteria and genetic engineering could not work. Phenotypes span the entire field of complexity as they comprise polynucleotide molecules, viroids, virus particles, bacterial cells, protist cells, fungi, plants, animals, animal societies, and eventually humans. In the future, any scholarly study of biological complexity and its origin will have to deal with structures, properties and functions of phenotypes, and must be concerned with genotype-phenotype mappings.

At present, realistic modeling of relations between genotypes and phenotypes is confined to the evolution of RNA molecules in the test tube. Simplified versions of RNA phenotypes, the so-called secondary structures, have been analysed

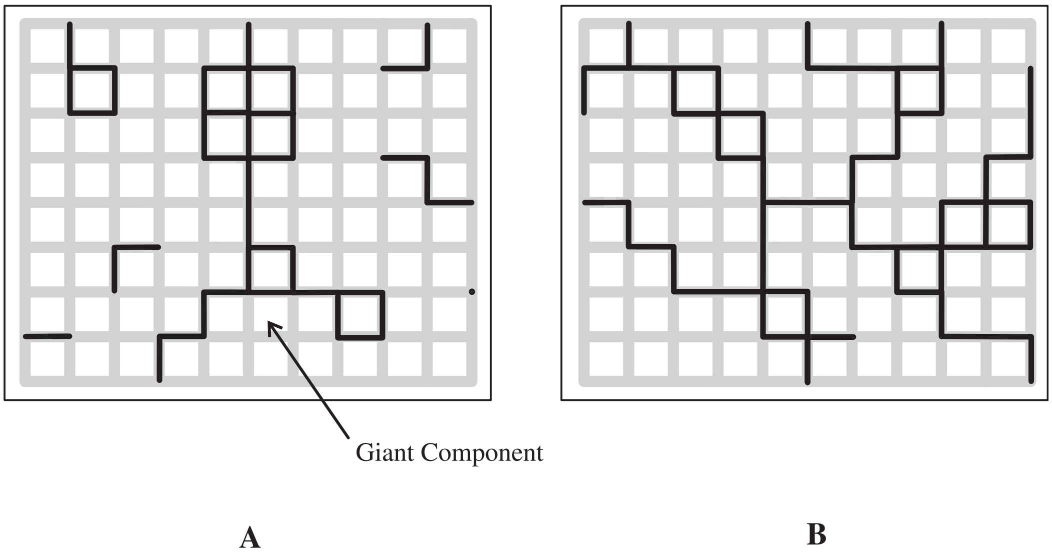

Figure 3: Connectivity in neutral networks. A neutral network is a set in sequence space that contains all sequences folding into a given structure. These sequences form the nodes of a graph that is derived from the set by connecting all nodes corresponding to pairs of sequences with Hamming distance one. A neutral network consists of many components if the average fraction of neutral neighbours of the respective structure in sequence space (λ) is below a threshold value (λ_{cr}). Random graph theory predicts in this case that the network consists of one giant component and many smaller ones (**A**, for the purpose of illustration the sequence space is represented by a two-dimensional grid). If λ exceeds the threshold value (**B**) the network is connected and spans the entire sequence space.

in great detail [12, 15, 16, 29]. The mapping of RNA genotypes into phenotypes has been found to be highly redundant. Many sequences form the same secondary structure and give rise to a high degree of neutrality in sequence space. In addition, these neutral sequences, i.e., sequences forming the same (common) secondary structure, are distributed almost randomly in sequence space. Systematic investigations have led to the formulation of two concepts that are highly relevant for understanding evolutionary processes [29]. These are the existence of **neutral networks** spanning the whole sequence space (Figure 3), and **shape-space covering** by a small section of sequence space (Figure 4). In order to find a given (common) structure one need only screen a spherical environment around

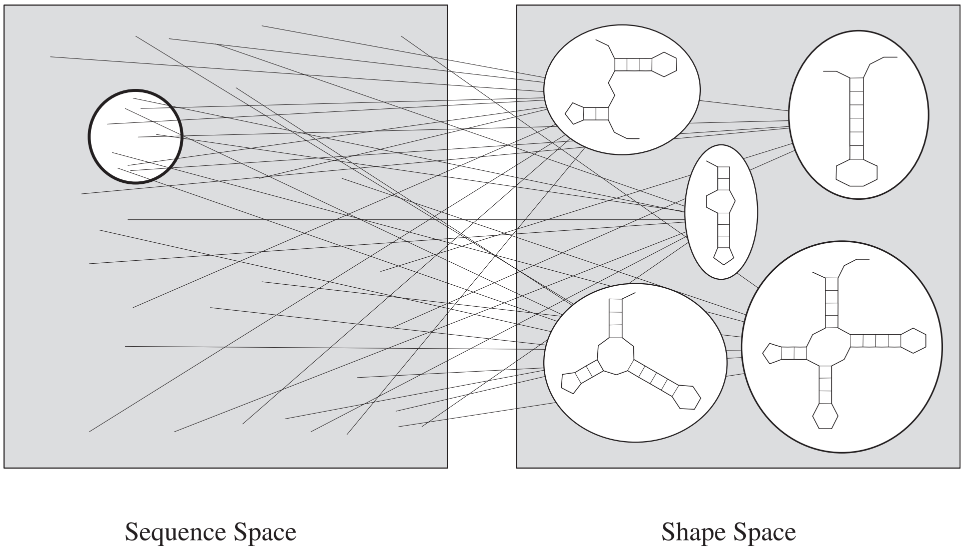

Figure 4: Shape space covering. Only a (relatively small) spherical environment around any arbitrarily chosen reference sequence has to be searched in order to find RNA sequences for every (common) secondary structure. In the sketch shown above points in sequence space (corresponding to sequences) are connected by straight lines with the secondary structures they form.

an arbitrarily chosen reference sequence that is much smaller than the entire sequence space. In fact, these properties of RNA sequence-to-structure mappings are changing the conventional view of optimization as they make evolutionary searching processes much easier than previously believed. The same is true for *in vitro* evolution applied to biotechnology [27].

DARWINIAN DYNAMICS

The conventional Darwinian view sees evolutionary optimization as hill-climbing on a fitness landscape. These landscapes are built upon sequence space by assigning a numerical (fitness) value to every genotype. Recent developments focus on their mathematical characterization and statistical analysis (For reviews see [31, 32]). The primary active products derived from the genomes are biopolymers, RNA molecules and proteins. From biophysics we know that point mutations (single nucleotide exchanges in DNA) may have all kinds of effects ranging from drastic changes in properties and functions to no change at all. Translated to fitness values, this would imply that in the neighborhood of a typical genotype we shall find genotypes of lower and higher fitness, as well as genotypes having essentially the same fitness and hence being selectively neutral.

Fitness landscapes are higly complex and rather bizarre objects with many local optima of greatly varying height. Darwinian evolution on such rugged fitness landscapes has to solve a tantalizing problem: going uphill soon ends in some minor peak, so optimization by (strictly) adaptive walks is doomed to end at (very) low improvements. The previously mentioned results on the properties of RNA-folding landscapes show a way out of this (local) optimization dilemma.

Neutral sequences form extended networks in sequence space [29] and facilitate optimization through adaptive walks (Figure 5). In the absence of more fit genotypes a population drifts in random-walk-like manner on the network and examines its current neighborhood for new genotypes of higher fitness [19]. Whenever the population arrives at a point that allows a transition to the neutral network of fitter variants, it will jump from the current network to the one with higher fitness. If the networks extend over sufficiently large areas in sequence space the populations will also be able to reach distant fitness peaks. Optimization follows a combined mechanism [28]: adaptive walks leading to minor peaks are supplemented by random drift along networks that enable populations to escape from evolutionary traps, thus allowing them to migrate to areas of higher fitness values. In this fashion, the global fitness optimum may eventually be reached.

Evolution following Darwin's principle on a fitness landscape, no matter whether or not it occurs via neutral networks, performs a search for the best suited forms within a given class. Optimization will make molecules, cells, organisms or societies more efficient. They become better at exploiting the available resources and they will increase fitness since they can afford to have more progeny with the same amount of resources. Thus Darwin's principle is nature's recipe

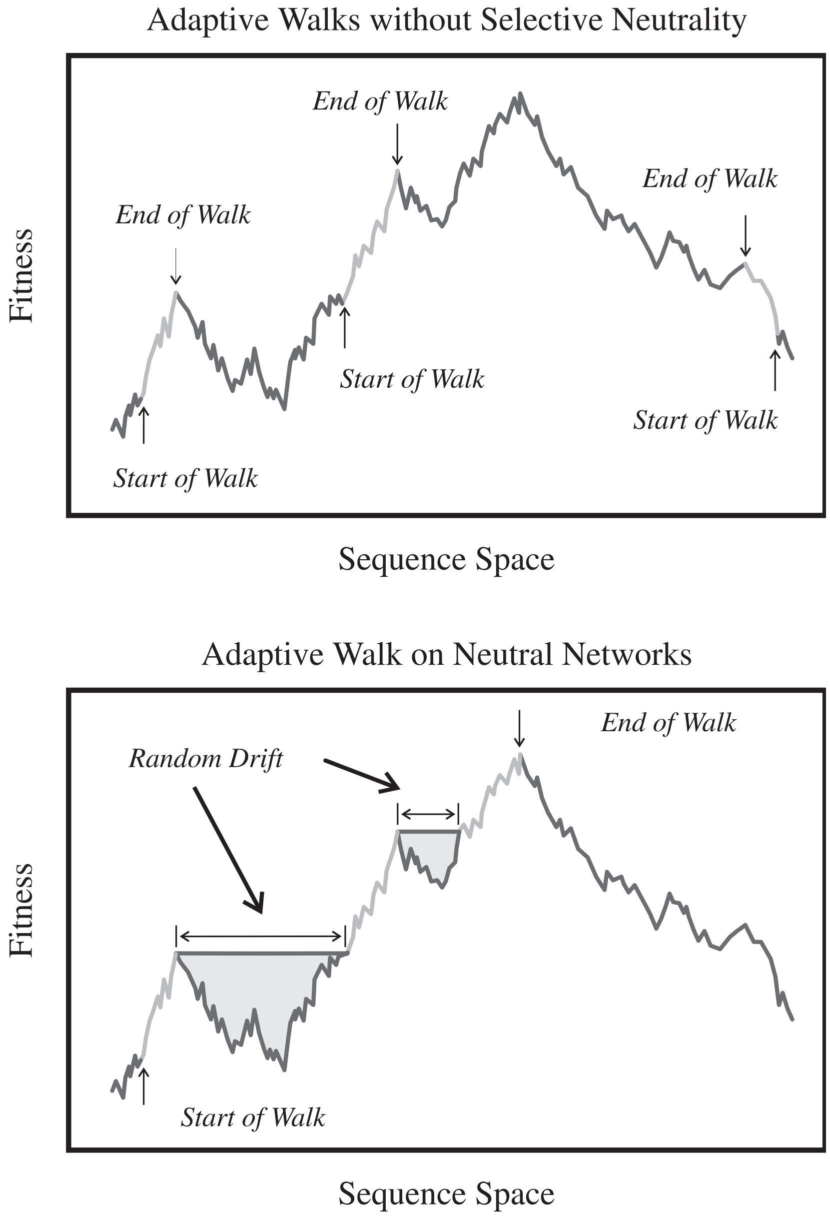

Figure 5: Evolutionary optimization in sequence space through adaptive walks of populations. Adaptive walks allow to choose the next step arbitrarily from all directions where fitness is (locally) non-decreasing. Populations can bridge over narrow valleys with widths of a few point mutations. In absence of selective neutrality (upper part) they are, however, unable to span longer Hamming distances and thus will approach only the next major fitness peak. Populations on rugged landscapes with extended neutral networks evolve by a combination of adaptive walks and random drift at constant fitness along the network (lower part). Eventually, populations reach the global maximum of the fitness landscape.

to deal with limited supplies. Optimization, however, does not lead to radical innovation. Going into the development of unknown properties and functions is too risky in times of scarcity, and those who make such attempts are not likely to survive. Darwinian evolution will not be able to generate the major transitions and jumps in complexity that we initially addressed.

An engineer is always tempted to design new versions of machinery from scratch as this allows him to avoid all errors previously made. If a machine gets more and more complex, design from scratch becomes too expensive and even human engineers cannot escape the need to retain older versions and build upon them. The development of computer operating systems serves as a good example, the well-known DOS being an extreme case of add-on design. To control disks is one of its least important tasks nowadays. The build-upon-latest-version principle has been characterized as "tinkering" by the French biologist François Jacob [20]. Nature is the most efficient known tinkerer and therefore also makes a mistake every now and then. A famous example is our vertebrate eye, where the nerve comes out on the wrong side of the retina, a mistake that was not made in the design of insect and cephalopode eyes, which originate from the same (primitive) genetic photosensoric system [24], as those of vertebrates. To make the point clear, I think there is nothing wrong with tinkering. On the contrary, it turns out that ingenious tinkering does indeed represent a building principle that is superior to the fully rational approach if you have to design for the unknown. Biological evolution is, after all, an excursion into an unknown future. The tinkerer is flexible and ready to change strategies if need be. Examples from nature could be ennumerated endlessly. Let me mention just one, the phylogenetic history of the human body plan from fish via amphibia to mammals walking on four legs and finally to the two-legged human construction. One principle of design that is especially well suited for tinkering uses modular construction: an enormous variety of different things can be assembled from a few modules – if one uses only the right set (Anyone who doubts this point should take a rest and play with LEGO).

A modular building principle of biopolymers has been suggested by Walter Gilbert [5, 14] in his **exon shuffling** approach to protein design. A limited number of modules, say between ten and some hundred thousand, have been postulated that are thought to represent elementary folding units of polypeptide chains into stable structures. The coding sequences of these units are assumed to correspond to the (primordial and, in some cases, present day) exons, which have been optimized initially to yield stable structures and simple catalysts. These elementary structural units were then combined through exon shuffling to form the proteins of present-day shapes and sizes. Modular construction is indeed found almost everywhere in nature. In the case of higher hierachical units, it is easy if not trivial, to recognize the modules: the cells in colonies and multicellular organisms or the individuals in animal and human societies. The flexibility provided by modular construction and the build-upon-latest-version strategy seems to be nature's way of coping with the unpredictable.

INNOVATION IN BIOLOGY

Lastly, we are left with the problem of radical innovation. The mechanism of evolution is built upon chemical kinetics of replication and mutation [6, 8]. Various scenarios have been derived by applying different types of replication. Independent replication led to Darwinian behavior based on competition and resulted in the concept of the molecular quasispecies [8, 7]. If replicating molecules become mutually dependent on each other, population dynamics changes and individuals in populations no longer compete. In the simplest case of cyclic catalytic coupling, they form a hypercycle that represents the simplest way of suppressing competition in the sense of molecular symbiosis [9]. The hypercycle scenario is indeed a good candidate for a mechanism leading to radical innovation [33]. Motivated by

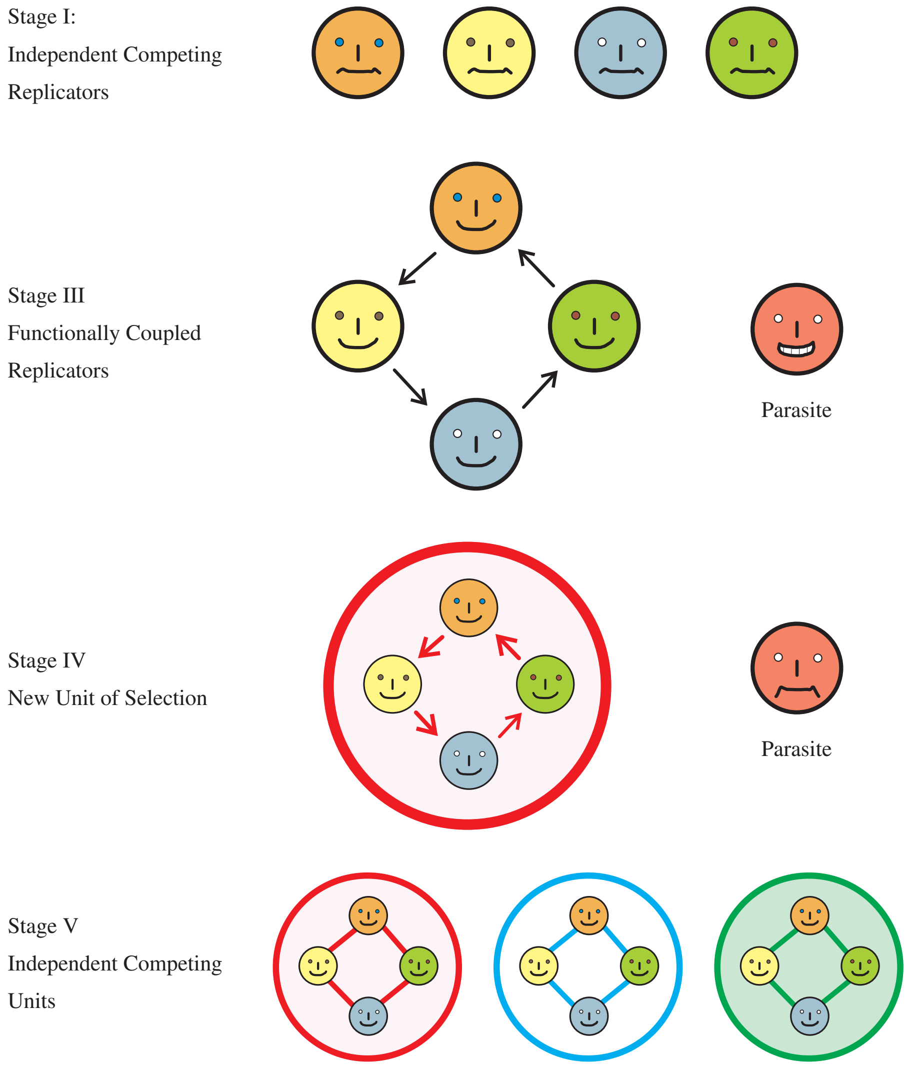

Figure 6: A simple model for the formation of higher hierarchical units. Major jumps in biological evolution are (often) accompanied by the formation of new units that integrate previous competitors into a larger composite entity (see text and [11]). Independent competitors (Stage I) are integrated by hypercycle-like catalytic couplings into a functional unit that is endangered by exploitation through parasites (Stage III). A new unit of selection is created by the fomation of boundaries separating the cooperating partners from the rest of the world (Stage IV). In this way the system is protected against straightforward parasitic exploitation. Further integration, for example by ligation of previously indepedent genetic information carriers into a single genome, variation through mutation and selection brings Darwinian evolution back to the system, now occurring on the level of the new (higher hierarchical) units.

the integrating capacities of hypercycles we proposed a model for innovation in evolution more than twelve years ago [10, 11]. It is based on the kinetic equations of replication and can be best characterized as a symbiontic mechanism. The transition to the next higher hierarchical level occurs in five steps (Figure 6):

(1) Independent replicators compete for resources, showing Darwinian evolution, and optimize their own individual benefits.

(2) Mutual dependence of reproductive success reduces and eventually eliminates competition between different replicators. These replicators might, for example, be members of the same quasispecies.

(3) Dynamical coupling of reproduction within a group of replicators causes them to grow together and to form a functional unit that integrates the different functions of several phenotypes.

(4) Spatial or other forms of integration join the replicators and create a new **unit of selection** at the next hierarchical level.

(5) Integration creates a new class of individuals that evolve as autonomous units by a Darwinian mutation and selection mechanism. Eventually, the genotypes of the initially independent replicators become joined and form a single, large genome.

This model is able to create complex systems of a hierarchical nature. An essential feature of systems produced by the model as well as by natural systems is a rather high degree of autonomy of the lower level subunits. In terms of our initial model of evolutionary dynamics, complexity increases in the genotype-phenotype mapping since the new genotype is much longer and the new phenotype has many more functions than its smaller precursors. The best example of the symbiotic mechanism of integration is the eukaryotic cell. Most biologists nowadays think that it is composed of previously independent cellular units: the pre-eukaryote, the pre-mitochondrion, and in case of plant cells, the pre-chloroplast, a (close) relative to present day cyanobacteria. The cell organelles seem to be the last phase of step (5). Most of the previous mitochondrial genes, for example, have been transferred to the cellular nucleus with a few remaining with the organelle and thereby providing a basis for a partial autonomy of the mitochondrion.

As long as the new and larger units have not been individualized by introducing (physical or communicational) barriers that allow members and non-members of the organization to be distinguished, they are subject to exploitation by parasites. Parasites are individuals (usually, but not necessarily, from the lower hierarchical level) that take advantage of the novel functions of the community without contributing to common resources. Such barriers are well known in nature: cell membranes, cell walls, animal skins, the barks of trees, as well as signals and languages that are only understood by the members of a (animal) society. The question then is how competitors can be forced to stay away from parasitic exploitation of common resources, and thus avoid the "tragedy of the commons" [13]. Who is going to pay the costs for good behavior? An answer might be that integration occurs primarily when resources are abundant and cheap. Then the extra costs for not competing would be small and could be overcompensated by even minor benefits for being cooperative.

Taken as given that modular design has determined progress in evolution from the very first days, we can visualize a rather simple genetic mechanism for making simple things more complex: gene and genome duplication. A replication error known as **insertion** causes a piece of DNA to be copied twice. The cost problem becomes even more evident when we consider this gene duplication as a mechanism for creating a basis for innovation in evolution [33]. Parts of the genome are duplicated and the genes of the second set are free to develop novel functions since they are not required for ordinary cellular life. This mechanism is a kind of short-cut for the symbiontic mechanism, as it avoids the excursion into population dynamics in step (2) and (3). Here the new variant with a larger genome has to outcompete the previously optimized smaller variant. Initial benefits from developing new functions will certainly be very small. The chances for survival of the new variant are good, however, when resources are cheap, since the cost of having a larger genome will be negligible. Precisely the same argument holds for the third mechanism for becoming more complex, which consists in the development of a somatic cell line leading to cell differentiation and multicellular organisms [33]. The reduction in fitness caused by making cells that do not contribute to reproduction will be small when the extra costs for additional cell divisions are low; in other words, when resources are cheap.

Radical innovation leading to an increase of complexity in evolution thus requires that (at least) two conditions are fulfilled. The internal structure of the organism has to have an intrinsic capacity for making genotypes and their unfolding more complex: molecules or cells have to be able to form catalytic or symbiontic interactions (symbiontic mechanism), genomes require a mechanism to become longer (gene duplication mechanism), cells must have a communication system that allows mutual interaction (epigenetic mechanism). Molecular biology tells us that all these prerequisites are and presumably were fulfilled (long) before the major evolutionary jumps occured. The second condition is an abundance of resources. In times when resources are scarce, fuzzy optimization pays. The variant that saves a little more in everyday life has more resources to spend on reproduction and will thus have more progeny. Times of austerity, in my opinion, are not the times of the great jumps in complexity. Major transitions in evolution occur only when energy or some other growth-determining resource are fairly cheap. Then the extra costs for developing something new are low and the pioneering variants have a chance to develop their fitness improving functions under low selective pressure.

Three temporal characteristics of terrestrial environments were mentioned in the paper's subtitle: **scarcity** and **abundance** of resources, as well as **unpredictability**. In summary, we have argued that nature uses optimization to deal with scarcity, she takes advantage of abundance to create innovation, and her recipe to master unpredictablity is tinkering and modular design.

REFERENCES

[1] W. B. Arthur. On the evolution of complexity. In G. Cowan, D. Pines, and D. Meltzer, editors, *Complexity: Methaphors, Models, and Reality*, volume XIX of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 65–78. Addison-Wesley Publ. Co., Redwood City, CA, 1994.

[2] C. H. Bennett. Dissipation, information, computational complexity, and the definition of organization. In D. Pines, editor, *Emerging Syntheses in Sciences*, volume I of *Santa Fe Institute Studies in the Sciences of Complexity*, pages 215–233. Addison-Wesley Publ. Co., Redwood City, CA, 1987.

[3] M. C. Boerlijst and P. Hogeweg. Spiral wave structure in pre-biotic evolution - hypercycles stable against parasites. *Physica D*, 48:17–28, 1991.

[4] T. Dobzhansky, F. J. Ayala, G. L. Stebbins, and J. W. Valentine. *Evolution.* W.H. Freeman & Co., San Francisco, CA, 1977.

[5] W. F. Doolittle. Genes in pieces: Were they ever together? *Nature*, 272:581–582, 1978.

[6] M. Eigen. Selforganization of matter and the evolution of biological macromolecules. *Naturwissenschaften*, 58:465–523, 1971.

[7] M. Eigen, J. McCaskill, and P. Schuster. The molecular quasispecies - An abridged account. *J. Phys. Chem.*, 92:6881–6891, 1988.

[8] M. Eigen and P. Schuster. The hypercycle. A principle of natural self-organization. Part A: Emergence of the hypercycle. *Naturwissenschaften*, 64:541–565, 1977.

[9] M. Eigen and P. Schuster. The hypercycle. A principle of natural self-organization. Part B: The abstract hypercycle. *Naturwissenschaften*, 65:7–41, 1978.

[10] M. Eigen and P. Schuster. The hypercycle. A principle of natural self-organization. Part C: The realistic hypercycle. *Naturwissenschaften*, 65:341–369, 1978.

[11] M. Eigen and P. Schuster. Stages of emerging life - Five principles of early organization. *J. Mol. Evol.*, 19:47–61, 1982.

[12] W. Fontana, D. A. Konings, P. Stadler, and P. Schuster. Statistics of RNA secondary structures. *Biopolymers*, 33:1389–1404, 1993.

[13] S. A. Frank. Mutual policing and repression of competition in the evolution of cooperative groups. *Nature*, 377:520–522, 1995.

[14] W. Gilbert. Why genes in pieces? *Nature*, 271:501, 1978.

[15] W. Grüner, R. Giegerich, D. Strothmann, C. Reidys, J. Weber, I. Hofacker, P. Stadler, and P. Schuster. Analysis of RNA sequence structure maps by exhaustive enumeration. I. Neutral networks. *Mh.Chem.*, 127:355–374, 1996.

[16] W. Grüner, R. Giegerich, D. Strothmann, C. Reidys, J. Weber, I. Hofacker, P. Stadler, and P. Schuster. Analysis of RNA sequence structure maps by exhaustive enumeration. II. Structure of neutral networks and shape space covering. *Mh.Chem.*, 127:375–389, 1996.

[17] P. Hogeweg and B. Hesper. Energy directed folding of RNA sequences. *Nucleic Acids Research*, 12:67–74, 1984.

[18] M. A. Huynen. Exploring phenotype space through neutral evolution. J. Mol. Evol., in press, 1996.

[19] M. A. Huynen, P. F.Stadler, and W. Fontana. Smoothness within ruggedness: The role of neutrality in adaptation. *Proc. Natl. Acad. Sci. USA*, 93:397–401, 1996.

[20] F. Jacob. *The Possible and the Actual.* Pantheon Books, New York, 1982.

[21] S. A. Kauffman. *The Origins of Order. Self-Organization and Selection in Evolution.* Oxford University Press, Oxford, UK, 1993.

[22] D. Konings and P. Hogeweg. Pattern analysis of RNA secondary structure. Similarity and consensus of minimal-energy folding. *J. Mol. Biol.*, 207:597–614, 1989.

[23] J. Maynard Smith and E. Szathmáry. *The Major Transitions in Evolution.* W. H. Freeman, Oxford, UK, 1995.

[24] D. E. Nilsson. Old genes for new eyes. *Current Biology*, 6:39–42, 1996.

[25] W. Schnabl, P. F. Stadler, C. Forst, and P. Schuster. Full characterization of a strange attractor. Chaotic dynamics on low-dimensional replicator systems. *Physica D*, 48:65–90, 1991.

[26] P. Schuster. Artificial life and molecular evolutionary biology. In F. Morán, A. Moreno, J. J. Morelo, and P. Chacón, editors, *Advances in Artificial Life*, volume 929 of *Lecture Notes in Artificial Intelligence*, pages 3–19. Springer-Verlag, Berlin, 1995.

[27] P. Schuster. How to search for RNA structures. Theoretical concepts in evolutionary biotechnology. *Journal of Biotechnology*, 41:239–257, 1995.

[28] P. Schuster. The role of neutral mutations in the evolution of RNA molecules. In S. Suhai, editor, *Computational Methods In Genome Research*. Plenum Press, New York, 1996.

[29] P. Schuster, W. Fontana, P. Stadler, and I. Hofacker. From sequences to shapes and back: A case study in RNA secondary structures. *Proc.Roy.Soc.(London)B*, 255:279–284, 1994.

[30] P. Schuster and K. Sigmund. Replicator dynamics. *J. Theor. Biol.*, 100:533–538, 1983.

[31] P. Schuster and P. Stadler. Landscapes: Complex optimization problems and biopolymer structures. *Computers Chem.*, 18:295–314, 1994.

[32] P. F. Stadler. Towards a theory of landscapes. In R. Lopéz-Peña, R. Capovilla, R. García-Pelayo, H. Waelbroeck, and F. Zertuche, editors, *Complex Systems and Binary Networks*, pages 77–163, Berlin, New York, 1995. Springer Verlag. SFI preprint 95-03-030.

[33] E. Szathmáry and J. Maynard Smith. The major evolutionary transitions. *Nature*, 374:227–232, 1995.

[34] S. Wright. The roles of mutation, inbreeding, crossbreeding and selection in evolution. In D. F. Jones, editor, *Int. Proceedings of the Sixth International Congress on Genetics*, volume 1, pages 356–366, 1932.

Peter Schuster works mainly on theoretical aspects of molecular evolution and structural biology. He did his PhD at the university of Vienna and was post-doc with Manfred Eigen in Göttingen. Since 1973 he is professor of theoretical chemistry in Vienna. In 1992 - 1995 he was founding director of the Institute for Molecular Biotechnology in Jena where he is still the head of the Department of Molecular Evolutionary Biology. Since 1991 he is member of the external faculty of the Santa Fe Institute.