Review Article

Evolvability and Speed of Evolutionary Algorithms in Light of Recent Developments in Biology

Ting Hu[1, 2] and Wolfgang Banzhaf[1]

[1] *Department of Computer Science, Memorial University, St. John's, NL, Canada A1B 3X5*
[2] *Norris Cotton Cancer Center, Dartmouth-Hitchcock Medical Center, Lebanon, NH 03756, USA*

Correspondence should be addressed to Ting Hu, Ting.Hu@Dartmouth.edu

Biological and artificial evolutionary systems exhibit varying degrees of evolvability and different rates of evolution. Such quantities can be affected by various factors. Here, we review some evolutionary mechanisms and discuss new developments in biology that can potentially improve evolvability or accelerate evolution in artificial systems. Biological notions are discussed to the degree they correspond to notions in Evolutionary Computation. We hope that the findings put forward here can be used to design computational models of evolution that produce significant gains in evolvability and evolutionary speed.

1. Introduction

The field of *Evolutionary Computation* (EC) has seen enormous progress since it was founded in the Sixties and Seventies of the 20th century [1–11], inspired by the evolutionary processes observed in the living world.

In EC, candidate solutions to optimization or learning problems are represented by structures similar to gene sequences and their phenotypic expressions. The ensemble of such solutions is referred to as a population. Evolutionary operators, such as mutation, recombination, and selection, are applied to this population. Solutions gradually improve by repeating a variation-selection cycle through numerous iterations of the evolutionary process. Essentially a search method, EC, often produces well-performing solutions to complex optimization and learning problems arising from various areas, to the point where its problem solving capability mirrors or even exceeds that of humans [12]. Yet, EC is not without weaknesses, and new algorithmic variants are constantly being introduced, studied, and applied.

The fundamental idea of EC was gleaned from biology, and more specifically, from Darwin's theory of evolution by natural selection [13] as embodied in the Neo-Darwinian synthesis [14, 15]. In the past decades, however, knowledge of natural evolution has improved profoundly in biology. This progress has, to a large degree, not been incorporated yet into computational models of evolution and therefore cannot be harvested for applications. We have argued that adopting new knowledge about natural evolution generated in areas such as molecular genetics, cell biology, developmental biology, and evolutionary biology would substantially benefit EC [16, 17].

The question then arises what the most important and revolutionary discoveries are in biology in recent times, and how they can be sufficiently abstracted to provide material for computational models. As the number of scientists working in the areas mentioned above is now higher than at any other time of the past and can be estimated to be well over a million, it becomes nontrivial to select those aspects of evolution that will have the most impact in computational models. A number of books have appeared in recent years that provide some guidance in this quest (see, e.g., [18–25]).

Here we restrict ourselves to mainly review the concepts of evolvability and the speed of evolution. This is motivated by the fact that EC approaches often suffer from progressive slow-down of evolutionary speed. While under some circumstances appreciated as convergence to a global optimum, for many real-world tasks convergence and the

corresponding slow-down of fitness improvements as well as the reduction in the diversity of solutions is more a predicament than an advantage. This is especially true for difficult problems where there is no hope to find an optimal solution, but where good solutions would already provide a benefit in the application. As a result, the development of systems that show continued evolutionary potential, *open-ended evolution*, as it has been termed, has gained prominence.

Open-ended evolution is a hallmark of Life. Thus, one alternative route to explore this topic in computation is via an Artificial Life approach [26]. And so a number of artificial systems have been designed in the meantime with the aim to simulate organic life *in silico*, such as *Tierra* [27], *Avida* [28], and *Evita* [29]. In these systems, computer code is regarded as "digital organisms" with CPU time the "energy" resource and memory the "material" resource. Digital organisms evolve through interactions with their neighbors and competition for resources. Fitness is not an explicit notion in these systems. However, this is still only an initial step towards understanding and realizing open-ended evolution in artificial systems, as Bedau and Brown [30] report that the long-term capability to adaptation seems to be missing from these systems compared to real organisms (for an example of long-term capabilities in simple organisms, see [31]); that is, this type of artificial evolution lacks evolvability in the "long-term".

In EC, the situation is even exacerbated by the existence of an explicitly defined fitness function, often in the form of a simple scalar. However, while it is the holy grail of computational models of evolution to achieve continued evolutionary potential, which has—to the best of our knowledge—not been reached to date, progress has still been achieved by studying more limited concepts like that of evolvability. In a nutshell, the hope is that by relating properties of natural evolutionary systems to mechanisms used by Nature to achieve them, we might learn enough to design algorithmic mechanisms that exhibit similar features. So let us start by looking more closely at *evolvability* and the *rate of evolution*.

1.1. Evolvability. In the process of evolution, genotypic variation explores new evolutionary material, the corresponding phenotypic variation provides adaptive characteristics, and stabilization operators like selection preserve improvements over previous generations. The cooperation of these activities is what allows evolution to work. Thus, the core mechanism of evolution is to assemble the forces of these operations that yield adaptive improvements implying the evolvability of an evolutionary system. A growing number of efforts have been dedicated to understanding [25, 32–38] and enhancing [39–45] evolvability.

While the concept of evolvability is still very much under discussion, we will adopt a definition that is equally applicable to natural and artificial systems.

Definition 1. Evolvability is the capability of a system to generate adaptive phenotypic variation and to transmit it via an evolutionary process.

Altenberg [46] describes evolvability from the viewpoint of EC as the ability of a genetic operator or representation scheme to produce offspring fitter than their parents. In biology, Kirschner and Gerhart [47] consider evolvability as an organism's capacity to generate heritable and selectable phenotypic variation. An explicit comparison between evolvability in biological and computational systems has been performed by Wagner and Altenberg [45]. In their view, evolvability must be seen as the ability of random variants to produce occasional improvements, which would depend critically on the plasticity of the genotype-phenotype map. The authors emphasize "variability" determined by the genotype-phenotype map as the *propensity to vary*, rather than variation itself. Marrow [48] suggests that evolvability means the capability to evolve, and this characteristic should be relevant to both natural and artificial evolutionary systems. He discusses a number of important contributions on this topic in both biology and EC and raise some open questions for further research.

Recently, a growing number of evolutionary biologists and computer scientists have shown interest in this topic. In an evolutionary system, many properties of a population are considered related to evolvability, including facilitation of extradimensional bypass and robustness against genotypic variation [49, 50], redundancy, flexibility during developmental processes [47], and mutation rate adaptation [51]. The notion of *evolutionary capacitance* has also been used in this context ([52], and references therein).

The detection and measurement of evolvability is an intriguing and nontrivial problem. Phenotypic fitness is directly observable and serves as a selection criterion. However, as a *potential* to generate better fitness and a *capability* for adaptive evolution, evolvability is a different type of observable, which is more difficult to observe and to quantify. Although a formal methodology on measuring evolvability has not yet been agreed upon in the literature, some empirical methods have been proposed nevertheless.

Nehaniv [53] proposes the perspective of using evolutionary system complexity to describe and measure evolvability. He defines the *exhibited evolvability* as an observable outcome generated by evolvability and measures evolvability by the rate of increasing complexity of evolutionary entities in an evolutionary system. Wagner proposes to simply measure the number of nonneutral 1-step mutation variations in a biological system of particular relevance to RNA evolution in order to quantify evolvability [54]. As one can see, Nehaniv's definition entailing more complex entities will in general also lead to a larger number of nonneutral 1-step variations and thus increase this measure of evolvability.

Another perspective on evolvability is provided by Earl and Deem [55] who suggest that evolvability can be selected for by variation in the environment. By observing genetic changes in protein evolution, they find that rapid or dramatic environmental change generates strong selection pressure for evolvability. Thus, high evolvability can be detected and favored by such selection pressure. For an artificial evolutionary system Reisinger et al. concur when they propose an indirect encoding representation to improve evolvability [56, 57]. A gradually changing fitness function

is designed to measure evolvability of representations and to evolve a population that is adaptive under different environments. Furthermore, as the pace of change of the fitness function increases, stronger selection pressure for evolvability is imposed.

1.2. Rate of Evolution. Related to the theme of evolvability is that of the rate of evolution. Evolvability defines how likely a system can generate adaptive phenotypic variations whereas the *rate of evolution* describes how fast this evolutionary process can proceed. The rate of evolution is a fascinating topic in evolutionary biology and has caused many debates already since Darwin's time. Darwin himself held the view of *phyletic gradualism*, hypothesizing that most evolution occurs uniformly, gradually moulded by selective conditions. Others were of a different opinion, and Eldredge and Gould proposed the theory of *punctuated equilibria* [58]. According to this idea, evolution occurs through bursts of innovation followed by long periods of stasis, a major challenge to Darwin's orthodoxy.

Definition 2. Rate of Evolution is a quantitative measure of the changes observable in an evolutionary system over generational (or otherwise appropriately defined) time-scales.

In biology, the rate of evolution has different definitions and measures depending on the underlying objects examined, for instance, gene sequences, proteins, organisms, and so forth. In molecular biology, the rate of evolution usually describes the rate of mutants being preserved as advantageous, that is, those that can generate phenotypic improvements. This is observed by looking at the fixation of alleles in genes. Biologists use the k_a/k_s ratio to measure the rate of gene sequences evolution [59–62]. It is known that some changes to a gene sequence may lead to differences in the amino acid sequence of an encoded protein while others will not, due to the degenerate code employed for translation. Therefore, such a measure can be used to compare two homologous protein-coding gene sequences of related species. The k_a/k_s ratio resulting from a measurement of the number of nonsynonymous (amino acid) substitutions per nonsynonymous site (k_a) to the number of synonymous substitutions per synonymous site (k_s) characterizes the rate of evolution between these two sequences. Since k_s measures neutral evolution (without considering functional improvements under selection pressure), the k_a/k_s ratio reflects the amount of *adaptive* evolution against the *background* amount of variation. Note that this is an approximation since there are nonsynonymous changes in amino acid sequences that do not change the function of the protein in which they appear.

In case $k_a/k_s > 1$, fixation of nonsynonymous substitutions is faster than that of synonymous substitutions, which means that *positive selection* fixes amino acid changes faster than silent changes. Mostly, however, one finds $k_a/k_s < 1$, the case where deleterious substitutions are eliminated by *purifying selection* (negative selection), and the rate of fixation of amino acid changes is smaller than the background rate of

variation. If $k_a = k_s$, the fixation of these two types of changes is at the same rate, a special case indicating, for example, pseudogenes. To summarize, measuring a large k_a/k_s ratio suggests that adaptation has been generated (and fixed) at a high rate. This measurement has been widely applied in the analysis of adaptive molecular evolution and is accepted as a general method for measuring the rate of gene sequence evolution in biology.

Other than at the molecular level, Worden has defined the concept of *genetic information in the phenotype* (GIP) in his work on the speed limit for evolution [63]. GIP is meant to be a measure of the amount of genetic information expressed in observable phenotype, and he uses the rate of increasing GIP to describe the rate of evolution. He proposes that GIP measurement can be applied in both biology and EC.

As we can see from these examples, both phenotypic effects and genotypic effects have to be taken into account when measuring the rate of evolution.

In artificial systems used for EC, the goal of evolution is much more specific than in nature: to find the solutions to a given problem. The rate of evolution in EC, therefore, usually refers to the speed of solving a specific problem, for example, to the speed of fitness improvements or the speed of approaching a fixed objective. The ability to define explicit phenotypic fitness is one of the most distinguishing features that differentiate EC from natural evolution. For the measurement of the rate of evolution, however, it offers a trap: to go entirely phenotypic, since in order to investigate the performance of a computational model, the rate of evolution is mostly measured by the speed of fitness function improvements. Other ad hoc methods are also utilized in EC, like the efficiency of algorithms and CPU time.

Another method, however at a deeper level than simple fitness function improvement, deserves mentioning: Bedau and Packard [64], for instance, propose a method for visualizing evolutionary adaptation. This method is useful to identify and measure the capability of creating adaptation during evolutionary processes. It is based on calculating evolutionary activity statistics of components in an evolutionary system, such as the numbers of particular genes (or alleles) in each generation and the persistence of these genes (or alleles) during evolution. During a decade of extensive development, the notion of *evolutionary activity* has been applied to various scales of genetic components, including alleles, allele tokens, phenotypic equivalence classes of alleles, and whole genotypes, in both artificial evolutionary systems and in the biosphere. In their more recent work, these authors emphasize two aspects for evolutionary adaptation: the *extent* and the *intensity* of evolutionary activity [65, 66]. The extent of evolutionary activity refers to how much of an adaptive structure is present in an evolutionary system, while the intensity concerns the capability of generating new adaptive structures. The measures of *cumulative evolutionary activity* and *mean cumulative evolutionary activity* characterize the extent of a system's evolutionary adaptation. On the other hand, *new activity* is a measure of the intensity of a system's evolutionary adaptation. Evolutionary activity can

be quantified and visualized during evolutionary adaptation. Its derivative is the concentration of a component's current presence, and its second derivative can be argued to reflect the rate of evolution at a particular time. Evolutionary activity is also claimed to be a straightforward method for studying evolvability [65]. The argument is that, since a system with high evolvability can create highly adaptive variation, the quantification of evolvability can be achieved by measuring the levels of extent and intensity of evolutionary activity.

1.3. Observations. It has been observed both in the paleontological record [67] and, more recently, through studies of molecular evolutionary systems [68] that the rate of evolution in biological systems is greatly varying. At times *selective sweeps* pass through a population that all but wipe out certain less advantageous alleles, while at other times seemingly nothing happens in terms of evolutionary changes. Thus we can legitimately speak of an acceleration of evolution under certain conditions, and of a slow-down under others.

In EC, on the other hand, the state of the art can be summarized by the observation that under most conditions, algorithms tend to show exponential decay in progress toward an optimum with often a painfully slow convergence for a large part of runs, or, alternatively, a *premature convergence* of the algorithm to the detriment of the produced solution, resulting in a stagnation of the search algorithm before it has reached an acceptable outcome. This has been realized to be related to the *record dynamics* shown in many natural and human systems [69, 70]. Record dynamics refers to the slow-down of records, for instance, in competition sports events, where after some time records become more and more difficult to break, due to the unchanging human physiology and the limitations this physiology imposes on achieving certain targets.

Contrast that with the world of natural evolution, where there is always a way to beat previous opponents, and to evolve in another direction that allows to increase fitness in some unforeseen ways. Surely, the implicit definition of fitness plays an important role here, as it allows enormous flexibility in achieving function. Further, the fact that the environment is permanently changing can be expected to be a key contributor to the evolvability in natural environments. Finally, the ability of living tissue (an intentionally vague term) to assemble in a hierarchical fashion, starting from atoms and molecules upward into ecosystems, provides building blocks and interactions of great richness that allows evolution to progress at different speeds, and notably to accelerate under favorable conditions.

A number of detailed observations on the factors that can accelerate evolution in the living world have been made in the past. Simon [71] raises the "nearly completely decomposable" property in multicellular organisms and proposes it to be an important property that can lead to faster fitness increases. In research on yeast genes, Gu et al. [72] report rapid evolution of gene expression and regulatory divergence after gene duplication. Gene (and segmental) duplication events contribute substantially to genomic and organismal evolution, since they provide abundant material for mutation and selection to generate new gene functions in a modular way. By studying the recent nucleotide substitutions in human evolution, Hawks et al. [73] find that, as a population becomes more adapted to its current environment, the rate of adaptive evolution slows down. However, a growing population size can provide the potential for rapid adaptive innovation. Thus, enlarging the population size and changing environmental conditions can both promote the rate of adaptive evolution. Kashtan et al. [74] confirm in a recent report that a varying environment can speed up evolution in an artificial evolutionary system. Other properties and techniques on the acceleration of evolution have been also investigated in biology and computing.

This review discusses evolvability and methods for accelerating artificial evolution by drawing ideas from complex natural systems. Notions from biology are introduced and their potential in designing new algorithms in EC is discussed. The review is organized along the work-flow of evolutionary algorithms. Section 2 starts with the characteristics of populations; variation operations are investigated in Section 3, separated into genotypic variation, phenotypic variation, and the transformations between them. Selection is discussed in Section 4, together with notions of fitness. The review concludes with a summary in Section 5.

2. Population

The general idea of EC is to adopt mechanisms of evolution from nature. In Darwin's theory of evolution, both the notion of variation and of natural selection are based on natural populations. However, populations simulated in a computer are usually simplified from their natural counterparts. A notable difference between natural and simulated population systems is that no identical individuals exist in a natural population, whereas this is allowed and most often the case with simulated populations. Tiny variances are considered an essential aspect of natural populations as they lead to the large diversity in natural evolution via amplifying effects produced by selection. Hence, more details should be taken into account also in computational populations that will ultimately allow a better differentiation of individuals. Because the representation chosen for individuals and the size parameter of a population can affect the performance of a computational model, it is an essential step in EC to determine these features of the simulated population.

2.1. Representation. The first step for setting up evolution with a population is to decide on the representation of evolutionary individuals. Each individual should be encoded as a candidate solution to a given problem, which subsequently determines the search space of the algorithm. Therefore, choosing a representation is important because it predicates the input to the search process that should produce a satisfactory output. Here, we highlight a two biological mechanisms, a protection mechanism for robust information preservation, and a communication mechanism for information interaction between different molecules.

2.1.1. Robustness and Redundancy. Living systems may seem wasteful and luxurious to computer scientists. The most distinguishing aspects of biology compared to other natural sciences are complexity and diversity, which are indeed of central concern to biologists. In the face of cruel competitive circumstances, organisms show great redundancy and resilience. Redundancy exists at different levels in natural organisms, including the genomic, transcriptomic, and phenotypic levels, all for the benefit of the robustness of the organism.

We adopt Wagner's definition for robustness here.

Definition 3 (robustness). The robustness of a biological or engineering system is its capability to continue functioning in the face of genetic or environmental perturbations [25].

In biology, the genome of an organism is defined as the information encoded in DNA sequences and inherited from generation to generation. The double helix structure of DNA sequences itself is a form of protective redundancy of genetic information. Genomes carry genes and other noncoding DNA sequences. A gene is a string of base pairs grouped by a function that is embodied in a protein or polypeptide (protein fragment). Noncoding DNA sequences, formerly called "junk DNA", are not expressed as proteins, although they might be transcribed into RNA and involved in manufacturing proteins or controlling that process. All in all, genes are only quite small a fraction of the entire genome [75], with more than 98% of the human genome, for instance, being noncoding DNA sequences [76]. Furthermore, even a gene sequence itself is divided into exons and introns, where exons directly determine the protein amino acid sequence but introns do not. Nevertheless, these noncoding DNA sequences are not useless. Recent biological discoveries show that they play an important role in the regulation of gene transcription [77]. Regulation mechanisms will be discussed later in Section 3.3.1. Wren et al. [78] find that tandem-repeat polymorphisms in genes are quite common, and that such polymorphisms can enhance the ability of some genes to respond rapidly to fluctuating selection pressure. The mechanism of gene duplication will be discussed in detail in Section 3.1.1. Moreover, diploid organisms have two copies of each chromosome, one copy inherited from each parent. Recent research has also found that a large number of DNA segments appear in more than two copies. *Copy Number Variations* (CNVs) in human and other mammalian genomes discovered lately account for a substantial amount of genetic variation other than single nucleotide polymorphisms (SNPs) [79–82]. CNVs and SNPs are considered to substantially contribute to genotypic variation, a phenomenon that will be discussed in detail later in Section 3.1.1.

Further down the line toward the phenotype is the transcriptome which describes the set of all transcribed RNAs in cells. In the human transcriptome, the proportion of transcribed nonprotein-coding sequences is large and shows great complexity [83]. Substantially more DNA is transcribed than is translated, and only a small proportion of mRNAs are translated into proteins. The rest is called *noncoding RNA* or

TABLE 1: Mechanisms responsible for creating redundancy and antiredundancy at the cellular level. (Adapted from Krakauer and Plotkin [86].)

Redundancy	Antiredundancy
Gene duplication	Overlapping reading frames
Neutral codon usage	Nonconservative codon bias
—	Gene silencing
Polyploidy	Haploidy
Multiple regulatory elements for n genes	Single regulatory element for n genes
Chaperone and heat shock proteins	—
Checkpoint genes promoting repair	Checkpoint genes inducing apoptosis
Telomerase induction	Loss of telomerase
Dominance	Incomplete dominance
Autophagy	—
mRNA surveillance	—
Bulk transmission	Bottlenecks in transmission
Molecular quality control	—
tRNA suppressor molecules	—
Modularity	—
Multiple organelle copies	Single organelle copies
Parallel metabolic pathways	Serial metabolic pathways
Correlated gene expression	Uncorrelated gene expression
DNA error repair	Loss of error repair

ncRNA. About 98% of all transcribed sequences in humans are of this type [84]. Although many of the functions of these noncoding sequences are unclear, the high complexity of the transcriptome hints at its importance in the mechanisms of organizing gene expression in a robust way [85].

Krakauer and Plotkin [86] go further and propose the new concept of *antiredundancy*. In their opinion antiredundancy emerges as does redundancy in cells, and natural organisms would be able to modify the redundancy properties of genotypes during evolution. Table 1 shows a summary of observed mechanisms responsible for both redundancy and antiredundancy at the cellular level. Mechanisms for redundancy mask the phenotypic effect of mutations and allow mutants to stay in populations, while mechanisms for antiredundancy enhance the efficiency of local selection to remove damaged components.

Going even further, we finally arrive at the phenotype: redundancy at the phenotypic level lies in an organism's robustness against intrinsic or environmental changes. With low robustness, a species will gradually decline and finally go extinct due to lethal mutations because random mutations in the genome usually cause deleterious changes with a potential to destroy the offspring.

It seems that robustness and evolvability have a contradictory relationship to each other. When a system has

high robustness in its genome, it can be tolerant to intrinsic or environmental changes, but that should leave it less evolvable, as variation would be masked, and vice versa. In recent contributions, Wagner [50, 54] resolves this apparent contradiction. He distinguishes robustness and evolvability as quantities at both the genotypic and the phenotypic levels. If one considers genotype, the more robust a genetic sequence is, the less innovation this sequence will produce. However, robustness and evolvability are characteristics of an entire system and if investigated at phenotypic level show a strong correlation. A system with high phenotypic robustness harbors a great number of "neutral" variations that have no functional effects. These neutral variations do not change phenotypic function during relatively static evolutionary periods but may be able to generate adaptation later under certain genetic or environmental changes. Thus, a system with high phenotypic robustness simply masks changes but provides great potential for phenotypic innovation in the future, for example, if conditions change and previously neutral changes suddenly have an effect. This is the core of the argument that high robustness and high evolvability are in fact correlated in nature [54], and this has been supported in many subsequent research [87–91]. Specifically, Draghi et al. [92] go further and for first time quantify the effects of robustness/neutrality on adaptation in an evolving population. They suggest a complex relationship between robustness and evolvability, which depends on the topology of the genotype network. Their results indicate that if the genotype space has no epistatic effects, a more robust population will have less evolvability. With epistasis, on the other hand, they find a nonmonotonic relationship between robustness and evolvability, that is, evolvability is the highest at an intermediate level of robustness.

Redundancy is wide-spread in natural organisms as an efficient protection mechanism against internal or environmental changes, whereas in EC models components that do not seem to be immediately relevant are often considered superfluous. In recent years, however, representation redundancy has arisen as a by-product of computational evolution and has attracted increasing interest from EC researchers.

Definition 4 (representations redundancy). In genetic and evolutionary algorithms, representations are redundant if the number of genotypes exceeds the number of phenotypes [93].

Rothlauf and Goldberg [93] examine the effects of redundant representations on the performance of an EC system both theoretically and empirically and propose that redundant representations can increase the reliability and efficiency of EC models. Specifically in genetic programming, representation redundancy is usually identified as *introns* (or noneffective, neutral code) [1] in programs. Researchers have investigated both the positive and negative effects of introns [94–97], and a positive relation between neutral code and evolvability in genetic programming has been suggested. The important role of redundancy in evolvability has now been realized. We might, therefore, consider designing protective

redundancy into our algorithms to make them resilient against changes while improving adaptivity. Such capabilities certainly complicate the algorithms but may be worthwhile if the resulting robustness can generate higher evolvability when applying intense pressure to produce adaptive responses. Evolution might even be accelerated because the system has a quick and robust reply to evolutionary pressures. With the growth of computational power available today ideas like these can be more easily explored than before.

2.1.2. Molecular Interaction. Natural living systems are remarkably diverse starting from so simple organisms as bacteria to highly complex creatures such as primates. This diversity is not the result of vastly different chemical constituents of organisms. In fact, many species carry out similar metabolic, cell division and replication processes under similar assembly principles [98]. The differences that distinguish species are caused by the arrangement and distribution of basic building blocks [99] and molecular interactions contribute significantly to these organizational mechanisms.

Molecular interactions in a cell happen between the same type of molecules, such as protein-protein interactions, or between different types of molecules, such as protein-DNA or RNA-protein interactions. Signals can also be sent between and responded to by cells in multicellular organisms. Molecular interactions can be triggered by energy supply, for example, in *metabolic pathways*, chains of interactions catalyzed by enzymes, or triggered by external stimuli, for example, *signaling pathways* that enable communication through the cell membrane [100]. Proteins are not only a product enabling various organismal structures but also work as control factors in various processes from the synthesis of a cell, metabolism, gene regulation, to sexual reproduction.

Metabolism is a key process to maintain the growth and reproduction of cells. The metabolic network of a cell is an elaborated set of numerous chemical reactions catalyzed by enzymes [101]. Different types and amounts of enzymes are produced according to different energy supplies, and these enzymes will determine different metabolic pathways by their catalysis. In the process of gene expression, the function achieved can be controlled by molecular interactions. For instance, the process of how a parsimonious bacterium responds to food supplies during metabolism shows a simple genetic switch mediated by molecular interactions. Since the metabolic pathways of bacteria are much simpler than those of multicellular organisms, the regulation of gene expression is more easily understandable in bacteria. The phenomenon of *enzyme induction* [22] describes the adaptation of a bacterium to material supplies by producing varying amounts of enzyme. What triggers this production and how does this mechanism work? The Jacob-Monod model (shown in Figure 1) first described the regulation mechanism of inhibiting or repressing genes by inhibitory proteins, called *repressors* in bacteria. The binding of lactose to a repressor enables the production of RNAs by removing the repressor from its binding sites on the gene sequence where
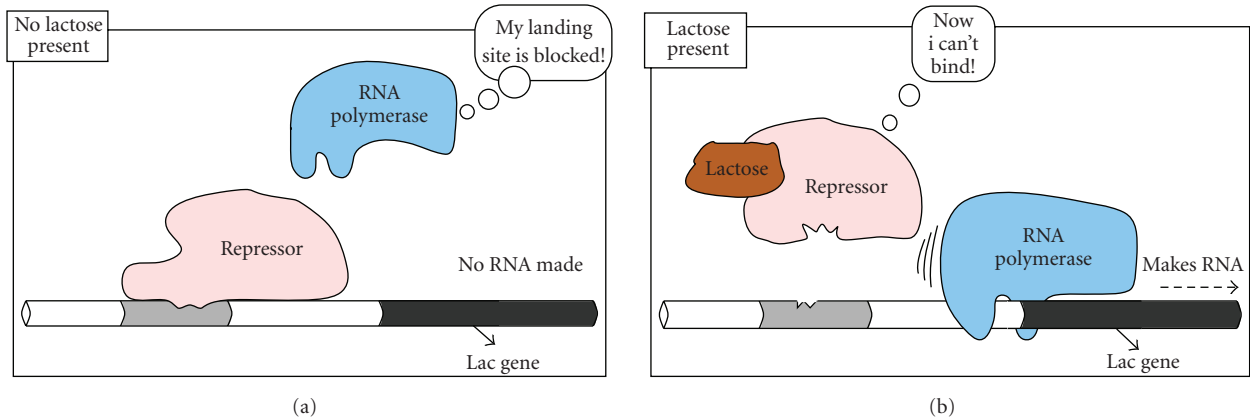
FIGURE 1: The genetic switch in the Jacob-Monod model. A specific repressor protein acts as a switch. When it binds to a DNA site near the gene encoding beta-galactosidase, the RNA polymerase protein cannot bind nor can it synthesize RNA from the gene. The gene is turned off. When lactose is present, it binds to the repressor and keeps it from the DNA site. The gene turns on. (Adapted from Kirschner and Gerhart [22].)

RNA polymerase can bind. However, this is not a simple on-off switch model. The continuity lies in the binding duration which determines the rate of protein synthesis. Therefore, if more sugar is absorbed during metabolism, more protein is synthesized by RNA translation. This simple sugar metabolism model captures the mechanism of how a repressor affects gene function. The enzyme here works as a trigger for the protein synthesis process under various molecular interactions. In addition, most enzyme effects are sensitive to ambient temperature [102], which is an important parameter to control metabolic interactions.

Signaling and cellular responses to signals are complex. These responses are controlled by a plethora of positive and negative feedback loops. The presence of feedback complicates the simple picture of a linear pathway but is an essential part of the signaling process [98]. This makes signaling pathways involving molecular or cellular communication a network-like structure, with complex regulatory processes at work. The cellular infrastructure of eukaryotic organisms is only a few times larger than that of bacteria, but the complexity of their signaling network control differs greatly, by orders of magnitude. The linkage between various parts of the gene expression apparatus in eukaryotic organisms is weakened by a far less precisely defined control than that found in prokaryotic cells [47]. For instance, geometric requirements for binding sites are significantly relaxed in eukaryotic gene regulation. A repressor does not have to bind at the exact position of a target but needs only to bind in the neighborhood. By lowering constraints for cooperation, such a weak linkage also enables potential interactions between different gene sequences. Signaling between cells is possible only after a sufficiently large number of repressors participate simultaneously. A single signal may incur a very complex response [49]. Allosteric proteins, which have multiple sites for interaction, also make gene expression more flexible because they have different sites for different functions. Regulatory decisions on which genes are transcribed when, where, and under what circumstances

makes eukaryotic cells well conserved but enormously adaptive to generate new phenotypes in changing environments [103].

Computational models have already been used to analyze and understand complex multi-input/output and higher-order signaling systems have been examined in bioinformatics [104]. In contrast, current EC models are mostly limited to representing evolutionary material based on the infrastructure of natural organisms, while disregarding the vast potential of interaction mechanisms for regulation and signaling at both the molecular and cellular levels. The absence of such mechanisms in EC, however, points to significant research opportunities in this area.

2.2. Population Size. After the encoding of an individual is determined, a population is set up. Several features of a computational population are tightly connected to its evolutionary capabilities, the most relevant of which is population size itself.

In nature, different species have different population sizes, a characteristic that plays an important role in evolution. In the living world it is common that smaller groups constituting species evolve faster, though smaller groups have a higher probability of becoming extinct, while species with larger populations evolve slower and can stay unchanged for relatively long periods. However, neither a small nor a large population size is unconditionally beneficial in evolution. The relation between them should be understood in different scenarios.

The study of *population genetics* was formulated by Fisher [105], Haldane [106], and Wright [107]. It focuses on gene frequency changes in populations under the effects of natural selection, mutation, genetic drift, and population size fluctuation. In this field, scientists have examined the role of population size in molecular evolution using mathematical analysis. The rate of molecular evolution is usually measured by the nonsynonymous to synonymous substitution ratio k_a/k_s, discussed in Section 1.2.

Decades ago Kimura [108] proposed a strong dependency of the rate of molecular evolution on population size. More recently, Gillespie [109, 110] has conjectured that there is only a very weak dependency on population size. Somewhat in the middle between these opinions, Ohta [111] finds population size to be related to the rate of evolution under particular assumptions regarding mutation types. The *nearly neutral theory* of molecular evolution proposed by Kimura and Ohta [112, 113] predicts that there is a substantial number of nearly neutral mutations (including slightly deleterious and slightly advantageous ones) in molecular evolution, and that these contribute to evolution by providing potential for future phenotypic innovation. Ohta [111] predicts that population size affects the rate of evolution under various mutation scenarios. If most mutations are deleterious, a smaller population can evolve faster, because the chance of a slightly deleterious mutant being favored by selection is greater within a smaller population and these nearly neutral mutations bring genetic variation and may further trigger phenotypic innovation. In contrast, if mutations are mostly advantageous, the rate of evolution in a larger population is greater. If most mutations are neutral, the evolution rate is nearly independent of population size. Since in general random mutations are more deleterious than advantageous in natural systems, species with a small population size usually evolve faster.

A number of studies focus on testing the relation between population size and evolution rate by using comparisons. Island endemic species usually have small population sizes because they are restricted to a limited geographical region. Woolfit and Bromham [114] study species on islands in support of the effect of population size on the rate of molecular evolution. They compare island endemic species to closely related species on a nearby mainland and find that island endemic species have a significantly higher nonsynonymous to synonymous nucleotide substitution ratio than their counterparts on the mainland. This result indicates that a decrease in the population size will lead to an increase in the rate of evolution. Wright et al. [115] study tropical species which are generally regarded to have a rapid molecular evolution rate due to several factors such as latitude and climate. It is believed that tropical organisms possess great species richness and dynamics with small but highly diverse populations [116]. However, there are also exceptions in that increasing population size can accelerate evolution as well. By studying the recent rapid molecular evolution in human genomes, Hawks et al. [73] suggest that if a population is highly adapted to a current environment, evolution will become stagnant. Under these circumstances a growing population size can provide the potential for rapid adaptive innovation. Thus, enlarging the population size under chaotic environments can promote the rate of adaptive evolution.

Population size is also involved in research on genome robustness. Visser et al. [85] postulate that the population size should be sufficiently large for selection to be effective to evolve the robustness of a system. Small populations have difficulty to achieve this robustness. In a different study Krakauer and Plotkin [86] find, however, that small populations will also favor evolving robustness by increasing genetic drift pressure and a buffering mechanism of hiding mutations from being reduced by selection. This hypothesis is supported by Elena et al. [117]. Among the different authors, there is agreement that the effect of population size, either large or small, varies in different models.

In general, the size of populations in EC is orders of magnitude lower than the size of populations of many naturally occurring species, especially those of simpler organisms like bacteria. The commonly adopted population size in EC varies from tens to thousands, with a few exceptions. Genetic Programming generally uses relatively large populations, due to the more complex and nonlinear fitness landscape than can be found in other branches of the EC family. However, the size of GP populations run is limited by resource constraints in the range of hundreds of thousands. Some of these algorithms have to run on parallel machines or on GPUs, since the evaluation of a large population of individuals requires enormous computational power (see, e.g., [118–120]). An order of magnitude like that of humankind, a billion individuals, is unheard of in EC approaches, which already points to a vast potential for doing research on EC in the future. A whole landscape of EC methods might emerge with populations that are large.

As a result of the use of small population sizes in the EC community, efforts have been dedicated to the optimization of population size [121], since a high correlation between population size and the performance of an EC algorithm is presumed. The challenge is that adapting population size is problem-specific and to date it is still unclear how to estimate the relation among various EC parameters. In general, current work on this topic concentrates on two tasks: (i) initializing a proper population size prior to a run, and (ii) adjusting population size during a run. Most theoretical work on population size initialization is based on Goldberg's component decomposition approach and the notion of *Building Blocks* [122, 123]. With many other publications, these contributions propose to choose the population size according to the "hardness" of a specific problem. They state a general principle in setting population size: the more difficult a problem is, the more diversity is required and the larger the population should be.

In the meantime, it has been found that even for a specific problem the requirements for population size may differ during different stages of evolution. As a result, empirical methods for adjusting population size during a run have been proposed, such as the *Genetic Algorithm with Variable Population Size* (GAVaPS) proposed by Arabas et al. [124], the *parameter-less GA* by Harik and Lobo [125], the *Adaptive Population size Genetic Algorithm* (APGA) by Back et al. [126], and the *Population Resizing on Fitness Improvement GA* (PRoFIGA) by Eiben et al. [127]. However, mechanisms for dynamically adjusting population size in EC are much simpler than those found in nature, in that a fluctuating population size still has little to do with mutation and selection patterns in different evolutionary stages. This relation requires further exploration as it seems to be a promising indicator for population size adjustment during a run.

3. Variation

Mutation and recombination operators are a main aspect of evolvability, since they generate the necessary variation among individuals that later can be acted on by cumulative selection processes. Due to the complex mapping process from genotypic to morphological level in biology, genotypic and phenotypic variation will be discussed separately.

3.1. Genotypic Variation. Genotypic variation generally means changes to DNA sequences in both protein-coding and noncoding regions in the form of point mutation and gene rearrangement. Gene sequences are highly conserved against lethal changes that would likely lead to destructive consequences otherwise because a tiny mutant at the genetic level can cause a great change in function [22]. In contrast, changes to the regulatory or noncoding part of sequences are considered more able to increase adaptability and plasticity of a system. In this section, we will discuss the general form of mutation first and then gene duplication as the most important form of rearrangement, followed by a comparison between point mutation and gene rearrangement.

3.1.1. Mutation. Although there can be many definitions of mutation, we here adopt one that emphasizes the primary difference to recombination, namely, that works with material from just one individual organism.

Definition 5 (mutation). Mutation is the process that creates new genetic material from the addition or multiplication of stochasticity in various forms to some original genetic material of an organism.

Point Mutation. Searching for the essential driving force of evolution has been a central topic in evolutionary biology. Since Darwin declared that natural selection is the main force of evolution, controversies have arisen on different aspects of this explanation. In modern biology, the two main schools of thought are selectionism and neutralism [128]. Some scientists argue that genotypic variation is maintained by selection, which is the central perspective of *neo-Darwinians*. Other evolutionists insist that high genotypic variation can be explained as a result of neutral mutations. In either case, mutation is accepted as a major mechanism to generate genotypic variation.

Mutation can happen anywhere on a DNA sequence, that is, in either coding or noncoding regions, and may consequently cause functional, regulatory, or structural changes, or no changes at all. The neutralist hypothesis is that the majority of observed sequence variation stored in the population is neutral. This is due to the compensating mechanisms of biological systems [128]. Most new mutations are deleterious, a few are advantageous, and many are neutral. However, most of the extant polymorphism observed in populations is the neutral variants. Deleterious mutations have been purged and advantageous mutations have swept through the populations.

What triggers mutation and what is the relation between mutation and selection? Does selection pressure indeed generate new mutations or simply allow existing mutants to be fixed faster than before? Research on mutation under selection has received wide interest since Darwin's time, but controversies have arisen regarding the effect of selection pressure on mutation, and different models have been proposed in the meantime [129]. It is now believed that it is impossible to separate any form of mutation from the effect of selection. In order to investigate "directed" mutation pathways Roth and Andersson [130] define *adaptive mutations* as fitter mutations that arise under selective conditions. In subsequent work [131–133], they propose a gene duplication-amplification model to study the mutagenesis stimulated by enhancement of selection. In addition, a recent study by Weinreich et al. [134] on the effects of Darwinian selection on random mutation argues that environmental selection can make some multistep mutation pathways unaccessible. By studying "five point mutations" in a lactamase allele that can increase bacterial resistance to an antibiotic, several mutation pathways are in principle possible for these mutations. After calculating the different probabilities of these pathways, their experimental results show that under intramolecular interactions that increase the fitness of proteins, only a small number of pathways are really accessible. This is quite an interesting result because mutations might be channeled by some unknown fitness-increasing principle(s) and the resulting proteins might be reproducible and even predictable. These feedback and interaction mechanisms may reduce the harm that mutations could bring to an organism. This point of view also conforms to Kirschner and Gerhart's definition of evolvability [47], which they define as "the ability to reduce the potential deleterious mutations and the ability to reduce the number of mutations needed to produce phenotypically novel traits". If mutations can be channeled, fewer changes might be needed to generate a required adaptation and, therefore, evolvability would be improved by this reduction in cost of mutations.

In EC, mutation is regarded as an important exploratory operator. Artificial evolutionary search should be good at both exploring suitable genetic novelty and maintaining successive improvements. Holland [7] discusses this principle as the tension between "exploration" and "exploitation". The mutation rate is important to keep this balance, and it has already been studied as an evolvable parameter contributing to evolvability. Bedau et al. [51, 135] divide evolutionary adaptation conceptually into two stages: the novelty stage, where an evolving system enhances its adaptability against a changing environment, and the memory stage, where the evolving system is building up this adaptability through incremental improvements. By providing a simple two-dimensional model, Bedau et al. postulate that the mutation rate should increase during the novelty stage and decrease during the memory stage. This fluctuation of mutation rates is able to keep the balance between evolutionary novelty and memory and thus increases the evolvability of adaptive systems.

However, compared to natural evolutionary systems, genotypic variation in computational systems is somewhat
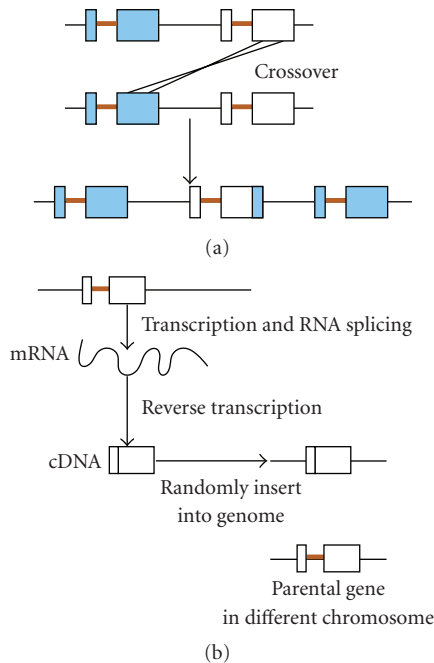
(a)

(b)

FIGURE 2: Two common modes of gene duplication. (a) Unequal crossover, which results in a recombination event in which the two recombining sites lie at nonidentical locations in the two parental DNA molecules. (b) Retroposition, which occurs when a message RNA (mRNA) is retrotranscribed to complementary DNA (cDNA) and then inserted into the genome. Squares represent exons and bold lines represent introns. (Adapted from Zhang [139].)

arbitrary and not as adaptive. First, the fixation process of mutations is not simulated appropriately in most EC algorithms, because all changes to individual sequences are mostly translated into phenotypic properties. Recovery or repair mechanisms are usually not applied to individuals suffering deleterious mutations, which make those individuals unfavored during the selection process. Second, the selection-driven mutation pathways found in natural systems are an interesting direction to explore for computational models and should be considered in future research in EC.

Gene Duplication. Gene duplication is an important mechanism creating new genes and new genetic subsystems. This mechanism has been recognized to generate abundant genetic material and contributes substantially to biological evolution [136]. A large number of duplicate genes have been discovered to exist in vertebrate genomes [137], and a repeated number of whole genome duplications have been established as key events in evolutionary history [138]. In modern biology, gene duplication and its subsequent function-specialized divergence are widely believed to be a major reason for functional novelty.

Gene duplication is usually generated by unequal crossover or retroposition [139] (see Figure 2). Unequal crossover is similar to but different from normal crossover that occurs when two chromosomes exchange a proportion of DNA at the same locus in base pair sequences. Unequal crossover happens if this exchange occurs in

different loci, with the consequence that duplicate genes appear in one chromosome while the other turns out to contain *pseudogenes*. Retroposition happens when an mRNA is retrotranscribed into a complementary DNA (cDNA) and then inserted into the original genome. Besides such gene duplication, duplication at other scales in cells has been discovered recently [138, 140], including segmental duplication and whole-genome duplication. Here, we only consider gene duplication. The main products of gene duplication are called *paralogous genes*, a type of homologous genes. Homologous genes have two main categories, *paralogs* and *orthologs*. Paralogs are results of gene duplication and code for proteins with different functions. Orthologs are the products of speciation events and the proteins they code for serve similar functions.

Once a gene duplication has occurred, a complex fixation process on the duplicate genes takes place. *Purifying selection* and *gene conversion* are the main pressures affecting the survival of duplicate genes [141]. Most duplicate genes become pseudogenes after one or more mutations disable them and no promoting function is yielded. However, multiple copies of identical genes can, after duplication, promote functional redundancy against fatal changes. The process of *pseudogenization* is reported to occur in the early stages of a rapid evolution [142] process, with evidence of many pseudogenes found in the human genome. Other duplicate genes are changed by selection pressure and functional divergence. Subfunctionalization and neofunctionalization are the two main mechanisms of functional divergence [139]. In *subfunctionalization* of two gene duplicates, shown in Figure 3, each copy adopts a different aspect of the function of the original gene. Both copies will be stably maintained because both aspects of the function are indispensable. Subfunctionalization leads to functional specialization by dividing multifunctional genes once the newly emerged genes perform better. Alternatively, some relatively new function can also evolve after gene duplication [143], and this process is called *neofunctionlization*. This has been termed the Dykhuizen-Hartl Effect [144] earlier, where a random mutation is preserved in the duplicated gene by reducing selection pressure due to functional redundancy that results from gene duplication. Such mutations may accumulate and induce a genetic function change depending on conditions of the (dynamic) environment. New adaptive functions may thus be generated and later preserved during evolution. By possibly creating novel functions and allowing evolution under fewer constraints, neofunctionlization is an important consequence of gene duplication.

In brief, the mechanism of gene duplication contributes substantially to genomic and organismal evolution. It provides abundant material for mutation and selection and allows to specialize function or generate completely new functions. The acceleration of protein sequence evolution after gene duplication has recently been confirmed in research on yeast genes by Gu et al. [72]. The authors use an additive expression distance between duplicate genes to measure the rate of expression divergence, and rapid evolution of gene expression as well as regulatory divergence after gene duplication is observed.
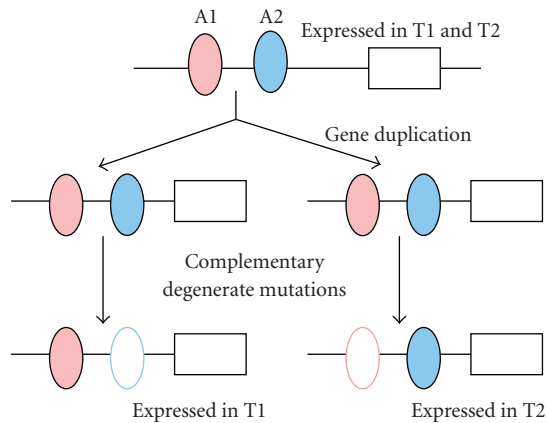
FIGURE 3: Division of expression after gene duplication. Squares represent genes, closed ovals represent cis-acting elements that regulate gene transcription, and open ovals represent deactivated cis-elements. Consider a gene that is expressed in tissues T1 and T2, with a cis-acting regulatory element A1 controlling the expression in T1 and A2 controlling the expression in T2. Following gene duplication, one daughter gene might lose the A1 element whereas the other gene might lose A2, so that each is expressed in only one of the two tissues. (Adapted from Zhang [139].)

One key idea how gene duplication can speed up evolution is Altenberg's *constructional selection* [145, 146]. The idea is that gene duplication enriches the genome with genes that are good at increasing fitness when duplicated. This is a second-order effect that can be considered to contribute to evolvability. For a more general review, see [147].

In summary, the mechanism of gene duplication can considerably increase evolvability of a system by reducing the cost of mutations. In EC, the idea of using gene duplication and deletion operators was proposed some time ago. Those operators are in general based on the method of variable-length genotypes and are executed with predefined duplication or deletion probability [46, 148–151]. Unfortunately, so far only application-oriented work has appeared with different representations [152], and a common framework for this concept is missing. More details of gene duplication in biology should be taken into account to benefit computational evolution. In particular, the question of how gene duplication reduces the limitations of mutation and selection and in the process promotes evolvability needs to be studied. Is there a way to implement functional specialization and innovation through gene duplication in EC?

Point Mutation versus Gene Rearrangement. A point mutation occurs when a base on a DNA sequence is changed into another base at the same locus. Gene rearrangement is a change in the order of a DNA sequence on a chromosome. This change can be an inversion, translocation, addition, or deletion of genes. Earlier research focused mostly on *Single Nucleotide Polymorphisms* (SNPs) in genomes due to the enormous complexity of genetic sequence analysis, but gene rearrangements have always been believed to contribute to

evolvability, possibly even more than simple point mutations [153]. Recent development of technology has now facilitated the shift in focus from a locus-based analysis to a genome-wide assessment of genotypic variation [79, 154].

Genetic rearrangements rather than point mutations can maintain the connective information carried by gene sequences. Because genes form networks of functional control, rearrangement is better able to preserve internal structures. Genetic changes are highly constrained by gene sequences and gene rearrangements occur far more frequently than point mutations.

The ubiquity of *Copy Number Variations* (CNVs) has been realized recently in mammalian genomes by different groups of biologists, such as Sebat et al. [81], Iafrate et al. [80], and Tuzun et al. [82]. CNV is regarded as a predominant type of genotypic variation leading to vast phenotypic diversity in mammalia. CNVs show that large segments of DNA, with sizes from thousand to millions of base pairs, can vary in copy number of genes. This variation can lead to protein dosage differentiation in the expression of genes, and CNV is therefore regarded as being responsible for a significant proportion of phenotypic variation [79]. The mechanisms that create CNV have not yet been clearly understood, but some hypotheses have been proposed in the literature. Fredman et al. [155] and Shaw and Lupski [156] propose that CNV might be the result of large segmental gene duplications or nonhomologous recombination events.

Recent bioinformatics research uses statistical and computational tools to analyze chromosomal evolution by a comparison of genome-rearrangements between sequences of related species [157]. Although the biochemical mechanisms of gene rearrangement are still far from being fully understood, we believe it is time to start using such rearrangement operations in computational models in EC. Particularly, the recent discovery of CNVs requires attention by computer scientists, in order to achieve similar benefits in EC.

3.1.2. Recombination. Recombination has been considered both as an exploratory and as a stabilizing operation in biology and in EC. Here we emphasize the origin of the genetic material being used for new combinations. Due to the size of search spaces, both effects are possible.

Definition 6 (recombination). Recombination is a process that generates combinations of existing genetic material from a multitude of organismic sources.

Recombination is regarded as an important force shaping genomes and phenotypes. Since some highly efficient and accurate computational methods can be used in biology, analysis of gene recombination has made much progress by way of comparing aligned genome sequences. These comparisons facilitate a better understanding of several aspects of genetic and evolutionary biology, notably genotypic and phenotypic variation and genome structures [158].

Recombination exchanges genetic material between two DNA sequences swapping strands between one or multiple crossover points. Recombination can occur on homologous

or nonhomologous sequences. The former is more prominent in research because it is more common and efficient in generating adaptation in nature. Generally, research on recombination focuses on prevalent eukaryotic organisms rather than prokaryotes, which do not have the sex property. Unequal crossover is fairly rare and may lead to duplication or loss of some genes (discussed in Section 3.1.1) and other results [159]. Combination events can take place between different gene sequences, as in *intergenic recombination*, or between alleles on the same gene sequence, as in *intragenic recombination* [158]. Despite various forms of recombination, their outcome is crossover at one or multiple points and a swapping of fragments of genetic sequences.

Kondrashov [160] proposes in his *deterministic mutation hypothesis* that sexual recombination can remove deleterious genes. It is generally believed that most nonneutral mutations are slightly deleterious. Kondrashov suggests that sexual recombination can distinguish individuals with cumulative, slightly deleterious mutations, and the ensuing selection pressure can eventually remove those disadvantaged mutations. Further, Hadany and Beker [161] strengthen this perspective in their research on the evolution of obligatory sex. Their model supports that sexual recombination offers both short-term and long-term advantages to sexually reproducing individuals and has a positive effect on the physiological fitness of an organism.

The rate of recombination can significantly affect the rate of adaptation. It is usually higher than the rate of mutation, which implies that recombination introduces much less lethality to an evolutionary population than mutation. Instead, it advances evolution remarkably by stabilizing adaptive traits from parents to offspring. This contributes to evolvability in the same way as other purifying selection does because the bounds on epistatic interactions between loci get progressively strengthened through selection over generations. By drawing a recombination map of the human genome, Kong et al. [162] discovered that recombination rates vary in different regions of the genome. This variation is due to such functional features as gene density, other gene properties, and frequency of sequence repetitions. Recombination rates are also different in autosomes between different sexes. Recombination contributes to producing both genotypic and phenotypic variation and is able to repair DNA double strand breaks. *Sexual reproduction* is an important outcome of recombination.

In EC, recombination operations are considered an essential search strategy. Chromosome coding is much more flexible in computation than in nature, and thus, various recombination techniques have been proposed and studied, including double-parent and multiparent crossover [163], fixed-length chromosome and variable-length chromosome crossover [164, 165], and homologous and nonhomologous crossover [166–168]. High recombination rates are usually also adopted in computation because of its perceived efficiency in generating beneficial genetic and phenotypic variation. Elsewhere, adaptive recombination rates are proposed to strike a balance between exploration and exploitation [169]. In most of these adaptive recombination rate schemes, modification of recombination rates is based on fitness

value. Different from natural recombination mechanisms, most adaptive recombination rate proposals simply react to the current status of the search, in order to escape from local optima. However, rate adaptation in biology is much more complex and suggests other models for computation. For instance, the rate may vary among different individuals or in different modules serving subfunctions in the genome. Such function-specific recombination rates could also consider the method of "compartmentalization" for modularity (Section 3.2.2). The notion of *epistatic clustering* in contributing to evolution of evolvability has recently been studied [170]. Genetic linkage patterns between different loci are claimed to affect recombination rates, and the simultaneous optimization of different recombination rates on different traits would be realized by a method called epistatic clustering. Evolvability would be improved through coevolution of trait clustering and recombination mechanisms.

3.2. Phenotypic Variation.
As mentioned in Section 2.1.2, despite their vast phenotypic differences, metabolic processes and cell structures in bacteria and humans are quite similar [22]. What, then, makes humans so different morphologically from other organisms? It is the regulation and reuse of these structural elements in different combinations that generates different complex phenotypic outcomes. Unfortunately, the relation between genotypic variation and phenotypic variation is still not fully clarified in current biological opinion. Since selection acts on phenotypes rather than on genotypes, phenotypic variation should be used to explain the immense diversity among organisms. Here, we discuss several aspects of phenotypic variation. We leave the discussion of the mapping process between genotype and phenotype that controls the direction of phenotypic changes resulting from genotypic variation to Section 3.3.1.

3.2.1. Conservation and Relaxation.
According to Kirschner and Gerhart, evolution possesses two important features: conservation at the molecular level and relaxation at the anatomical and physiological level [22]. By conservation it is meant that the genetic components of organisms tend to maintain relatively stable structures; relaxation refers to the less constrained phenotypic diversification of organisms. The authors state that conservation on the genotypic level reduces the constraints on the phenotypic level.

In Darwin's evolutionary theory, all organisms have evolved from the same ancestor. After primal initialization and evolution, genetic structures of organisms are highly conserved during the course of billions of years [101]. This can well explain why the number of human genes is only a few times that of bacterial genomes but significant anatomical and behavioral differences exist between them. The surprisingly small number of genes in humans and other complex organisms demonstrates that the great diversity and complexity at the anatomical and physiological levels have to rely on and organize/reuse limited genetic material. When certain organisms need to improve their adaptivity in order to survive in a new environment, the regulation system only has to recombine existing mechanisms for the

generation of adaptive functions, which requires little or no new genetic material [47]. Not only are gene sequences highly conserved, but also the *core processes* of coordination of the genetic material are well conserved since the time they initially emerged [22]. These conserved core processes are used repeatedly for different purposes and functions under different circumstances, at different times, with different genetic material. The Baldwin Effect [171] explains that phenotypic variation is not generated out of the blue but through regulation of existing components in organisms: mutation simply stabilizes and extends what has already existed to improve somatic adaptability towards external stimulations.

This conservation mechanism can efficiently prevent lethal changes in genotypes and is an economic method to increase the adaptability of organisms. New material is not needed to adapt to changing environments, but few modifications will suffice.

Functional innovation is heavily constrained due to molecular interactions among various genetic components that are involved to produce a specific trait. If the participation of more genetic components is needed, it becomes harder for functions to change. In fact, relatively little genetic material is required to generate all proteins of organisms. Under selection pressure from a changing environment, organisms have to yield adaptive phenotypic traits to survive, however, and the highly conserved core processes mentioned above are used repeatedly to generate new cooperation among the conserved genetic material, bringing about fitter function and behavior. Relying more on the combinatorics of components is equivalent to relaxing phenotypic variation.

The relaxation on phenotypic variation has been highlighted as the notion of "deconstraint" in Kirschner and Gerhart's [47] research on evolvability which studies the mapping from genotype to phenotype. Enhancing phenotypic variability under changing environmental conditions allows nature more evolvability. Not only can deleterious changes be avoided, but also nonlethal genetic and phenotypic variation is indeed the material from which innovation can be generated.

Turning again to EC: what is the role of conservation and relaxation in EC? First, an economic use of genomes or building blocks can help to conserve genetic information. Second, it can be assumed that by reducing the constraints on changes to a phenotype the exploratory capability of a computational system to find better solutions can be enhanced. How such a process can be implemented in actual systems is presently unknown, but a worthwhile line of inquiry.

3.2.2. Modularity. Modularity is a widespread structural property of complex systems. It has attracted considerable interest from studies of both natural and artificial evolutionary systems and is regarded as strongly related to evolvability [45] and the acceleration of evolution [71].

Modularity exists at various levels, for example, at the level of gene expression or embryonic development. Here we adopt the definition of modularity proposed by Simon [172] in his research on hierarchies in complex systems.

Definition 7 (modularity). In a complex system, modularity refers to the property that a loose horizontal coupling exits between the entities at the same level of this system [172].

Simon [71, 173] further defines that "a system is *nearly decomposable* if it consists of a hierarchy of components, such that, at any level of the hierarchy, the rate of interaction within components at that level is much higher than the rates of interactions between different components". Although this "Near Decomposability (ND)" is attributed to a *vertical* separation while modularity describes the separable property of components horizontally at the same level, they seem closely related in that they both describe how a complex system is decomposed into subsystems.

The modularity property of genotype-phenotype mappings has been extensively studied in gene expression. It reduces harmful pleiotropic effects of gene expression and can lead to adaptive phenotypic variation. *Pleiotropy* is a general property of genotypic variation, expressing the fact that one change at the genetic level can cause a multitude of functional changes at the phenotypic level. Pleiotropy can generate both advantageous and disadvantageous results. Pleiotropy can sometimes generate unexpectedly improved functions but can also be harmful or even fatal to evolutionary systems [174]. Since a gene can affect multiple functions, optimizing one particular function at the phenotypic level inevitably incurs side-effects on other functions. Bonner [175] proposes the notion of "gene nets" by grouping gene actions and their products into discrete units during evolution. In general, for a given organism, the mapping from genotype to phenotype can be divided into modules such that the sets of genes in one module only affect the functions in that same module. The mapping is therefore decomposed into groups of independent "submappings". Bonner finds that the phenomenon of gene nets becomes increasingly prevalent as organisms become more complex. Wagner and Altenberg [45] investigate modularity in genotype-phenotype mappings from both perspectives, biology and EC. They interpret modularity as a means for dividing phenotypic traits into different "compartments" to reduce interference among different optimization modules. With such modularity, optimization of a function in one module has no effect on functions in other modules. As a result, pleiotropy can be confined to a known set of functions during evolution. Figure 4 shows a simple example of this idea of modular separation.

Wagner and Altenberg [45] further propose that modularity results from evolutionary modifications in natural organisms. In their view, the evolution of modularity follows two mechanisms, dissociation and integration. Dissociation is the suppression of pleiotropic effects by disconnecting interactions between different modules, while integration is realized by strengthening of pleiotropic connections among traits in the same modules. Both mechanisms are driven by selection pressure.

Thus, modularity can be conceptualized as an evolutionary mechanism to promote evolvability. It reduces the interdependence of disjoint components and consequently reduces the chance of pleiotropic damage by mutation [47].
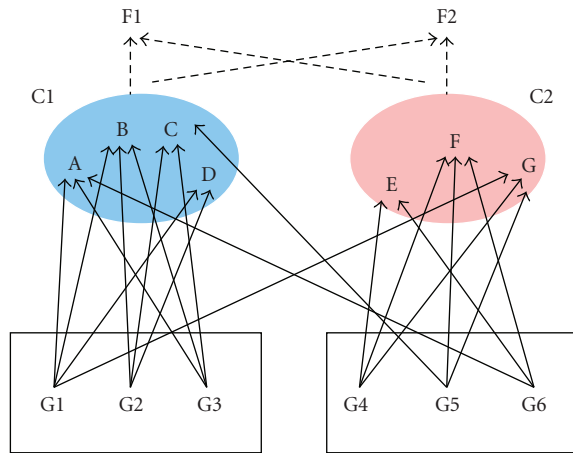
FIGURE 4: Example of a modular representation. Complexes C1 = {A, B, C, D} and C2 = {E, F, G} serve to functions F1 and F2. Each character complex has a primary function, F1 for C1 and F2 for C2. Only weak influences exist of C1 on F2 and vice versa. The genetic representation is modular because the pleiotropic effects of the genes M1 = {G1, G2, G3} have primarily pleiotropic effects on the characters in C1 and M2 = {G4, G5, G6} on the characters in complex C2. There are more pleiotropic effects on the characters within each complex than between them. (Adapted from Wagner and Altenberg [45].)

It allows genotypic variation and selection to affect separate features in a complex system and to evolve various functions without interference [176]. Subsystems as part of an entire system can evolve faster to optimize their local subfunctions individually, by decreasing crosstalk between genetic changes. In a study of encoding schemes in EC by Kazadi et al. [177], a *compartment* is defined similar to a module in the genotype-phenotype mapping, and such compartmentalization at different levels is claimed to contribute to the acceleration of evolution. In RNA research, Manrubia and Briones [178] propose that the increase of molecule length and subsequent increase in functional complexity could be mediated by modular evolution. They find that short replicating RNA sequences with a small population size can be assembled in a modular way and can create complex multifunctional molecules faster than conventional evolution of complex individuals toward multiple optima.

Modularity in general has been widely used in computer science and engineering by subdividing complex entities into smaller components to yield higher computational efficiency and has similarly played a key role in EC from the outset. In fact, it can be argued that the building block hypothesis is at its core an argument about modularity. However, complex genotype-phenotype maps and other mechanisms (like growth and development) to generate modularity in EC are relative newcomers and it is expected that studying these mechanisms in biology will result in more sophisticated means to produce modularity in EC. Since modularity is the most universal property of phenotypes in natural systems, there is ample ground to expect that the economic and sophisticated mechanisms used by Nature through regulating

and reusing relatively simple genotypic material will be a major force in shaping complex phenotypes also in EC.

3.2.3. Facilitated Variation. Kirschner and Gerhart [22] emphasize that variation is much less random at the phenotypic level of organisms than at the genotypic level, where genetic mutations show considerable randomness. Since phenotypic variation should be favored by selection via modifying existing evolutionary components, they call this variation *facilitated.*

Kirschner and Gerhart summarize three principles of facilitated variation. It serves (i) to reduce lethal pleiotropic effects, (ii) to increase phenotypic variation in light of a given number of genetic changes, and (iii) to improve genetic diversity in evolutionary populations (by reducing lethality). Evolution is not so much affected by the content of genetic and protein structures but by regulation capabilities to organize and reuse these functional parts and to decide the targets of such regulation. The core processes instead are conserved being built in a special way, only to be linked together under new circumstances like time, place, and the number of genetic components that may participate in generating new phenotypic variation. It is clear that only adaptive phenotypic variation can be maintained during evolution, and the relevant product proteins mostly will have multiple functions for various adaptive requirements under selection.

Variation in EC systems seems to be more random than that in natural evolution. Despite the limitations in recognizing these phenomena in biology, we should explore methods to reduce randomness in computational models in order to make evolving processes more "intelligent" and to facilitate the discovery of good solutions. Some steps have been taken in EC literature in this light. Researchers have designed more sophisticated techniques to improve the adaptation of algorithms. One idea was developed for Evolutionary Strategies first and later applied to other branches in EC [179–182]. Further, the evolution of "smarter" operators for EC in a higher-level evolutionary process has been examined in metaevolution [183–186]. A more recent contribution looking at the effect of changing environments on variation in a computational framework has been the GA of Parter et al. [187].

3.3. Transformation from Genotype to Phenotype. It is at the intersection between genotypes and phenotypes where most of the mechanisms reside that allow for facilitated variation. A subject of much study both in natural and artificial systems has been the genotype-phenotype map. In recent years, epigenetic effects, long suspected to have enormous influence on the final expression of the phenotype, have assumed center stage in biology. Epigenetics [188] is a rapidly developing and prominent research topic, both in relation to the development of healthy phenotypes as well as those who show deficiencies. This will constitute the second part of this section. Finally, epigenetics and a consequence of the amplified power of expression regulation through epigenetics are the mechanisms of development

of multicellular organisms. These are discussed in the last subsection here, concluding the transformation of genotypes into phenotypes.

3.3.1. Genotype-Phenotype Mapping. In EC, mapping from genotype to phenotype is often an encoding process, especially in evolutionary algorithms and evolutionary strategies, where the mapping mechanism is used in most cases to directly calculate a fitness function of an individual. However, in nature, the mapping process is much more complex, typically from highly conserved genotypic information to greatly divergent polymorphism in phenotypes. The fundamental process in biological genotype-phenotype mapping is gene expression, and the most important mechanism in this process is regulation of gene expression, which will be discussed next. Since research on transcriptional regulation has discovered increasing evidence that RNA plays an important role in gene expression, the transcriptome, that is, the set of all transcribed RNAs, will be reviewed then.

Regulation of Gene Expression. In biology, the core processes (Section 3.2.1) of organisms are responsible for generating anatomy and behavior using genetic and cell materials. These core processes include metabolism, gene expression, and interaction among molecules and cells [22], which are well conserved but still under exploration. Regulation of gene expression is the most important mechanism among the core processes to facilitate organismal novelties in evolution. Kirschner and Gerhart highlight the characteristics of "conservation" and "economy" in regulatory core processes in [22].

Scientists have been trying to understand the process of gene expression for decades. In 1956, Crick proposed the *Central Dogma* of molecular biology, as shown in Figure 5, which describes the transmission of genetic information from DNA to protein. The circular arrow around DNA symbolizes that a DNA is a template for self-replication. The arrow from DNA to RNA indicates that an RNA is transcribed on a DNA template, and the arrow from RNA to protein signifies that a protein is translated on an RNA template.

Subsequent biological research revealed that the process of gene expression is much more complex than such a linear flow and involves a considerable number of complex regulation operations. The Central Dogma was challenged by discoveries of proteins playing an important role in regulation of gene expression and, most recently, the noncoding RNA control of chromosome architecture proposed by Mattick [189]. In this section, we concentrate on gene expression regulation by proteins and will discuss RNA effects in next section.

Recall the discussion of genome redundancy in Section 2.1.1. Coding regions on genetic sequences that can be expressed into proteins only occupy a small portion of the entire genome in eukaryotic cells. This discovery indicated that a huge number of regulatory elements exist in genomes that participate in generating adaptation in evolution according to changes in environments. Although
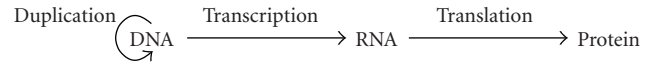

Figure 5: Central Dogma. The Central Dogma of biology by Crick holds that genetic information normally flows from DNA to RNA to protein, which involves the mechanisms of gene replication, transcription and translate.

living systems have evolved for billions of years, regulatory core processes in various organisms have remained mostly unchanged despite species divergence. By comparison of related species from the same ancestors, such as humans and chimpanzees, at both the molecular and organismal levels King and Wilson [153] had already found in 1975 that genetic structures in these two species are almost the same while at the organismal level, the anatomy, physiology, behavior and ecology of these two species are significantly different. This suggested to them that the complex adaptive evolution is produced by a combination and multiple utilization of similar, highly conserved genetic components under the control of regulatory systems.

A key step in the regulation of gene expression is transcription. Studies there are concentrated on two primary components: *promotors* and *transcription factors*. Promotors, also known as cis-regulatory sequences, are responsible for regulatory transcription. Cis-regulatory sequences are noncoding DNA sequences which determine when and where "their" genes are transcribed by regulating access of polymerase to transcription start sites. Transcription factors are proteins interacting with these cis-regulatory sequences by binding to certain sites on DNA sequences. Readers interested in more details are referred to Wray et al. [77]. Transcription factors act either as activators or as repressors of gene expression. For example, if a transcription factor A binds to a site on a DNA sequence that is responsible for generating protein B, then this factor A is regarded as a repressor to protein B. In addition, as a protein itself, factor A also has its template gene sequence. If another transcription factor C can bind to this site and represses the generation of protein A, C acts as a repressor to A but in turn as an activator to the expression of protein B. These activators and repressors can work together as a network of logic control. Promotors usually contain a number of binding sites for transcription factors, where each site can only be occupied by one factor at a time. These binding sites occupy, however, only a small fraction of sequences and are distributed unevenly. Some binding sites of different functions can overlap. Furthermore, binding affinities of different materials are important for regulation as well. On the other hand, most transcription factors have numerous target genes and use priorities in binding with any of them [77]. This sophisticated network endows the regulation system with high robustness and plasticity necessary for evolution of capabilities of organisms.

Kauffman [190, 191] holds a long standing opinion that gene regulation networks are dynamical systems and that many phenotypic traits are encoded in the *dynamical*

attractors of these systems. Dynamical attractors refer to cyclic trajectories of the transformations of states of these networks and their study provides clues to the behavior and properties of gene regulatory networks. Kauffman's point of view—namely, that the topology of a gene regulatory network largely decides cell types, cell fates, and functional states of the cell—has been supported by a number of more recent studies [192–195]. Meanwhile, simplifying computational models has been proposed to study dynamical attractors. Aldana et al. [196] model gene activities using random Boolean networks (RBN) with varying topologies. They report that a network with scale-free output topology and operating close to the critical regime (neither ordered nor chaotic) possesses the greatest robustness and evolvability compared to networks with other topologies and acting in different dynamical regimes. Further support comes from [197–201], which again confirms Wagner's argument that high robustness and high evolvability can coexist in natural systems (see Section 2.1.1).

Evolution of cis-regulatory sequences as noncoding sequences is considerably different from that of protein-coding sequences and is less understood. King and Wilson [153] suggest that protein-coding sequences are highly conserved during evolution since they were synthesized. It is mutations on promoters that causes most morphological variation. Research on the evolution of transcriptional regulation has become mainstream in molecular biology in recent years [77]. In particular, Roderiguez-Trelles et al. [75] find that significant substitution rate differences exist among different promotors, and even some neighboring cis-regulatory promotors involved in the same regulatory network can have different evolution rates. Moreover, Stone and Wray [202] propose that local point mutations on binding sites can lead to rapid evolution in gene expression, which indicates their potential of accelerating evolution. Wagner [203] points out that other simple changes such as gene duplication and deletion of promotors can also result in rapid evolution in gene regulatory networks. By comparing genomes, Fondon and Garner [204] discover that gene-associated tandem repeat expansions and contractions exist and give rise to rapid morphological evolution. In their experimental research, a tandem repeat mutation shows both elevated purity and intensive length polymorphism among different dog breeds. Mutations on noncoding sequences can modify regulation of the target genes, the length of coding loci to transcribe, and the occurrence conditions. Furthermore, they also result in morphological variation and accelerated phenotypic evolution.

Since the mechanisms of regulation of gene expression can well explain many phenomena in evolvability and rapid evolution in living systems, research on artificial regulatory networks has now started in computer science. Several models of artificial evolution regulatory networks have been proposed such as Banzhaf et al. [205–208], Chavoya and Duthen [209], Mattiussi and Floreano [210], and Nehaniv [211]. These artificial models intend to generate regulatory behavior akin to that of natural systems. However, these research efforts are still in their early stages, and more
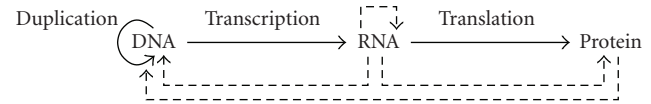

FIGURE 6: Eukaryotic genetic system. Expression of genes is not with an irreversible linear flow in eukaryotes, but involves frequent feedback and interactions among different molecules including DNA, RNA, and protein, as the dotted lines shown here.

work on evolvability and dynamics in artificial regulatory networks is necessary.

The Transcriptome. The transcriptome, or collection of transcripts, refers to all RNAs produced in a single or a group of cells, working as an intermediate component of gene expression. In high-level eukaryotes such as humans, most regions of the transcriptome are not translated into protein. What necessitates the existence of such a large number of RNAs in the transcriptome of high-level eukaryotes? Regulatory function is one answer to this question. Although regulation of gene expression starts with the transcription step, these transcribed but nontranslated sequences or noncoding RNA sequences act as regulators for translation in gene expression and currently attract increasing interest in biological research [83, 212].

An RNA is not just a temporary medium between genes and proteins as described in the Central Dogma. In high-level eukaryotes, the information transmission from DNA to protein is not a one-way process but involves many functionalities of the transcriptome. The new perspective of gene expression proposed by Mattick [189, 212] can be described in Figure 6. Compared to a prokaryotic genetic system, an eukaryotic system has a parallel control mechanism with multiple outputs and information transfers. Rather than a simple medium of gene expression, RNA metabolism and interaction have been discovered playing an important role in gene expression regulation.

Mattick [84] proposes that noncoding RNAs participate extensively in gene expression regulation, being present in about 98% of all transcriptional outputs in eukaryotes. In research on the human transcriptome, Frith et al. [83] find that noncoding RNAs play an important role in generating phenotypic variation. Noncoding RNAs can be classified into two categories: introns and other noncoding RNAs.

Regulation of the transcriptome shows contributions to evolvability and rapid evolution. Introns, an important category of noncoding RNAs, are found more susceptible to mutations than their neighboring protein-coding exons. Rather than having no function, as thought previously, it was found that introns do have influence on regulation (see, e.g., [213]). The fewer constraints imposed on introns by selection offer flexibility to generate new functions and rapid protein sequence evolution during the process of regulation, especially in connection with alternative splicing. The evolution of RNA communication networks may also accelerate the evolution of gene expression, as observed by Mattick [84]. These RNA communication networks,
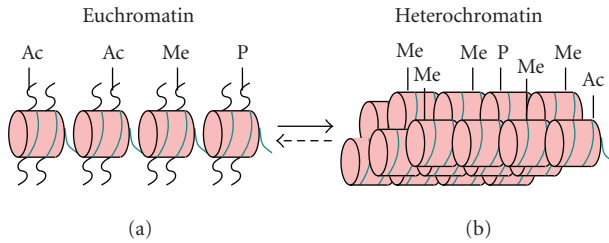
Euchromatin Heterochromatin

(a) (b)

FIGURE 7: Euchromatin and Heterochromatin. Histone tails have three types of modification including acetylation (Ac), phosphorylation (p), and methylation (Me). Euchromatin (a) is the loosely packed state that most histone tails are attached by acetyl groups. Heterochromatin (b) is the tightly packed state that most histone tails are attached by methyl groups. (Adapted from Jenuwein and Allis [218].)

which describe interaction among different layers of RNA signaling, provide a sophisticated regulatory architecture, enabling DNA-DNA, DNA-RNA, or RNA-RNA communication, DNA methylation, chromatin generation, and RNA translation.

Compared to natural systems, the genotype-phenotype mapping in EC is rather primitive still and a transcriptome is mostly missing in algorithms. The complex RNA parallel information transfer framework inspires various applications. Based on what computational models have already achieved with artificial regulatory networks, more mechanisms should be implemented, especially the newly discovered powerful mechanisms of transcriptome regulation (see a step in this direction here [214]).

3.3.2. Epigenetic Mechanism. Epigenetics has become a new research direction in evolutionary biology [21]. Literally, "epi"-genetic control lies in the regulation of gene expression *without* changing the DNA sequence itself; so it is "beyond the conventional genetic" control. Epigenetic regulation arises during the processes of organism development and cell proliferation, triggered by intrinsic signals or environmental stimulations [215]. Epigenetic changes are heritable in the short term from cell generation to cell generation, and these stable alterations do not involve mutations on DNA sequences. Epigenetic regulation of DNA expression lies at the heart of many complex and long-term human diseases [216].

Previous research in genetics mostly focused on the sequential information carried by DNA. However, DNA sequences are coiled up in cells in intimate complexes with the help of so-called histone proteins. A DNA sequence wrapped with histones comprises a *nucleosome. Chromatin* is the complex of nucleosomes in the nucleus of cells which participates in the control process of gene expression. The chromatin composition varies according to cell type and response to internal and external signals. The different composition of chromatin may affect expression and thus change the produced proteins even in the absences of DNA sequence modification [217].

The main mechanisms of epigenetic control are DNA methylation and histone modification [215]. Modifications to chromatin, either on the DNA sequence itself (*DNA methylation*) or on its surrounding proteins (*histone modification*), affect gene expression and can be inherited from cell generation to cell generation during cell division. DNA methylation is a chemical addition to DNA sequences. Genes with methyl marks are repressed in expression, despite their unchanged DNA content [219]. In histone modification, the tails of histone proteins are modified by different molecular attachments, for examole, acetyl, phosphoryl, and methyl groups (see Figure 7). If acetyl groups are attached to the histone tails of a chromatin, it will be loosely packed, a state called *euchromatin*. In euchromatin, DNA is readable and can be transcribed into RNA and later translated into proteins. In contrast, if methyl groups are attached to histone tails, chromatin is tightly compressed, a state called *heterochromatin*. In the heterochromatin state, genes are inaccessible to the transcriptional machinery such as RNA polymerase or to transcription factors, and genes are prevented from being transcribed [220]. Other mechanisms recognized to be responsible for epigenetic regulation of gene expression include chromatin remodeling, histone variant composition, and noncoding RNA regulation. A discussion of these mechanisms can be found by Allis et al. [188].

The key feature of epigenetic mechanisms is their ability to coordinate internal and environmental signals which can collaborate to modify protein production [215]. The underlying interactions involve various molecules, such as DNA, RNA, and proteins, but the extensive feedback between these molecules is still beyond our current understanding.

We believe that epigenetics opens up a new field in evolvability studies for both biology and EC. Sophisticated epigenetic feedback networks suggest a new structure for EC compared to the linear flow of computation usually employed in the literature. For instance, in dynamic optimization problems, not all genes responsible for different subfunctions need to be expressed all the time. We anticipate that a "controller switch" can be integrated into the genotype allowing short-term changes, where fragments of the genome can be turned on and off in response to external feedback. Such a mechanism for repression of expression has barely been used in computation. Similar multilayer adaptive encoding schemes have been proposed, for example, the *messy Genetic Algorithm* (mGA) [164] that combines short building blocks to form variable-length chromosomes to increasingly cover all features of a problem, or *diploid Genetic Algorithm*, for example, [221] using a two-chromosome representation to adapt phenotypic variation in dynamic environments. However, existing work has not embedded the organizational epigenetic control in algorithms that would allow significant flexibility in changing environments. We anticipate that epigenetic mechanisms will play a crucial role in increasing the evolvability of EC algorithms.

3.3.3. Development. Evolutionary developmental biology with the subject of the relation between evolution and development, nick-named *evo-devo*, has arisen as a productive

research direction which tries to unify concepts that have been separated for a long number of years. The developmental viewpoint provides crucial clues to many puzzles and controversies that have arisen in genetics and evolutionary biology in the past [222]. Vice versa, evolution is key to understanding the developmental mechanisms that have shaped multicellular life [223].

Definition 8 (development). Development is the process by which a multicellular organism unfolds its phenotype, starting from a fertilized single-cell stadium (the zygote), to a mature multicellular stadium through a defined sequence of stages that are under the control of its genome and heavily influenced by its environment.

West-Eberhard [224] discusses the relation of development and evolution and suggests that it is important to reexamine major themes of evolutionary biology in the light of development. Molecular biology has extensively investigated evolution on the genotypic level, studying the mechanisms of gene expression and protein formation, the effect of mutations on genes, and other questions. It is, however, development which produces the multicellular phenotypes and their variation that ultimately is screened by selection. So in order to examine the effect of a mutation on the evolution of multicellular organisms, one has to look at the effects of this mutation in development. Development emphasizes the time-dimension of an organism and the continuity of phenotypic changes in its interaction with the environment.

A major focus of the field of evo-devo is to study the role of *phenotypic plasticity*, or *developmental plasticity* in evolution [225–228]. Phenotypic plasticity is the phenotypic responsiveness of an organism to environmental input, and it is the most universal property of the phenotype of organisms. Organisms can alter their form, status, behavior, movement, or other features in response to environmental stimuli. These changes mostly will not involve any modifications of their genome. This flexibility is a result of the development process, with a complex mapping from genotype to phenotype.

The effect of phenotypic plasticity on the rate of evolution is a subject of debate [229]. It either can accelerate evolution since new and adaptive alternative phenotypes are generated to match the current environment or can also be considered to delay the rate of genetic changes since this flexibility is able to provide adaptation to an organism without the need to modify its genotype. The role of phenotypic plasticity in evolution depends on which level of evolution is studied, and under what conditions [227].

It is, however, clear that both major properties of the phenotype, its plasticity and its modularity (see Section 3.2.2), are the result of the hierarchical organization of the development process producing the phenotype from the genotype. These characteristics both contribute a great deal to the evolvability of living systems [230]. Kirschner and Gerhart [47, 231] hold the opinion that plasticity and modularity contribute to evolvability due to their advantages, first at providing adaptation at the individual level, and

secondly at benefiting a population's ability to diversify and persist.

The importance of the developmental point of view seems to be partially realized by the EC community. A new area named generative and developmental systems has emerged and attracted studies. *Artificial or computational embryogeny* was first introduced to simulate the development process *in silico* (see, e.g., [232–234]). More recently, inspired by the complex mapping from genotype to phenotype, computer scientists have started to allow more freedom and scalability when representing individuals, a topic known as *indirect encoding*. With an indirect encoding scheme, a genotype does not map directly to units of structure in its phenotype, but a growth or developmental process is allowed in this mapping [235]. Various encoding methods have been proposed using, for example, hierarchical grammars [236, 237], or simulating cell chemical processes [238, 239]. Indirect encoding schemes have shown advantages over traditional one-to-one direct encodings [240, 241]. Indirect encoding is a first step to simulate biological development in computational systems by allowing more freedom and complexity in the genotype-phenotype mapping, but it is by no means the full story of development. As evolutionary developmental biology continues to produce new insights, it will be imperative for the EC community to increase its efforts to design new algorithms that are inspired by evo-devo.

4. Selection

Although Darwin's theory of evolution being directed primarily by natural selection has been the subject of much argument, selection is an extremely important operation to stabilize the functional traits already generated by some exploratory operations [128]. Selection mechanisms are divided into two types by their effects on different stages of evolution. First, positive selection enhances the fixation of advantageous alleles thus improving the diversity in early stages of evolution [139, 143]. Second, negative selection, also known as stabilizing selection or purifying selection, occurs at later stages of evolution when genetic diversity decreases when such selection eliminates deleterious alleles and only stabilizes specific traits [141]. The balance between selection and diversity of an evolutionary population has been a critical problem, and the dynamic pressure and some consequences of selection are still under active investigation. In general, selection pressure is produced by two factors, the environment and mating competition, both of which will be discussed next.

4.1. Environmental Selection. Environmental selection originates in the external surroundings and enforces the adaptivity of organisms to survive. Since Darwin environmental selection has received extensive attention in evolutionary biology. Natural selection is an extremely important driving force for adaptive evolution in natural populations [242].

The first response of a living organisms to a changing environment is *somatic adaptation*. A simple example of

somatic adaptation is human temperature compensation [22]. When the external temperature increases above normal, humans will sweat to adapt to this new environment. Shivering will occur if temperature falls below a normal value. Somatic adaptation happens directly as an organismal reaction to a changing environment and is not fixed in morphological structures as an evolutionary change unless some deeper adaptations caused by somatic changes can increase the survivability of an organism. Organisms have plenty of latent traits within their somatic adaptability; so they have a fairly high tolerance to changes in their environment [22]. Somatic adaptation mechanisms can only adjust existing functions to external changes. However, if evolution acts for a long time under environmental selection, changes may be stabilized by mutations in the germ-line after somatic adaptation has tested them through promotion of survivability of the organism.

Selection can act at different levels depending on its targets [47]. These might be individual selection, individual-and-clade selection, or clade selection. At the individual level, the selection process has the fewest constraints since it directly affects phenotypic function fitness, and fewer mutational changes are required for a new adaptive trait. An individual can also interact with others in a clade, such as through recombination, and survive under selective pressure as a member in this clade. At the highest level, selection can happen on the level of an entire clade given large environmental impact, and the entire clade can, as a whole, escape from extinction. Some small groups of the lineage might go extinct, but the entire line will be able to survive even if it might have to go through population bottlenecks. This idea has drawn more attentions in some subsequent works [243, 244].

Interactions between different species may also cause environmental changes. Phillips and Shine [245] report an interesting phenomenon on species invasion. Toxic cane toads induced morphological changes among a species of snakes in Australia. Generally, native natural ecosystems can be devastated by the invasion of new species. At upon the arrival of an invasive species, the number of native organisms may decrease. However, as these native organisms adapt towards the invaders, the impact of the invasion declines and a new balance is achieved. Morphological changes are fixed subsequently. Complex natural ecosystems possess communities with highly frequent and dense interactions between species as well as between species-specific functional traits within a species.

Environmental selection is now widely accepted as contributing significantly to natural evolution, and has entered the mainstream of studies in evolvability. As a potential to generate adaptation, evolvability is difficult to observe and to select for. However, there is increasing research arguing that evolvability is selectable and environmental selection can improve the evolution of evolvability. In the real world, the environment is changing constantly and fixes beneficial mutations, and there is a growing acceptance that a changing environment is a key ingredient to studying evolvability. Selection pressure is a critical operator to control an evolutionary process. Earl and Deem [55] suggest that selection

pressure is increasingly strong when the environment becomes uncertain. Dramatic environmental changes lead to selection for better evolvability. They consider evolvability as a selectable trait, and facilitating environmental changes can be a method to accelerate evolution. A recent simulation by Kashtan et al. [74] in a biologically realistic setting also suggests that varying environments may accelerate natural evolution. In their work, different scenarios of temporarily changing optima were used. Kashtan and Alon [246] report that a goal that varies in a modular way can speed up evolution. Other work [247] takes into account the effect of the rate of environmental change. By observing the dynamics of *adaptive walks* under scenarios of varying speeds, they find that environments with varying rates of change have noticeably different effects on the fixation of beneficial mutations, the substitution time required, and the final phenotypic variation.

In EC, selection strategies are considered affecting search capability significantly during an evolutionary process. Different selection strategies have been proposed and the dynamics of selection pressure has been studied extensively [248, 249]. Since the effects of environmental selection on the evolution of evolvability have been recognized, further research on the dynamics of selection is required. Moreover, somatic adaptation might be considered when applying selection. Group-based selection methods should also be studied for varying selection pressure, so that a balance between the development of a minority and of the entire population can be dynamically achieved.

4.2. Sexual Selection. Sexual selection was proposed by Darwin as the pressure away from the possibility of mating failure. Two forms of sexual selection pressure are met by mature high-level animals: the battle between male individuals who fight, and the competition through mating choice made by females. Fisher [105] proposed a *runaway process*, where a male trait and female preference for it can both evolve dramatically over time until finally checked by severe counter-selection. In modern biology, scientists pay much attention to these sex-based competitions that can generate and evolve several kinds of traits in high-level organisms. For instance, Kirkpatrick and Ravigne [250] find that some secondary sexual characteristics among individuals of the same sex can trigger rapid speciation.

Sexual selection happens at the interspecies level and affects reproductive fitness of individuals. Reproductive fitness is the probability of successfully generating offspring. Sexual selection has two main forms: intrasexual selection and intersexual selection. Intrasexual selection is known as the combat between competitive male individuals, and usually occurs in the form of a fight. Intersexual selection is based on the choice made by the opposite sex. Male secondary sexual characteristics and female mating preferences can affect each other and evolve cooperatively [251]. This joint selection pressure, combined with natural selection, is a powerful force for rapid evolution.

Recent research in biology has connected sexual selection to the acceleration of evolution. Colegrave [252] finds that

the rate of adaptation can be increased by *sex mechanisms* because sexual selection allows a rapid adaptive response under changing conditions by fixing beneficial mutations. Swanson and Vacquier [253] observe that rapid evolution emerges in reproductive proteins. This rapid evolution is forced by three main selective factors: sperm competition, sexual selection, and sexual conflict. Sperm competition is quite fierce in that each sperm will compete with billions of others to fuse with the only egg, and this competition exists in multiple steps for the sperm. Sexual selection happens when different eggs have varying affinities for a special allele of a sperm-surface protein, and only the egg with the highest affinity is most likely to bind to this sperm. Sexual conflict means that only one egg can be fused with the sperm to avoid polyspermy such that only one embryo is fertilized. These types of mechanisms add considerable selection pressure to reproductive proteins and thus trigger rapid evolution in certain regions of these proteins.

The concept of *mating choice* was already applied in EC decades ago by Miller [254, 255]. Some coevolutionary algorithms have been proposed to simulate mechanisms from sexual selection by constructing subgroups which can affect each other cooperatively to evolve in parallel. As more and more knowledge has been accumulated by biologists on the complex process of sexual selection, especially on the advantages that sex mechanisms contribute to the acceleration of speciation and evolution, this knowledge should be better incorporated in EC.

4.3. Fitness Evaluation. Fitness evaluation measures behavior or function of individuals or species. In nature, fitness of an individual or species is implicit and subject to natural selection, whereas in EC, fitness is mostly based on numerical values of an individual as solution to a given problem, and this fitness is explicit.

Definition 9. Fitness is the measure to quantify an evolutionary individual/component with regard to its ability to survive and reproduce in a certain environment.

In nature, adaptable species survive by passing different challenges, and less fit species may become extinct during evolution. Adaptability lies not only in the currently existing adaptivity to the environment but also in the capability to generate more adapted offspring. In essence, fitness of natural organisms is implicit and is subject to natural selection. Empirically, biologists use mathematical methods to quantify fitness. *Individual fitness* usually refers to the viability of an individual, that is, its probability to survive [256]. Moreover, individuals having more offspring can be considered as fitter ones since their genetic information is more likely to be preserved. Other than at the individual level, in classic population genetics literature [257], the *genotype fitness* quantifies the frequency changes of a genotype in a population during transformation from one generation to the next. Various measures have been proposed in the biological literature (see [258, 259] for detailed reviews). The above implicit fitness in natural organisms emphasizes evolvability under intricate pressures from interactions

among evolutionary components, internal or external to these organisms during a long, continuing evolutionary process [170]. In reality, the fitness of individuals in a system can vary a great deal. Moreover, a large-scale quality differentiation exists in almost every natural evolutionary system, and these vastly diverse evolution systems exhibit substantial evolvability. Since selection and evaluation act directly on observable phenotypic functions but evolvability only provides the potential for better functions, selection and evaluation for evolvability are not observable directly.

Since EC has been widely applied in many areas of industry and academia, fitness evaluation arises as a difficult problem because it is usually very CPU-intensive. In the current literature, two main methods of fitness evaluation are employed, absolute fitness and relative fitness. Absolute fitness of each individual usually refers to its value of a specified fitness function. Relative fitness compares different individuals and gives a rank to each individual to produce a record of winners. This latter method is good at suppressing exceptionally good individuals, thus, helping an evolutionary system to escape from premature convergence. In fact, evaluating the fitness of each individual is usually difficult for many optimization problems in the real world because explicit fitness can be hard to define and expensive to calculate. As a result, *fitness approximation* has been proposed with differing levels of approximation, including "problem approximation", "functional approximation", and "evolutionary approximation". Jin [260] has surveyed these approaches. They are sensitive to training data and to varying constraints of different models; so a common framework would be required. Moreover, Reisinger and Miikkulainen [56] propose an evolvable representation and an evaluation strategy to exert indirect selection pressure on evolvability. In their work, a systematically changing fitness function is adopted according to a special evolvable representation that can reflect efficiently how genetic changes restructure phenotypic variation. Thus, evolvability can be evaluated through the way such a systematic structure can expand in phenotypes. These approaches might provide a good starting point to simulate the implicit adaptive fitness evaluation from nature, a method that has good prospects for detecting evolvability in EC.

5. Conclusion

Since Darwin proposed his theory of natural evolution based on heritable variation and natural selection, an enormous research effort has been dedicated to revealing the intricacies of the processes involved. In modern biology, a host of details about mechanisms of evolution and factors that can affect evolution have been revealed. Besides understanding the history of evolution, biologists are currently paying attention to the capability of organisms to evolve and to the evolution of such capability in an open-ended natural evolutionary process. Varying evolution rates among different species or different regions of genetic material in an organism attract researchers' interest under the aspect of the acceleration of evolution. Meanwhile, in artificial evolutionary systems,

one is also working on improving the power of systems by studying more intelligent and adaptive mechanisms.

Evolvability, as the capability to generate adaptation by producing fitter offspring via evolutionary operations, has received considerable interest in recent research in both biology and EC. Substantial work has been published on this topic in both areas, and we have tried to cover many of the factors that contribute to evolvability in this review. After some phenomena of rapid evolution were found in nature, acceleration of evolution also became an hot research topic, not the least because an increase in the speed of artificial evolution would greatly benefit applications.

More is to come. As a result of the unrelenting progress in biology, a lively discussion has now ensued as to what a gene really is [261–264]. There were times when the notion of a gene was simple. But with the advent of alternative splicing, intron activity during regulation, iRNA, and other wonderful intricacies [265], the simple life of geneticists seems to be over. In this review, our perspective was much more limited by discussing only two very narrow aspects of evolution, its speed and the issue of evolvability. We hope that the ideas discussed here can inspire new methods and applications of EC.

Acknowledgments

References

[1] W. Banzhaf, P. Nordin, R. E. Keller, and F. D. Francone, *Genetic Programming—An Introduction: On the Automatic Evolution of Computer Programs and Its Applications*, Morgan Kaufmann, San Francisco, CA, USA, 1998.

[2] K. A. De Jong, *Evolutionary Computation: A Unified Approach*, MIT Press, Cambridge, Mass, USA, 2006.

[3] A. Eiben and J. Smith, *Introduction to Evolutionary Computing*, Springer, Berlin, Germany, 2003.

[4] D. B. Fogel, *Evolutionary Computation: Toward a New Philosophy of Machine Intelligence*, IEEE Press, Orlando, Fla, USA, 1995.

[5] L. J. Fogel, A. J. Owens, and M. J. Walsh, *Artificial Intelligence through Simulated Evolution*, John Wiley & Sons, New York, NY, USA, 1966.

[6] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, Mass, USA, 1989.

[7] J. H. Holland, *Adaptation in Natural and Artificial Systems*, University of Michigan Press, Ann Arbor, Mich, USA, 1975.

[8] J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection*, MIT Press, Cambridge, Mass, USA, 1992.

[9] M. Mitchell, *An Introduction to Genetic Algorithms*, MIT Press, Cambridge, Mass, USA, 1998.

[10] I. Rechenberg, *Evolution Strategy*, Frommann-Holzboog, Stuttgart, Germany, 1973.

[11] H.-P. Schwefel, *Evolution and Optimum Seeking: The Sixth Generation*, John Wiley & Sons, New York, NY, USA, 1995.

[12] "Annual "HUMIES" Awards for Human Competitive Results," Produced by Genetic and Evolutionary Computation held by GECCO: Genetic and Evolutionary Computation Conference, ACM, New York, NY, USA, since 2004, http://www.humancompetitive.org.

[13] C. Darwin, *On the Origin of Species by Means of Natural Selection*, John Murray, London, UK, 1859.

[14] S. J. Gould, *The Structure of Evolutionary Theory*, Harvard University Press, Cambridge, Mass, USA, 2002.

[15] E. Mayr and W. B. Provine, Eds., *The Evolutionary Synthesis*, Harvard University Press, Cambridge, Mass, USA, 2nd edition, 1998.

[16] W. Banzhaf, G. Beslon, S. Christensen, et al., "Guidelines: from artificial evolution to computational evolution: a research agenda," *Nature Reviews Genetics*, vol. 7, no. 9, pp. 729–735, 2006.

[17] J. Timmis, M. Amos, W. Banzhaf, and A. Tyrrell, "Going back to our roots: second generation biocomputing," *International Journal of Unconventional Computing*, vol. 2, no. 4, pp. 349–378, 2006.

[18] S. Carroll, J. Grenier, and S. Weatherbee, *From DNA to Diversity*, Blackwell, Oxford, UK, 2005.

[19] G. Cochran and H. Harpending, *The 10,000 Year Explosion*, Basic Books, New York, NY, USA, 2009.

[20] G. Forgacs and S. A. Newman, *Biological Physics of the Developing Embryo*, Cambridge University, Cambridge, UK, 2005.

[21] E. Jablonka and M. J. Lamb, *Evolution in Four Dimensions: Genetic, Epigenetic, Behavioral, and Symbolic Variation in the History of Life*, MIT Press, Cambridge, Mass, USA, 2005.

[22] M. W. Kirschner and J. C. Gerhart, *The Plausibility of Life*, Yale University Press, New Haven, Conn, USA, 2005.

[23] M. Novak, *Evolutionary Dynamics*, Harvard University Press, Cambridge, Mass, USA, 2006.

[24] S. R. Palumbi, *The Evolution Explosion*, Norton, New York, NY, USA, 2002.

[25] A. Wagner, *Robustness and Evolvability in Living Systems*, Princeton University Press, Princeton, NJ, USA, 2005.

[26] M. A. Bedau, J. S. McCaskill, N. H. Packard, et al., "Open problems in artificial life," *Artificial Life*, vol. 6, no. 4, pp. 363–376, 2000.

[27] T. S. Ray, "An approach to the synthesis of life," in *Artificial Life II*, pp. 371–408, Addison-Wesley, Reading, Mass, USA, 1991.

[28] C. Adami and C. T. Brown, "Evolutionary learning in the 2D artificial life system "avida"," in *Artificial Life IV*, pp. 377–381, MIT press, Cambridge, Mass, USA, 1994.

[29] M. A. Bedau, E. Snyder, and C. T. Brown, "A comparison of evolutionary activity in artificial evolving systems and in the biosphere," in *Proceedings of the 4th European Conference on Artificial Life*, pp. 125–134, Springer, 1997.

[30] M. A. Bedau and C. T. Brown, "Visualizing evolutionary activity of genotypes," *Artificial Life*, vol. 5, no. 1, pp. 17–35, 1999.

[31] J. E. Barrick, D. S. Yu, S. H. Yoon, et al., "Genome evolution and adaptation in a long-term experiment with *Escherichia coli*," *Nature*, vol. 461, no. 7268, pp. 1243–1247, 2009.

[32] P. Alberch, "From genes to phenotype: dynamical systems and evolvability," *Genetica*, vol. 84, no. 1, pp. 5–11, 1991.

[33] J. F. Y. Brookfield, "Evolution: the evolvability enigma," *Current Biology*, vol. 11, no. 3, pp. R106–R108, 2001.

[34] A. J. R. Carter, J. Hermisson, and T. F. Hansen, "The role of epistatic gene interactions in the response to selection and the evolution of evolvability," *Theoretical Population Biology*, vol. 68, no. 3, pp. 179–196, 2005.

[35] J. L. Hendrikse, T. E. Parsons, and B. Hall-grimsson, "Evolvability as the proper focus of evolutionary developmental biology," *Evolution and Development*, vol. 9, no. 4, pp. 393–401, 2007.

[36] D. Houle, "Comparing evolvability and variability of quantitative traits," *Genetics*, vol. 130, no. 1, pp. 195–204, 1992.

[37] K. F. Liem, "Key evolutionary innovations, differential diversity, and symecomorphosis," in *Evolutionary Innovations*, pp. 147–170, University of Chicago Press, Chicago, Ill, USA, 1990.

[38] P. D. Sniegowski and H. A. Murphy, "Evolvability," *Current Biology*, vol. 16, no. 19, pp. R831–R834, 2006.

[39] G. Bell, "Evolutionary genetics: the evolution of evolution," *Heredity*, vol. 94, no. 1, pp. 1–2, 2005.

[40] R. Dawkins, "The evolution of evolvability," in *Artificial Life: The Quest for a New Creation*, pp. 201–220, Addison-Wesley, Reading, Mass, USA, 1989.

[41] A. G. Jones, S. J. Arnold, and R. Burger, "The mutation matrix and the evolution of evolvability," *Evolution*, vol. 61, no. 4, pp. 727–745, 2007.

[42] L. Partridge and N. H. Barton, "Evolving evolvability," *Nature*, vol. 407, no. 6803, pp. 457–458, 2000.

[43] M. Pigliucci, "Perspectives: is evolvability evolvable?" *Nature Reviews Genetics*, vol. 9, pp. 75–82, 2008.

[44] M. Radman, I. Matic, and F. Taddei, "Evolution of evolvability," *Annals of the New York Academy of Sciences*, vol. 870, pp. 146–155, 1999.

[45] G. P. Wagner and L. Altenberg, "Complex adaptations and the evolution of evolvability," *Evolution*, vol. 50, no. 3, pp. 967–976, 1996.

[46] L. Altenberg, "The evolution of evolvability in genetic programming," in *Advances in Genetic Programming*, pp. 47–74, MIT Press, Cambridge, Mass, USA, 1994.

[47] M. Kirschner and J. Gerhart, "Evolvability," *Proceedings of the National Academy of Sciences (USA)*, vol. 95, no. 15, pp. 8420–8427, 1998.

[48] P. Marrow, "Evolvability: evolution, computation, and biology," in *Proceedings of the Genetic and Evolutionary Computation Conference Workshop on Evolvability (GECCO '99)*, pp. 30–33, Morgan Kaufmann, San Francisco, CA, USA, 1999.

[49] M. Conrad, "The geometry of evolution," *BioSystems*, vol. 24, no. 1, pp. 61–81, 1990.

[50] A. Wagner, "Robustness, evolvability, and neutrality," *FEBS Letters*, vol. 579, no. 8, pp. 1772–1778, 2005.

[51] M. A. Bedau and N. H. Packard, "Evolution of evolvability via adaptation of mutation rates," *BioSystems*, vol. 69, no. 2-3, pp. 143–162, 2003.

[52] J. Masel, "Evolutionary capacitance may be favored by natural selection," *Genetics*, vol. 170, no. 3, pp. 1359–1371, 2005.

[53] C. L. Nehaniv, "Measuring evolvability as the rate of complexity increase," in *Proceedings of the 7th Artificial Life Workshop*, pp. 55–57, MIT Press, 2000.

[54] A. Wagner, "Robustness and evolvability: a paradox resolved," *Proceedings of the Royal Society B*, vol. 275, no. 1630, pp. 91–100, 2008.

[55] D. J. Earl and M. W. Deem, "Evolvability is a selectable trait," *Proceedings of the National Academy of Sciences of USA*, vol. 101, no. 32, pp. 11531–11536, 2004.

[56] J. Reisinger and R. Miikkulainen, "Selecting for evolvable representations," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '06)*, vol. 2, pp. 1297–1304, ACM, 2006.

[57] J. Reisinger, K. O. Stanley, and R. Miikkulainen, "Towards an empirical measure of evolvability," in *Proceedings of the Genetic and Evolutionary Computation Conference Workshop Program (GECCO '05)*, pp. 257–264, ACM, 2005.

[58] N. Eldredge and S. J. Gould, "Punctuated equilibria: an alternative to phyletic gradualism," in *Models in Paleobiology*, pp. 82–115, Freeman Cooper, 1972.

[59] W. H. Li, C. I. Wu, and C. C. Luo, "A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes," *Molecular Biology and Evolution*, vol. 2, no. 2, pp. 150–174, 1985.

[60] T. Miyata and T. Yasunaga, "Molecular evolution of mRNA: a method for estimating evolutionary rates of synonymous and amino acid substitutions from homologous nucleotide sequences and its application," *Journal of Molecular Evolution*, vol. 16, no. 1, pp. 23–36, 1980.

[61] M. Nei and T. Gojobori, "Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions," *Molecular Biology and Evolution*, vol. 3, no. 5, pp. 418–426, 1986.

[62] Z. Yang and J. R. Bielawski, "Statistical methods for detecting molecular adaptation," *Trends in Ecology and Evolution*, vol. 15, no. 12, pp. 496–503, 2000.

[63] R. P. Worden, "A speed limit for evolution," *Journal of Theoretical Biology*, vol. 176, no. 1, pp. 137–152, 1995.

[64] M. A. Bedau and N. H. Packard, "Measurement of evolutionary activity, teleology, and life," in *Artificial Life II*, pp. 431–461, Addison-Wesley, Reading, Mass, USA, 1992.

[65] M. A. Bedau, "Quantifying the extent and intensity of adaptive evolution," in *Proceedings of the Genetic and Evolution Computation Conference Workshop Program (GECCO '99)*, pp. 34–37, Morgan Kaufmann, San Francisco, CA, USA, 1999.

[66] M. J. Raven and M. A. Bedau, "General framework for evolutionary activity," in *Proceedings of the 7th European Conference on Artificial Life*, pp. 676–685, Springer, 2003.

[67] G. G. Simpson, *Tempo and Mode in Evolution*, Columbia University Press, New York, NY, USA, 1944.

[68] Z. D. Blount, C. Z. Borland, and R. E. Lenski, "Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli*," *Proceedings of the National Academy of Sciences (USA)*, vol. 105, no. 23, pp. 7899–7906, 2008.

[69] L. P. Oliveira, H. J. Jensen, M. Nicodemi, and P. Sibani, "Record dynamics and the observed temperature plateau in the magnetic creep-rate of type-II superconductors," *Physical Review B*, vol. 71, Article ID 104526, 2005.

[70] P. Sibani, M. R. Schmidt, and P. Alstrøm, "Fitness optimization and decay of extinction rate through biological evolution," *Physical Review Letters*, vol. 75, no. 10, pp. 2055–2058, 1995.

[71] H. A. Simon, "Near decomposability and the speed of evolution," *Industrial and Corporate Change*, vol. 11, no. 3, pp. 587–599, 2002.

[72] X. Gu, Z. Zhang, and W. Huang, "Rapid evolution of expression and regulatory divergences after yeast gene duplication," *Proceedings of the National Academy of Sciences (USA)*, vol. 102, no. 3, pp. 707–712, 2005.

[73] J. Hawks, E. T. Wang, G. M. Cochran, H. C. Harpending, and R. K. Moyzis, "Recent acceleration of human adaptive evolution," *Proceedings of the National Academy of Sciences (USA)*, vol. 104, no. 52, pp. 20753–20758, 2007.

[74] N. Kashtan, E. Noor, and U. Alon, "Varying environments can speed up evolution," *Proceedings of the National Academy of Sciences (USA)*, vol. 104, no. 34, pp. 13711–13716, 2007.

[75] F. Rodriguez-Trelles, R. Tarrio, and F. J. Ayala, "Evolution of cis-regulatory regions versus codifying regions," *International Journal of Developmental Biology*, vol. 47, no. 7-8, pp. 665–673, 2003.

[76] S. Asthana, W. S. Noble, G. Kryukov, C. E. Grant, S. Sunyaev, and J. A. Stamatoyannopoulos, "Widely distributed noncoding purifying selection in the human genome," *Proceedings of the National Academy of Sciences (USA)*, vol. 104, no. 30, pp. 12410–12415, 2007.

[77] G. A. Wray, M. W. Hahn, E. Abouheif, et al., "The evolution of transcriptional regulation in eukaryotes," *Molecular Biology and Evolution*, vol. 20, no. 9, pp. 1377–1419, 2003.

[78] J. D. Wren, E. Forgacs, J. W. Fondon III, et al., "Repeat polymorphisms within gene regions: phenotypic and evolutionary implications," *American Journal of Human Genetics*, vol. 67, no. 2, pp. 345–356, 2000.

[79] J. L. Freeman, G. H. Perry, L. Feuk, et al., "Copy number variation: new insights in genome diversity," *Genome Research*, vol. 16, no. 8, pp. 949–961, 2006.

[80] A. J. Iafrate, L. Feuk, M. N. Rivera, et al., "Detection of large-scale variation in the human genome," *Nature Genetics*, vol. 36, no. 9, pp. 949–951, 2004.

[81] J. Sebat, B. Lakshmi, J. Troge, et al., "Large-scale copy number polymorphism in the human genome," *Science*, vol. 305, no. 5683, pp. 525–528, 2004.

[82] E. Tuzun, A. J. Sharp, J. A. Bailey, et al., "Fine-scale structural variation of the human genome," *Nature Genetics*, vol. 37, no. 7, pp. 727–732, 2005.

[83] M. C. Frith, M. Pheasant, and J. S. Mattick, "The amazing complexity of the human transcriptome," *European Journal of Human Genetics*, vol. 13, no. 8, pp. 894–897, 2005.

[84] J. S. Mattick, "Non-coding RNAs: the architects of eukaryotic complexity," *EMBO Reports*, vol. 2, no. 11, pp. 986–991, 2001.

[85] J. A. de Visser, J. Hermisson, G. P. Wagner, et al., "Perspective: evolution and detection of genetic robustness," *Evolution*, vol. 57, no. 9, pp. 1959–1972, 2003.

[86] D. C. Krakauer and J. B. Plotkin, "Redundancy, antiredundancy, and the robustness of genomes," *Proceedings of the National Academy of Sciences (USA)*, vol. 99, no. 3, pp. 1405–1409, 2002.

[87] J. D. Bloom, S. T. Labthavikul, C. R. Otey, and F. H. Arnold, "Protein stability promotes evolvability," *Proceedings of the National Academy of Sciences (USA)*, vol. 103, no. 15, pp. 5869–5874, 2006.

[88] B. C. Daniels, Y.-J. Chen, J. P. Sethna, R. N. Gutenkunst, and C. R. Myers, "Sloppiness, robustness, and evolvability in systems biology," *Current Opinion in Biotechnology*, vol. 19, no. 4, pp. 389–395, 2008.

[89] S. F. Elena and R. Sanjuan, "The effect of genetic robustness on evolvability in digital organisms," *BMC Evolutionary Biology*, vol. 8, no. 1, article 284, 2008.

[90] R. C. McBride, C. B. Ogbunugafor, and P. E. Turner, "Robustness promotes evolvability of thermotolerance in an RNA virus," *BMC Evolutionary Biology*, vol. 8, no. 1, article 231, 2008.

[91] R. Wilds, S. A. Kauffman, and L. Glass, "Evolution of complex dynamics," *Chaos*, vol. 18, no. 3, Article ID 033109, 2008.

[92] J. A. Draghi, T. L. Parsons, G. P. Wagner, and J. B. Plotkin, "Mutational robustness can facilitate adaptation," *Nature*, vol. 463, no. 7279, pp. 353–355, 2010.

[93] F. Rothlauf and D. E. Goldberg, "Redundant representations in evolutionary computation," *Evolutionary Computation*, vol. 11, no. 4, pp. 381–415, 2003.

[94] W. B. Langdon, T. Soule, R. Poli, and J. A. Foster, "The evolution of size and shape," in *Advances in Genetic Programming*, vol. 3, pp. 163–190, MIT Press, Cambridge, Mass, USA, 1999.

[95] S. Luke, "Code growth is not caused by introns," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '00)*, pp. 228–235, Morgan Kaufmann, San Francisco, CA, USA, 2000.

[96] P. Nordin, F. Francone, and W. Banzhaf, "Explicitly defined introns and destructive crossover in genetic programming," in *Advances in Genetic Programming*, vol. 2, pp. 111–134, MIT Press, Cambridge, Mass, USA, 1996.

[97] T. Yu and J. Miller, "Neutrality and the evolvability of boolean function landscape," in *Proceedings of the 4th European Conference on Genetic Programming (EuroGP '01)*, vol. 2038 of *Lecture Notes in Computer Science*, pp. 204–217, Springer, 2001.

[98] M. Beckerman, *Molecular and Cellular Signaling*, Springer, Berlin, Germany, 2005.

[99] F. Jacob, "Evolution and tinkering," *Science*, vol. 196, no. 4295, pp. 1161–1166, 1977.

[100] J. Cohen, "Bioinformatics—an introduction for computer scientists," *ACM Computing Surveys*, vol. 36, no. 2, pp. 122–158, 2004.

[101] J. D. Watson, N. H. Hopkins, J. W. Roberts, J. A. Steitz, and A. M. Weiner, *Molecular Biology of the Gene*, Benjamin-Cummings, 5th edition, 2004.

[102] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson, *Molecular Biology of the Cell*, Garland, New York, NY, USA, 4th edition, 2002.

[103] M. Z. Ludwig, N. H. Patel, and M. Kreitman, "Functional analysis of eve stripe 2 enhancer evolution in Drosophila, rules governing conservation and change," *Development*, vol. 125, no. 5, pp. 949–958, 1998.

[104] S. R. Neves and R. Iyengar, "Modeling of signaling networks," *BioEssays*, vol. 24, no. 12, pp. 1110–1117, 2002.

[105] R. A. Fisher, *Genetical Theory of Natural Selection*, Clarendon, Oxford, UK, 1930.

[106] J. B. S. Haldane, *The Causes of Evolution*, Harper and Row, New York, NY, USA, 1932.

[107] S. Wright, "Evolution in Mendelian populations," *Genetics*, vol. 16, no. 2, pp. 97–159, 1931.

[108] M. Kimura, "Evolutionary rate at the molecular level," *Nature*, vol. 217, no. 5129, pp. 624–626, 1968.

[109] J. H. Gillespie, "The role of population size in molecular evolution," *Theoretical Population Biology*, vol. 55, no. 2, pp. 145–156, 1999.

[110] J. H. Gillespie, "Is the population size of a species relevant to its evolution?" *Evolution*, vol. 55, no. 11, pp. 2161–2169, 2001.

[111] T. Ohta, "Population size and rate of evolution," *Journal of Molecular Evolution*, vol. 1, no. 4, pp. 305–314, 1972.

[112] T. Ohta, "The nearly neutral theory of molecular evolution," *Annual Review of Ecology and Systematics*, vol. 23, no. 1, pp. 263–286, 1992.

[113] T. Ohta and M. Kimura, "On the constancy of the evolutionary rate of cistrons," *Journal of Molecular Evolution*, vol. 1, no. 1, pp. 18–25, 1971.

[114] M. Woolfit and L. Bromham, "Population size and molecular evolution on islands," *Proceedings of the Royal Society B*, vol. 272, no. 1578, pp. 2277–2282, 2005.

[115] S. Wright, J. Keeling, and L. Gillman, "The road from Santa Rosalia: a faster tempo of evolution in tropical climates," *Proceedings of the National Academy of Sciences (USA)*, vol. 103, no. 20, pp. 7718–7722, 2006.

[116] G. C. Stevens, "The latitudinal gradient in geographical range: how so many species coexist in the tropics," *The American Naturalists*, vol. 133, no. 2, pp. 240–256, 1989.

[117] S. F. Elena, C. O. Wilke, C. Ofria, and R. E. Lenski, "Effects of population size and mutation rate on the evolution of mutational robustness," *Evolution*, vol. 61, no. 3, pp. 666–674, 2007.

[118] W. Banzhaf, S. Harding, W. Langdon, and G. Wilson, "Accelerating genetic programming through graphics processing units," in *Genetic Programming Theory and Practice VI*, pp. 229–248, Springer, Berlin, Germany, 2009.

[119] H. Juillé and J. B. Pollack, "Massively parallel genetic programming," in *Advances in Genetic Programming*, vol. 2, pp. 339–357, MIT Press, Cambridge, Mass, USA, 1996.

[120] W. B. Langdon, "A many threaded CUDA interpreter for genetic programming," in *Proceedings of the 13th European Conference on Genetic Programming (EuroGP '10)*, vol. 6021 of *Lecture Notes in Computer Science*, pp. 146–158, Springer, 2010.

[121] F. G. Lobo and C. F. Lima, "A review of adaptive population sizing schemes in genetic algorithms," in *Proceedings of the Workshops on Genetic and Evolutionary Computation (GECCO '05)*, pp. 228–234, ACM, 2005.

[122] D. E. Goldberg, "Sizing populations for serial and parallel genetic algorithms," in *Proceedings of the 3rd International Conference on Genetic Algorithms*, pp. 70–79, Morgan Kaufmann, San Francisco, CA, USA, 1989.

[123] D. E. Goldberg, K. Deb, and J. H. Clark, "Genetic algorithms, noise, and the sizing of populations," *Complex Systems*, vol. 6, no. 4, pp. 333–362, 1992.

[124] J. Arabas, Z. Michalewicz, and J. Mulawka, "GAVaPS—a genetic algorithm with varying population size," in *Proceedings of the 1st IEEE Conference on Evolutionary Computation (CEC '94)*, vol. 1, pp. 73–78, IEEE, 1994.

[125] G. R. Harik and F. G. Lobo, "A parameter-less genetic algorithm," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '99)*, pp. 258–267, Morgan Kaufmann, San Francisco, CA, USA, 1999.

[126] T. Back, A. E. Eiben, and N. A. L. van der Vaart, "An empirical study on GAs "with-out parameters"," in *Proceedings of the 6th International Conference on Parallel Problem Solving from Nature (PPSN '00)*, vol. 1917 of *Lecture Notes in Computer Science*, pp. 315–324, Springer, 2000.

[127] A. E. Eiben, E. Marchiori, and V. A. Valkó, "Evolutionary algorithms with on-the-fly population size adjustment," in *Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN '04)*, vol. 3242 of *Lecture Notes in Computer Science*, pp. 41–50, Springer, 2004.

[128] M. Nei, "Selectionism and neutralism in molecular evolution," *Molecular Biology and Evolution*, vol. 22, no. 12, pp. 2318–2342, 2005.

[129] J. R. Roth, E. Kugelberg, A. B. Reams, E. Kofoid, and D. I. Andersson, "Origin of mutations under selection: the adaptive mutation controversy," *Annual Review of Microbiology*, vol. 60, pp. 477–501, 2006.

[130] J. R. Roth and D. I. Andersson, "Adaptive mutation: how growth under selection stimulates Lac+ reversion by increasing target copy number," *Journal of Bacteriology*, vol. 186, no. 15, pp. 4855–4860, 2004.

[131] H. Hendrickson, E. S. Slechta, U. Bergthorsson, D. I. Andersson, and J. R. Roth, "Amplification-mutagenesis: evidence that "directed" adaptive mutation and general hypermutability result from growth with a selected gene amplification," *Proceedings of the National Academy of Sciences (USA)*, vol. 99, no. 4, pp. 2164–2169, 2002.

[132] E. Kugelberg, E. Kofoid, A. B. Reams, D. I. Andersson, and J. R. Roth, "Multiple pathways of selected gene amplification during adaptive mutation," *Proceedings of the National Academy of Sciences (USA)*, vol. 103, no. 46, pp. 17319–17324, 2006.

[133] E. S. Slechta, K. L. Bunny, E. Kugelberg, E. Kofoid, D. I. Andersson, and J. R. Roth, "Adaptive mutation: general mutagenesis is not a programmed response to stress but results from rare coamplification of dinB with lac," *Proceedings of the National Academy of Sciences (USA)*, vol. 100, no. 22, pp. 12847–12852, 2003.

[134] D. M. Weinreich, N. F. Delaney, M. A. DePristo, and D. L. Hartl, "Darwinian evolution can follow only very few mutational paths to fitter proteins," *Science*, vol. 312, no. 5770, pp. 111–114, 2006.

[135] M. A. Bedau and R. Seymour, "Adaptation of mutation rates in a simple model of evolution," in *Complex Systems: Mechanism of Adaptation*, pp. 179–183, IOS Press, Amsterdam, The Netherlands, 1994.

[136] S. Ohno, *Evolution by Gene Duplication*, Springer, Berlin, Germany, 1970.

[137] M. Nei, "Gene duplication and nucleotide substitution in evolution," *Nature*, vol. 221, no. 5175, pp. 40–42, 1969.

[138] J. Spring, "Genome duplication strikes back," *Nature Genetics*, vol. 31, no. 2, pp. 128–129, 2002.

[139] J. Zhang, "Evolution by gene duplication: an update," *Trends in Ecology and Evolution*, vol. 18, no. 6, pp. 292–298, 2003.

[140] J. A. Bailey, Z. Gu, R. A. Clark, et al., "Recent segmental duplications in the human genome," *Science*, vol. 297, no. 5583, pp. 1003–1007, 2002.

[141] M. Nei and A. P. Rooney, "Concerted and birth-and-death evolution of multigene families," *Annual Review of Genetics*, vol. 39, no. 15, pp. 121–152, 2005.

[142] P. M. Harrison, H. Hegyi, S. Balasubramanian, et al., "Molecular fossils in the human genome: identification and analysis of the pseudogenes in chromosomes 21 and 22," *Genome Research*, vol. 12, no. 2, pp. 272–280, 2002.

[143] J. Zhang, H. F. Rosenberg, and M. Nei, "Positive Darwinian selection after gene duplication in primate ribonuclease genes," *Proceedings of the National Academy of Sciences (USA)*, vol. 95, no. 7, pp. 3708–3713, 1998.

[144] D. Dykhuizen and D. L. Hartl, "Selective neutrality of 6PGD allozymes in *E. coli* and the effects of genetic background," *Genetics*, vol. 96, no. 4, pp. 801–817, 1980.

[145] L. Altenberg, "Knowledge representation in the genome: new genes, exons, and pleiotropy," *Genetics*, vol. 110, p. s41, 1985.

[146] L. Altenberg, "Genome growth and the evolution of the genotype-phenotype map," in *Proceedings of the Evolution and Biocomputation Workshop*, vol. 899 of *Lecture Notes in Computer Science*, pp. 205–259, Springer, 1995.

[147] D. Sankoff, "Gene and genome duplication," *Current Opinion in Genetics and Development*, vol. 11, no. 6, pp. 681–684, 2001.

[148] F. Gruau, "Genetic micro programming of neural networks," in *Advances in Genetic Programming*, pp. 495–518, MIT Press, Cambridge, Mass, USA, 1994.

[149] J. R. Koza, "Gene duplication to enable genetic programming to concurrently evolve both the architecture and work-performing steps of a computer program," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, pp. 695–717, Morgan Kaufmann, San Francisco, CA, USA, 1995.

[150] K. Schmitt, "Using gene duplication and gene deletion in evolution strategies," Tech. Rep., Collaborative Research Center 531, University of Dortmund, Dortmund, Germany, 2004.

[151] M. Schutz, "Other operations: gene duplication and deletion," in *Handbook of Evolutionary Computation*, pp. 8–15, Oxford University Press, Oxford, UK, 1997.

[152] R. Calabretta, S. Nolfi, D. Parisi, and G. P. Wagner, "Duplication of modules facilitates the evolution of functional specialization," *Artificial Life*, vol. 6, no. 1, pp. 69–84, 2000.

[153] M. C. King and A. C. Wilson, "Evolution at two levels in humans and chimpanzees," *Science*, vol. 188, no. 4184, pp. 107–116, 1975.

[154] J. H. Moore, F. W. Asselbergs, and S. M. Williams, "Bioinformatics challenges for genome-wide association studies," *Bioinformatics*, vol. 26, no. 4, pp. 445–455, 2010.

[155] D. Fredman, S. J. White, S. Potter, E. E. Eichler, et al., "Complex SNP-related sequence variation in segmental genome duplications," *Nature Genetics*, vol. 36, no. 8, pp. 861–866, 2004.

[156] C. J. Shaw and J. R. Lupski, "Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease," *Human Molecular Genetics*, vol. 13, no. 1, pp. R57–R64, 2004.

[157] D. Sankoff, "Rearrangements and chromosomal evolution," *Current Opinion in Genetics and Development*, vol. 13, no. 6, pp. 583–587, 2003.

[158] D. Posada, K. A. Crandall, and E. C. Holmes, "Recombination in evolutionary genomics," *Annual Review of Genetics*, vol. 36, pp. 75–97, 2002.

[159] G. P. Smith, "Evolution of repeated DNA sequences by unequal crossover," *Science*, vol. 191, no. 4227, pp. 528–535, 1976.

[160] A. S. Kondrashov, "Deleterious mutations and the evolution of sexual reproduction," *Nature*, vol. 336, no. 6198, pp. 435–440, 1988.

[161] L. Hadany and T. Beker, "Sexual selection and the evolution of obligatory sex," *BMC Evolutionary Biology*, vol. 7, no. 1, article 245, 2007.

[162] A. Kong, D. F. Gudbjartsson, J. Sainz, et al., "A high-resolution recombination map of the human genome," *Nature Genetics*, vol. 31, no. 3, pp. 241–247, 2002.

[163] A. E. Eiben, "Multiparent recombination in evolutionary computing," in *Advances in Evolutionary Computing: Theory and Applications*, pp. 175–192, Springer, Berlin, Germany, 2003.

[164] D. E. Goldberg, B. Korb, and K. Deb, "Messy genetic algorithms: motivation, analysis, and first results," *Complex Systems*, vol. 3, no. 5, pp. 493–530, 1989.

[165] I. Harvey and I. Harvey, "The SAGA cross: the mechanics of recombination for species with variable length genotypes," in *Proceedings of the 2nd International Conference on Parallel Problem Solving from Nature (PPSN '92)*, pp. 269–278, North-Holland, 1992.

[166] W. B. Langdon, "Size fair and homologous tree genetic programming crossovers," *Genetic Programming and Evolvable Machines*, vol. 1, no. 1, pp. 95–119, 2000.

[167] P. Nordin, W. Banzhaf, and F. Francone, "Efficient evolution of machine code for cisc architectures using blocks and homologous crossover," in *Advances in Genetic Programming*, vol. 3, pp. 275–299, MIT Press, Cambridge, Mass, USA, 1999.

[168] R. Poli and N. F. Mcphee, "Exact schema theory for gp and variable-length gas with homologous crossover," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '01)*, pp. 104–111, Morgan Kaufmann, San Francisco, CA, USA, 2001.

[169] M. Srinivas and L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 24, no. 4, pp. 656–667, 1994.

[170] J. W. Pepper, "The evolution of evolvability in genetic linkage patterns," *BioSystems*, vol. 69, no. 2-3, pp. 115–126, 2003.

[171] G. G. Simpson, "The Baldwin effect," *Evolution*, vol. 7, no. 2, pp. 110–117, 1953.

[172] H. A. Simon, "The organization of complex systems," in *Hierarchy Theory: The Challenge of Complex Systems*, pp. 1–27, George Braziller, 1973.

[173] H. A. Simon, "The architecture of complexity," *Proceedings of the American Philosophical Society*, vol. 106, pp. 467–482, 1962.

[174] L. Altenberg, "The schema theorem and Price's theorem," in *Fundations of Genetic Algorithms*, vol. 3, pp. 23–49, MIT Press, Cambridge, Mass, USA, 1995.

[175] J. T. Bonner, *The Evolution of Complexity by Means of Natural Selection*, Princeton University, Princeton, NJ, USA, 1988.

[176] J. E. G. Leigh, "Levels of selection, poten-tial conflicts, and their resolution: the role of the "common good"," in *Levels of Selection in Evolution*, pp. 15–30, Princeton University Press, Princeton, NJ, USA, 1999.

[177] S. T. Kazadi, D. Lee, R. Modi, J. Sy, and W. Lue, "Levels of compartmentalization in artificial evolution," in *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO '00)*, pp. 841–848, Morgan Kaufmann, San Francisco, CA, USA, 2000.

[178] S. C. Manrubia and C. Briones, "Modular evolution and increase of functional complexity in replicating RNA molecules," *RNA*, vol. 13, no. 1, pp. 97–107, 2007.

[179] T. Back, "Self-adaptation in genetic algorithms," in *Proceedings of the 1st European Conference on Artificial Life: Toward a Practice of Autonomous Systems*, pp. 263–271, MIT Press, 1992.

[180] N. Hansen and A. Ostermeier, "Completely derandomized self-adaptation in evolution strategies," *Evolutionary Computation*, vol. 9, no. 2, pp. 159–195, 2001.

[181] R. Hinterding, Z. Michalewicz, and A. Eiben, "Adaptation in evolutionary computation: a survey," in *Proceedings of the IEEE Conference on Evolutionary Computation (CEC '97)*, pp. 65–69, IEEE, 1997.

[182] H. Kita, "A comparison study of self-adaptation in evolution strategies and real-coded genetic algorithms," *Evolutionary Computation*, vol. 9, no. 2, pp. 223–241, 2001.

[183] B. Edmonds, "Meta-genetic programming: co-evolving the operators of variation," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 9, pp. 13–29, 2001.

[184] W. Kantschik, P. Dittrich, M. Brameier, and W. Banzhaf, "Empirical analysis of different levels of meta-evolution,"

in *Proceedings of the Congress on Evolutionary Computation (CEC '99)*, vol. 3, pp. 2086–2093, IEEE Press, 1999.

[185] L. Spector and A. Robinson, "Genetic programming and autoconstructive evolution with the push programming language," *Genetic Programming and Evolvable Machines*, vol. 3, no. 1, pp. 7–40, 2002.

[186] J. Tavares, P. Machado, A. Cardoso, F. B. Pereira, and E. Costa, "On the evolution of evolutionary algorithms," in *Proceedings of the 7th European Conference on Genetic Programming (EuroGP '04)*, vol. 3003 of *Lecture Notes in Computer Science*, pp. 389–398, Springer, 2004.

[187] M. Parter, N. Kashtan, and U. Alon, "Facilitated variation: how evolution learns from past environments to generalize to new environments," *PLoS Computational Biology*, vol. 4, no. 11, article e1000206, 2008.

[188] C. D. Allis, T. Jenuwein, and D. Reinberg, Eds., *Epigenetics*, Cold Spring Harbor Laboratory, Woodbury, NY, USA, 2007.

[189] J. S. Mattick, "Challenging the dogma: the hidden layer of non-protein-coding RNAs in complex organisms," *BioEssays*, vol. 25, no. 10, pp. 930–939, 2003.

[190] S. A. Kauffman, "Metabolic stability and epigenesis in randomly constructed genetic nets," *Journal of Theoretical Biology*, vol. 22, no. 3, pp. 437–467, 1969.

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

[192] R. Albert and H. G. Othmer, "The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*," *Journal of Theoretical Biology*, vol. 223, no. 1, pp. 1–18, 2003.

[193] C. Espinosa-Soto, P. Padilla-Longoria, and E. R. Alvarez-Buylla, "A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles," *Plant Cell*, vol. 16, no. 11, pp. 2923–2939, 2004.

[194] S. Huang, G. Eichler, Y. Bar-Yam, and D. E. Ingber, "Cell fates as high-dimensional attractor states of a complex gene regulatory network," *Physical Review Letters*, vol. 94, no. 12, Article ID 128701, 2005.

[195] S. Huang and D. E. Ingber, "Shape-dependent control of cell growth, differentiation, and apoptosis: switching between attractors in cell regulatory networks," *Experimental Cell Research*, vol. 261, no. 1, pp. 91–103, 2000.

[196] M. Aldana, E. Balleza, S. Kauffman, and O. Resendiz, "Robustness and evolvability in genetic regulatory networks," *Journal of Theoretical Biology*, vol. 245, no. 3, pp. 433–448, 2007.

[197] E. Balleza, E. R. Alvarez-Buylla, A. Chaos, S. Kauffman, I. Shmulevich, and M. Aldana, "Critical dynamics in genetic regulatory networks: examples from four kingdoms," *PLoS One*, vol. 3, no. 6, article e2456, 2008.

[198] M. Nykter, N. D. Price, M. Aldana, et al., "Gene expression dynamics in the macrophage exhibit criticality," *Proceedings of the National Academy of Sciences (USA)*, vol. 105, no. 6, pp. 1897–1900, 2008.

[199] R. Serra, M. Villani, A. Graudenzi, and S. A. Kauffman, "Why a simple model of genetic regulatory networks describes the distribution of avalanches in gene expression data," *Journal of Theoretical Biology*, vol. 246, no. 3, pp. 449–460, 2007.

[200] R. Serra, M. Villani, and A. Semeria, "Genetic network models and statistical properties of gene expression data in knock-out experiments," *Journal of Theoretical Biology*, vol. 227, no. 1, pp. 149–157, 2004.

[201] I. Shmulevich, S. A. Kauffman, and M. Aldana, "Eukaryotic cells are dynamically ordered or critical but not chaotic," *Proceedings of the National Academy of Sciences (USA)*, vol. 102, no. 38, pp. 13439–13444, 2005.

[202] J. R. Stone and G. A. Wray, "Rapid evolution of cis-regulatory sequences via local point mutations," *Molecular Biology and Evolution*, vol. 18, no. 9, pp. 1764–1770, 2001.

[203] A. Wagner, "Evolution of gene networks by gene duplications: a mathematical model and its implications on genome organization," *Proceedings of the National Academy of Sciences (USA)*, vol. 91, no. 10, pp. 4387–4391, 1994.

[204] J. W. Fondon III and H. R. Garner, "Molecular origins of rapid and continuous morphological evolution," *Proceedings of the National Academy of Sciences (USA)*, vol. 101, no. 52, pp. 18058–18063, 2004.

[205] W. Banzhaf, "Artificial regulatory networks and genetic programming," in *Genetic Programming Theory and Practice*, pp. 43–62, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2003.

[206] W. Banzhaf, "On the dynamics of an artificial regulatory networks," in *Proceedings of the 7th European Conference on Artificial Life: Advances in Artificial Life*, vol. 2801 of *Lecture Notes in Artificial Intelligence*, pp. 217–227, Springer, 2003.

[207] W. Banzhaf and P. D. Kuo, "Network motifs in natural and artificial transcriptional regulatory networks," *Journal of Biological Physics and Chemistry*, vol. 4, no. 2, pp. 85–92, 2004.

[208] P. Kuo, W. Banzhaf, and A. Leier, "Network topology and the evolution of dynamics in an artificial genetic regulatory network model created by whole genome duplication and divergence," *BioSystems*, vol. 85, no. 3, pp. 177–200, 2006.

[209] A. Chavoya and Y. Duthen, "Evolving an artificial regulatory network for 2D cell patterning," in *Proceedings of the IEEE Symposium on Artificial Life (ALIFE '07)*, pp. 47–53, IEEE, Honolulu, Hawaii, USA, 2007.

[210] C. Mattiussi and D. Floreano, "Analog genetic encoding for the evolution of circuits and networks," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 5, pp. 596–607, 2007.

[211] C. L. Nehaniv, "Self-replication, evolvability and asynchronicity in stochastic worlds," in *Stochastic Algorithms: Foundations and Applications*, pp. 126–169, Springer, Berlin, Germany, 2005.

[212] J. S. Mattick, "RNA regulation: a new genetics?" *Nature Reviews Genetics*, vol. 5, no. 4, pp. 316–323, 2004.

[213] C. P. Mack and G. K. Owens, "Regulation of smooth muscle α-actin expression in vivo is dependent on CArG elements within the 5′ and first intron promoter regions," *Circulation Research*, vol. 84, no. 7, pp. 852–861, 1999.

[214] A. Gonçalves and E. Costa, "A computational model of gene regulatory networks and its topological properties," in *Proceedings of the 11th International Conference on the Simulation and Synthesis of Living Systems (ALIFE '08)*, pp. 204–211, 2008.

[215] R. Jaenisch and A. Bird, "Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals," *Nature Genetics*, vol. 33, pp. 245–254, 2003.

[216] S. B. Baylin and J. G. Herman, "DNA hypermethylation in tumorigenesis: epigenetics joins genetics," *Trends in Genetics*, vol. 16, no. 4, pp. 168–174, 2000.

[217] C. D. Allis, T. Jenuwein, and D. Reinberg, "Overview and concepts," in *Epigenetics*, pp. 23–61, Cold Spring Harbor Laboratory, Woodbury, NY, USA, 2007.

[218] T. Jenuwein and C. D. Allis, "Translating the histone code," *Science*, vol. 293, no. 5532, pp. 1074–1080, 2001.

[219] A. P. Wolffe and M. A. Matzke, "Epigenetics: regulation through repression," *Science*, vol. 286, no. 5439, pp. 481–486, 1999.

[220] S. I. S. Grewal and D. Moazed, "Heterochromatin and epigenetic control of gene expression," *Science*, vol. 301, no. 5634, pp. 798–802, 2003.

[221] A. S. Uyar and A. E. Harmanci, "A new population based adaptive domination change mechanism for diploid genetic algorithms in dynamic environments," *Soft Computing*, vol. 9, no. 11, pp. 803–814, 2005.

[222] W. Arthur, "The effect of development on the direction of evolution: toward a twenty-first century consensus," *Evolution and Development*, vol. 6, no. 4, pp. 282–288, 2004.

[223] R. A. Raff, *The Shape of Life: Genes, Development, and the Evolution of Animal Form*, University of Chicago Press, Chicago, Ill, USA, 1996.

[224] M. West-Eberhard, *Developmental Plasticity and Evolution*, Oxford University Press, Oxford, UK, 2003.

[225] A. D. Bradshaw, "Unravelling phenotypic plasticity—why should we bother?" *New Phytologist*, vol. 170, no. 4, pp. 644–648, 2006.

[226] T. J. Garland and S. A. Kelly, "Phenotypic plasticity and experimental evolution," *The Journal of Experimental Biology*, vol. 209, no. 12, pp. 2344–2361, 2006.

[227] T. D. Price, A. Qvarnstr, and D. E. Irwin, "The role of phenotypic plasticity in driving genetic evolution," *Proceedings of the Royal Society B*, vol. 270, no. 1523, pp. 1433–1440, 2003.

[228] F. Valladares, D. Sanchez-Gomez, and M. A. Zavala, "Quantitative estimation of phenotypic plasticity: bridging the gap between the evolutionary concept and its ecological applications," *Journal of Ecology*, vol. 94, no. 6, pp. 1103–1116, 2006.

[229] W. T. Wcislo, "Behavioral environments and evolutionary change," *Annual Review of Ecology and Systematics*, vol. 20, pp. 137–169, 1989.

[230] J. Draghi and G. P. Wagner, "Evolution of evolvability in a developmental model," *Evolution*, vol. 62, no. 2, pp. 301–315, 2008.

[231] J. C. Gerhart and M. W. Kirschner, *Cells, Embryos, and Evolution*, Blackwell Science, Oxford, UK, 1997.

[232] P. Bentley and S. Kumar, "Three ways to grow designs: a comparison of embryogenies for an evolutionary design problem," in *Proceedings of the Genetic and Evolutionary Computation Conference Workshop on Evolvability (GECCO '99)*, pp. 35–43, Morgan Kaufmann, San Francisco, CA, USA, 1999.

[233] F. Gruau, D. Whitley, and L. Pyeatt, "A comparison between cellular encoding and direct encoding for genetic neural networks," in *Proceedings of the 1st Annual Conference on Genetic Programming (GECCO '96)*, pp. 81–89, MIT Press, Cambridge, Mass, USA, 1996.

[234] G. S. Hornby and J. B. Pollack, "Creating high-level components with a generative representation for body-brain evolution," *Artificial Life*, vol. 8, no. 3, pp. 223–246, 2002.

[235] K. O. Stanley, "Generative and developmental systems," in *Proceedings of the Annual Conference Companion on Genetic and Evolutionary Computation Conference (GECCO '09)*, pp. 3335–3354, ACM, New York, NY, USA, 2009.

[236] G. S. Hornby and J. B. Pollack, "The advantages of generative grammatical encodings for physical design," in *Proceedings of the IEEE Conference on Evolutionary Computation (CEC '01)*, vol. 1, pp. 600–607, IEEE, 2001.

[237] S. K. Sharma and G. W. Irwin, "Fuzzy coding of genetic algorithms," *IEEE Transactions on Evolutionary Computation*, vol. 7, no. 4, pp. 344–355, 2003.

[238] W. Banzhaf and J. Miller, "The challenge of complexity," in *Frontiers in Evolutionary Computation*, pp. 243–260, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

[239] D. Roggen and D. Federici, "Multi-cellular development: is there scalability and robustness to gain?" in *Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN '04)*, vol. 3242 of *Lecture Notes in Computer Science*, pp. 391–400, 2004.

[240] S. Harding and W. Banzhaf, "Artificial development," in *Organic Computing*, pp. 201–220, Springer, Berlin, Germany, 2008.

[241] S. Harding, J. F. Miller, and W. Banzhaf, "Self modifying cartesian genetic programming: fibonacci, squares, regression and summing," in *Proceedings of the 12th European Conference on Genetic Programming (EuroGP '09)*, vol. 5481 of *Lecture Notes in Computer Science*, pp. 133–144, Springer, Tubingen, Germany, 2009.

[242] A. P. Hendry, "Evolutionary biology: the power of natural selection," *Nature*, vol. 433, no. 7027, pp. 694–695, 2005.

[243] L. Ancel-Meyers, F. D. Ancel, and M. Lachmann, "Evolution of genetic potential," *PLoS Computational Biology*, vol. 1, no. 3, pp. 236–243, 2005.

[244] A. Crombach and P. Hogeweg, "Evolution of evolvability in gene regulatory networks," *PLoS Computational Biology*, vol. 4, no. 7, article e1000112, 2008.

[245] B. L. Phillips and R. Shine, "Adapting to an invasive species: toxic cane toads induce morphological change in Australian snakes," *Proceedings of the National Academy of Sciences (USA)*, vol. 101, no. 49, pp. 17150–17155, 2004.

[246] N. Kashtan and U. Alon, "Spontaneous evolution of modularity and network motifs," *Proceedings of the National Academy of Sciences (USA)*, vol. 102, no. 39, pp. 13773–13778, 2005.

[247] S. Collins, J. de Meaux, and C. Acquisti, "Adaptive walks toward a moving optimum," *Genetics*, vol. 176, no. 2, pp. 1089–1099, 2007.

[248] T. Blickle and L. Thiele, "A comparison of selection schemes used in evolutionary algorithms," *Evolutionary Computation*, vol. 4, no. 4, pp. 361–394, 1996.

[249] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms 1*, pp. 69–93, Morgan Kaufmann, San Mateo, CA, USA, 1991.

[250] M. Kirkpatrick and V. Ravigne, "Speciation by natural and sexual selection: models and experiments," *The American Naturalist*, vol. 159, no. 3, pp. S22–S35, 2002.

[251] R. Lande, "Models of speciation by sexual selection on polygenic traits," *Proceedings of the National Academy of Sciences (USA)*, vol. 78, no. 6, pp. 3721–3725, 1981.

[252] N. Colegrave, "Sex releases the speed limit on evolution," *Nature*, vol. 420, no. 6916, pp. 664–666, 2002.

[253] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.

[254] G. F. Miller, "Exploiting mate choice in evolutionary computation: sexual selection as a process of search, optimization, and diversification," in *Selected Papers from AISB Workshop on Evolutionary Computing*, pp. 65–79, Springer, Berlin, Germany, 1994.

[255] G. F. Miller, *The Mating Mind: How Sexual Choice Shaped Human Nature*, Anchor, Harpswell, Me, USA, 2001.

[256] J. H. Gillespie, *Population Genetics: A Concise Guide*, Johns Hopkins University Press, Baltimore, MD, USA, 2004.

[257] J. F. Crow and M. Kimura, *An Introduction to Population Genetics Theory*, Harper and Row, New York, NY, USA, 1970.

[258] H. A. Orr, "Fitness and its role in evolutionary genetics," *Nature Reviews Genetics*, vol. 10, no. 8, pp. 531–539, 2009.

[259] E. Sober, "The two faces of fitness," in *Thinking about Evolution: Historical, Philosophical, and Political Perspectives*, pp. 309–321, Cambridge University Press, Cambridge, UK, 2001.

[260] Y. Jin, "A comprehensive survey of fitness approximation in evolutionary computation," *Soft Computing*, vol. 9, no. 1, pp. 3–12, 2005.

[261] M. B. Gerstein, C. Bruce, J. S. Rozowsky, et al., "What is a gene, post-ENCODE? History and updated definition," *Genome Research*, vol. 17, no. 6, pp. 669–681, 2007.

[262] T. R. Gingeras, "Origin of phenotypes: genes and transcripts," *Genome Research*, vol. 17, no. 6, pp. 682–690, 2007.

[263] K. Scherrer and J. Jost, "The gene and the genon concept: a functional and information-theoretic analysis," *Molecular Systems Biology*, vol. 3, p. 87, 2007.

[264] P. F. Stadler, S. J. Prohaska, C. V. Forst, and D. C. Krakauer, "Defining genes: a computational framework," *Theory in Biosciences*, vol. 128, no. 3, pp. 165–170, 2009.

[265] E. Birney, J. A. Stamatoyannopoulos, A. Dutta, et al., "Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project," *Nature*, vol. 447, no. 7146, pp. 799–816, 2007.