The Physics–Biology Continuum Refutes Darwinism: Evolution is Directed by the Homeostasis-Dependent Bidirectional Relation between Genome and Phenotype

Didier Auboeuf
ENS de Lyon, Univ Claude Bernard, CNRS UMR 5239, INSERM U1210, Laboratory of Biology and Modelling of the Cell, 46 Allée d'Italie, Site Jacques Monod, F-69007, Lyon, France; Mail: didier.auboeuf@inserm.fr

Abstract

The physics–biology continuum relies on the fact that life emerged from prebiotic molecules. Here, I argue that life emerged from the physical coupling between the synthesis of nucleic acids and the synthesis of amino acid polymers. Owing to this physical coupling, amino acid polymers (or proto-phenotypes) maintained the physicochemical parameter equilibria (proto-homeostasis) in the immediate environment of their encoding nucleic acids (or proto-genomes). This protected the proto-genome physicochemical integrity (i.e., atomic composition) from environmental physicochemical stresses, and therefore increased the probability of reproducing the proto-genome without variation. From there, genomes evolved depending on the biological activities they generated in response to environmental fluctuations. Thus, a genome generating an internal environment whose physicochemical parameters guarantee homeostasis and genome integrity has a higher probability to be reproduced without variation and therefore to reproduce the same phenotype in offspring. Otherwise, the genome is modified by the imbalances of the internal physicochemical parameters it generates, until new emerging biological activities maintain homeostasis. In sum, evolution depends on feedforward and feedback loops between genome and phenotype, since the internal physicochemical conditions that a genome generates in response to environmental fluctuations in turn either guarantee the stability or direct the variation of the genome.

Introduction

Living organisms are made of cells, which comprise molecules that are composed of atoms. This hierarchical organization implies that life emerged from prebiotic molecules and therefore that biological laws emerge from physical laws. For example, reproduction, one of the two main properties of living organisms, emerges from the replication of nucleic acids (RNA or DNA); in turn, replication emerges from base pairing—i.e., from the geometric arrangement of atoms in nucleotides—as demonstrated by James Watson and Francis Crick [1,2]. Of course, living organism reproduction cannot be reduced to nucleotide physical properties, yet it still depends on base pairing. The second property of living organisms that corresponds to autonomy (or self-maintenance) relies on internal physicochemical parameter equilibria that depend on molecular physicochemical properties, such as hydropathy, folding, and catalytic activities, among many others [3-7].

The fact that life emerged from prebiotic molecules, implying that biological laws emerge from physical laws, defines the notion of physics–biology continuum. This notion is not equivalent to the reductionist point of view that prevailed in the 20th century. Reductionism means that life could be explained simply by knowing the underlying molecular properties and processes. However, as emphasized notably by Carl Woese and Denis Noble [8,9], biology cannot be reduced to chemistry and physics because i) the properties of a given organization scale

of living organisms (e.g., a cell) cannot be reduced to the sum of the properties of the lower organization scale (e.g., molecules); and ii) an upper organization scale can modify the properties of the lower organization scale. The first point corresponds to upward causation and to the notion of emergence, which relies on the fact that the properties of a system depend on the dynamic interactions between its elements, implying that the properties of a system are more complex than just the sum of the properties of its elements ([8-10]). The second point corresponds to downward causation, which can be illustrated by protein- and nucleic acid-biochemical modifications resulting from cellular activities that can modify in a dynamic and environment-depending manner the atomic composition—and therefore the properties— of cellular polymers ([9-11]). Since cellular molecules can be chemically modified by the cell, then the properties of a cell cannot be predicted just by knowing the intrinsic or initial properties of molecules. Even though biology cannot be reduced to physics and chemistry, the hierarchical organization of living organisms nonetheless implies that life emerged from pre-biotic molecules establishing a continuum between physics and biology.

This physics–biology continuum implies establishing a continuum between inert matter and living organisms, which may seem incongruous for some of our contemporaries. Nevertheless, it must be recalled that it was incongruous to a majority of scientists of the 19th century to establish a continuum between all forms of life, including humans. Yet, this is what some evolutionary scientists of the 19th century, in particular Jean-Baptiste de Lamarck, Alfred Russel Wallace and Charles Darwin, succeeded in doing by providing evidence of evolution [12,13]. Just as scientists of the 19th century unified all life forms within biology (a term coined by Lamarck), those of the 21st century will unify physics and biology by establishing a continuum between the different forms that matter can adopt.

The first objective of this article is to show that the unification of physics and biology leads to the refutation of Darwinism and the derived theories of the 20th century, such as the Modern Synthesis mostly based on the notions of random mutations and natural selection. The second objective is to show that the physics–biology continuum allows us to define evolution as the consequence of the bidirectional relationship between the physicochemical integrity of genomes and the physicochemical parameter equilibria (i.e., homeostasis) of the internal environment generated by genomes. Several clarifications are necessary for these objectives.

First, Darwinism is part of the modern scientific culture. As such, it is often presented as the Theory of Evolution, suggesting that evolution and Darwinism are synonymous. However, evolution is a fact, while Darwinism and the derived theories are only explanations of the course of evolution. Challenging Darwinism or the Modern Synthesis does not mean questioning evolution (e.g., [15-18]).

Second, it is not possible to easily summarize the current theories of evolution deriving from Darwinism due to the many amendments that have been made during the 20th century and that are still being proposed (e.g., [14,15,19]). Natural selection and chance (i.e., random mutations) remain however two central explanatory principles of current evolutionary theories. In this setting, the reality of natural selection and chance in biology is not to be denied. Natural selection is a fact, in the sense that an organism not adapted to its environment dies and does not reproduce. Similarly, chance (or stochasticity) is central to biology, as every biological activity depends on the Brownian motion of molecules that are not predetermined to meet (e.g.,

[16,20]). Questioning Darwinism and its derivatives does not mean questioning the reality of natural selection or chance, but it does mean questioning their explanatory power of evolution.

Third, the current evolutionary theories derive from Darwinism, which was built on observations made for complex multicellular organisms that were classified by naturalists like Carl von Linné (1707–1778) according to macroscopic phenotypic traits. Somehow, Darwinism was born in an attempt to explain living organism classification according to the principle that phenotypic traits can provide advantages or disadvantages with respect to the number of descendants or fitness [13]. According to this principle, evolution would rely on a unidirectional principle in which random mutations (genome variations) induce phenotypic variations, but in which the phenotype has no effect on genome variations. However, a living organism cannot be reduced to the sum of phenotypic traits nor to the number of its descendants. As defined by Claude Bernard and Walter Cannon during the second half of the 19th century and first half of the 20th, respectively [3,21], a living organism is composed of an internal environment whose physiochemical parameters are maintained around physiological values (i.e., homeostasis), outside of which the organism loses its autonomy [3,5,6,21,22]. What is the role of homeostasis in evolution? Could homeostasis as a product of the phenotype—and therefore of the genome—influence genome variations?

I will address these questions starting from life origin. I will argue that life emerged from the physical coupling between the synthesis of nucleic acids (or proto-genomes) and the synthesis of amino acid polymers (or proto-phenotypes), which sealed a bidirectional relationship between genome and phenotype through homeostasis. I will next show that this bidirectional relationship explains the evolution of unicellular and multicellular organisms. In a general framework and in a historical perspective, this article extends on elements that I have already published [23-26].

Part 1: Definition of the genome–phenotype relationship within the physics–biology framework

The genome–phenotype relationship is commonly described as a unidirectional cause-and-effect relationship, according to Francis Crick's formula known as the Central Dogma of Molecular Biology (gene → product → function → phenotype) or according to the formula that summarizes current evolutionary theories (random mutations → phenotypes → natural selection) [27]. This genome–phenotype unidirectional relationship implies that the phenotype cannot modify the genome. Focusing on the physical nature of DNA as a polymer, I will show that the physics–biology continuum instead reveals a bidirectional relationship between genome and phenotype through homeostasis, with consequences on genome variation.

It became customary during the 20th century to describe DNA as a carrier of information, and to represent DNA as a virtual sequence of four letters symbolizing nucleotides. These symbolic representations have facilitated the spreading of erroneous notions, such as mutations being described as random errors, justifying in part the central place given to chance in current evolutionary theories (e.g., [28-31]). In this context, emphasizing the continuum between physics and biology means acknowledging the physical nature of biological objects, and in particular that of DNA, a polymer composed of atoms. In doing so, it becomes easy to refute the proposition that mutations are the consequence of random errors and instead to define mutations as the

consequence of the interactions between DNA physicochemical properties and physicochemical conditions of the internal environment of living organisms. Indeed, atoms composing DNA can chemically react with a diversity of molecules present around it. These chemical reactions modify the nature or the geometric arrangement of the atoms in nucleotides and therefore their interactions with other nucleotides leading to mutations during replication (e.g., [2,31-33]). For example, oxidation of guanine enhances its interaction with adenine rather than cytosine, leading to G:C>T:A mutations during replication [33]. Importantly, the frequency, localization and nature of DNA physicochemical modifications are not random but depend on intrinsic parameters (i.e., the physicochemical properties of nucleotide combinations) and extrinsic parameters (i.e., the physicochemical stresses to which nucleotides are exposed). Indeed, depending on their composition, some nucleotide combinations can adopt structures that are more or less sensitive to physicochemical stresses or that can be differentially affected by oxidation–reduction reactions, alkylation, pH or temperature variations, among many other physicochemical parameters (e.g., [34-40]). In addition, each physicochemical stress can induce a particular modification at the atomic scale, leading to particular mutational biases, signatures or spectra (e.g., [35-41]). Thus, DNA physicochemical modifications and therefore genomic variations during replication (i.e., mutations) depend on the interactions between nucleotide-dependent physicochemical properties and internal environment physicochemical parameters. But what determines the internal environment physicochemical parameters of living organisms?

Emphasizing the physics—biology continuum means acknowledging the fact that DNA is not an information support system but rather is a polymer producing other polymers (i.e., RNA and proteins), the activities of which determine internal physicochemical parameters. Importantly, cells produce RNA and proteins (using DNA as a template) in response to environmental fluctuations, which balances internal physicochemical parameters when organisms are adapted to their environment [42,43]. Thus, when a living organism is adapted to its environment, it produces biopolymers from the DNA template that maintain the internal physicochemical parameters (i.e., homeostasis), which increases the probability of preserving the DNA physiochemical integrity and consequently the probability of reproducing DNA molecules without variation (Figure 1A, environment A). Otherwise, the internal physicochemical imbalances triggered by environmental fluctuations challenge the physicochemical integrity of DNA molecules, increasing the probability of mutations (Figure 1A, environment B).

In summary, in response to environmental fluctuations, a genome generates internal physicochemical conditions that can either maintain the genome integrity and therefore its reproduction without variation, or trigger genome variations. I will show in the next section that this principle emerged from the physical coupling of nucleic and amino acid polymer syntheses at the origin of life. I will show how the physical coupling of nucleic and amino acid polymer syntheses triggered the emergence of successive and interdependent layers of internal physicochemical environments (Figure 1B) that correspond to active interfaces between environmental- and genome-variations.

Part 2: Physical coupling between nucleic and amino acid polymer syntheses at the origin of life

There is a consensus concerning the fact that the molecules at the origin of life, some 4 billion years ago, were nucleic acids (probably RNA-like molecules or proto-RNAs) and amino acid polymers (probably peptides or proto-proteins) (e.g., [44-47]). This is supported by the fact that their complementary properties are the physicochemical foundations of the two main properties of living organisms: reproduction (relying on nucleic acid replication) and homeostasis (relying on protein activities).

Nucleic acid replication depends on two parameters: efficiency and "fidelity". Replication efficiency corresponds to the kinetics of stable covalent bond formation between nucleotides, which influences the number of molecular descendants. Replication fidelity depends on transient interactions between nucleotides (base pairing) and corresponds to the probability that the nucleotide order of newly synthetized polymers is complementary to that of the templates. While current evolutionary theories mainly focus on these two parameters, a third parameter, namely polymer physicochemical integrity, is fundamental to understanding nucleic acid evolution. Polymer physicochemical integrity corresponds to polymer resistance to hydrolysis and to polymer chemical modifications by molecules from the environment that can change the atomic composition in nucleotides.

Acknowledging the physicochemical integrity of nucleic acids as a main parameter of polymer evolution has two major interdependent consequences. First, depending on their physicochemical properties, some nucleotide combinations within a polymer (e.g., a proto-RNA) have a higher probability to be affected by some physicochemical stresses (see Part 1). This means that some nucleotide combinations have a higher probability than others to undergo mutations when exposed to certain physicochemical stresses (Figure 2A, 1). Second, and as a consequence, new nucleotide combinations arising from mutations have a higher probability to be reproduced in downstream replication cycles in the same environment when they change the physicochemical polymer properties, resulting in an increased polymer resistance or resilience to environmental physicochemical stresses that trigger their formation (Figure 2A, 2). Therefore, nucleic acid evolution was likely driven by feedforward and feedback loops, during which environmental physicochemical stresses induced nucleotide variations in “sensitive” regions, until new nucleotide combinations were no longer sensitive to the initial stresses. This principle explains the emergence of amino acid polymers and their relationship with nucleic acids.

The nucleotide-dependent physicochemical properties that can relax or counteract environmental physicochemical stresses are very limited, as the atomic composition and geometric arrangement of nucleotides is strongly constrained by replication itself. In contrast, amino acid polymers can "explore" and adopt a diversity of physicochemical properties and activities (e.g., [48-50]). For example, amino acids or small peptides can increase the physicochemical integrity of their interacting RNAs as well as the auto-replicative activity of ribozymes (i.e., RNA replicases), suggesting that amino acids or peptides (proto-proteins) chemically formed in the "primordial soup" could have contributed to simultaneously increasing the physicochemical integrity and replication of their interacting proto-RNAs (e.g., [48,50-56]). For example, P. van der Gulik and D. Speijer proposed that the interactions between proto-RNAs and aspartic acid (Asp)- and glycine (Gly)-containing peptides could have simultaneously enhanced proto-RNA polymerization while limiting their hydrolysis (due to

interactions between these peptides and magnesium ions, with the latter either catalyzing RNA polymerization or enhancing the reverse reaction, i.e., RNA hydrolysis) [54,55]. Thus, chemically generated small peptides in the "primordial soup" interacting preferentially (in the sense of statistical probability) with certain proto-RNA sequences could have increased the reproduction of their interacting proto-RNAs, notably by enhancing the proto-RNA physicochemical integrity. The physical interaction between proto-RNAs and proto-proteins according to their physicochemical affinity and complementarity likely contributed to establish the interdependency between these polymers, as supported by the genetic code organization and as proposed by G. Gamow, C. Woese and M. Yarus (e.g., [57-61]).

Because of the role of amino acids and peptides as cofactors of proto-RNA replication and stability, proto-RNAs interacting with amino acids or peptides that they themselves produce would have had a greater probability of being stabilized and replicated by reducing the random availability of cofactors (Figure 2B, 1 vs. 2). Supporting this notion, the genetic code organization shows that: i) proto-RNAs were likely directly involved in the biogenesis of some amino acids, as pointed out by J. Wong and M. Di Giulio (e.g., [60-64]); and ii) RNAs have a high propensity to interact with their encoded proteins, as notably shown by B. Zagrovic (e.g., [65]). However, the fact that proto-RNAs produced amino acid or peptide cofactors is unlikely to be enough to explain a strict interdependency between nucleic acids and amino acid polymers.

Indeed, proto-RNAs that did not produce peptide cofactors could have acted as parasites and used the cofactors produced by other proto-RNAs, which ultimately would have led to a dead-end. Supporting this notion, modeling and experimental approaches demonstrate that the free (i.e., random) diffusion of RNA replicases produced by an RNA template leads to ending its replication cycles, since randomly diffusing replicases can be "titrated" by mutants (or parasites), becoming smaller and losing their ability to produce an active replicase (e.g., [66-72]). Likewise, the free diffusion of peptide cofactors produced by some proto-RNAs and used by parasite proto-RNAs (i.e., that do not produce these cofactors) could have made peptides useless for the proto-RNAs producing them. Some physical compartmentalization must have been required to establish a tight interdependency between proto-RNAs and their encoded proteins.

In this setting, several authors, including C. Woese, K. Gordon, Y. Gounaris, B. Francis, A. Altstein, and C. Carter, have proposed that nucleic and amino acid polymer synthesis was physically coupled at the origin of life, as is still the case in prokaryotes (e.g., [44,73-81]). For example, proto-tRNAs composed of two or three nucleotides attached to an amino acid or a peptide could have been used simultaneously in RNA and protein synthesis, if: i) the nucleotides of the proto-tRNAs interacting with the template were directly incorporated into the nascent RNAs; and ii) the attached amino acids or peptides served as cofactors for nucleic acid polymerization while being simultaneously incorporated into the nascent peptides (see [23] for more details).

The key role of the physical coupling between nucleic and amino acid polymer synthesis in the emergence of life can be highlighted by two interdependent consequences. First, the physical coupling would have limited the free and random diffusion of protein cofactors, as nascent and neo-synthetized proteins would have had a higher probability to only enhance replication and the physicochemical integrity of their encoding RNA proto-genomes (Figure 2B, panel 2 vs 3). In a way, the physical coupling between RNA and protein synthesis

would have played the role of a rudimentary compartmentalization. Indeed, the physical coupling between RNA and protein synthesis implies that an RNA proto-genome producing a protein cofactor that increases only its own physicochemical integrity and/or replication has a greater probability of being accurately and efficiently reproduced. This gives the proto-genome molecular descendants a greater probability of reproducing the same protein cofactors, and so on. Conversely, when an RNA proto-genome produced an inefficient protein cofactor (e.g., impacting only its encoding RNA), this directly decreased the replication cycles of only the encoding RNA.

Second, the physical coupling between nucleic and amino acid polymer synthesis led to the emergence of a rudimentary proto-intracellular environment. Indeed, the physical proximity of newly synthetized proteins to the RNA proto-genome template being replicated buffered environmental physicochemical stresses on the proto-genome before, during, and after replication (Figure 2B, panel 3). For example, Asp- and Gly-containing proteins produced on a proto-genome could trap metal ions such as magnesium, thereby decreasing the probability of neo-synthesized RNA hydrolysis [54,55]. Likewise, the local accumulation of neo-synthesized RNA-protein complexes in the proximity of a replicating proto-genome would have favored the formation of aqueous phase without membranes (i.e., condensed liquid droplets) formed by RNA–protein complexes (e.g., [50,70,82-84]). Therefore, a molecular complex self-organization relying on the coupling of RNA and protein synthesis could have triggered the emergence of a rudimentary intracellular environment (or proto-intracellular environment) insofar it creates a physical “barrier” between the proto-genome and the environment (Figure 1B). A complementary step has been achieved through the production by proto-genomes of amphiphilic proteins that enhance lipid-depend formation of proto-membranes (e.g., [83,85-87]). Both membrane-less and membrane-dependent compartmentalization would have led to the emergence of a stable intracellular environment that protected the physicochemical integrity of the proto-genome from environmental fluctuations (Figure 1B). Cells could thus have emerged because a proto-genome producing proteins (biological activities) that generate a rudimentary intracellular-like environment have in turn protected the proto-genome physicochemical integrity, increasing the probability of reproducing the same proto-genome and therefore the same proteins (or phenotype) in offspring.

In summary, in a prebiotic world, environmental physicochemical stresses could have induced chemical modifications of certain nucleotide combinations depending on their physicochemical properties. These chemical modifications resulted in mutations during replication (Figure 2A). A consequence of this targeted phenomenon is that new variants that changed the physicochemical properties and increased the polymer resilience with respect to the initial stresses had a greater probability of being replicated during downstream replication cycles in the same environment (Figure 2C, level 1). In addition, the physical coupling between RNA and protein synthesis allowed the emergence of a rudimentary proto-intracellular environment that protected the genome physicochemical integrity by buffering or counteracting the initial stresses (Figure 2C, level 2). In other words, genomes and their deriving biological activities evolved depending on i) nucleic acid-dependent physicochemical properties in regards to environmental physicochemical stresses, and ii) the emergence of biological activities relaxing the initial stresses at the origin of genome variations. I will show in the next section that this principle explains the evolution of unicellular organisms.

Part 3: Feedforward and feedback loops between physiological and genetic adaptation in unicellular evolution

The first cellular organisms appeared probably more than 3.5 billions years ago (e.g., [83,85-89]). The cell is delimited by a membrane that isolates, from the environment, the intracellular compartment containing one or more DNA molecules. The DNA molecules of unicellular organisms have the property of being templates for both reproduction (DNA replication) and the cellular physiological adaptation in response to environmental fluctuations, through the synthesis of RNAs and their translation into proteins, as demonstrated by F. Crick, J. Watson, F. Jacob, and F. Monod [42,43]. This DNA property of unicellular organisms allows us to highlight the homeostasis-dependent bidirectional relationship between physiological adaptation (the phenotype) and genetic adaptation (i.e., genome evolution).

The physiological adaptation of unicellular organisms relies on a feedback loop during which genomic regions, solicited by fluctuations of environmental physicochemical parameters, produce RNA polymers that can be translated into proteins that maintain the intracellular physicochemical parameter equilibria (cellular homeostasis) [43] (Figure 3A, loop 1). However, if the initial intracellular equilibrium is not re-established and imbalance persists, the same genomic regions are constantly solicited, which can challenge their physicochemical integrity and potentially lead to mutations during replication. I have already listed numerous experimental data obtained from all forms of life [23-26] that show for example that the sustained transcriptional activity of a genomic region can induce a local physical instability of DNA (via DNA breaks), conflicts between RNA and DNA polymerases (transcription/replication conflicts), and/or the formation of single-stranded DNA that is exposed to reactive molecules of the intracellular environment and that can undergo chemical modifications (e.g., [90-99]). These observations demonstrate that the cell´s experience (i.e., the transcriptional response to physicochemical stresses) leaves physicochemical footprints (breaks or chemical modifications) on specific genomic regions (Figure 3A, 2). These footprints then increase the probability of mutations during replication. This notion is equivalent to what several authors have already proposed, including J. Cairns, B. Wright, S. Rosenberg, M. Saier, J. Shapiro, A. Danchin, A. Yona, and D. Nobles, who emphasized a continuum between stress cellular response/physiological adaptation and genetic adaptation (e.g., [100-107]).

If mutations are triggered by environment-dependent cellular stresses (i.e., “the cellular experience”), then mutations are not random, in terms of: i) frequency and location, ii) nature, and iii) cellular phenotype. Accordingly, i) the load and the location of mutations are not random but depend, for example, on gene transcriptional activity in all forms of life (e.g., [90-99]); ii) each physicochemical stress to which a cell is exposed (e.g., temperature, metabolite availability, oxygen) is associated with a mutational bias, -spectrum, or -signature for all forms of life (e.g., [41,94,108-111]); and iii) mutations are not random with respect to cellular phenotypes, for two main reasons. First, mutations occur in genomic locations (e.g., genes) that are involved in the phenotype being challenged by environmental stresses, and second, the mutational process only ends when new genetics variants relax the initial stress (Figure 3A, loop 3).

What does relaxation of the initial stress in unicellular organisms mean? New genetic variants can relax or counteract the initial stress at two levels. First, new genetic variants can increase the resistance or resilience of modified genomic regions in terms of physicochemical properties (Figure 3B, level 1). This means that genetic

variants that change the physicochemical properties of a challenged-DNA genomic region, such that this region becomes physically or chemically resistant to the initial stresses, have a higher probability of being reproduced in downstream replication cycles. This implies that the frequency of genetic variants can evolve in unicellular organisms without generating phenotypic traits that would be under the control of natural selection but rather, depend on their effects on the physicochemical properties of DNA challenged by environmental-dependent physicochemical stresses.

Second, new genetic variants can generate new biological activities (e.g., new protein activities) in offspring that can counteract the initial constraints and contribute to homeostasis challenged by environmental stresses (Figure 3B, level 2). By re-establishing homeostasis in offspring, such genetic variants maintain the physicochemical integrity of the descendant DNA and therefore have a greater probability of being reproduced during the offsprings´ reproduction. This implies that the frequency of genetic variants evolves based on their impact on homeostasis rather than on descendant number.

To illustrate the feedforward and feedback loops described in Figure 3B, I will provide two examples among several others [23] and emphasize a link between the modifications of DNA physical properties and the emergence of cellular phenotypes. The first example corresponds to the link between gene GC-content and gene expression level. As mentioned above, sustained transcriptional activation can lead to DNA breaks and/or transcription-replication conflicts. Because of the physical properties of A-T vs. G-C base stacking interactions, AT-rich regions are particularly sensitive to transcriptional-dependent physical stresses, due to their lower flexibility as compared to GC-rich regions. Accordingly, GC-rich genes are often more highly expressed than AT-rich genes in many species (e.g., [112-120]). Remarkably, DNA damages generated under transcription-induced physical stresses or under transcription-replication conflicts induce AT>GC mutations. For example, DNA breaks can trigger homologous recombination, which favors GC over AT nucleotides because of the physical properties of T:G mismatches, through a phenomenon called “GC-biased gene conversion” (e.g., [121-126]). Further, transcription-replication conflicts can trigger adenine deamination, resulting in A:T>G:C mutations, and the incorporation of GC instead of AT nucleotides can be favored in early phases of replication [95,127]. Collectively, these observations support a model where transcriptionally over-stimulated genes (for example, in response to environmental fluctuations) can mutate and become enriched in GC nucleotides over replication cycles. As a high GC-content increases DNA resistance to transcription-dependent physical stresses, we can conclude that transcription-dependent stresses on a genomic region lead to mutations that locally increase the GC-content, which in turn increases locally the genome´s resistance to physical constraints (Figure 3B, level 1). Importantly, a high gene GC-content increases gene expression efficiency for various physical reasons and at multiple levels, including transcription, RNA maturation, and translation (e.g., [112-120,128-136], and see [23] for more details). Thus, transcription-dependent physical stresses can induce AT>GC mutations (i.e., increase gene GC-content) across replication cycles, which in turn increases DNA resistance to the initial constraints in offspring (Figure 3B, level 1). Simultaneously, it increases the product amount of genes stimulated by the initial stress, which can in turn relax the initial constraint in offspring (Figure 3B, level 2). Of note, other mutational processes that affect

transcriptionally-stressed genes, including gene duplication and (retro)transposon insertion, can have a similar transcription-dependent relaxing effects (see [23] for details).

The convergence between physical DNA resistance to transcription-dependent physical stresses (Figure 3B, level 1) and the efficiency of gene product synthesis (Figure 3B, level 2) may seem fortuitous. In reality, this convergence is a direct consequence of the physical coupling between the synthesis of nucleic and amino acid polymers (see Part 2) that has imposed that nucleic and amino acid polymers and their synthesis processes have co-evolved and are co-adapted to the same physicochemical constraints as this is still visible in the genetic code organization (e.g., [60,61, 115,137-142]; and see [23] for more details).

Eukaryogenesis provides a second example illustrating the link between the complexity of genome organization and the complexity of cellular biological activities. L. Margulis proposed that eukaryotic cells originate about 2.4 billion years ago from the internalization by Archaea of proteobacteria (the ancestors of mitochondria) as a consequence of the oxygen rise on Earth [147,148]. The use of oxygen by proto-mitochondria for energy production induced the increase in the concentration of intracellular oxygen derivatives (or reactive oxygen species [ROS]) that can react with DNA and induce genomic instability. This is believed to have played a major role in the emergence of eukaryote features. For example, intracellular ROS production by proto-mitochondria could have stimulated the expansion of eukaryotic introns, which in turn protected exonic coding sequences from oxidation, as introns are larger than exons and have a higher probability of reacting with and thereby trapping ROS [149-152]. Supporting this possibility, intronic sequences varied at a much higher rate than exonic ones, and introns are richer in AT nucleotides than exons, which is a signature of GC nucleotide oxidation inducing GC>AT mutations (e.g., [33,135,153,154]). Moreover, histones bind to GC-rich exonic regions, which are more flexible than AT-rich regions, further protecting exons from oxidation (e.g., [112,149,153,155-157]). To summarize, intracellular ROS production by proto-mitochondria could have induced proto-eukaryote genome instability, which next led to the emergence of genomic features (e.g., intron expansion, exon-histone associations) that protected exonic coding sequences from oxidation (Figure 3B, level 1). In this model, eukaryote genomic organization emerged during evolution not as a consequence of selective advantage, but simply because the ROS-induced genome reorganization increased the physiochemical integrity of genomes, in particular that of exonic coding sequences, which in turn increased the probability of reproducing the same genome organization across generations despite intracellular ROS production.

In addition to protecting exonic coding sequences from oxidation, intron expansion and DNA-histone association (i.e., genome organization complexification) have allowed the diversification of biological activities encoded by eukaryotic genomes (i.e., phenotype complexification). For example, the presence of introns in eukaryotic genes allows alternative splicing, which generates a diversity of protein isoforms from a single gene and greatly contributes to the diversification of eukaryotic metabolic enzymes encoded by a single genome (e.g., [154,158,159]). Moreover, DNA–histone interactions led to the emergence of chromatin- and metabolic-dependent co-regulation of gene network expression, which increases the complexity of functional networks such as metabolic pathways (e.g., [157,160-165]). Thus, the raise of intracellular ROS triggered eukaryote genome reorganization that i) directly buffered the ROS-dependent genomic instability (Figure 3B, level 1), and

ii) contributed to the diversification of oxygen-dependent metabolic pathways, which in turn limited the ROS-dependent genomic instability (e.g., [148,161,166,167]) (Figure 3B, level 2).

The model depicted in Figure 3B can be challenged experimentally and is suitable for predictions. For example, a unicellular organism could be exposed to variations in a specific physicochemical parameter beyond physiological conditions. Repeating this experiment on several individual cells or cell populations would make it possible to analyze changes in gene expression and epigenetic modifications within single cells while sacrificing other cells for genetic analysis. A first prediction is that the occurrence of mutations in a cell population is statistically reproducible and predictable from local changes in chromatin, RNA biogenesis, and/or gene expression, which reflect the physiological adaptation of the organism. If the locations and nature of mutations can be predicted from footprints left on DNA/RNA in a cell experience–dependent manner, then mutations are not random errors but are the consequence of cell activities involved in physiological adaptation. The repeatability of genetic variations would therefore not be the consequence of natural selection but the consequence of the repeatability of mutational processes depending on cell physiological adaptation processes. A second prediction is that mutations whose frequency increases after several replication cycles induce relaxation of the initial stress i) by conferring new physicochemical properties to transcriptionally-challenged genomic regions (Figure 3B, level 1) and ii) by potentially allowing the emergence of new biological activities that more or less directly relax or counteract the initial stress (Figure 3B, level 2). These possibilities can be experimentally tested by artificially generating the predicted or identified genetic variations that should prevent the appearance of further mutations before exposing genetically modified cells to mutagenic stresses.

In summary, when a unicellular organism is physiologically adapted to its environment, its genome produces polymers (RNAs and proteins) that maintain homeostasis, and thus the physicochemical integrity of the genome and its reproduction without variation (Figure 3A, 1). When the organism is not adapted, the internal imbalances induced by the environmental stresses leave "footprints" on transcriptionally solicited genomic regions, resulting in mutations during replication (Figure 3A, 2). This mutational process stops when i) new genetic variants confer locally new physicochemical properties to the DNA making it resistant to the initial stresses (Figure 3B, Level 1) and ii) eventually allow the emergence of new biological activities in offspring that counteract the initial stresses (Figure 3B, Level 2). The homeostasis-dependent bidirectional relationship between physiological adaptation and genome variation in unicellulars is based on the fact that the same DNA molecules serve both physiological adaptation and reproduction (Figure 3A). What about multicellular organisms where physiological adaptation depends on somatic cells, while reproduction depends on germ cells?

Part 4: Feedforward and feedback loops between somatic cell–dependent homeostasis and germ cell-DNA physicochemical integrity in multicellular organisms

The ancestors of today´s multicellular organisms appeared 2 billion years ago and were probably cellular colonies (or aggregates) comprising eukaryotic cells coming from the division of a single cell (e.g., [168-172]). The formation of cellular colonies had two main consequences. First, it induced cellular specialization through the differential expression of genes, as different cells are exposed to different physicochemical constraints

depending on their location within the colony (e.g., [168,169,173-176]). Second, it led to the emergence of an intercellular space resulting from cell proximity and exchanges [177], which can be considered as the ancestor of the internal environment of multicellular organisms (proto-internal environment, Figure 1B). Both cellular specialization and the formation of an intercellular space allowed cellular colonies to be resilient to a wider range of environmental physicochemical stresses since for example they allowed to buffer environmental fluctuations [169,173-175,177]. As a consequence, cellular specialization and the formation of an intercellular space protected the cellular DNA physicochemical integrity from a wider range of environmental stresses. This means that colony formation at the origin of multicellular organisms did not represent a selective advantage. Rather, this phenotype (i.e., colony formation) was simply reproduced because it maintains the intra- and inter-cellular physicochemical equilibria (i.e., homeostasis), which maintains the integrity of the underlying genome, increasing in turn the probability that the phenotype is itself reproduced across generations (Figure 4A, 1 vs. 2).

However, the increase in cell number within individual colonies could have increased the concentration of potential genotoxic metabolic waste products (e.g., ROS) into the intercellular space, thereby triggering instability of genomes, including those of the proto-zygotes (i.e., cells that give rise to new colonies). This genome instability led to the emergence of meiosis, as the molecular ancestors of proteins involved in meiosis were involved in DNA repair, and as meiosis allows the removal of oxidized nucleotides via multiple phenomena, as shown by C. Berstein, H. Berstein, R. Michod, E. Horandl, and D. Speijer [178-181]. For example, DNA homologous recombination during meiosis corresponds initially to a DNA repair process and favors GC over AT nucleotides, which counteracts GC>AT mutations induced by GC-oxidation (e.g., [121,179,182-184]). Therefore, meiosis did not emerge because of a potential selective advantage but because it increased the probability of reproducing without variation a genome that gives rise to the same phenotype in offspring (Figure 4A, 3).

One consequence of meiosis is the formation of haploid cells (or gametes) derived from germ cells that do not participate directly to the organism's metabolic activities, but that benefit from the formation of a stable internal environment generated by somatic cell activities. This corresponds to Weismann's germ-soma barrier [185] and to the concept of division of labors, stating that germ cells are excluded from potentially genotoxic metabolic activities while being protected from environmental fluctuations thanks to somatic cell activities [174,180,186-189]. In other words, somatic cell activities (i.e., the phenotype of multicellular organisms) generate a stable internal environment through homeostasis, which protects germ cell DNA from environmental-dependent stresses. This increases the probability of reproducing the DNA molecules without variation giving rise to the same phenotype after fecundation (Figure 4A, 3).

As a direct consequence, if somatic cells do not maintain homeostasis, then the physicochemical integrity of germ cell DNA is challenged, increasing the probability of *de novo* mutations in germ cells. Several observations support this possibility. First, metabolic disorders (e.g., diabetes and obesity) in humans and in model organisms are associated with a global imbalance of sugar- and lipid-derivatives that result in germ cell DNA damages, including DNA oxidation, alkylation, and fragmentation (e.g., [190-196]). These observations demonstrate that the physicochemical integrity of germ cell DNA depends on homeostasis. Second, the location and nature of *de novo* mutations in germ cells are not random but depend on: i) the gene transcriptional activity

in germ cells; ii) the local chromatin organization of germ cells; and iii) the gene position with respect to replication origins (e.g., [197-205]). These observations are identical to those made in unicellular organisms (see Part 3), supporting a model whereby mutational processes in germ cells are directed by physicochemical parameters of the germ cell microenvironment produced by somatic cell activities (Figure 4B, 1).

In addition to triggering mutations in germ cell through internal physicochemical parameter imbalances, somatic cells may have a direct effect on *de novo* mutations in germ cells, through: i) somatic cell–dependent epigenetic changes within targeted loci of germ cell chromatin, and ii) somatic cell–produced small RNAs that are picked up by germ cells (Figure 4B, 2). These two processes have been referred as to inter- and trans-generational epigenetic transmission, as they occur in the germ cells of the F0 generation and can impact the development of the F1, F2, and F3 generations [206-215]. For example, dietary imbalances in parents can induce somatic cell–dependent chromatin modifications within targeted loci of germ cells, which then impacts the expression of epigenetically modified genes during the development of descendants (e.g., [207-209,212,216]). Importantly, as epigenetic marks and small RNAs can locally influence the rate and nature of mutations through a variety of processes, somatic-dependent epigenetic modifications of germ cells might direct germ cell DNA mutations (e.g., [206,211,217-219]), although this possibility is still under debate [220]. Of note, somatic cells that are challenged by environmental fluctuations produce circular DNA fragments and RNA molecules that are secreted within extracellular vesicles. These somatic cell–secreted molecules could be incorporated into germ cell DNA either directly or after reverse transcription, which would be similar to Horizontal Gene Transfer occurring widely in unicellular organisms and would represent another way by which somatic cells can modify germ cell DNA sequence [107,213,221,222].

To summarize, somatic cells can generate a stable internal environment contributing to maintain the physicochemical integrity of germ cell DNA, and therefore its reproduction without variation. This increases the probability of transmitting the same genome and generating the same phenotype in the next generations. However, when somatic cells are challenged by environmental changes, the internal environment—including the germ cell microenvironment—can be unbalanced and can challenge the physicochemical integrity of germ cell DNA, leading to *de novo* mutations. Therefore, a causal link (feedforward loop) can be established between environmental fluctuations that challenge somatic-dependent physicochemical equilibria of the internal environment, which in turn challenges germ cell DNA physicochemical integrity and triggers *de novo* mutations. This description breaks with the Darwinism-derived theories, which consider *de novo* mutations to be random errors independent from somatic cell activities and the environment.

The next question concerns the fate of *de novo* mutations (i.e., new genetic variants) across several generations. While current evolutionary theories propose that genetic variant frequency depends on the fitness of individuals (i.e., the number of descendants), it will be shown below that genetic variant frequency across generations depends on two interdependent feedback loops acting at the parental germ cell population level (Figure 4C, 1st generation) or at the descendant somatic cell population level (Figure 4C, 2nd generation).

What is the germ cell population, and what is its contribution to multicellular organism evolution? After a few divisions of the fertilized egg, a group of so-called primordial germ cells give rise to a dynamic cell

population through a few dozens to several hundred cell divisions, prior to meiosis and the formation of gametes [222-225]. Very importantly, among all produced germ cells, only a small number gives rise to mature gametes, and only a rather small proportion of gametes eventually gives rise to descendants after fertilization [223-227]. As gametes giving rise to offspring come from a cell population, this implies that the probability of transmission of genetic variants between two generations depends on their frequency within the parental germ cell population. Accordingly, a new genetic variant appearing early during germ cell differentiation has a greater probability of being found in a large number of gametes and therefore of being transmitted to the next generation [198,202,228,229]. Thus, a central question in evolution of multicellular organisms concerns the parameters that influence genetic variant frequency in the parental germ cell population.

As physicochemical stresses in the germ cell microenvironment induce *de novo* mutations, new genetic variants that counteract or relax the microenvironment-dependent initial stresses have a greater probability of being reproduced during the replication cycles of germ cells, and thus of being frequent in this population (Figure 4C, 1st generation). This principle is similar to that described for unicellular organisms (see Part 3). It is supported by the fact that new genetic variants and new genes in multicellular organisms are often associated with germ cell functions, indicating that the interaction between germ cells and their microenvironment plays a driving role in species evolution (e.g., [230-239]). Importantly, chromatin germ cells undergo massive reorganization, implying that all genomic regions can be exposed to physicochemical imbalances of the germ cell microenvironment [232,240,241]. In addition, and as a consequence of their massive chromatin reorganization, germ cells express a large variety of transcripts, and therefore many new genes may first be expressed in germ cells before being eventually expressed in somatic cells in the course of evolution [230-232,235-239]. In this setting, genetic variants (i.e., alleles) that are only expressed in germ cells can have major consequences on speciation. Indeed, new alleles that change the germ cell phenotype can ultimately lead to the emergence of a subpopulation of individuals who can no longer reproduce with the rest of the population. Such isolated subpopulations necessarily lead to the impoverishment of genetic variants and potentially to the emergence of phenotypic traits that are specific to the subpopulation. This phenomenon explains the importance of population size in phenotype evolution [242].

The feedback loop at the level of parental germ cells (Figure 4C, 1st generation) is not sufficient to explain the evolution of the frequency of new genetic variants across several generations when their cognate alleles are expressed in descendant somatic cells. To address this issue, four main possibilities must be taken into account. A first possibility is that the expression of new alleles in descendant somatic cells creates drastic imbalances outside of physiological ranges (early after fecundation, during development, or in adulthood) that lead to death without further descendants. This would correspond to the so-called negative natural selection, leading to allele elimination from the genetic pool.

A second possibility is that new alleles that induce more-or-less mild imbalances during development are permanently and selectively repressed in descendant somatic cells. For example, recessive pathological alleles have no effect when they co-exist with wild-type alleles. This can in part be explained by monoallelic repression that corresponds to the permanent repression of selected alleles in somatic cells through epigenetic

modifications initiated during development [243-245]. Importantly, repression of selected alleles can be transmitted from one generation to the other through a phenomenon called imprinting [246-252]. In this setting, imprinting could be associated with the elimination of repressed alleles from the genetic pool through a phenomenon called transmission ratio distortion. Transmission ratio distortion is an exception to Mendel's laws of equal segregation of parental alleles, as it results in the preferential transmission from F1 to F2 of alleles coming from one of the two F0 parents [253,254]. Although hypothetical, this phenomenon that is more frequent than previously anticipated, could depend on allele-specific epigenetic marks [180,229,253-259]. If proven, the coupling between imprinting and transmission ratio distortion could result in the elimination from the genetic pool -during gametogenesis, fecundation, or just after fecundation- of alleles that are epigenetically repressed across several generations.

A third possibility is that new alleles that induce imbalances of the internal environment in descendants escape the selected repression during development. Such alleles may induce imbalances in the descendant germ cell microenvironment, which in turn can trigger other rounds of mutational processes in descendant germ cells. This means that new alleles (i.e., new genetic variants) cannot be transmitted stably through several generations or can lead to the appearance of new mutants as long as they do not contribute to maintaining homeostasis.

Finally, the expression of new alleles in descendant somatic cells may contribute to maintaining the organismal homeostasis—i.e., the physicochemical parameter equilibrium of the germ cell microenvironment. In that case, the physicochemical integrity of germ cell DNA would be maintained, giving the germ cell DNA a higher probability to be transmitted without variation to the third generation (Figure 4C, 2nd generation).

How can new alleles that appear in parental germ cells and that maintain parental germ cell homeostasis (Figure 4C, 1st generation) allow directly the emergence in descendant somatic cells of new biological activities that maintain homeostasis at the organismal level in descendants (Figure 4C, 2nd generation)? The answer to this question is probably simpler than it seems. As the formation of a new allele is induced by internal physicochemical stresses and can only be frequent in the parental germ cell population when its expression relaxes or counteracts the initial stresses (Figure 4C, 1st generation), then the descendant somatic cells that are exposed during development to the same internal stresses can express these alleles and thus relax or counteract the stresses (Figure 4C, 2nd generation). As a consequence, descendant somatic cells expressing the new alleles contribute to maintaining the equilibria of the internal environment, which in turn allows the accurate replication of the descendant germ cell genome and therefore its transmission to the 3rd generation. Importantly, descendant somatic cells that express new alleles during development may acquire new features and/or undergo supplementary division cycles. This could drive the emergence of new cell populations during development and therefore the emergence of new phenotypic traits in adulthood. These phenotypic traits and the underlying genetic variants would next evolve depending on their effects on homeostasis. Eventually, natural selection could also differentially filter genetic variants having a similar effect on homeostasis in different individuals but allowing the emergence of phenotypic traits having different impacts on descendant number.

The model described in Figure 4C can be challenged experimentally and is suitable for predictions. Multicellular organisms could be exposed to variations in a specific physicochemical parameter in order to

measure the consequences on internal physicochemical parameters, in particular in the germ cell microenvironment. Thereafter, analysis of gene expression, epigenetic marks, and RNA content of individual germ cells and/or present in extracellular vesicles produced by somatic cells would make it possible to analyze how germ cells get physiologically impacted by these variations, eventually in a somatic cell-depending manner. At the same time, sequencing of single germ cell genomes across gametogenesis would make it possible to define the link between germ cell microenvironment variations and genetic variations in the germ population. A first prediction is that the occurrence of *de novo* mutations in germ cells is statistically reproducible and predictable from local changes in chromatin, RNA uptake or biogenesis, and/or gene expression in germ cells. A second prediction is that the frequency of new genetic variants in the germ cell population must be explainable in terms of their effects on releasing or counteracting physicochemical stresses generated by the germ cell microenvironment either because of their effects on DNA physicochemical properties or because of their effects on germ cell homeostasis (Figure 4C, 1st generation). It would next be necessary to study how the genetic variants transmitted from one generation to the next potentially allow the emergence of new somatic cell activities with respect to the initial environmental stresses. A third prediction is that new genetic variants allow the emergence within somatic cells of molecular innovations that counteract or absorb environmental physicochemical constraints at the whole organism level through homeostasis, either during development or in adulthood (Figure 4C, 2nd generation). These possibilities can be experimentally tested by artificially generating the predicted or identified genetic variations that should prevent the appearance of further mutations before exposing genetically modified organisms to mutagenic stresses.

In summary, maintenance of internal physicochemical parameter equilibria thanks to somatic cell activities despite environmental fluctuations guarantees the physicochemical integrity of germ cell DNA–thus its reproduction without variation (Figure 4A, 3). However, when organisms are not adapted to their environment, the resulting internal disequilibria trigger genomic variations (Figure 4B). The frequency of the resulting genetic variants over several generations depends on two successive and interdependent feedback loops. The first loop depends on whether parental germ cell genetic variants allow to counteract the physicochemical stresses of the parental internal environment (Figures 4C, 1st generation). This feedback loop has an impact on the frequency of genetic variants within the parental germ population and therefore on the probability of their transmission to the descendants. The second feedback loop depends on the potential consequences of genetic variants on descendant somatic cell activities (Figure 4C, 2nd generation). If the descendant somatic cells expressing new genetic variants maintain homeostasis, then the DNA of the descendants' germ cells is stable and is transmitted to the third generation without variation.

To conclude, a genome—in unicellular or multicellular organisms—generates biological activities in response to environmental fluctuations, which either create a stable internal environment that guarantees the physicochemical integrity of the genome and thus its reproduction without variation, or which instead create an internal environment whose physicochemical conditions direct genomic variations, inducing the appearance of new genetic variants. The frequency of these genetic variants over several generations then depends on their effects in counteracting the internal imbalances induced by environmental fluctuations.

Discussion

Current theories of evolution cannot be summarized easily, due to the many amendments that have been made over the past century (e.g., [14,15,19]). Nevertheless, two major explanatory principles—random mutations (or chance) and natural selection—are still presented to the general public as the scientific explanations of evolution. However, chance and natural selection cannot explain evolution because they generate ruptures (or discontinuities) between: i) physical and biological laws; ii) different physiological states of individuals; and iii) life and its evolution.

By definition, chance does not explain how mutations (or genetic variants) are physically generated. Eluding somehow this fundamental question, biologists have introduced the notion of errors. However, this notion is meaningless, as biological objects (such as DNA) are physical objects composed of atoms, and the physicochemical reactions triggering DNA variations do not make "mistakes" (see Part 1). Therefore, the notion of errors creates a rupture or discontinuity between physical and biological laws. In addition, placing chance at the center of an evolutionary theory of living organisms would be tantamount to explaining the existence of any physical object simply as the product of chance. If it can be said that the elements that compose a physical object are not predetermined to meet each other, science does not aim to explain why physical objects exist but how they are formed–that is, according to which laws and parameters. The same applies to genetic variants. If the notion of chance does not explain how genetic variants are formed, neither does natural selection, as natural selection is a filter acting once genetic variants are formed. Consequently, the current evolutionary theories based on chance and natural selection elude through the notion of "errors", one of the most important question to be addressed to understand evolution that concerns the physicochemical laws driving genetic variant formation.

By definition, natural selection cannot be a universal explanation of evolution since it only acts in two extreme situations: negative natural selection, which acts when imbalances of internal physicochemical parameters are such that individuals have no descendants; and positive natural selection, which can only explain the frequency of a very small proportion of genetic variants as natural selection can only act when two genetic variants have the same global effect on homeostasis while inducing the emergence of some phenotypic traits that impact the descendant number. Therefore, natural selection cannot explain what happens when genetic variants do not affect the number of descendants. This led to the notion of neutral natural selection, which is particularly meaningless when combined with the notions of random mutations. The very limited explanatory power of natural selection relies on the fact that the role that is attributed to natural selection in evolution is based on a wrong assumption. Indeed, to assert that natural selection has an effect on the frequency of genetic variants across several generations, it is necessary to assume that any phenotypic trait (and thus the underlying genetic variant) that allows an individual to have more descendants is necessarily transmitted to the offspring. However, this statement is false, as there is no relationship between the descendant number and the similarity between descendants and parents. For example, the frequency of genetic variants coming from individuals with very few but genetically identical descendants increases as compared to genetic variants coming from individuals with numerous but genetically distinct descendants. Similarly, while an individual may survive in an

environment thanks to a phenotypic trait, there is no guarantee that its DNA does not mutate during replication and reproduction, resulting in the loss of the advantageous phenotypic trait and the appearance of a potential new one in descendants. Natural selection is a fact, yet it has a very limited explanatory power, given the fact that advantageous traits in the Darwinian sense (and the underlying genetic variants) are not necessarily transmitted across generations, as there is no relation between the number of descendants and the fact that descendants share common traits with their parents.

Finally, current theories of evolution cannot explain evolution because they generate a rupture between life and its evolution. Current evolutionary theories are based on a cause–effect unidirectional principle (random mutations → phenotypes → natural selection). However, life relies on bidirectional relationships, with the effect impacting the cause. Indeed, life relies on the balance of internal physicochemical parameters that are constantly challenged by environmental fluctuations and that can be maintained only thanks to feedforward and feedback loops. Since life relies on feedforward and feedback loops, so does evolution. Genetic variants can only be reproduced during replication if they directly or indirectly counteract or relax the physicochemical stresses that triggered their formation (Figures 2C, 3B and 4C). In other words, biological activities (i.e., a phenotype), as a product of genetic variants, can only be reproduced across generations if they maintain the stability of the underlying genetic variants. This principle is perfectly explained from the moment that the products (e.g., proteins) of nucleic acids (proto-genomes) have contributed to maintaining the proto-genome physicochemical integrity (i.e., atomic composition, see Part 2). This bidirectional relationship between two polymers triggers the emergence of homeostasis, which is the result of the biogenesis of polymers from genome whose physicochemical integrity they guarantee. The bidirectional relationship between genome (physicochemical integrity) and phenotype (or internal physicochemical parameter equilibria) explains the evolution of increasingly complex biological activities triggering the emergence of successive layers of internal and interdependent environments, which ultimately and collectively contribute to maintaining the physiochemical integrity of the genome that generate these internal environments (Figure 1B).

To conclude, while notions such as errors, chance, or natural selection popularized evolution and therefore the continuum between all forms of life, these notions are now meaningless since evolution implies that living organisms emerged from prebiotic molecules. This implies in turn a continuum between physics and biology. The physics-biology continuum means that a genome (i.e., a polymer) generates other polymers (i.e., RNA, proteins) whose activities generate the internal physicochemical conditions on which genome stability and variation depend.

Figure legends

Figure 1:

A. A phenotype adapted to the environment (environment A) maintains, through the synthesis of biopolymers (RNAs and proteins), the physicochemical parameters of the internal/intracellular environment (homeostasis), which preserves the physicochemical integrity of DNA (stability). Therefore, an adapted phenotype increases the probability of reproducing the same DNA molecules during replication without variation. When the phenotype is not adapted to the environment (environment B), it induces physicochemical parameter imbalances in the internal/intracellular environment, challenging the physicochemical integrity of DNA (instability) and leading to mutations during replication.

B. Evolution of living organisms originated 4 billion years ago with the synthesis of amino acid polymers (proteins) from nucleic acids (e.g., RNAs), leading to the emergence of a self-organized proto-intracellular environment. From there, higher biological organization levels (i.e., different biological-dependent environments) emerged that evolved depending on their effects on the physicochemical integrity of the underlying genome that is constantly challenged by environmental fluctuations.

Figure 2:

A. Each local nucleotide combination (symbolized by different colors) has particular physicochemical properties, making it sensitive to certain physicochemical stresses. Therefore, depending on its composition and physicochemical properties, a nucleotide combination can have a higher probability of being modified when exposed to a particular physicochemical stress (1), which results in local nucleotide composition variations during replication (mutation). Consequently, a change in nucleotide composition that increases the resistance or resilience of the new combination to the initial environmental stresses has a higher probability of being replicated in downstream replication cycles in the same environment (2).

B. The molecules at the origin of life were probably self-replicating, RNA-like polymers (proto-RNAs). The physicochemical integrity (i.e., atomic composition) of proto-RNAs was constantly challenged by environmental fluctuations, leading to high random variability during replication (1). RNAs (or proto-genomes) producing protein cofactors (proto-phenotypes) that contribute to increasing the proto-genome physicochemical integrity had a higher probability of being replicated without variations (2). By being synthesized during RNA replication (3, coupling), proteins increased the physicochemical integrity and replication of the proto-genome (RNA) from which they directly derived. A consequence of the physical coupling of RNA and protein synthesis is the emergence of a self-organizing rudimentary proto-intracellular environment that contributes to maintaining the physicochemical integrity of the proto-genome.

C. Environmental physicochemical stresses can induce genetic variations—that is, the formation of new genetic variants (A and B)—in a proto-genome. The genetic variants that have a higher probability of being reproduced during subsequent replication cycles are those that generate new physicochemical properties of the proto-genome that make it resistant or resilient to the initial stresses (level 1), or that generate new physicochemical properties of proteins that counteract or relax the initial stresses (level 2). Consequently, the frequency of genetic variants across replication cycles depends on their feedback effects on the physicochemical stresses that trigger their formation.

Figure 3:

A. Environmental fluctuations can induce changes in the intracellular environment, which in turn induce gene transcription of RNA, whose translation into proteins (gene products) allows a return to the initial intracellular environment equilibrium (loop 1). As gene transcription generates physical stresses on DNA, a return to intracellular equilibrium releases the initial DNA constraints, thereby increasing the probability of reproducing

DNA without variations during replication. If the variations in environmental physicochemical parameters exceed physiological values, the biogenesis of gene products does not re-establish the initial cellular equilibria; DNA is therefore subject to sustained stresses, leaving marks or footprints (breaks or chemical modifications) on DNA (2, DNA damages). These footprints induce mutations during replication. When mutations confer new properties to DNA and/or generate new gene products that release the initial stresses on DNA, then the mutational process stops, and the new DNA is replicated without variation over reproductive cycles (loop 3).
B. Environmental physicochemical stresses can induce mutations or new genetic variants (A and B). The variants that induce the emergence of new physicochemical properties of the genome that make it resistant or resilient to the initial stresses have a greater probability of being reproduced in subsequent generations (level 1). Among these genetic variants, those that allow the emergence of new biological activities (phenotypes) that counteract or relax the initial constraints have a greater probability of being reproduced in subsequent generations (level 2). Consequently, the frequency of genetic variants across generations depends on their feedback effects on the physicochemical stresses that trigger their formation.

Figure 4:
A. The DNA molecule of each individual in a population of unicellular organisms is exposed to a variety of environmental physicochemical stresses, increasing the probability of mutations in each genome (1). The formation of cellular colonies resulting from the division of a single cell (proto-zygote) limits the exposure of each cell to different stresses, thus increasing the probability of reproducing the same genome and therefore the same phenotype across generations (2). Increasing the cellular density within a colony can increase the concentration of potentially genotoxic metabolites, such as ROS, in the intercellular space, which challenges the DNA physicochemical integrity and therefore decreases the probability of generating the same phenotype in subsequent generations (3, red broken lines). Cell metabolite genotoxicity led to the emergence of meiosis, which preserves the physicochemical integrity of DNA molecules of cells that are specialized in reproduction (germ cells giving rise to gametes). Meiosis increases the probability of reproducing the genome without variation, and thus increases the probability of reproducing the same phenotype across generations (3).
B. Environmentally-challenged somatic cells have an effect on the formation of *de novo* mutations in the germ cell population. First, somatic cell activities can generate physicochemical parameter imbalances in the internal environment and thus in the germ cell microenvironment. These imbalances can induce *de novo* mutations in germ cells (1). Second, somatic cells can also induce epigenetic modifications (chromatin modification and small RNAs) within germ cells that could trigger *de novo* mutations within targeted loci (2).
C. Environmental physicochemical stresses can challenge the activity of somatic cells, which can lead to imbalances in the internal environment and thus in the germ cell microenvironment. These imbalances can lead to *de novo* mutations in germ cells (A and B). The resulting genetic variants that relax the initial stresses from the parental germ cell microenvironment have a greater probability of being reproduced during gametogenesis (1^{st} generation). Consequently, they have a higher probability of being frequent among parental gametes and therefore of being passed on to the next generation. If these genetic variants are expressed in descendant somatic cells after fertilization and counteract or relax the initial stresses (2^{nd} generation), they maintain the internal physicochemical parameter equilibria allowing the replication without variations during gametogenesis of the descendant germ cell genome that is transmitted to the third generation. Consequently, the frequency of genetic variants across generations depends on their feedback effects at the germ and somatic cell level on the physicochemical stresses that trigger their formation.

References

1. Watson, J.D.; Crick, F.H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **1953**, *171*, 737-738, doi:10.1038/171737a0.
2. Watson, J.D.; Crick, F.H. Genetical implications of the structure of deoxyribonucleic acid. *Nature* **1953**, *171*, 964-967, doi:10.1038/171964b0.
3. Bernard, C. *Introduction à l'étude de la médecine expérimentale*; Flammarion, Ed. 1865.
4. Schrödinger, E. *What is life?*; Press, C.U., Ed. 1944.
5. Varela, F.G.; Maturana, H.R.; Uribe, R. Autopoiesis: the organization of living systems, its characterization and a model. *Curr Mod Biol* **1974**, *5*, 187-196, doi:10.1016/0303-2647(74)90031-8.
6. Noble, D. Claude Bernard, the first systems biologist, and the future of physiology. *Exp Physiol* **2008**, *93*, 16-26, doi:10.1113/expphysiol.2007.038695.
7. Milo, R.; Phillips, R. *Cell biology by the numbers*; Science, G., Ed. 2016.
8. Woese, C.R. A new biology for a new century. *Microbiol Mol Biol Rev* **2004**, *68*, 173-186, doi:10.1128/MMBR.68.2.173-186.2004.
9. Noble, D. A theory of biological relativity: no privileged level of causation. *Interface Focus* **2012**, *2*, 55-64, doi:10.1098/rsfs.2011.0067.
10. Trewavas, A. A brief history of systems biology. "Every object that biology studies is a system of systems." Francois Jacob (1974). *Plant Cell* **2006**, *18*, 2420-2430, doi:10.1105/tpc.106.042267.
11. Noble, R.; Tasaki, K.; Noble, P.J.; Noble, D. Biological Relativity Requires Circular Causality but Not Symmetry of Causation: So, Where, What and When Are the Boundaries? *Front Physiol* **2019**, *10*, 827, doi:10.3389/fphys.2019.00827.
12. Lamarck, J.-B. *Philosophie zoologique, ou exposition des considérations relatives à l'histoire naturelle des animaux* Paris, France: Flammarion ed.; 1809.
13. Darwin, C. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life*; Russel, L.J., Ed. 1859.
14. Huxley, J. *Evolution: the modern synthesis*; London, U.A.U., Ed. 1942.
15. Laland, K.; Uller, T.; Feldman, M.; Sterelny, K.; Muller, G.B.; Moczek, A.; Jablonka, E.; Odling-Smee, J.; Wray, G.A.; Hoekstra, H.E., et al. Does evolutionary theory need a rethink? *Nature* **2014**, *514*, 161-164, doi:10.1038/514161a.
16. Noble, D. Evolution viewed from physics, physiology and medicine. *Interface Focus* **2017**, *7*, 20160159, doi:10.1098/rsfs.2016.0159.
17. Shapiro, J.A. No genome is an island: toward a 21st century agenda for evolution. *Ann N Y Acad Sci* **2019**, *1447*, 21-52, doi:10.1111/nyas.14044.
18. Diogo, R. Quasi-religious Belief in Darwin and Darwinism: "Straw-Men" Scientist Believers Everywhere. *Current Molecular Biology Reports* **2020**, *6*, 16–31.
19. Corning, P.A. Beyond the modern synthesis: A framework for a more inclusive biological synthesis. *Prog Biophys Mol Biol* **2020**, *153*, 5-12, doi:10.1016/j.pbiomolbio.2020.02.002.
20. Elowitz, M.B.; Levine, A.J.; Siggia, E.D.; Swain, P.S. Stochastic gene expression in a single cell. *Science* **2002**, *297*, 1183-1186, doi:10.1126/science.1070919.
21. Cannon, W. Organization for physiological homoestasis. *Physiol. Rev.* **1929**, *9*, 399-431.
22. Billman, G.E. Homeostasis: The Underappreciated and Far Too Often Ignored Central Organizing Principle of Physiology. *Front Physiol* **2020**, *11*, 200, doi:10.3389/fphys.2020.00200.
23. Auboeuf, D. Physicochemical Foundations of Life that Direct Evolution: Chance and Natural Selection are not Evolutionary Driving Forces. *Life (Basel)* **2020**, *10*, doi:10.3390/life10020007.
24. Auboeuf, D. Alternative mRNA processing sites decrease genetic variability while increasing functional diversity. *Transcription* **2018**, *9*, 75-87, doi:10.1080/21541264.2017.1373891.
25. Auboeuf, D. Genome evolution is driven by gene expression-generated biophysical constraints through RNA-directed genetic variation: A hypothesis. *Bioessays* **2017**, *39*, doi:10.1002/bies.201700069.
26. Auboeuf, D. Putative RNA-Directed Adaptive Mutations in Cancer Evolution. *Transcription* **2016**, 10.1080/21541264.2016.1221491, 0, doi:10.1080/21541264.2016.1221491.
27. Crick, F. Central dogma of molecular biology. *Nature* **1970**, *227*, 561-563, doi:10.1038/227561a0.
28. Luria, S.E.; Delbruck, M. Mutations of Bacteria from Virus Sensitivity to Virus Resistance. *Genetics* **1943**, *28*, 491-511.
29. Gregory, T. Understanding Natural Selection: Essential Concepts and Common Misconceptions. *Evolution: Education and Outreach* **2009**, *2*, 156-175.
30. Lynch, M.; Ackerman, M.S.; Gout, J.F.; Long, H.; Sung, W.; Thomas, W.K.; Foster, P.L. Genetic drift, selection and the evolution of the mutation rate. *Nat Rev Genet* **2016**, *17*, 704-714, doi:10.1038/nrg.2016.104.
31. Tomkova, M.; Schuster-Bockler, B. DNA Modifications: Naturally More Error Prone? *Trends Genet* **2018**, *34*, 627-638, doi:10.1016/j.tig.2018.04.005.

32. Marnett, L.J.; Plastaras, J.P. Endogenous DNA damage and mutation. *Trends Genet* **2001**, *17*, 214-221, doi:10.1016/s0168-9525(01)02239-9.
33. Cooke, M.S.; Evans, M.D.; Dizdaroglu, M.; Lunec, J. Oxidative DNA damage: mechanisms, mutation, and disease. *FASEB J* **2003**, *17*, 1195-1214, doi:10.1096/fj.02-0752rev.
34. Singer, B.; Kusmierek, J.T. Chemical mutagenesis. *Annu Rev Biochem* **1982**, *51*, 655-693, doi:10.1146/annurev.bi.51.070182.003255.
35. Boulikas, T. Evolutionary consequences of nonrandom damage and repair of chromatin domains. *J Mol Evol* **1992**, *35*, 156-180, doi:10.1007/BF00183227.
36. Zhao, J.; Bacolla, A.; Wang, G.; Vasquez, K.M. Non-B DNA structure-induced genetic instability and evolution. *Cell Mol Life Sci* **2010**, *67*, 43-62, doi:10.1007/s00018-009-0131-2.
37. Chu, X.L.; Zhang, B.W.; Zhang, Q.G.; Zhu, B.R.; Lin, K.; Zhang, D.Y. Temperature responses of mutation rate and mutational spectrum in an Escherichia coli strain and the correlation with metabolic rate. *BMC Evol Biol* **2018**, *18*, 126, doi:10.1186/s12862-018-1252-8.
38. Schroeder, J.W.; Yeesin, P.; Simmons, L.A.; Wang, J.D. Sources of spontaneous mutagenesis in bacteria. *Crit Rev Biochem Mol Biol* **2018**, *53*, 29-48, doi:10.1080/10409238.2017.1394262.
39. Suarez-Villagran, M.Y.; Azevedo, R.B.R.; Miller, J.H., Jr. Influence of Electron-Holes on DNA Sequence-Specific Mutation Rates. *Genome Biol Evol* **2018**, *10*, 1039-1047, doi:10.1093/gbe/evy060.
40. Duan, C.; Huan, Q.; Chen, X.; Wu, S.; Carey, L.B.; He, X.; Qian, W. Reduced intrinsic DNA curvature leads to increased mutation rate. *Genome Biol* **2018**, *19*, 132, doi:10.1186/s13059-018-1525-y.
41. Liu, H.; Zhang, J. Yeast Spontaneous Mutation Rate and Spectrum Vary with Environment. *Curr Biol* **2019**, *29*, 1584-1591 e1583, doi:10.1016/j.cub.2019.03.054.
42. Crick, F.H. On protein synthesis. *Symp Soc Exp Biol* **1958**, *12*, 138-163.
43. Jacob, F.; Monod, J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **1961**, *3*, 318-356, doi:10.1016/s0022-2836(61)80072-7.
44. Woese, C.R. The fundamental nature of the genetic code: prebiotic interactions between polynucleotides and polyamino acids or their derivatives. *Proc Natl Acad Sci U S A* **1968**, *59*, 110-117, doi:10.1073/pnas.59.1.110.
45. Orgel, L.E. Evolution of the genetic apparatus. *J Mol Biol* **1968**, *38*, 381-393, doi:10.1016/0022-2836(68)90393-8.
46. Gilbert, W. The RNA world. *Nature* **1986**, *319*, 618.
47. Spitzer, J.; Pielak, G.J.; Poolman, B. Emergence of life: Physical chemistry changes the paradigm. *Biol Direct* **2015**, *10*, 33, doi:10.1186/s13062-015-0060-y.
48. Noller, H.F. The driving force for molecular evolution of translation. *RNA* **2004**, *10*, 1833-1837, doi:10.1261/rna.7142404.
49. Doig, A.J. Frozen, but no accident - why the 20 standard amino acids were selected. *FEBS J* **2017**, *284*, 1296-1305, doi:10.1111/febs.13982.
50. Wachowius, F.; Attwater, J.; Holliger, P. Nucleic acids: function and potential for abiogenesis. *Q Rev Biophys* **2017**, *50*, e4, doi:10.1017/S0033583517000038.
51. Poole, A.M.; Jeffares, D.C.; Penny, D. The path from the RNA world. *J Mol Evol* **1998**, *46*, 1-17, doi:10.1007/pl00006275.
52. Szathmary, E. The origin of the genetic code: amino acids as cofactors in an RNA world. *Trends Genet* **1999**, *15*, 223-229.
53. Wieczorek, R.; Dorr, M.; Chotera, A.; Luisi, P.L.; Monnard, P.A. Formation of RNA phosphodiester bond by histidine-containing dipeptides. *Chembiochem* **2013**, *14*, 217-223, doi:10.1002/cbic.201200643.
54. van der Gulik, P.T. On the Origin of Sequence. *Life (Basel)* **2015**, *5*, 1629-1637, doi:10.3390/life5041629.
55. van der Gulik, P.T.; Speijer, D. How amino acids and peptides shaped the RNA world. *Life (Basel)* **2015**, *5*, 230-246, doi:10.3390/life5010230.
56. Tagami, S.; Attwater, J.; Holliger, P. Simple peptides derived from the ribosomal core potentiate RNA polymerase ribozyme function. *Nat Chem* **2017**, *9*, 325-332, doi:10.1038/nchem.2739.
57. Gamow, G. Possible Relation between Deoxyribonucleic Acid and Protein Structures. *Nature* **1954**, *173*, 318-318, doi:10.1038/173318a0.
58. Woese, C.R.; Dugre, D.H.; Saxinger, W.C.; Dugre, S.A. The molecular basis for the genetic code. *Proc Natl Acad Sci U S A* **1966**, *55*, 966-974, doi:10.1073/pnas.55.4.966.
59. Yarus, M. The Genetic Code and RNA-Amino Acid Affinities. *Life (Basel)* **2017**, *7*, doi:10.3390/life7020013.
60. Koonin, E.V.; Novozhilov, A.S. Origin and Evolution of the Universal Genetic Code. *Annu Rev Genet* **2017**, *51*, 45-62, doi:10.1146/annurev-genet-120116-024713.
61. Kun, A.; Radvanyi, A. The evolution of the genetic code: Impasses and challenges. *Biosystems* **2018**, *164*, 217-225, doi:10.1016/j.biosystems.2017.10.006.

62. Wong, J.T. A co-evolution theory of the genetic code. *Proc Natl Acad Sci U S A* **1975**, *72*, 1909-1912, doi:10.1073/pnas.72.5.1909.
63. Di Giulio, M.; Medugno, M. The historical factor: the biosynthetic relationships between amino acids and their physicochemical properties in the origin of the genetic code. *J Mol Evol* **1998**, *46*, 615-621, doi:10.1007/pl00006341.
64. Copley, S.D.; Smith, E.; Morowitz, H.J. A mechanism for the association of amino acids with their codons and the origin of the genetic code. *Proc Natl Acad Sci U S A* **2005**, *102*, 4442-4447, doi:10.1073/pnas.0501049102.
65. Polyansky, A.A.; Hlevnjak, M.; Zagrovic, B. Analogue encoding of physicochemical properties of proteins in their cognate messenger RNAs. *Nat Commun* **2013**, *4*, 2784, doi:10.1038/ncomms3784.
66. Eigen, M.; Biebricher, C.K.; Gebinoga, M.; Gardiner, W.C. The hypercycle. Coupling of RNA and protein biosynthesis in the infection cycle of an RNA bacteriophage. *Biochemistry* **1991**, *30*, 11005-11018, doi:10.1021/bi00110a001.
67. Boerlijst, M.H., P. Spiral wave structure in pre-biotic evolution: hypercycles stable against parasites. *Physica D: Nonlinear Phenomena* **1991**, *48*, 17-28.
68. Szabo, P.; Scheuring, I.; Czaran, T.; Szathmary, E. In silico simulations reveal that replicators with limited dispersal evolve towards higher efficiency and fidelity. *Nature* **2002**, *420*, 340-343, doi:10.1038/nature01187.
69. Takeuchi, N.; Hogeweg, P. Multilevel selection in models of prebiotic evolution II: a direct comparison of compartmentalization and spatial self-organization. *PLoS Comput Biol* **2009**, *5*, e1000542, doi:10.1371/journal.pcbi.1000542.
70. Pressman, A.; Blanco, C.; Chen, I.A. The RNA World as a Model System to Study the Origin of Life. *Curr Biol* **2015**, *25*, R953-963, doi:10.1016/j.cub.2015.06.016.
71. Bansho, Y.; Furubayashi, T.; Ichihashi, N.; Yomo, T. Host-parasite oscillation dynamics and evolution in a compartmentalized RNA replication system. *Proc Natl Acad Sci U S A* **2016**, *113*, 4045-4050, doi:10.1073/pnas.1524404113.
72. Ichihashi, N.; Yomo, T. Constructive Approaches for Understanding the Origin of Self-Replication and Evolution. *Life (Basel)* **2016**, *6*, doi:10.3390/life6030026.
73. Gordon, K.H. Were RNA replication and translation directly coupled in the RNA (+protein?) World? *J Theor Biol* **1995**, *173*, 179-193, doi:10.1006/jtbi.1995.0054.
74. Kunin, V. A system of two polymerases--a model for the origin of life. *Orig Life Evol Biosph* **2000**, *30*, 459-466, doi:10.1023/a:1006672126867.
75. Penny, D. An Interpretive Review of the Origin of Life Research. *Biology and Philosophy* **2005**, *20*, 633-671, doi:10.1007/s10539-004-7342-6.
76. Proshkin, S.; Rahmouni, A.R.; Mironov, A.; Nudler, E. Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science* **2010**, *328*, 504-508, doi:10.1126/science.1184939.
77. Gounaris, Y. An evolutionary theory based on a protein-mRNA co-synthesis hypothesis. *Journal of Biological Research-Thessaloniki* **2011**, *15*, 3-16.
78. Morgens, D.W. The protein invasion: a broad review on the origin of the translational system. *J Mol Evol* **2013**, *77*, 185-196, doi:10.1007/s00239-013-9592-x.
79. Francis, B.R. The Hypothesis that the Genetic Code Originated in Coupled Synthesis of Proteins and the Evolutionary Predecessors of Nucleic Acids in Primitive Cells. *Life (Basel)* **2015**, *5*, 467-505, doi:10.3390/life5010467.
80. Altstein, A.D. The progene hypothesis: the nucleoprotein world and how life began. *Biol Direct* **2015**, *10*, 67, doi:10.1186/s13062-015-0096-z.
81. Carter, C.W. What RNA World? Why a Peptide/RNA Partnership Merits Renewed Experimental Attention. *Life (Basel)* **2015**, *5*, 294-320, doi:10.3390/life5010294.
82. Lin, Y.H.; Forman-Kay, J.D.; Chan, H.S. Theories for Sequence-Dependent Phase Behaviors of Biomolecular Condensates. *Biochemistry* **2018**, *57*, 2499-2508, doi:10.1021/acs.biochem.8b00058.
83. Matveev, V.V. Cell theory, intrinsically disordered proteins, and the physics of the origin of life. *Prog Biophys Mol Biol* **2019**, *149*, 114-130, doi:10.1016/j.pbiomolbio.2019.04.001.
84. Staněk, D.; Fox, A.H. Nuclear bodies: news insights into structure and function. *Curr Opin Cell Biol* **2017**, *46*, 94-101, doi:10.1016/j.ceb.2017.05.001.
85. Schrum, J.P.; Zhu, T.F.; Szostak, J.W. The origins of cellular life. *Cold Spring Harb Perspect Biol* **2010**, *2*, a002212, doi:10.1101/cshperspect.a002212.
86. Taran, O.; Chen, C.; Omosun, T.O.; Hsieh, M.C.; Rha, A.; Goodwin, J.T.; Mehta, A.K.; Grover, M.A.; Lynn, D.G. Expanding the informational chemistries of life: peptide/RNA networks. *Philos Trans A Math Phys Eng Sci* **2017**, *375*, doi:10.1098/rsta.2016.0356.
87. Schreiber, A.; Huber, M.C.; Schiller, S.M. Prebiotic Protocell Model Based on Dynamic Protein Membranes Accommodating Anabolic Reactions. *Langmuir* **2019**, *35*, 9593-9610, doi:10.1021/acs.langmuir.9b00445.

88. Schopfa J W, K.A.B., Czajac A D, Tripathic A B. Evidence of Archean life: Stromatolites and microfossils. *Precambrian Research* **2007**, *158*, 141-155.
89. Knoll, A.H.; Nowak, M.A. The timetable of evolution. *Science advances* **2017**, *3*, e1603076, doi:10.1126/sciadv.1603076.
90. Davis, B.D. Transcriptional bias: a non-Lamarckian mechanism for substrate-induced mutations. *Proc Natl Acad Sci U S A* **1989**, *86*, 5005-5009, doi:10.1073/pnas.86.13.5005.
91. Saier, M.H., Jr. Did adaptive and directed mutation evolve to accelerate stress-induced evolutionary change? *J Mol Microbiol Biotechnol* **2011**, *21*, 5-7, doi:10.1159/000332746.
92. Wright, B.E.; Schmidt, K.H.; Minnick, M.F. Kinetic models reveal the in vivo mechanisms of mutagenesis in microbes and man. *Mutat Res* **2013**, *752*, 129-137, doi:10.1016/j.mrrev.2012.12.003.
93. Gaillard, H.; Herrera-Moyano, E.; Aguilera, A. Transcription-associated genome instability. *Chem Rev* **2013**, *113*, 8638-8661, doi:10.1021/cr400017y.
94. Sekowska, A.; Wendel, S.; Fischer, E.C.; Norholm, M.H.H.; Danchin, A. Generation of mutation hotspots in ageing bacterial colonies. *Sci Rep* **2016**, *6*, 2, doi:10.1038/s41598-016-0005-4.
95. Sankar, T.S.; Wastuwidyaningtyas, B.D.; Dong, Y.; Lewis, S.A.; Wang, J.D. The nature of mutations induced by replication-transcription collisions. *Nature* **2016**, *535*, 178-181, doi:10.1038/nature18316.
96. Pannunzio, N.R.; Lieber, M.R. RNA Polymerase Collision versus DNA Structural Distortion: Twists and Turns Can Cause Break Failure. *Mol Cell* **2016**, *62*, 327-334, doi:10.1016/j.molcel.2016.03.034.
97. Sebastian, R.; Oberdoerffer, P. Transcription-associated events affecting genomic integrity. *Philos Trans R Soc Lond B Biol Sci* **2017**, *372*, doi:10.1098/rstb.2016.0288.
98. Hull, R.M.; Cruz, C.; Jack, C.V.; Houseley, J. Environmental change drives accelerated adaptation through stimulated copy number variation. *PLoS Biol* **2017**, *15*, e2001333, doi:10.1371/journal.pbio.2001333.
99. Correa, R.; Thornton, P.C.; Rosenberg, S.M.; Hastings, P.J. Oxygen and RNA in stress-induced mutation. *Curr Genet* **2018**, *64*, 769-776, doi:10.1007/s00294-017-0801-9.
100. Cairns, J.; Overbaugh, J.; Miller, S. The origin of mutants. *Nature* **1988**, *335*, 142-145, doi:10.1038/335142a0.
101. Rosenberg, S.M. Evolving responsively: adaptive mutation. *Nat Rev Genet* **2001**, *2*, 504-515, doi:10.1038/35080556.
102. Wright, B.E. Stress-directed adaptive mutations and evolution. *Mol Microbiol* **2004**, *52*, 643-650, doi:10.1111/j.1365-2958.2004.04012.x.
103. Danchin, E.; Pocheville, A. Inheritance is where physiology meets evolution. *J Physiol* **2014**, *592*, 2307-2317, doi:10.1113/jphysiol.2014.272096.
104. Yona, A.H.; Frumkin, I.; Pilpel, Y. A relay race on the evolutionary adaptation spectrum. *Cell* **2015**, *163*, 549-559, doi:10.1016/j.cell.2015.10.005.
105. Noble, R.; Noble, D. Was the Watchmaker Blind? Or Was She One-Eyed? *Biology (Basel)* **2017**, *6*, doi:10.3390/biology6040047.
106. Saier, M.H., Jr.; Kukita, C.; Zhang, Z. Transposon-mediated directed mutation in bacteria and eukaryotes. *Front Biosci (Landmark Ed)* **2017**, *22*, 1458-1468.
107. Shapiro, J.A. Living Organisms Author Their Read-Write Genomes in Evolution. *Biology (Basel)* **2017**, *6*, doi:10.3390/biology6040042.
108. Lee, H.; Popodi, E.; Tang, H.; Foster, P.L. Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. *Proc Natl Acad Sci U S A* **2012**, *109*, E2774-2783, doi:10.1073/pnas.1210309109.
109. Shewaramani, S.; Finn, T.J.; Leahy, S.C.; Kassen, R.; Rainey, P.B.; Moon, C.D. Anaerobically Grown Escherichia coli Has an Enhanced Mutation Rate and Distinct Mutational Spectra. *PLoS Genet* **2017**, *13*, e1006570, doi:10.1371/journal.pgen.1006570.
110. Maharjan, R.P.; Ferenci, T. A shifting mutational landscape in 6 nutritional states: Stress-induced mutagenesis as a series of distinct stress input-mutation output relationships. *PLoS Biol* **2017**, *15*, e2001477, doi:10.1371/journal.pbio.2001477.
111. Deatherage, D.E.; Kepner, J.L.; Bennett, A.F.; Lenski, R.E.; Barrick, J.E. Specificity of genome evolution in experimental populations of Escherichia coli evolved at different temperatures. *Proc Natl Acad Sci U S A* **2017**, *114*, E1904-E1912, doi:10.1073/pnas.1616132114.
112. Vinogradov, A.E. DNA helix: the importance of being GC-rich. *Nucleic Acids Res* **2003**, *31*, 1838-1844.
113. Babbitt, G.A.; Alawad, M.A.; Schulze, K.V.; Hudson, A.O. Synonymous codon bias and functional constraint on GC3-related DNA backbone dynamics in the prokaryotic nucleoid. *Nucleic Acids Res* **2014**, *42*, 10915-10926, doi:10.1093/nar/gku811.
114. Dans, P.D.; Faustino, I.; Battistini, F.; Zakrzewska, K.; Lavery, R.; Orozco, M. Unraveling the sequence-dependent polymorphic behavior of d(CpG) steps in B-DNA. *Nucleic Acids Res* **2014**, *42*, 11304-11320, doi:10.1093/nar/gku809.

115. Goncearenco, A.; Ma, B.G.; Berezovsky, I.N. Molecular mechanisms of adaptation emerging from the physics and evolution of nucleic acids and proteins. *Nucleic Acids Res* **2014**, *42*, 2879-2892, doi:10.1093/nar/gkt1336.
116. Harteis, S.; Schneider, S. Making the bend: DNA tertiary structure and protein-DNA interactions. *Int J Mol Sci* **2014**, *15*, 12335-12363, doi:10.3390/ijms150712335.
117. Reymer, A.; Zakrzewska, K.; Lavery, R. Sequence-dependent response of DNA to torsional stress: a potential biological regulation mechanism. *Nucleic Acids Res* **2018**, *46*, 1684-1694, doi:10.1093/nar/gkx1270.
118. Gul, I.S.; Staal, J.; Hulpiau, P.; De Keuckelaere, E.; Kamm, K.; Deroo, T.; Sanders, E.; Staes, K.; Driege, Y.; Saeys, Y., et al. GC Content of Early Metazoan Genes and Its Impact on Gene Expression Levels in Mammalian Cell Lines. *Genome Biol Evol* **2018**, *10*, 909-917, doi:10.1093/gbe/evy040.
119. Versteeg, R.; van Schaik, B.D.; van Batenburg, M.F.; Roos, M.; Monajemi, R.; Caron, H.; Bussemaker, H.J.; van Kampen, A.H. The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. *Genome Res* **2003**, *13*, 1998-2004, doi:10.1101/gr.1649303.
120. Urrutia, A.O.; Hurst, L.D. The signature of selection mediated by expression on human genes. *Genome Res* **2003**, *13*, 2260-2264, doi:10.1101/gr.641103.
121. Duret, L.; Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu Rev Genomics Hum Genet* **2009**, *10*, 285-311, doi:10.1146/annurev-genom-082908-150001.
122. Pessia, E.; Popa, A.; Mousset, S.; Rezvoy, C.; Duret, L.; Marais, G.A. Evidence for widespread GC-biased gene conversion in eukaryotes. *Genome Biol Evol* **2012**, *4*, 675-682, doi:10.1093/gbe/evs052.
123. Imhof, P.; Zahran, M. The effect of a G:T mispair on the dynamics of DNA. *PLoS One* **2013**, *8*, e53305, doi:10.1371/journal.pone.0053305.
124. Yin, Y.; Yang, L.; Zheng, G.; Gu, C.; Yi, C.; He, C.; Gao, Y.Q.; Zhao, X.S. Dynamics of spontaneous flipping of a mismatched base in DNA duplex. *Proc Natl Acad Sci U S A* **2014**, *111*, 8043-8048, doi:10.1073/pnas.1400667111.
125. Lassalle, F.; Périan, S.; Bataillon, T.; Nesme, X.; Duret, L.; Daubin, V. GC-Content evolution in bacterial genomes: the biased gene conversion hypothesis expands. *PLoS Genet* **2015**, *11*, e1004941, doi:10.1371/journal.pgen.1004941.
126. Weissman, J.L.; Fagan, W.F.; Johnson, P.L.F. Linking high GC content to the repair of double strand breaks in prokaryotic genomes. *PLoS Genet* **2019**, *15*, e1008493, doi:10.1371/journal.pgen.1008493.
127. Sueoka, N. Wide intra-genomic G+C heterogeneity in human and chicken is mainly due to strand-symmetric directional mutation pressures: dGTP-oxidation and symmetric cytosine-deamination hypotheses. *Gene* **2002**, *300*, 141-154, doi:10.1016/s0378-1119(02)01046-6.
128. Tilgner, H.; Knowles, D.G.; Johnson, R.; Davis, C.A.; Chakrabortty, S.; Djebali, S.; Curado, J.; Snyder, M.; Gingeras, T.R.; Guigo, R. Deep sequencing of subcellular RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res* **2012**, *22*, 1616-1625, doi:10.1101/gr.134445.111.
129. Raghavan, R.; Kelkar, Y.D.; Ochman, H. A selective force favoring increased G+C content in bacterial genes. *Proc Natl Acad Sci U S A* **2012**, *109*, 14504-14507, doi:10.1073/pnas.1205683109.
130. Ran, W.; Kristensen, D.M.; Koonin, E.V. Coupling between protein level selection and codon usage optimization in the evolution of bacteria and archaea. *mBio* **2014**, *5*, e00956-00914, doi:10.1128/mBio.00956-14.
131. Gao, N.; Lu, G.; Lercher, M.J.; Chen, W.H. Selection for energy efficiency drives strand-biased gene distribution in prokaryotes. *Sci Rep* **2017**, *7*, 10572, doi:10.1038/s41598-017-11159-3.
132. Halder, B.; Malakar, A.K.; Chakraborty, S. Nucleotide composition determines the role of translational efficiency in human genes. *Bioinformation* **2017**, *13*, 46-53, doi:10.6026/97320630013046.
133. Du, M.Z.; Zhang, C.; Wang, H.; Liu, S.; Wei, W.; Guo, F.B. The GC Content as a Main Factor Shaping the Amino Acid Usage During Bacterial Evolution Process. *Front Microbiol* **2018**, *9*, 2948, doi:10.3389/fmicb.2018.02948.
134. Quandt, E.M.; Traverse, C.C.; Ochman, H. Local genic base composition impacts protein production and cellular fitness. *PeerJ* **2018**, *6*, e4286, doi:10.7717/peerj.4286.
135. Lemaire, S.; Fontrodona, N.; Claude, J.-B.; Polvèche, H.; Aubé, F.; Modolo, L.; Bourgeois, C.F.; Mortreux, F.; Auboeuf, D. Interplay between gene nucleotide composition bias and splicing. *bioRxiv 605832* **2019**.
136. Mordstein, C.; Savisaar, R.; Young, R.S.; Bazile, J.; Talmane, L.; Luft, J.; Liss, M.; Taylor, M.S.; Hurst, L.D.; Kudla, G. Codon Usage and Splicing Jointly Influence mRNA Localization. *Cell systems* **2020**, *10*, 351-362.e358, doi:10.1016/j.cels.2020.03.001.
137. Hickey, D.A.; Singer, G.A. Genomic and proteomic adaptations to growth at high temperature. *Genome Biol* **2004**, *5*, 117, doi:10.1186/gb-2004-5-10-117.
138. Zeldovich, K.B.; Berezovsky, I.N.; Shakhnovich, E.I. Protein and DNA sequence determinants of thermophilic adaptation. *PLoS Comput Biol* **2007**, *3*, e5, doi:10.1371/journal.pcbi.0030005.

139. Bender, A.; Hajieva, P.; Moosmann, B. Adaptive antioxidant methionine accumulation in respiratory chain complexes explains the use of a deviant genetic code in mitochondria. *Proc Natl Acad Sci U S A* **2008**, *105*, 16496-16501, doi:10.1073/pnas.0802779105.
140. Elser, J.J.; Acquisti, C.; Kumar, S. Stoichiogenomics: the evolutionary ecology of macromolecular elemental composition. *Trends Ecol Evol* **2011**, *26*, 38-44, doi:10.1016/j.tree.2010.10.006.
141. Babbitt, G.A.; Coppola, E.E.; Mortensen, J.S.; Ekeren, P.X.; Viola, C.; Goldblatt, D.; Hudson, A.O. Triplet-Based Codon Organization Optimizes the Impact of Synonymous Mutation on Nucleic Acid Molecular Dynamics. *J Mol Evol* **2018**, *86*, 91-102, doi:10.1007/s00239-018-9828-x.
142. Granold, M.; Hajieva, P.; Tosa, M.I.; Irimie, F.D.; Moosmann, B. Modern diversification of the amino acid repertoire driven by oxygen. *Proc Natl Acad Sci U S A* **2018**, *115*, 41-46, doi:10.1073/pnas.1717100115.
143. White, K.A.; Ruiz, D.G.; Szpiech, Z.A.; Strauli, N.B.; Hernandez, R.D.; Jacobson, M.P.; Barber, D.L. Cancer-associated arginine-to-histidine mutations confer a gain in pH sensing to mutant proteins. *Sci Signal* **2017**, *10*, doi:10.1126/scisignal.aam9931.
144. Szpiech, Z.A.; Strauli, N.B.; White, K.A.; Ruiz, D.G.; Jacobson, M.P.; Barber, D.L.; Hernandez, R.D. Prominent features of the amino acid mutation landscape in cancer. *PLoS One* **2017**, *12*, e0183273, doi:10.1371/journal.pone.0183273.
145. Tsuber, V.; Kadamov, Y.; Brautigam, L.; Berglund, U.W.; Helleday, T. Mutations in Cancer Cause Gain of Cysteine, Histidine, and Tryptophan at the Expense of a Net Loss of Arginine on the Proteome Level. *Biomolecules* **2017**, *7*, doi:10.3390/biom7030049.
146. Son, H.; Kang, H.; Kim, H.S.; Kim, S. Somatic mutation driven codon transition bias in human cancer. *Sci Rep* **2017**, *7*, 14204, doi:10.1038/s41598-017-14543-1.
147. Sagan, L. On the origin of mitosing cells. *J Theor Biol* **1967**, *14*, 255-274, doi:10.1016/0022-5193(67)90079-3.
148. Speijer, D. Alternating terminal electron-acceptors at the basis of symbiogenesis: How oxygen ignited eukaryotic evolution. *Bioessays* **2017**, *39*, doi:10.1002/bies.201600174.
149. Friedman, K.; Heller, A. On the Non-Uniform Distribution of Guanine in Introns of Human Genes: Possible Protection of Exons against Oxidation by Proximal Intron Poly-G Sequences. *J. Phys. Chem. B* **2001**, *105*, 11859-11865.
150. Martin, W.; Koonin, E.V. Introns and the origin of nucleus-cytosol compartmentalization. *Nature* **2006**, *440*, 41-45, doi:10.1038/nature04531.
151. Coros, C.J.; Piazza, C.L.; Chalamcharla, V.R.; Belfort, M. A mutant screen reveals RNase E as a silencer of group II intron retromobility in Escherichia coli. *RNA* **2008**, *14*, 2634-2644, doi:10.1261/rna.1247608.
152. Belfort, M. Mobile self-splicing introns and inteins as environmental sensors. *Curr Opin Microbiol* **2017**, *38*, 51-58, doi:10.1016/j.mib.2017.04.003.
153. Amit, M.; Donyo, M.; Hollander, D.; Goren, A.; Kim, E.; Gelfman, S.; Lev-Maor, G.; Burstein, D.; Schwartz, S.; Postolsky, B., et al. Differential GC content between exons and introns establishes distinct strategies of splice-site recognition. *Cell Rep* **2012**, *1*, 543-556, doi:10.1016/j.celrep.2012.03.013.
154. Mukherjee, D.; Saha, D.; Acharya, D.; Mukherjee, A.; Chakraborty, S.; Ghosh, T.C. The role of introns in the conservation of the metabolic genes of Arabidopsis thaliana. *Genomics* **2018**, *110*, 310-317, doi:10.1016/j.ygeno.2017.12.003.
155. Enright, H.; Miller, W.J.; Hays, R.; Floyd, R.A.; Hebbel, R.P. Preferential targeting of oxidative base damage to internucleosomal DNA. *Carcinogenesis* **1996**, *17*, 1175-1177, doi:10.1093/carcin/17.5.1175.
156. Chen, X.; Chen, Z.; Chen, H.; Su, Z.; Yang, J.; Lin, F.; Shi, S.; He, X. Nucleosomes suppress spontaneous mutations base-specifically in eukaryotes. *Science* **2012**, *335*, 1235-1238, doi:10.1126/science.1217580.
157. Drinnenberg, I.A.; Berger, F.; Elsasser, S.J.; Andersen, P.R.; Ausio, J.; Bickmore, W.A.; Blackwell, A.R.; Erwin, D.H.; Gahan, J.M.; Gaut, B.S., et al. EvoChromo: towards a synthesis of chromatin biology and evolution. *Development* **2019**, *146*, doi:10.1242/dev.178962.
158. Mattick, J.S.; Gagen, M.J. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol* **2001**, *18*, 1611-1630, doi:10.1093/oxfordjournals.molbev.a003951.
159. Kozlovski, I.; Siegfried, Z.; Amar-Schwartz, A.; Karni, R. The role of RNA alternative splicing in regulating cancer metabolism. *Hum Genet* **2017**, *136*, 1113-1127, doi:10.1007/s00439-017-1803-x.
160. Hintze, A.; Adami, C. Evolution of complex modular biological networks. *PLoS Comput Biol* **2008**, *4*, e23, doi:10.1371/journal.pcbi.0040023.
161. Jeltsch, A. Oxygen, epigenetic signaling, and the evolution of early life. *Trends Biochem Sci* **2013**, *38*, 172-176, doi:10.1016/j.tibs.2013.02.001.
162. Tsochatzidou, M.; Malliarou, M.; Papanikolaou, N.; Roca, J.; Nikolaou, C. Genome urbanization: clusters of topologically co-regulated genes delineate functional compartments in the genome of Saccharomyces cerevisiae. *Nucleic Acids Res* **2017**, *45*, 5818-5828, doi:10.1093/nar/gkx198.

163. Hajheidari, M.; Koncz, C.; Bucher, M. Chromatin Evolution-Key Innovations Underpinning Morphological Complexity. *Front Plant Sci* **2019**, *10*, 454, doi:10.3389/fpls.2019.00454.
164. Dobon, B.; Montanucci, L.; Pereto, J.; Bertranpetit, J.; Laayouni, H. Gene connectivity and enzyme evolution in the human metabolic network. *Biol Direct* **2019**, *14*, 17, doi:10.1186/s13062-019-0248-7.
165. Dai, Z.; Ramesh, V.; Locasale, J.W. The evolving metabolic landscape of chromatin biology and epigenetics. *Nat Rev Genet* **2020**, *21*, 737-753, doi:10.1038/s41576-020-0270-8.
166. Raymond, J.; Segre, D. The effect of oxygen on biochemical networks and the evolution of complex life. *Science* **2006**, *311*, 1764-1767, doi:10.1126/science.1118439.
167. Koch, L.G.; Britton, S.L. Aerobic metabolism underlies complexity and capacity. *J Physiol* **2008**, *586*, 83-95, doi:10.1113/jphysiol.2007.144709.
168. Kaiser, D. Building a multicellular organism. *Annu Rev Genet* **2001**, *35*, 103-123, doi:10.1146/annurev.genet.35.102401.090145.
169. Grosberg, R.K.; Strathmann, R.R. The Evolution of Multicellularity: A Minor Major Transition? *Annual Review of Ecology, Evolution, and Systematics* **2007**, *38*, 621-654, doi:10.1146/annurev.ecolsys.36.102403.114735.
170. El Albani, A.; Bengtson, S.; Canfield, D.E.; Bekker, A.; Macchiarelli, R.; Mazurier, A.; Hammarlund, E.U.; Boulvais, P.; Dupuy, J.J.; Fontaine, C., et al. Large colonial organisms with coordinated growth in oxygenated environments 2.1 Gyr ago. *Nature* **2010**, *466*, 100-104, doi:10.1038/nature09166.
171. Newman, S.A. Why are there eggs? *Biochem Biophys Res Commun* **2014**, *450*, 1225-1230, doi:10.1016/j.bbrc.2014.03.132.
172. Niklas, K.J.; Newman, S.A. The many roads to and from multicellularity. *J Exp Bot* **2020**, *71*, 3247-3253, doi:10.1093/jxb/erz547.
173. Furusawa, C.; Kaneko, K. Origin of multicellular organisms as an inevitable consequence of dynamical systems. *Anat Rec* **2002**, *268*, 327-342, doi:10.1002/ar.10164.
174. Ispolatov, I.; Ackermann, M.; Doebeli, M. Division of labour and the evolution of multicellularity. *Proc Biol Sci* **2012**, *279*, 1768-1776, doi:10.1098/rspb.2011.1999.
175. Varahan, S.; Walvekar, A.; Sinha, V.; Krishna, S.; Laxman, S. Metabolic constraints drive self-organization of specialized cell groups. *Elife* **2019**, *8*, doi:10.7554/eLife.46735.
176. Newman, S.A. Cell differentiation: What have we learned in 50 years? *J Theor Biol* **2020**, *485*, 110031, doi:10.1016/j.jtbi.2019.110031.
177. Bich, L.; Pradeu, T.; Moreau, J.F. Understanding Multicellularity: The Functional Organization of the Intercellular Space. *Front Physiol* **2019**, *10*, 1170, doi:10.3389/fphys.2019.01170.
178. Bernstein, H.; Byerly, H.C.; Hopf, F.A.; Michod, R.E. Genetic damage, mutation, and the evolution of sex. *Science* **1985**, *229*, 1277-1281, doi:10.1126/science.3898363.
179. Bernstein, H.; Bernstein, C.; Richard E. Michod. Meiosis as an Evolutionary Adaptation for DNA Repair. *DNA Repair, Inna Kruman, IntechOpen* **2011**, DOI: 10.5772/25117. , doi:DOI: 10.5772/25117. .
180. Horandl, E.; Speijer, D. How oxygen gave rise to eukaryotic sex. *Proc Biol Sci* **2018**, *285*, doi:10.1098/rspb.2017.2706.
181. Speijer, D.; Lukes, J.; Elias, M. Sex is a ubiquitous, ancient, and inherent attribute of eukaryotic life. *Proc Natl Acad Sci U S A* **2015**, *112*, 8827-8834, doi:10.1073/pnas.1501725112.
182. Bengtsson, B.O. Biased conversion as the primary function of recombination. *Genetical research* **1986**, *47*, 77-80, doi:10.1017/s001667230002454x.
183. Marais, G. Biased gene conversion: implications for genome and sex evolution. *Trends Genet* **2003**, *19*, 330-338, doi:10.1016/S0168-9525(03)00116-1.
184. Long, H.; Sung, W.; Kucukyildirim, S.; Williams, E.; Miller, S.F.; Guo, W.; Patterson, C.; Gregory, C.; Strauss, C.; Stone, C., et al. Evolutionary determinants of genome-wide nucleotide composition. *Nat Ecol Evol* **2018**, *2*, 237-240, doi:10.1038/s41559-017-0425-y.
185. Weismann, A. *Essay Upon heridity*; Oxford: 1889.
186. Bell, G. *The origin and early evolution of germ cells as illustrated by the Volvocales in n The Origin and Evolution of Sex*; Halvorson, H.O.M., A., editors, Ed. 1985.
187. Michod, R.E. Evolution of individuality during the transition from unicellular to multicellular life. *Proc Natl Acad Sci U S A* **2007**, *104 Suppl 1*, 8613-8618, doi:10.1073/pnas.0701489104.
188. Goldsby, H.J.; Knoester, D.B.; Ofria, C.; Kerr, B. The evolutionary origin of somatic cells under the dirty work hypothesis. *PLoS Biol* **2014**, *12*, e1001858, doi:10.1371/journal.pbio.1001858.
189. Wahl, M.E.; Murray, A.W. Multicellularity makes somatic differentiation evolutionarily stable. *Proc Natl Acad Sci U S A* **2016**, *113*, 8362-8367, doi:10.1073/pnas.1608278113.
190. Aitken, R.J.; Baker, M.A.; Sawyer, D. Oxidative stress in the male germ line and its role in the aetiology of male infertility and genetic disease. *Reprod Biomed Online* **2003**, *7*, 65-70, doi:10.1016/s1472-6483(10)61730-0.

191. Karimi, J.; Goodarzi, M.T.; Tavilani, H.; Khodadadi, I.; Amiri, I. Increased receptor for advanced glycation end products in spermatozoa of diabetic men and its association with sperm nuclear DNA fragmentation. *Andrologia* **2012**, *44 Suppl 1*, 280-286, doi:10.1111/j.1439-0272.2011.01178.x.
192. Vecoli, C.; Montano, L.; Andreassi, M.G. Environmental pollutants: genetic damage and epigenetic changes in male germ cells. *Environ Sci Pollut Res Int* **2016**, *23*, 23339-23348, doi:10.1007/s11356-016-7728-4.
193. Pourmasumi, S.; Sabeti, P.; Rahiminia, T.; Mangoli, E.; Tabibnejad, N.; Talebi, A.R. The etiologies of DNA abnormalities in male infertility: An assessment and review. *Int J Reprod Biomed (Yazd)* **2017**, *15*, 331-344.
194. Liu, Y.; Ding, Z. Obesity, a serious etiologic factor for male subfertility in modern society. *Reproduction* **2017**, *154*, R123-r131, doi:10.1530/rep-17-0161.
195. Nevin, C.; McNeil, L.; Ahmed, N.; Murgatroyd, C.; Brison, D.; Carroll, M. Investigating the Glycating Effects of Glucose, Glyoxal and Methylglyoxal on Human Sperm. *Sci Rep* **2018**, *8*, 9002, doi:10.1038/s41598-018-27108-7.
196. Ricci, E.; Al-Beitawi, S.; Cipriani, S.; Alteri, A.; Chiaffarino, F.; Candiani, M.; Gerli, S.; Viganó, P.; Parazzini, F. Dietary habits and semen parameters: a systematic narrative review. *Andrology* **2018**, *6*, 104-116, doi:10.1111/andr.12452.
197. McVicker, G.; Green, P. Genomic signatures of germline gene expression. *Genome Res* **2010**, *20*, 1503-1511, doi:10.1101/gr.106666.110.
198. Michaelson, J.J.; Shi, Y.; Gujral, M.; Zheng, H.; Malhotra, D.; Jin, X.; Jian, M.; Liu, G.; Greer, D.; Bhandari, A., et al. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* **2012**, *151*, 1431-1442, doi:10.1016/j.cell.2012.11.019.
199. Francioli, L.C.; Polak, P.P.; Koren, A.; Menelaou, A.; Chun, S.; Renkens, I.; Genome of the Netherlands, C.; van Duijn, C.M.; Swertz, M.; Wijmenga, C., et al. Genome-wide patterns and properties of de novo mutations in humans. *Nat Genet* **2015**, *47*, 822-826, doi:10.1038/ng.3292.
200. Rahbari, R.; Wuster, A.; Lindsay, S.J.; Hardwick, R.J.; Alexandrov, L.B.; Turki, S.A.; Dominiczak, A.; Morris, A.; Porteous, D.; Smith, B., et al. Timing, rates and spectra of human germline mutation. *Nat Genet* **2016**, *48*, 126-133, doi:10.1038/ng.3469.
201. Acuna-Hidalgo, R.; Veltman, J.A.; Hoischen, A. New insights into the generation and role of de novo mutations in health and disease. *Genome Biol* **2016**, *17*, 241, doi:10.1186/s13059-016-1110-1.
202. Goldmann, J.M.; Veltman, J.A.; Gilissen, C. De Novo Mutations Reflect Development and Aging of the Human Germline. *Trends Genet* **2019**, *35*, 828-839, doi:10.1016/j.tig.2019.08.005.
203. Gonzalez-Perez, A.; Sabarinathan, R.; Lopez-Bigas, N. Local Determinants of the Mutational Landscape of the Human Genome. *Cell* **2019**, *177*, 101-114, doi:10.1016/j.cell.2019.02.051.
204. Roychowdhury, T.; Abyzov, A. Chromatin organization modulates the origin of heritable structural variations in human genome. *Nucleic Acids Res* **2019**, *47*, 2766-2777, doi:10.1093/nar/gkz103.
205. Xia, B.; Yan, Y.; Baron, M.; Wagner, F.; Barkley, D.; Chiodin, M.; Kim, S.Y.; Keefe, D.L.; Alukal, J.P.; Boeke, J.D., et al. Widespread Transcriptional Scanning in the Testis Modulates Gene Evolution Rates. *Cell* **2020**, *180*, 248-262 e221, doi:10.1016/j.cell.2019.12.015.
206. Clarke, H.J.; Vieux, K.F. Epigenetic inheritance through the female germ-line: The known, the unknown, and the possible. *Semin Cell Dev Biol* **2015**, *43*, 106-116, doi:10.1016/j.semcdb.2015.07.003.
207. Sharma, U.; Rando, O.J. Metabolic Inputs into the Epigenome. *Cell Metab* **2017**, *25*, 544-558, doi:10.1016/j.cmet.2017.02.003.
208. Sales, V.M.; Ferguson-Smith, A.C.; Patti, M.E. Epigenetic Mechanisms of Transmission of Metabolic Disease across Generations. *Cell Metab* **2017**, *25*, 559-571, doi:10.1016/j.cmet.2017.02.016.
209. Donkin, I.; Barres, R. Sperm epigenetics and influence of environmental factors. *Mol Metab* **2018**, *14*, 1-11, doi:10.1016/j.molmet.2018.02.006.
210. Sun, W.; von Meyenn, F.; Peleg-Raibstein, D.; Wolfrum, C. Environmental and Nutritional Effects Regulating Adipose Tissue Function and Metabolism Across Generations. *Adv Sci (Weinh)* **2019**, *6*, 1900275, doi:10.1002/advs.201900275.
211. Legoff, L.; D'Cruz, S.C.; Tevosian, S.; Primig, M.; Smagulova, F. Transgenerational Inheritance of Environmentally Induced Epigenetic Alterations during Mammalian Development. *Cells* **2019**, *8*, doi:10.3390/cells8121559.
212. Sarker, G.; Sun, W.; Rosenkranz, D.; Pelczar, P.; Opitz, L.; Efthymiou, V.; Wolfrum, C.; Peleg-Raibstein, D. Maternal overnutrition programs hedonic and metabolic phenotypes across generations through sperm tsRNAs. *Proc Natl Acad Sci U S A* **2019**, *116*, 10547-10556, doi:10.1073/pnas.1820810116.
213. Sciamanna, I.; Serafino, A.; Shapiro, J.A.; Spadafora, C. The active role of spermatozoa in transgenerational inheritance. *Proc Biol Sci* **2019**, *286*, 20191263, doi:10.1098/rspb.2019.1263.
214. Morgan, C.P.; Chan, J.C.; Bale, T.L. Driving the Next Generation: Paternal Lifetime Experiences Transmitted via Extracellular Vesicles and Their Small RNA Cargo. *Biol Psychiatry* **2019**, *85*, 164-171, doi:10.1016/j.biopsych.2018.09.007.

215. Bohacek, J.; Rassoulzadegan, M. Sperm RNA: Quo vadis? *Semin Cell Dev Biol* **2020**, *97*, 123-130, doi:10.1016/j.semcdb.2019.07.005.
216. Zhang, Y.; Shi, J.; Rassoulzadegan, M.; Tuorto, F.; Chen, Q. Sperm RNA code programmes the metabolic health of offspring. *Nat Rev Endocrinol* **2019**, *15*, 489-498, doi:10.1038/s41574-019-0226-2.
217. Skinner, M.K.; Guerrero-Bosagna, C.; Haque, M.M. Environmentally induced epigenetic transgenerational inheritance of sperm epimutations promote genetic mutations. *Epigenetics* **2015**, *10*, 762-771, doi:10.1080/15592294.2015.1062207.
218. Danchin, E.; Pocheville, A.; Huneman, P. Early in life effects and heredity: reconciling neo-Darwinism with neo-Lamarckism under the banner of the inclusive evolutionary synthesis. *Philos Trans R Soc Lond B Biol Sci* **2019**, *374*, 20180113, doi:10.1098/rstb.2018.0113.
219. Guerrero-Bosagna, C. From epigenotype to new genotypes: Relevance of epigenetic mechanisms in the emergence of genomic evolutionary novelty. *Semin Cell Dev Biol* **2020**, *97*, 86-92, doi:10.1016/j.semcdb.2019.07.006.
220. Horsthemke, B. A critical view on transgenerational epigenetic inheritance in humans. *Nat Commun* **2018**, *9*, 2973, doi:10.1038/s41467-018-05445-5.
221. Kawamura, Y.; Sanchez Calle, A.; Yamamoto, Y.; Sato, T.A.; Ochiya, T. Extracellular vesicles mediate the horizontal transfer of an active LINE-1 retrotransposon. *J Extracell Vesicles* **2019**, *8*, 1643214, doi:10.1080/20013078.2019.1643214.
222. Spadafora, C. The "evolutionary field" hypothesis. Non-Mendelian transgenerational inheritance mediates diversification and evolution. *Prog Biophys Mol Biol* **2018**, *134*, 27-37, doi:10.1016/j.pbiomolbio.2017.12.001.
223. Sasaki, H.; Matsui, Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat Rev Genet* **2008**, *9*, 129-140, doi:10.1038/nrg2295.
224. Strome, S.; Updike, D. Specifying and protecting germ cell fate. *Nat Rev Mol Cell Biol* **2015**, *16*, 406-416, doi:10.1038/nrm4009.
225. Nelson, J.O.; Chen, C.; Yamashita, Y.M. Germline stem cell homeostasis. *Curr Top Dev Biol* **2019**, *135*, 203-244, doi:10.1016/bs.ctdb.2019.04.006.
226. Shaha, C.; Tripathi, R.; Mishra, D.P. Male germ cell apoptosis: regulation and biology. *Philos Trans R Soc Lond B Biol Sci* **2010**, *365*, 1501-1515, doi:10.1098/rstb.2009.0124.
227. Yadav, P.K.; Tiwari, M.; Gupta, A.; Sharma, A.; Prasad, S.; Pandey, A.N.; Chaube, S.K. Germ cell depletion from mammalian ovary: possible involvement of apoptosis and autophagy. *J Biomed Sci* **2018**, *25*, 36, doi:10.1186/s12929-018-0438-0.
228. Arnheim, N.; Calabrese, P. Germline Stem Cell Competition, Mutation Hot Spots, Genetic Disorders, and Older Fathers. *Annu Rev Genomics Hum Genet* **2016**, *17*, 219-243, doi:10.1146/annurev-genom-083115-022656.
229. Immler, S.; Otto, S.P. The Evolutionary Consequences of Selection at the Haploid Gametic Stage. *Am Nat* **2018**, *192*, 241-249, doi:10.1086/698483.
230. Kaessmann, H. Origins, evolution, and phenotypic impact of new genes. *Genome Res* **2010**, *20*, 1313-1326, doi:10.1101/gr.101386.109.
231. Zhang, Y.E.; Long, M. New genes contribute to genetic and phenotypic novelties in human evolution. *Curr Opin Genet Dev* **2014**, *29*, 90-96, doi:10.1016/j.gde.2014.08.013.
232. Werner, A.; Piatek, M.J.; Mattick, J.S. Transpositional shuffling and quality control in male germ cells to enhance evolution of complex organisms. *Ann N Y Acad Sci* **2015**, *1341*, 156-163, doi:10.1111/nyas.12608.
233. Simonti, C.N.; Capra, J.A. The evolution of the human genome. *Curr Opin Genet Dev* **2015**, *35*, 9-15, doi:10.1016/j.gde.2015.08.005.
234. Vandenbrouck, Y.; Lane, L.; Carapito, C.; Duek, P.; Rondel, K.; Bruley, C.; Macron, C.; Gonzalez de Peredo, A.; Coute, Y.; Chaoui, K., et al. Looking for Missing Proteins in the Proteome of Human Spermatozoa: An Update. *J Proteome Res* **2016**, *15*, 3998-4019, doi:10.1021/acs.jproteome.6b00400.
235. Nyberg, K.G.; Carthew, R.W. Out of the testis: biological impacts of new genes. *Genes Dev* **2017**, *31*, 1825-1826, doi:10.1101/gad.307496.117.
236. Kondo, S.; Vedanayagam, J.; Mohammed, J.; Eizadshenass, S.; Kan, L.; Pang, N.; Aradhya, R.; Siepel, A.; Steinhauer, J.; Lai, E.C. New genes often acquire male-specific functions but rarely become essential in Drosophila. *Genes Dev* **2017**, *31*, 1841-1846, doi:10.1101/gad.303131.117.
237. Zhang, J.Y.; Zhou, Q. On the Regulatory Evolution of New Genes Throughout Their Life History. *Mol Biol Evol* **2019**, *36*, 15-27, doi:10.1093/molbev/msy206.
238. Witt, E.; Benjamin, S.; Svetec, N.; Zhao, L. Testis single-cell RNA-seq reveals the dynamics of de novo gene transcription and germline mutational bias in Drosophila. *Elife* **2019**, *8*, doi:10.7554/eLife.47138.
239. Raices, J.B.; Otto, P.A.; Vibranovski, M.D. Haploid selection drives new gene male germline expression. *Genome Res* **2019**, *29*, 1115-1122, doi:10.1101/gr.238824.118.

240. Gregoire, M.C.; Massonneau, J.; Simard, O.; Gouraud, A.; Brazeau, M.A.; Arguin, M.; Leduc, F.; Boissonneault, G. Male-driven de novo mutations in haploid germ cells. *Mol Hum Reprod* **2013**, *19*, 495-499, doi:10.1093/molehr/gat022.
241. Cave, T.; Desmarais, R.; Lacombe-Burgoyne, C.; Boissonneault, G. Genetic Instability and Chromatin Remodeling in Spermatids. *Genes (Basel)* **2019**, *10*, doi:10.3390/genes10010040.
242. Lanfear, R.; Kokko, H.; Eyre-Walker, A. Population size and the rate of evolution. *Trends Ecol Evol* **2014**, *29*, 33-41, doi:10.1016/j.tree.2013.09.009.
243. Huber, C.D.; Durvasula, A.; Hancock, A.M.; Lohmueller, K.E. Gene expression drives the evolution of dominance. *Nat Commun* **2018**, *9*, 2750, doi:10.1038/s41467-018-05281-7.
244. Nag, A.; Savova, V.; Fung, H.L.; Miron, A.; Yuan, G.C.; Zhang, K.; Gimelbrant, A.A. Chromatin signature of widespread monoallelic expression. *Elife* **2013**, *2*, e01256, doi:10.7554/eLife.01256.
245. Massah, S.; Beischlag, T.V.; Prefontaine, G.G. Epigenetic events regulating monoallelic gene expression. *Crit Rev Biochem Mol Biol* **2015**, *50*, 337-358, doi:10.3109/10409238.2015.1064350.
246. Guilmatre, A.; Sharp, A.J. Parent of origin effects. *Clin Genet* **2012**, *81*, 201-209, doi:10.1111/j.1399-0004.2011.01790.x.
247. Patten, M.M.; Ross, L.; Curley, J.P.; Queller, D.C.; Bonduriansky, R.; Wolf, J.B. The evolution of genomic imprinting: theories, predictions and empirical tests. *Heredity (Edinb)* **2014**, *113*, 119-128, doi:10.1038/hdy.2014.29.
248. Kappil, M.; Lambertini, L.; Chen, J. Environmental Influences on Genomic Imprinting. *Curr Environ Health Rep* **2015**, *2*, 155-162, doi:10.1007/s40572-015-0046-z.
249. Rodrigues, J.A.; Zilberman, D. Evolution and function of genomic imprinting in plants. *Genes Dev* **2015**, *29*, 2517-2531, doi:10.1101/gad.269902.115.
250. Ben Maamar, M.; King, S.E.; Nilsson, E.; Beck, D.; Skinner, M.K. Epigenetic transgenerational inheritance of parent-of-origin allelic transmission of outcross pathology and sperm epimutations. *Dev Biol* **2020**, *458*, 106-119, doi:10.1016/j.ydbio.2019.10.030.
251. Thamban, T.; Agarwaal, V.; Khosla, S. Role of genomic imprinting in mammalian development. *J Biosci* **2020**, *45*.
252. Batista, R.A.; Kohler, C. Genomic imprinting in plants-revisiting existing models. *Genes Dev* **2020**, *34*, 24-36, doi:10.1101/gad.332924.119.
253. Huang, L.O.; Labbe, A.; Infante-Rivard, C. Transmission ratio distortion: review of concept and implications for genetic association studies. *Hum Genet* **2013**, *132*, 245-263, doi:10.1007/s00439-012-1257-0.
254. Fishman, L.; McIntosh, M. Standard Deviations: The Biological Bases of Transmission Ratio Distortion. *Annu Rev Genet* **2019**, 10.1146/annurev-genet-112618-043905, doi:10.1146/annurev-genet-112618-043905.
255. Brachet, E.; Sommermeyer, V.; Borde, V. Interplay between modifications of chromatin and meiotic recombination hotspots. *Biol Cell* **2012**, *104*, 51-69, doi:10.1111/boc.201100113.
256. Arbeithuber, B.; Betancourt, A.J.; Ebner, T.; Tiemann-Boege, I. Crossovers are associated with mutation and biased gene conversion at recombination hotspots. *Proc Natl Acad Sci U S A* **2015**, *112*, 2109-2114, doi:10.1073/pnas.1416622112.
257. Nadeau, J.H. Do Gametes Woo? Evidence for Their Nonrandom Union at Fertilization. *Genetics* **2017**, *207*, 369-387, doi:10.1534/genetics.117.300109.
258. Tock, A.J.; Henderson, I.R. Hotspots for Initiation of Meiotic Recombination. *Front Genet* **2018**, *9*, 521, doi:10.3389/fgene.2018.00521.
259. Immler, S. The sperm factor: paternal impact beyond genes. *Heredity (Edinb)* **2018**, *121*, 239-247, doi:10.1038/s41437-018-0111-0.