Networks Consolidate the Core Concepts of Evolution by Natural Selection

François Papale, Jordane Saget, Eric L Bapteste

Trends in Microbiology

CellPress
REVIEWS

Opinion

Networks Consolidate the Core Concepts of Evolution by Natural Selection

François Papale,[1,2] Jordane Saget,[2] and Éric Bapteste[2,*]

Microbiology has unraveled rich evidence of ongoing reticulate evolutionary processes and complex interactions both within and between cells. These phenomena feature real biological networks, which can logically be analyzed using network-based tools. It is thus not surprising that network sciences, a field independent from evolutionary biology and microbiology, have recently pervasively infused their methods into both fields. Importantly, network tools bring forward observations enhancing the understanding of three core evolutionary concepts: variation, fitness, and heredity. Consequently, our work shows how network sciences can enhance evolutionary theory by explaining the evolution by natural selection of a broad diversity of units of selection, while updating the popular figure of Darwin's tree of life with a comprehensive sketch of the networks of evolution.

Introduction: The Core Principles of Evolution by Natural Selection

In 1859, the theory of **evolution by natural selection (ENS)** (see Glossary) revolutionized the understanding of the history and diversity of life [1]. Biodiversity started to be explained as the result of the divergence of biological entities with respect to the populations of their last common ancestors by a general evolutionary process common to all life forms, including microbes. As summarized by Lewontin [2], ENS operates under three conditions: (i) that members of a population present phenotypic variation, (ii) that their different phenotypes translate into differences in survival and reproduction rates (differential fitness), and that (iii) these fitness differences are inheritable from parents to offspring. Therefore, the presence of variation within a population, the transmission of fitness-related variation from ancestors to descendants, and the causes of fitness differences are key conditions for evolution by natural selection.

Importantly, ENS was historically developed hand-in-hand with tree-thinking [3], wrongly suggesting that ENS requires and implies a simple genealogical continuity: the perpetuation of a lineage by descent with modification, thereby justifying the existence of a single tree of species relating all living beings, including microbes. Yet, many discoveries, especially from microbiology, also reported abundant reticulate processes and interactions involving the microbial world (Figure 1, Key Figure), which was even seen by some as a headlong challenge to (neo-)Darwinism and sometimes as a support to Lamarckian models of evolution [4–7].

While the evolution of real biological networks seemed to some to challenge the tree of life and the importance attributed by Darwin to the principle of divergence, we argue that the existence of networks is no danger to the hypothesis of ENS. Indeed, ENS can demonstrably apply to more complicated entities than simple (monogenomic) lineages, shedding evolutionary light on a variety of otherwise unexplained biological phenomena [8,9]. Network-based tools support a reformulation of the core concepts of ENS, not because these are false, but because a broader theoretical framework can be provided by network thinking (Box 1).

Highlights

Network sciences are introducing network-based models into all relevant biological fields, notably in microbiology and evolutionary biology.

The three core principles of evolution – variation, heredity, and differential fitness – crystalized in the 1970s, still serve as a conceptual benchmark for the theory of evolution by natural selection.

Network-based inquiries in microbiology tap into the complexity of the systems of interactions that underscore the evolutionary dynamics of cellular life, thereby augmenting our understanding of the three core principles of evolution.

Conversely, the improved understanding of evolutionary dynamics provided by networks informs microbiology in various ways.

The network-based understanding of evolution is often considered as anti-Darwinian in spirit, a claim that can be debunked using microbiology.

[1]Departement of Philosophy, University of Montreal, Montréal, QC, H3C 3J7, Canada
[2]Institut de Systématique, Evolution, Biodiversité (ISYEB), Sorbonne Université, CNRS, Museum National d'Histoire Naturelle, EPHE, Université des Antilles, 75005 Paris, France

*Correspondence: epbapteste@gmail.com (É. Bapteste).

Key Figure

A Network-Based Representation of the Evolution of Cellular Life

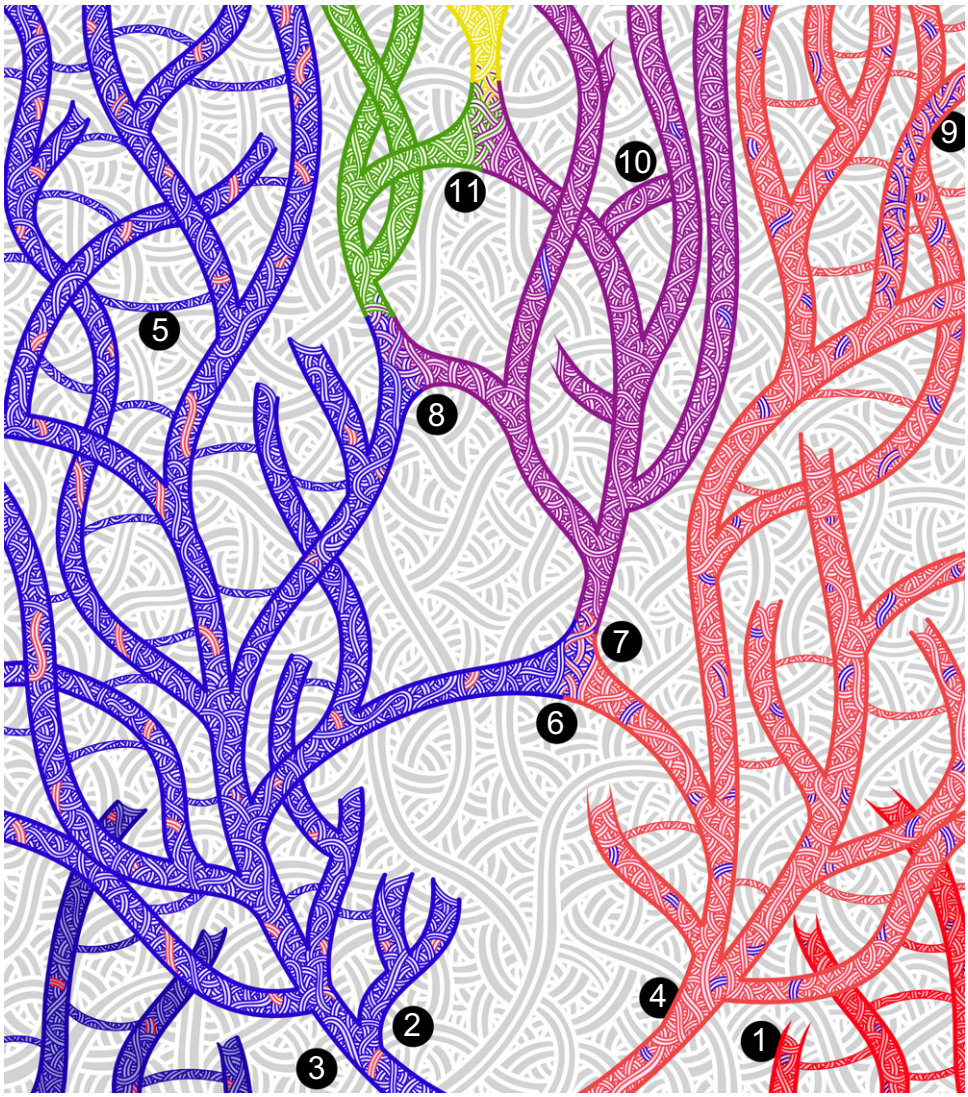

Figure 1. This is an artistic view of the type of processes involved in the evolution of cellular life. Blue lineages: bacteria. Red lineages: archaea. Purple lineages: nonphotosynthetic eukaryotes. Green lineages: primary photosynthetic eukaryotes. Yellow lineages: secondarily photosynthetic eukaryotes. Triple lines within cellular lineages correspond to lineages of gene families in interactions. Thread colors of gene families represent the origin of the gene family: blue for bacteria, red for archaea, other colors for eukaryotes. These colors can be combined in case of evolution of chimeric composite genes. Processes are indicated by circles; the numbers correspond to: (1) extinction, (2) divergence, (3) coalescence (illustrating how distinct lineages trace back to a single common ancestor), (4) intralineage molecular interactions (the networks of 'small branches' within bigger ones), (5) DNA transfer intra-domain, (6) primary endosymbiosis (mitochondria), (7) evolution of chimeric genes, (8) primary endosymbiosis (plastids), (9) massive inter-domain DNA transfer (Haloarchaea), (10) introgression, and (11) secondary endosymbiosis (plastids). The gray network background describes nongenetic interactions between biotic entities/environmental processes.

Glossary

Candidate phyla radiation (CPR): this refers to a group of newly discovered bacterial phyla that are composed of small-genome entities and whose discovery significantly increased the overall known bacterial biodiversity.

DPANN: *Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, Nanohaloarchaea*; DPANN is a superphylum of Archaea, grouping together a variety of phyla, among which are the five that form the acronym and the name of the superphylum.

Evolution by means of natural selection (ENS): refers to changes in the distribution of variation in a population, which is guided by the selective pressures of the environment, that is, natural selection.

Lateral gene transfer (LGT): this refers to the transfer of genetic material between organisms, which is not due to reproduction as traditionally construed and which happens between potentially phylogenetically unrelated entities.

Box 1. Traditional and Updated Interpretations of Evolution by Natural Selection

Traditionally, ENS describes an evolutionary process that involves independent lineages, competing for survival and reproduction. According to Lewontin, this requires three conditions:

- **Variation**
  Relevant individual differences or variations are traditionally associated with independent lineages. Network tools first provide much more complex descriptions of this already recognized variation. Second, they highlight that variation also arises in broader entities, such as multispecies consortia, through various biological interactions.
- **Differential Fitness**
  Due to variation, some units of selection survive and reproduce better than others. Bearers of variation are also bearers of fitness. Traditionally, only entities forming vertical lineages can be bearers of fitness. Network modeling, by contrast, shows that more complex entities can be bearers of fitness (such as squid–*Vibrio* associations). It also shows that fitness is a relational property that can be accurately assessed only by analyzing the interaction networks from which units of selection emerge (such as the topological stability of any squid–*Vibrio* interaction network, and the feedback loops within it) and the interaction networks in which units of selection are involved (such as the global or local topological stability of the ecological network into which squid–*Vibrio* associations figure as nodes).
- **Inheritance**
  Traditionally, inheritance involves the formation of vertical lineages – which can be traced thanks to the privileged transmission of traits from parents to their offspring. Network tools have widely transformed our understanding of inheritance patterns (e.g., the ubiquity of LGT, represented in phylogenetic networks). Network tools also expand our understanding of how inheritance is realized, by modeling the channels of transmission of different components, critical to assess the re-production of a unit of selection.

Consistently, an updated formulation of ENS, from the viewpoint of network modeling, will consider that ENS could explain the evolution of interaction networks from which units of selection – such as genes, organisms, multispecies consortia, symbiotic assemblages, etc. – emerge. It can do so if (i) populations of such interaction networks vary, (ii) in such a way that determines the capacity of interaction networks to persist and to be re-produced (differential fitness), (iii) via simple or complex transmission channels, re-producing at least the interaction subnetworks that were bearing the traits tied to differential fitness (inheritance).

The argument that a different formalization of evolutionary processes might allow for a more comprehensive investigation of evolution is not new [10–12]. Yet, this argument is usually motivated by independent suggestions to emphasize specific underappreciated biological processes studied by different biological disciplines [13–18]. By contrast, the development of the field of network sciences [19–21], an emerging transversal field that is largely independent from biological disciplines, by the types of findings that network tools allow, will strongly support a reformulation of Lewontin's three conditions for ENS. Network sciences borrow tools and practices from graph theory, information theory, computer science, and physics in order to analyze various aspects of networks and their dynamics. The input of this nascent field, in both evolutionary biology and microbiology, is increasingly noticeable [22–51]. Crucially, the development of network sciences shows that the modes of generation of variation, the modes of transmission of that variation, and the causes of fitness differences can be better understood in the microbial world by giving greater room to network tools. Consequently, researchers will be able to analyze a richer diversity of candidate selective units, because what can be said to vary, to have fitness, and to be featured in hereditary interactions, can be a much more complex set of lineages, involved in much more complex interactions, than the traditional focus on lineages suggests. Sets of (related or unrelated) entities, whose components are demonstrably connected by reticulate processes or interactions, can also figure on the list of what natural selection can explain.

Networks Highlight Additional Sources of Variation

In the classic theory of ENS, the presence of variations matters, but not their origins. Network tools uphold this view, while providing sharper descriptions of 'what is variation'.

Consider analyses of microbiology and transcriptomics with network tools. Network analyses of gene evolution, of gene regulation, and of protein–protein interactions [22,23,25–52] are currently

enriching the understanding of the reticulate processes and interactions that lead to a diversity of, for instance, organismal phenotypes. These network tools have shown that variation within a lineage obtains either when the components (nodes) of the modeled interaction networks change, for example, when new genes or new proteins evolve, or when the interactions between components (edges) change, for example, when new associations of genes and/or proteins (corresponding to new edges in the interaction networks) evolve. For example, domain combination events (or exon shuffling) create novel genes from genetic material within an organismal lineage. This universal process has been modeled using oriented domain (or exon) networks, describing the order in which domains (or exons) succeed to one another in genes or in proteins of different organisms [52–55]. Organisms emerge from interactions of modular infraorganismal entities into which only network modeling can provide satisfying insights. Thus, even the typical units of selection (e.g., organisms or genes) can be thought of as emerging from interaction networks. A gene emerges from a complex network involving domains, an organism emerges from a complex network involving genes. A comprehensive understanding of the sources of this variation warrants the use of network-based 'tools'.

Moreover, this network-based perspective on the sources of variation within traditional units of selection warrants an expansion of the concept of evolutionarily relevant variation, which invites an explanation of a broader set of units of selection by ENS (Box 2). Thus, network tools bring forward **lateral gene transfer (LGT)** [56–60], introducing genetic material from a nondirect parent into a recipient host (Figure 1), as the most studied source of polyphyletically generated genetic variations. Two kinds of network tools are nowadays involved in cutting-edge inquiries on LGT: gene-sharing networks [33,61–65] and bipartite gene–genome networks [29,66,67]. Both stress the importance of LGT in a broad diversity of entities: within the cellular world, within mobile elements, and between cellular organisms and mobile elements, showing that lineages within the microbial world are comprised of modular, phylogenetically mosaic members, as illustrated by the discoveries of pangenomes [68,69] and the fluidity of microbial genomes [57,58,63,70]. Evolutionarily relevant variation arises at the intersection of complex phylogenetic trajectories and through actual networks of interactions involving independent lineages.

Within the eukaryotic world, endosymbioses (Figure 1) – at the origin of the eukaryotic cells [71] as well as at the origins of primary, secondary, and tertiary photosynthetic eukaryotes [72–74] – offer other notorious examples of variations with foreign sources that led to the emergence of novel units of selection. Moreover, within phylogenetically composite protists, some traits evolved via additional reticulate processes and interactions between components from different lineages [75]. For example, the microbial eye of *Nematodinium* [76] is a structure which arises from interactions between secondarily acquired plastids and mitochondria, hence, *within* an organism yet between components with distinct origins. While neo-Darwinism would treat entities like *Nematodinium* as coherent variational units arising from phylogenetic unity, a network-based perspective shows that the variation at play is irreducibly the effect of actual networks of interactions at work across both space (networks of interactions) and time (phylogenetic networks). This rationale applies also to genes themselves, for example, S-genes in protists [41,77,78] (Figure 1). The process of endosymbiosis, and the subsequent processes of emergence of chimeric traits, can thus be better modeled and analyzed using network tools [23,77] because some significant variation originates from interactions between the lineage of interest and other lineages.

Moreover, significant variations caused by multispecies interactions can also arise in the absence of genetic introgression inside a single host lineage. We call the resulting effects coconstructed variations. Coconstructed variation arises when organisms interact functionally (again, not through any type of gene exchange) so that a new phenotype arises in at least one of the

Box 2. Diversity of the Units of Selection to Be Unraveled by Network Analyses

Network studies can unravel a diversity of sources of variation, a diversity of channels of inheritance, and unravel relational causes of fitness, as selectable units emerge from, or are part of, interaction networks. Thus, for any unit of selection, variations can be due to genetic changes in interaction networks: (i) from a single monophyletic source, (ii) from polyphyletic sources, or (iii) they can be coconstructed in the absence of introgression. Likewise, inheritance channels can be; (iv) simple (all components of the interaction network from which the unit of selection emerges are transmitted together), or (v) multiple (when different components of the interaction network from which the unit of selection emerges are transmitted separately); when multiple channels of inheritance exist (vi) the transmission of different components can be synchronized or (vii) nonsynchronized. Finally, the fitness of a selective unit may have (viii) internal causes (being the result of stabilizing processes, such as feedback loops, between the components of interaction networks) and/or (ix) external causes (such as the stability of the interaction networks to which the selective unit is connected). As network tools provide evidence for (ii–iii), and (v–ix), ENS could explain the evolution of entities presenting a broad range of combinations of variation (i–iii), inheritance (iv–vii), and fitness (viii–ix) modalities.

For example, in the case of bacterial clones, organismal variation, inheritance, and fitness only involve related entities. ENS generates traditional units of selection, that is, monogenomic individuals within populations. By contrast, in the case of bacterial cells harboring new laterally acquired genes, organismal variation originates from polyphyletic sources, which are inherited by vertical descent inside their host lineage, a combination known as introgressive descent [8], and the resulting phenotype is selected within microbial communities. ENS generates units of selection that are still individual organisms within populations, yet these individuals are genetic chimera, merging genetic information from multiple sources. Network analyses of the sources of variations (monophyletic vs polyphyletic) can then determine the ratio of monogenomic:merger individuals evolved by ENS in microbial communities.

Another example is *Chlorochromatium aggregatum*, a phototrophic bacterial consortium in which partners from two species synchronize their cellular division [123]; in this case, variation-bearing fitness originates from polyphyletic sources, is inherited through multiple 'synchronized' channels by vertical descent, and the resulting phenotype is selected across groups bearing the traits. ENS generates units of selection that differ from standard organisms, usually called egalitarian collectives of entities.

Yet, the case of multispecies consortia, such as squid–*Vibrio* associations, in which an association of partners from two species is repeatedly reconstructed, or in the case of hosts with horizontally transmitted symbionts, are different, because variation-bearing fitness originates from polyphyletic sources, is re-produced through multiple 'desynchronized' channels by vertical descent, and the resulting phenotype is selected across groups bearing the traits. ENS here generates transient re-produced collectives of entities; such units of selection differ from the above-mentioned traditional organisms and from egalitarian collectives of entities. Network analyses of the channels of inheritance (synchronized vs desynchronized) can then determine the ratio of egalitarian collectives of entities: transient re-produced collectives of entities within an environment. Network analyses can also encourage description of transient re-reproduced collectives of individuals as recurring patterns in interaction networks, because then the units of selection bearing the fitness-related traits are interaction networks, which beget similar networks by re-construction.

ENS could explain even more unusual units of selection. In the case of the nitrogen cycle, variation-bearing fitness originates from polyphyletic sources and is re-produced through multiple desynchronized channels by vertical descent and by introgressive descent in at least some of the entities whose associations generate a selectable trait. The resulting phenotype is selected among populations of nitrogen cycles. Here, ENS directly generates functional patterns of interactions as selective units, also known as 'songs' in the ITSNTS ('It's the song not the singer' theory) models [124]. These patterns of interactions are re-produced across evolutionary time, although some of their components (e.g., the microbial taxa contributing to nitrogen cycling) can change. Network analyses of the channels of inheritance (songs re-produced by 'desynchronized vertical descent' vs songs re-produced by 'desynchronized vertical + introgressive descent') can determine the ratio of songs in which components do not change to songs in which components change on the planet.

Finally, in the case of obligate syntrophic consortia of archaeal methanotrophs and sulfate-reducing bacteria in thermodynamically challenging conditions [125], or in the case of density-dependent bacterial and bacteriophage interactions, variation-bearing fitness originates from polyphyletic sources, is inherited through multiple channels, and the fitness is conditioned by a feedback loop affecting the reproduction/survival of the associated entities producing the selectable traits. Then, ENS generates a remarkable kind of interaction networks, which we call evosystems. Network analyses of the causes of fitness (absence of, or facultative feedback loops between associated entities vs obligate feedback loops between associated entities) can determine the ratio of general songs to evosystems.

Therefore, ENS could explain the evolution of; (i) individual entities (be they genetic mosaics or not) to, (ii) egalitarian collectives of entities, (iii) transient re-produced collectives of entities, (iv) songs, and (v) evosystems.

interacting entities. The bioluminescence of interacting *Vibrio fischeri* and *Euprymna scolopes* is a coconstructed trait that can presumably vary across a population of interacting squids and microbes, and accordingly stands as evolutionarily relevant coconstructed variation [79].

Abundant cases of coconstructed variations have been described based on studies of symbiotic associations of microbes with plant or animal hosts [80,81]. Holobionts have been proposed, and rightly criticized [82,83], as the paroxysmal situation in which some multispecies symbiotic associations could lead to emergent, coconstructed selectable traits [84]. Importantly, the generation of coconstructed variation cannot be accurately described by focusing only on host lineages. Hence the single-lineage-associated forms of variation are here highly insufficient. Yet, multipartite networks offer a method to highlight and to analyze such multilevel and multispecific associations [85,86]. For instance, microbiome–microbiota–eukaryotic hosts tripartite networks, simultaneously representing the distribution of microbial genes within microbial taxa, the distribution of microbial genes within eukaryotic hosts, and the distribution of microbial taxa within eukaryotic hosts, provide a framework to identify microbial genes involved in the coconstruction of holobiont variation, irrespective of what microbial taxa carry these genes. Thus, the ability to extract energy from carbohydrates in individual mammals depends on their gut microbial genes rather than on their gut microbial taxa [87,88], and their impact on the functional integration of the interacting host and microbes. The claim here is 'not' that holobionts are, by definition, units of selection because they can be modeled as networks. This is an empirical question. The claim is that, to solve this empirical problem, one is required to consider complex variation, which emerges from real networks of interactions.

Overall, networks do not challenge the notion that variation (*sensu* Lewontin) is mandatory for ENS. Yet, they provide a clearer picture of what 'is' variation. Network tools have thus enhanced our understanding of variation at all levels of organization: genes, microbes (viruses included), eukaryotes, and multispecies assemblages. And because network studies keep on accumulating evidence demonstrating the diversity of the causes of variation, network tools inexorably show that focusing on single lineage entities as 'the' single units of selection is increasingly problematic. Moreover, tracking with more accuracy the sources of variation opens the possibility to track additional actual selective units, larger than single lineages and involving multiple lineages (e.g., operationally defined, in interaction network models, as communities of nodes from multiple lineages). Thus, what appears to some as a Lamarckian (read non-Darwinian) process in a classical evolutionary perspective because it results from reticulated processes [89–93] (e.g., a laterally acquired gene in lineage A from lineage B) is reframed by network analysis of the phenomena in play. Acquired variation can instead be seen as internal variation with respect to some selective unit broader than organisms from a given lineage, that is, a broader network of interactions, such as gene exchange communities transiently arising from the interaction between independent phylogenetic lineages. Furthermore, as we show next, network tools allow us to test whether such transient interaction networks can have inheritable fitness-related traits, that is, whether they can be re-produced, making them *bona fide* units of selection.

Networks Highlight Additional Channels of Transmission

In the standard model of ENS, the inheritance of fitness-related variation is centered on lineage formation by vertical descent. By contrast, network-based models provide powerful explanations of complex nonvertical inheritance processes involving entities at various levels of organization, such as plasmids, phages, and chromosomes, and complex inheritance patterns, such as those related to introgressive descent [8], gene externalization [28,29,86], or autologous genes [94]. Accordingly, network-based methods make two distinct contributions to the understanding of heredity in the context of ENS. First, network tools allow for a better understanding of a variety

of channels of transmission and the phylogenetic patterns they create. Second, network tools highlight that interaction networks are featured among the entities that are more or less well reproduced from one generation to the next, thanks to heredity.

Transmission of variation can be realized by pathways much more complex than the ones suggested by the vertical replication-with-fidelity framework. It was recently shown that multipartite viruses (e.g., the faba bean necrotic stunt virus, from the genus *Nanovirus*, and the family Nanoviridae), while being considered unitary reproductive entities, see their reproduction being triggered by the interaction of genome segments present in different host cells [95]. This complexity of transmission roads is also true of variation arising in polyphyletic entities, which, by definition, involves networks of transmission [96,97]. Plasmid-encoded or organellar-encoded variations, for example, are known to display different modes and sometimes different channels of transmission from chromosomal-encoded or nuclear-encoded variations [86,96].

Yet, this plurality of modes and channels of transmission tends to be considered as a secondary issue, unduly complicating the description of inheritance without theoretical significance. Accordingly, few models of inheritance assume decoupled transmission roads of genes carried on entities from different levels of biological organization (e.g., viruses, plasmids, mitochondria, chloroplasts, chromosomes, distantly related cells within multispecies biofilms). Network studies, by including various interacting entities in inheritance models, can alleviate this neglect. Indeed, networks incorporate interacting components that are not all phylogenetically related, such as plasmids, phages, and chromosomes, even if these components belong to different levels of organization [29,33,66,98]. This alternative modeling stresses the diversity and importance of channels and modes of transmissions beyond direct inheritance from one (or two related) parent(s) to their offspring, pervasive in microbiology. It shows the many roads to LGT and the phylogenetic patterns these processes generate in nature, such as conjugal transfer of chromosomal DNA [99], nanotubes (between multiple neighbor species [100,101]), and a diversity of more or less host specific mobile elements [66,102–104]. Network tools, such as gene sharing and bipartite gene–genome network analyses, also highlight barriers to transmission, such as alternative genetic codes [29] in some taxa, or predict processes interfering with LGT rates [65] that add up to various types of barrier to gene transfer already characterized [105,106].

Remarkably, mobile genetic elements do not necessarily independently move DNA between cells; some also affect genetic transmission by interacting among themselves, a phenomenon that is best modeled using interaction networks. For example, some viruses are plasmid superspreaders [107]: these viruses release intact plasmids from infected cells. Similarly, *Arbitrium*, a viral peptide used by viruses for communication, determines whether phages will enter into the lytic or lysogenic phase, thus affecting the spread of their own genes and of those of their cellular hosts [108]. The system involving viruses has specific characteristics, such as density, and only by taking the whole system into account can we understand the patterns of transmission of viruses and their hosts.

In the case of endosymbioses too, transmission trespasses phylogenetic borders, since immediate descendants resulting from these introgressive events did not look like a single individual from their ancestral lineages. More generally, in case of symbioses and host–microbe interactions, a richer view on the diversity of channels and modes of transmission is required to track the transmission of traits, when these traits are coconstructed [97,109,110]. A temporal decoupling in the transmission of microbe and host genes typically produces complex evolutionary dynamics [91,92] in need of ENS-based interpretations.

The above examples illustrate that some channels of heredity are irreducibly reticulated: they form phylogenetic networks which can increasingly be studied using network-based tools, to enhance the understanding of the dynamics of heredity. The complexity of transmission channels and phylogenetic patterns means that the individuals that vary, and that realize heredity, can be themselves pictured as complex networks of interactions. Simply put, there is inheritance when interaction networks, defining a candidate unit of selection, are re-produced. For example, consider an entity emerging from an interaction network with a topology T, which is re-produced (via as many channels of transmission as a specific case requires) with a high degree of fidelity. Heredity is then realized by the one or many channels of transmission, which ensure the similarity between the 'parent' networks and their offspring. But what if an interaction network with the same topology T′ arises independently from T, without any connections between their channels of transmission, as complex as these may be. Would the existence of T and T′ be considered as an instance of heredity? What is at stake here is clearly the generalization of the Darwinian theory, beyond its traditional domain: if any reappearance of a topology is a realization of heredity, then certainly Darwinian evolution would be even more ubiquitous than usually construed. A conservative answer is that, for T and T′ to be the result of heredity, there must be some historical continuity between the two topologies, that is, some components of T and of T′, at least, must directly be related, with the components in T being the ancestors of the components in T′. The key to establish this historical continuity is precisely to track the transmission of components between interaction networks. The independent emergence of two identical prokaryotic syntrophic consortia, for example, in two distinct environments would be a case of convergence, not of heredity.

Therefore, important improvements to network-based tools are still needed to understand the dynamics of transmission, and the heredity of complex units of selection. Indeed, most of the network tools that are currently used lack directionality in time. Whereas network comparison can easily satisfy basic needs by comparing different network topologies (e.g., identifying common subgraphs T and T′ in interaction networks produced from time series), such practice does not yet represent and analyze the specific dynamics, that is, transmission channels, that link two common subgraphs in networks from two time slices.

Although implementing temporality within a single network representation of a unit of selection to capture transmission and inheritance across time slices of interaction networks from time series remains a challenge, the ubiquity of complex transmission of interacting components means that simple neo-Darwinian schemes of heredity and, consequently, heritability, deserve to be enriched as additional, more complicated units of selection, whose traits transmission follows multiple channels, and are exposed by network analyses.

Networks Specify the Context Required for Fitness-Based Selection

The third standard condition for ENS is differential fitness. Importantly, fitness is a relational property relating an individual to its ecological context, but this relational dimension is traditionally 'black boxed' and reduced to an individual's reproductive output. Network sciences start cracking into this black box by providing tools to dive into the complexity of fitness-determining dynamics. This rationale applies across levels of organization, as network-based tools provide insights into the fitness of gene lineages as well as organisms or communities.

At the organism level, fitness can be conceived as the outcome of the topology of a dynamic ecological network involving a community of organisms and their environment. For example, high biodiversity within a microbial community is correlated with high resistance to invasion by outside species [111,112], and recently the cause of this increased stability was investigated using network tools. This investigation showed that the versatile plant pathogen *Ralstonia*

Box 3. Alternative Approaches to Fitness via Networks

Network models known as constructive neutral evolution (CNE), or as presuppression, have been proposed to explain the persistence of an *a priori* costly trait for its bearer(s) (e.g., genes with a deleterious mutation), when this persistence involves selectively neutral, compensatory interactions. In short, deleterious changes in a component are presuppressed by their existing interaction with other components in an interaction network. This would be true within a lineage (e.g., by presuppression of disadvantageous mutations by interacting proteic partners [126–129]) and beyond [e.g., by presuppression of detrimental gene losses in microbial communities [130,131], or in the case in which CNE is also seen as an explanation for the evolution of metabolic hands-off involving nonautonomous individual lineages, as massively proposed for **candidate phyla radiation (CPR)** and **DPANN** [121,122]]. In these situations, entities with reduced individual fitness persist in the environment as members of larger interaction networks, thus delineating a space in which ENS on isolated organismal lineages is not the prime explanation for their traits' diversity and distribution. By contrast, the use of network models can also contribute to the unraveling of the selection of unexpected units, such as patterns of interaction with selectable functions. For example, networks of interaction can be considered units of selection when a real metabolic function sustained by a pattern of interactions between microbes (associated to various taxa) collectively fixing nitrogen is being re-produced so that the frequency of the function encoded by these interactions (nitrogen fixation [121,122]) increases independently from the taxa that sustain it because the microbial taxa involved in that function can be replaced by others taking their role in the real interaction network [124]. Importantly, only network tools could test and analyze the evolution of such complex functions by CNE or by ENS.

solanacearum must invade a community in the rhizosphere of a plant, say a tomato plant, and reach a certain threshold to become virulent for the plant. Thus, the actual resource competition network of the indigenous microbial community must be explored to understand the microbial community's resilience. Specifically, the connectance (measure of the capacity of the community to exploit available resources), the nestedness (measure of resource consumption overlap between generalists and specialists within the community), and the niche overlap (similarity of resource consumption between community members and invading species) have been shown to correlate with the chances of success of the invading pathogen using bipartite resource competition graphs [113]. Therefore, the structure of the interactions between microbial community members (i.e., the actual networks studied using network-based methods) determines the fitness of a potentially invading species. Moreover, there is a further payoff of recognizing the network-like nature of fitness. Not only does the specificity of the actual microbial network provide crucial information about the fitness of microbes, but it is also a highly influential aspect of the fitness of the macro-organisms related to the microbial community. The fitness of the unit of selection at one level of organization, in this case a tomato plant [113], is correlated with the capacity of a pathogen, a unit of selection at a lower level of organization, to invade a microbial community and spread in it. Hence, the fitness of entities traditionally conceived as belonging to different levels of selection cannot be analyzed independently, so that network-based tools must be mobilized to unpack the complexity of the interactions at play.

This situation echoes the vast literature on animal-gut microbiomes [114–116] and experiments in the animal world. For example, the addition of a new lineage in a community (hence of a new node in the ecological network), such as a stick bug on a nonnative bush, can profoundly and quickly redefine the fitness of the organisms across the entire community [117,118] by attracting generalist predators into their ecological network. It is only when the ecological network is stable over organismal generations that organismal trait fitness can have some cumulative effects.

Precisely, the stability of ecological networks can now be modeled via dynamic graphs representing ecological interactions between organisms, qua units of selections, and the evolution of these patterns of interactions could thus explicitly model the relational dimension of natural selection, providing additional modeling of fitness [119]. Consequently, analyses of the structure of interaction patterns using network models, such as co-occurrence networks currently produced to analyze -omics data from various environmental samples and time series [31,120],

could specify when, how, and at what level of organization could fitness, the third core concept of ENS, be effectively invoked to explain diversity. By contrast, the modeling of interactions between components, for example, with protein–protein interaction networks [42,49,50], with metabolic networks [121,122], or with ecological networks [26,51,120], could unravel cases where presuppression rather than ENS is expected to drive the evolution of diversity (Box 3).

Concluding Remarks

The use of network tools in various biological fields, among which microbiology stands out, reflects a methodological opportunity provided by the development of network sciences, rather than a deliberate choice of microbiology and evolutionary biology to renew their ontologies to welcome more units of selection. However, the multiplication of network representations and methods in microbiology may have deep collateral effects (see Outstanding Questions). Networks may become the next tools of choice in evolutionary biology, precisely because they provide an analytically tractable framework to handle a much more inclusive (and complex) ontology of entities and processes.

If we are right, an external discipline to evolutionary biology, network sciences, will, in the near future, fundamentally contribute to the conceptual development of the evolutionary theory by revising its core concepts (Box 1). This will happen in two (somewhat overlapping) stages. First, a technology-driven collision, correlated with the increase in computing power, is already taking place, testifying through the multiplication of network studies in microbiology that both fields are culturally permeable. Second, this multiplication of networks in microbiology will lead to a reformulation of ENS (Box 1). This reformulation will preserve the role of ENS as a central general process to understand microbial evolution, while extending the explanatory scope of ENS to units with increasingly complicated compositions and organizations, beyond simple monogenomic lineages (Box 2).

Acknowledgments

We would like to thank Professors Ford Doolittle, Frédéric Bouchard, Philippe Huneman, Philippe Lopez, Dr Andrew Watson, and Dr Aaron Novick for critical comments. The Social Sciences and Humanities Research Council of Canada provided financial support for this research.

Author Contributions

All the authors wrote the paper. J.S. and É.B. designed and realized Figure 1.

References

1. Darwin, C.A. (1859) *On the Origin of Species by Means of Natural Selection*, John Murray
2. Lewontin, R.C. (1970) The units of selection. *Annu. Rev. Ecol. Syst.* 1, 1–18
3. O'Hara, R.J. (1997) Population thinking and tree thinking in systematics. *Zool. Scr.* 26, 323–329
4. Forterre, P. (2012) Darwin's goldmine is still open: variation and selection run the world. *Front. Cell. Infect. Microbiol.* 2, 106
5. Merhej, V. and Raoult, D. (2012) Rhizome of life, catastrophes, sequence exchanges, gene creations, and giant viruses: how microbial genomics challenges Darwin. *Front. Cell. Infect. Microbiol.* 2, 113
6. Raoult, D. and Koonin, E.V. (2012) Microbial genomics challenge Darwin. *Front. Cell. Infect. Microbiol.* 2, 127
7. Woese, C.R. (2002) On the evolution of cells. *Proc. Natl. Acad. Sci. U. S. A.* 99, 8742–8747
8. Bapteste, E. *et al.* (2012) Evolutionary analyses of non-genealogical bonds produced by introgressive descent. *Proc. Natl. Acad. Sci. U. S. A.* 109, 18266–18272
9. Dupré, J. (2012) *Processes of Life: Essays in the Philosophy of Biology*, Oxford University Press
10. Godfrey-Smith, P. (2009) *Darwinian Populations and Natural Selection*, Oxford University Press
11. Laland, K. *et al.* (2014) Does evolutionary theory need a rethink? Yes, urgently. *Nature* 514, 163–164
12. Pigliucci, M., Müller, G.B., eds (2010) *Evolution – The Extended Synthesis*, MIT Press
13. Doolittle, W.F. and Bapteste, E. (2007) Pattern pluralism and the Tree of Life hypothesis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 2043–2049
14. Gould, S.J. (1989) *Wonderful Life. The Burgess Shale and the Nature of History*, Norton
15. Odling-Smee, F.J. *et al.* (2003) *Niche Construction: The Neglected Process in Evolution*, Princeton University Press
16. Sapp, J. (2009) *The New Foundations of Evolution: On the Tree of Life*, Oxford University Press
17. Walsh, D.M. (2015) *Organisms, Agency, and Evolution*, Cambridge University Press
18. West-Eberhard, M. (2003) *Developmental Plasticity and Evolution*, Oxford University Press
19. Barabási, A.L. (2016) *Network Science*, Cambridge University Press
20. Newman, M.E.J. (2018) *Networks* (2nd edn), Oxford University Press
21. Strogatz, S.H. (2001) Exploring complex networks. *Nature* 410, 268–276

Outstanding Questions

Could inclusive analyses of the sources of variations be simultaneously performed for all microbial genomes and all genomes of mobile elements to quantify the proportion of laterally acquired genes in pangenomes, not only for related cells, but also for related mobile elements, such as viruses and plasmids, and to classify hybrid mobile elements using multipartite networks?

How can we use networks to model the heredity of complex (highly polyphyletic) units of selection, such as holobionts and multispecies biofilms? Can we generate novel network tools depicting their transmission channels and classify these entities based on the resulting topologies?

How can directionality in time be implemented in network analysis of heredity? Is it possible to go beyond the comparison of the successive states of a network in a time series?

How can networks be used to model fitness variations and the stability of complex ecological networks of interactions based on -omics data from microbial communities?

Some adaptive traits result from transient interactions between entities and processes, operating at different time scales and levels of organization. Given this complex situation, how can multilevel temporal networks be developed to keep track of such adaptations?

How can the reformulation of ENS provided in this article (Box 1) apply to the evolution of entities such as ecosystems and Gaia?

22. Alvarez-Ponce, D. *et al.* (2013) Gene similarity networks provide tools for understanding eukaryote origins and evolution. *Proc. Natl. Acad. Sci. U. S. A.* 110, E1594–E1603
23. Alvarez-Ponce, D. *et al.* (2017) Position matters: network centrality considerably impacts rates of protein evolution in the human protein–protein interaction network. *Genome Biol. Evol.* 9, 1742–1756
24. Bapteste, E. and Huneman, P. (2018) Towards a dynamic interaction network of life to unify and expand the evolutionary theory. *BMC Biol.* Published online May 29, 2018. https://doi.org/10.1186/s12915-018-0531-6
25. Cancherini, D.V. *et al.* (2010) The role of exon shuffling in shaping protein–protein interaction networks. *BMC Genomics* 11, S11
26. Chaffron, S. *et al.* (2010) A global network of coexisting microbes from environmental and whole-genome sequence data. *Genome Res.* 20, 947–959
27. Conant, G.C. and Wolfe, K.H. (2006) Functional partitioning of yeast co-expression networks after genome duplication. *PLoS Biol.* 4, e109
28. Corel, E. *et al.* (2016) Network-thinking: graphs to analyze microbial complexity and evolution. *Trends Microbiol.* 24, 224–237
29. Corel, E. *et al.* (2018) Bipartite network analysis of gene sharings in the microbial world. *Mol. Biol. Evol.* 35, 899–913
30. Faria, J.P. *et al.* (2014) Genome-scale bacterial transcriptional regulatory networks: reconstruction and integrated analysis with metabolic models. *Brief. Bioinform.* 15, 592–611
31. Faust, K. and Raes, J. (2012) Microbial interactions: from networks to models. *Nat. Rev. Microbiol.* 10, 538–550
32. Gouy, A. *et al.* (2017) Detecting gene subnetworks under selection in biological pathways. *Nucleic Acids Res.* 45, e149
33. Halary, S. *et al.* (2010) Network analyses structure genetic diversity in independent genetic worlds. *Proc. Natl. Acad. Sci. U. S. A.* 107, 127–132
34. Halfon, M.S. (2017) Perspectives on gene regulatory network evolution. *Trends Genet.* 33, 436–447
35. Imbeault, M. *et al.* (2017) KRAB zinc-finger proteins contribute to the evolution of gene regulatory networks. *Nature* 543, 550–554
36. Koch, C. *et al.* (2017) Inference and evolutionary analysis of genome-scale regulatory networks in large phylogenies. *Cell Syst.* 4, 543–558.e8
37. Layeghifard, M. *et al.* (2017) Disentangling interactions in the microbiome: a network perspective. *Trends Microbiol.* 25, 217–228
38. Marbach, D. *et al.* (2012) Wisdom of crowds for robust gene network inference. *Nat. Methods* 9, 796–804
39. Martinez-Pastor, M. *et al.* (2017) Transcriptional regulation in Archaea: from individual genes to global regulatory networks. *Annu. Rev. Genet.* 51, 143–170
40. Mateos, J.L. *et al.* (2017) Divergence of regulatory networks governed by the orthologous transcription factors FLC and PEP1 in Brassicaceae species. *Proc. Natl. Acad. Sci. U. S. A.* 114, E11037–E11046
41. Meheust, R. *et al.* (2016) Protein networks identify novel symbiogenetic genes resulting from plastid endosymbiosis. *Proc. Natl. Acad. Sci. U. S. A.* 113, 3579–3584
42. Qin, H. *et al.* (2003) Evolution of the yeast protein interaction network. *Proc. Natl. Acad. Sci. U. S. A.* 100, 12820–12824
43. Raymond, J. and Segrè, D. (2006) The effect of oxygen on biochemical networks and the evolution of complex life. *Science* 311, 1764–1767
44. Ruprecht, C. *et al.* (2017) Phylogenomic analysis of gene co-expression networks reveals the evolution of functional modules. *Plant J. Cell Mol. Biol.* 90, 447–465
45. Ruprecht, C. *et al.* (2017) Beyond genomics: studying evolution with gene coexpression networks. *Trends Plant Sci.* 22, 298–307
46. Seshasayee, A.S.N. *et al.* (2006) Transcriptional regulatory networks in bacteria: from input signals to output responses. *Curr. Opin. Microbiol.* 9, 511–519
47. Shen-Orr, S.S. *et al.* (2002) Network motifs in the transcriptional regulation network of *Escherichia coli*. *Nat. Genet.* 31, 64–68
48. Stuart, J.M. *et al.* (2003) A gene-coexpression network for global discovery of conserved genetic modules. *Science* 302, 249–255
49. Szklarczyk, D. *et al.* (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452
50. Tamames, J. *et al.* (2007) Modular organization in the reductive evolution of protein–protein interaction networks. *Genome Biol.* 8, R94
51. Thébault, E. and Fontaine, C. (2010) Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science* 329, 853–856
52. de Souza, S.J. (2012) Domain shuffling and the increasing complexity of biological networks. *BioEssays* 34, 655–657
53. Pasek, S. *et al.* (2006) Gene fusion/fission is a major contributor to evolution of multi-domain bacterial proteins. *Bioinformatics* 22, 1418–1423
54. Patthy, L. (2003) Modular assembly of genes and the evolution of new functions. *Genetica* 118, 217–231
55. Wang, M. and Caetano-Anolles, G. (2009) The evolutionary mechanics of domain organization in proteomes and the rise of modularity in the protein world. *Structure* 17, 66–78
56. Doolittle, W.F. (1999) Phylogenetic classification and the universal tree. *Science* 284, 2124–2128
57. Nelson-Sathi, S. *et al.* (2012) Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of Haloarchaea. *Proc. Natl. Acad. Sci. U. S. A.* 109, 20537–20542
58. Zhaxybayeva, O. *et al.* (2006) Phylogenetic analyses of cyanobacterial genomes: quantification of horizontal gene transfer events. *Genome Res.* 16, 1099–1108
59. Skippington, E. and Ragan, M.A. (2011) Lateral genetic transfer and the construction of genetic exchange communities. *FEMS Microbiol. Rev.* 35, 707–735
60. Lopez-Garcia, P. *et al.* (2015) Bacterial gene import and mesophilic adaptation in archaea. *Nat. Rev. Microbiol.* 13, 447–456
61. Dagan, T. and Martin, W. (2007) Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proc. Natl. Acad. Sci. U. S. A.* 104, 870–875
62. Halary, S. *et al.* (2013) EGN: a wizard for construction of gene and genome similarity networks. *BMC Evol. Biol.* 13, 146
63. Kloesges, T. *et al.* (2011) Networks of gene sharing among 329 proteobacterial genomes reveal differences in lateral gene transfer frequency at different phylogenetic depths. *Mol. Biol. Evol.* 28, 1057–1074
64. Lima-Mendez, G. *et al.* (2008) Reticulate representation of evolutionary and functional relationships between phage genomes. *Mol. Biol. Evol.* 25, 762–777
65. Popa, O. *et al.* (2011) Directed networks reveal genomic barriers and DNA repair bypasses to lateral gene transfer among prokaryotes. *Genome Res.* 21, 599–609
66. Yutin, N. *et al.* (2013) Virophages, polintons, and transpovirons: a complex evolutionary network of diverse selfish genetic elements with different reproduction strategies. *Virol. J.* 10, 158
67. Iranzo, J. *et al.* (2016) The double-stranded DNA virosphere as a modular hierarchical network of gene sharing. *mBio* 7, e00978-16
68. McInerney, J.O. *et al.* (2017) Why prokaryotes have pangenomes. *Nat. Microbiol.* 2, 17040
69. Tettelin, H. *et al.* (2005) Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial 'pan-genome. *Proc. Natl. Acad. Sci. U. S. A.* 102, 13950–13955
70. Nakamura, Y. *et al.* (2004) Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat. Genet.* 36, 760–766
71. Martin, W.F. *et al.* (2015) Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 370, 20140330
72. Archibald, J.M. (2015) Genomic perspectives on the birth and spread of plastids. *Proc. Natl. Acad. Sci. U. S. A.* 112, 10147–10153

73. Lane, C.E. and Archibald, J.M. (2008) The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol. Evol.* 23, 268–275
74. Nowack, E.C. and Grossman, A.R. (2012) Trafficking of protein into the recently established photosynthetic organelles of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. U. S. A.* 109, 5340–5345
75. Booth, A. *et al.* (2016) The modern synthesis in the light of microbial genomics. *Annu. Rev. Microbiol.* 70, 279–297
76. Gavelis, G.S. *et al.* (2015) Eye-like ocelloids are built from different endosymbiotically acquired components. *Nature* 523, 204–207
77. Dorrell, R.G. *et al.* (2017) Chimeric origins of ochrophytes and haptophytes revealed through an ancient plastid proteome. *eLife* 6, e23717
78. Meheust, R. *et al.* (2018) Hundreds of novel composite genes and chimeric genes with bacterial origins contributed to haloarchaeal evolution. *Genome Biol.* 19, 75
79. Borges, R.M. (2017) Co-niche construction between hosts and symbionts: ideas and evidence. *J. Genet.* 96, 483–489
80. Chiu, L. and Gilbert, S.F. (2015) The birth of the holobiont: multi-species birthing through mutual scaffolding and niche construction. *Biosemiotics* 8, 191–210
81. McFall-Ngai, M. *et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U. S. A.* 110, 3229–3236
82. Moran, N.A. and Sloan, D.B. (2015) The Hologenome Concept: helpful or hollow? *PLoS Biol.* 13, e1002311
83. Douglas, A.E. and Werren, J.H. (2016) Holes in the hologenome: why host–microbe symbioses are not holobionts. *mBio* 7, e02099-15
84. Bordenstein, S.R. and Theis, K.R. (2015) Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol.* 13, e1002226
85. Ley, R.E. *et al.* (2008) Evolution of mammals and their gut microbes. *Science* 320, 1647–1651
86. Vigliotti, C. *et al.* (2018) Tracking the rules of transmission and introgression with networks. *Microbiol. Spectr.* Published online April 1, 2018. https://doi.org/10.1128/microbiolspec.MTBP-0008-2016
87. Doolittle, W.F. and Zhaxybayeva, O. (2010) Metagenomics and the units of biological organization. *Bioscience* 60, 102–112
88. Turnbaugh, P.J. *et al.* (2009) A core gut microbiome in obese and lean twins. *Nature* 457, 480–484
89. Koonin, E.V. (2014) Carl Woese's vision of cellular evolution and the domains of life. *RNA Biol.* 11, 197–204
90. Koonin, E.V. and Wolf, Y.I. (2009) Is evolution Darwinian or/and Lamarckian? *Biol. Direct* 4, 42
91. Osmanovic, D. *et al.* (2018) Darwinian selection of host and bacteria supports emergence of Lamarckian-like adaptation of the system as a whole. *Biol. Direct* 13, 24
92. Rosenberg, E. *et al.* (2009) The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework. *Environ. Microbiol.* 11, 2959–2962
93. Skinner, M.K. (2015) Environmental epigenetics and a unified theory of the molecular aspects of evolution: a neo-Lamarckian concept that facilitates neo-Darwinian evolution. *Genome Biol. Evol.* 7, 1296–1302
94. Popa, O. *et al.* (2017) Phylogenomic networks reveal limited phylogenetic range of lateral gene transfer by transduction. *ISME J.* 11, 543–554
95. Sicard, A. *et al.* (2019) A multicellular way of life for a multipartite virus. *eLife* 8, e43599
96. Lamm, E. (2018) Inheritance systems. In *Stanford Encyclopedia of Philosophy* (Zalta, E.N., ed.), Metaphysics Research Laboratory, Stanford University
97. van Opstal, E.J. and Bordenstein, S.R. (2015) Microbiome. Rethinking heritability of the microbiome. *Science* 349, 1172–1173
98. Campos, M. *et al.* (2015) A membrane computing simulator of trans-hierarchical antibiotic resistance evolution dynamics in nested ecological compartments (ARES). *Biol. Direct* 10, 41
99. Dordet-Frisoni, E. *et al.* (2014) Chromosomal transfers in mycoplasmas: when minimal genomes go mobile. *mBio* 5, e01958-14
100. Dubey, G.P. and Ben-Yehuda, S. (2011) Intercellular nanotubes mediate bacterial communication. *Cell* 144, 590–600
101. Baidya, A.K. *et al.* (2018) Bacterial nanotubes: a conduit for intercellular molecular trade. *Curr. Opin. Microbiol.* 42, 1–6
102. Fernandez-Lopez, R. *et al.* (2017) Towards a taxonomy of conjugative plasmids. *Curr. Opin. Microbiol.* 38, 106–113
103. Johnson, C.M. and Grossman, A.D. (2015) Integrative and conjugative elements (ICEs): what they do and how they work. *Annu. Rev. Genet.* 49, 577–601
104. McDaniel, L.D. *et al.* (2010) High frequency of horizontal gene transfer in the oceans. *Science* 330, 50
105. Labrie, S.J. *et al.* (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8, 317–327
106. McMahon, S.A. *et al.* (2009) Extensive DNA mimicry by the ArdA anti-restriction protein and its role in the spread of antibiotic resistance. *Nucleic Acids Res.* 37, 4887–4897
107. Keen, E.C. *et al.* (2017) Novel 'superspreader' bacteriophages promote horizontal gene transfer by transformation. *mBio* 8, e02115–e02116
108. Erez, Z. *et al.* (2017) Communication between viruses guides lysis–lysogeny decisions. *Nature* 541, 488–493
109. Gilbert, S.F. *et al.* (2015) Eco-Evo-Devo: developmental symbiosis and developmental plasticity as evolutionary agents. *Nat. Rev. Genet.* 16, 611–622
110. Moran, N.A. *et al.* (2008) Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* 42, 165–190
111. Dillon, R.J. *et al.* (2005) Diversity of locust gut bacteria protects against pathogen invasion. *Ecol. Lett.* 8, 1291–1298
112. van Elsas, J.D. *et al.* (2012) Microbial diversity determines the invasion of soil by a bacterial pathogen. *Proc. Natl. Acad. Sci. U. S. A.* 109, 1159–1164
113. Wei, Z. *et al.* (2015) Trophic network architecture of root-associated bacterial communities determines pathogen invasion and plant health. *Nat. Commun.* 6, 8413
114. Jha, A.R. *et al.* (2018) Gut microbiome transition across a lifestyle gradient in Himalaya. *PLoS Biol.* 16, e2005396
115. Lozupone, C.A. *et al.* (2012) Diversity, stability and resilience of the human gut microbiota. *Nature* 489, 220–230
116. Venturelli, O.S. *et al.* (2018) Deciphering microbial interactions in synthetic human gut microbiome communities. *Mol. Syst. Biol.* 14, e8157
117. Farkas, T.E. *et al.* (2013) Evolution of camouflage drives rapid ecological change in an insect community. *Curr. Biol.* 23, 1835–1843
118. Lallensack, R. (2018) How warp-speed evolution is transforming ecology. *Nature* 554, 19–21
119. Fragata, I. *et al.* (2019) Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* 34, 69–82
120. Lima-Mendez, G. *et al.* (2015) Ocean plankton. Determinants of community structure in the global plankton interactome. *Science* 348, 1262073
121. Castelle, C.J. and Banfield, J.F. (2018) Major new microbial groups expand diversity and alter our understanding of the tree of life. *Cell* 172, 1181–1197
122. Castelle, C.J. *et al.* (2018) Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat. Rev. Microbiol.* 16, 629–645
123. Wanner, G. *et al.* (2008) Ultrastructural characterization of the prokaryotic symbiosis in 'Chlorochromatium aggregatum'. *J. Bacteriol.* 190, 3721–3730
124. Doolittle, W.F. and Inkpen, S.A. (2018) Processes and patterns of interaction as units of selection: An introduction to ITSNTS thinking. *Proc. Natl. Acad. Sci. U. S. A.* 115, 4006–4014
125. DeLong, E.F. (2007) Life on the thermodynamic edge. *Science* 317, 327–328
126. Doolittle, W.F. *et al.* (2011) Comment on 'Does constructive neutral evolution play an important role in the origin of cellular complexity? *BioEssays* 33, 427–429
127. Gray, M.W. *et al.* (2010) Cell biology. Irremediable complexity? *Science* 330, 920–921
128. Lukes, J. *et al.* (2011) How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life* 63, 528–537
129. Stoltzfus, A. (2012) Constructive neutral evolution: exploring evolutionary theory's curious disconnect. *Biol. Direct* 7, 35
130. Morris, J.J. *et al.* (2012) The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *mBio* 3, e00036-12
131. Selosse, M.A. *et al.* (2014) Microbial priming of plant and animal immunity: symbionts as developmental signals. *Trends Microbiol.* 22, 607–613