



Evolution of microbes and viruses: a paradigm shift in evolutionary biology?

Eugene V. Koonin* and Yuri I. Wolf

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

Edited by:

Didier Raoult, Université de la Méditerranée, France

Reviewed by:

Kalliopi Georgiades, Unité de Recherche en maladies infectieuses tropicales émergentes, France
L. Aravind, National Institutes of Health, USA

*Correspondence:

Eugene V. Koonin, National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, NIH Bldg. 38a, Bethesda 20894, MD, USA.
e-mail: koonin@ncbi.nlm.nih.gov

When Charles Darwin formulated the central principles of evolutionary biology in the *Origin of Species* in 1859 and the architects of the Modern Synthesis integrated these principles with population genetics almost a century later, the principal if not the sole objects of evolutionary biology were multicellular eukaryotes, primarily animals and plants. Before the advent of efficient gene sequencing, all attempts to extend evolutionary studies to bacteria have been futile. Sequencing of the rRNA genes in thousands of microbes allowed the construction of the three-domain “ribosomal Tree of Life” that was widely thought to have resolved the evolutionary relationships between the cellular life forms. However, subsequent massive sequencing of numerous, complete microbial genomes revealed novel evolutionary phenomena, the most fundamental of these being: (1) pervasive horizontal gene transfer (HGT), in large part mediated by viruses and plasmids, that shapes the genomes of archaea and bacteria and call for a radical revision (if not abandonment) of the Tree of Life concept, (2) Lamarckian-type inheritance that appears to be critical for antiviral defense and other forms of adaptation in prokaryotes, and (3) evolution of evolvability, i.e., dedicated mechanisms for evolution such as vehicles for HGT and stress-induced mutagenesis systems. In the non-cellular part of the microbial world, phylogenomics and metagenomics of viruses and related selfish genetic elements revealed enormous genetic and molecular diversity and extremely high abundance of viruses that come across as the dominant biological entities on earth. Furthermore, the perennial arms race between viruses and their hosts is one of the defining factors of evolution. Thus, microbial phylogenomics adds new dimensions to the fundamental picture of evolution even as the principle of descent with modification discovered by Darwin and the laws of population genetics remain at the core of evolutionary biology.

Keywords: Darwin, modern synthesis, comparative genomics, tree of life, horizontal gene transfer

INTRODUCTION

Charles Darwin’s *On the Origin of Species* that appeared in London in 1859 (Darwin, 1859) was the first plausible, detailed account of biological evolution, after the simultaneous and independent brief outlines by Darwin and Alfred Russell Wallace that were published the previous year (Darwin, 1858; Wallace, 1858). Darwin did not discover evolution and did not even offer the first coherent description of evolution: exactly 50 years before the appearance of the *Origin*, the French botanist and zoologist Jean-Baptiste Lamarck published his magnum opus *Philosophie Zoologique* (Lamarck, 1809) in which he outlined his vision of the history of life in considerable detail. However, the cornerstone of Lamarck’s worldview was the purported intrinsic drive of evolving organisms toward “perfection,” a patently non-scientific, irrational idea. Moreover, Lamarck’s view of the role of evolution in the history of life was severely limited: he did not postulate deep common ancestry of life forms but rather believed in multiple acts of creation, perhaps a separate act for each species. Prescient ideas on evolutionary changes of organisms actually have been developed centuries before Lamarck and Darwin, most notably by the great Roman thinker Titus Lucretius Carus (2011).

However, the fact remains that it was Darwin’s first evolutionary synthesis that had launched the field of evolutionary biology in a sense close to the modern one and had remained central to biological thinking over the last 150 years inasmuch as “nothing in biology makes sense except in the light of evolution” (Dobzhansky, 1973). Darwin’s concept lacked the essential foundation in genetics for the obvious reason that mechanisms of heredity were unknown in his day. Hence Darwin’s deep concern over the so-called Jenkin nightmare, the objection to Darwin’s concept according to which beneficial changes would be “diluted” after several generations in the progeny of organisms in which they occurred. The genetic basis of evolution was established after the rediscovery of Mendel’s laws, with the development of population genetics in the first third of the twentieth century, primarily, through the pioneering work of Fisher, Wright, and Haldane (Fisher, 1930; Haldane, 1932). The new, advanced understanding of evolution, informed by theoretical and experimental work in genetics, was consolidated in the Modern Synthesis of evolutionary biology, usually, associated with the names of Dobzhansky, Julius Huxley, Mayr, and Simpson (Dobzhansky, 1937; Simpson, 1944). Apparently, the Modern Synthesis reached its mature form

during the 1959 centennial celebration for the *Origin* in Chicago (Tax and Callender, 1960; Browne, 2008).

Now, 50 years after the consolidation of the Modern Synthesis, evolutionary biology undoubtedly faces a new major challenge and, at the same time, the prospect of a new conceptual breakthrough (Rose and Oakley, 2007). If the Modern Synthesis can be succinctly described as Darwinism in the Light of Genetics (often referred to as neodarwinism), then, the new stage is Evolutionary Biology in the Light of Genomics and Microbiology. The combination of genomics and microbiology is indeed critical in the advent of this new age of evolutionary biology (Koonin and Wolf, 2008; Koonin, 2009a; Woese and Goldenfeld, 2009). Lamarck and Darwin (let alone Lucretius) were plainly unaware of the existence of genomes and microbes. The architects of the Modern Synthesis certainly knew about genomes and microbes “in principle” but, in the former case, did not know enough to incorporate information on genomes beyond the (important but limited) level of formal genetics, and in the latter case, did not realize the importance of microbes for understanding evolution at all.

In this article, we attempt to outline the key changes to the basic tenets of evolutionary biology brought about primarily by comparative and functional microbial genomics and argue that, in many respects, the genomic stage could be a more radical departure from the Modern Synthesis than the latter was from classic Darwinian concepts.

FROM THE TREE OF LIFE TO THE WEB OF GENE TREES

The famous sole illustration of the *Origin of Species* shows a Tree of Life (or more precisely, a series of trees presumably depicting the evolution of different divisions of organisms). Obviously, Darwin was not the first to use a tree to depict history. Before him, trees had been employed for many centuries to capture human genealogy, e.g., that of the Old Testament patriarchs as well as later monarchs. Darwin, however, was the first to make the crucial conceptual step by boldly proposing that the entire history of life could (at least in principle) be accurately represented by a tree growing from a single root. Darwin’s tree was a sheer scheme, without any attempt to assign real life forms to the branches but in just a few years Ernst Haeckel populated the tree by a huge variety of organisms, almost exclusively animals (Haeckel, 1997). Haeckel inferred the relationships between organisms reflected in the topology of his tree primarily on the data of comparative anatomy that was already advanced in his day. Over the next century, there was considerable progress in this field leading to improved resolution of the tree but qualitatively the situation has not changed. Phylogeny largely served as a tool for systematics, and the architects of the Modern Synthesis were much more interested in mechanisms of microevolution and speciation than in the course of macroevolution that is supposedly reflected in the Tree of Life. Although by mid-twentieth century microbiologists had realized full well that microbes possess genomes and can mutate, and accordingly, should evolve, in principle, similarly to animals and plants, all attempts to infer microbial evolution from morphological and physiological characters had been unqualified failures (Stanier and Van Niel, 1962).

The fortunes of phylogeny and microbial evolution changed abruptly in the late 1970s when Carl Woese and colleagues

realized that the nucleotide sequence of a universally conserved molecule, 16S rRNA, could be used to infer a universal phylogenetic tree (rather incredibly, from today’s vantage point, Woese’s original seminal work employed oligonucleotide maps of 16S RNA rather than sequences; however, the actual sequences became readily available shortly, and the main conclusions of the early studies stood) (Woese, 1987). Comparison of 16S RNA sequences had swiftly led to the discovery of a distinct domain of life, the Archaea, and its distinct phylogenetic affinity with the eukaryotes (Woese and Fox, 1977; Woese et al., 1990; Woese, 2004). Over the following few years, major phyla of Bacteria, Archaea and unicellular eukaryotes have been established (Woese, 1987), and the famous tripartite tree (**Figure 1**) emerged as the paradigm of the history of cellular life on earth which it more or less remains to this day (Woese et al., 1990; Pace, 1997, 2006, 2009). This was a veritable triumph of molecular phylogenetics and a dramatic departure from Haeckel’s Tree of Life. In Haeckel’s tree, Protista (unicellular eukaryotes) and Monera (bacteria) occupied unspecified positions near the root. For all purposes, these measly, tiny creatures were not considered important in the big picture of evolution. The tripartite tree of Woese and colleagues was a complete change of perspective. Now, two of the three domains of life were represented by prokaryotes (former Monera), and within the eukaryote domain, the majority of the phyla were represented by unicellular organisms (former Protista). The life forms formerly considered “important,” i.e., the complex multicellular organisms (animals and plants), represent only two among the numerous branches of eukaryotes. There is no denying the fact that the true biodiversity on this planet is the diversity of unicellular microbes.

In the 1980s, when the paradigmatic status of the three-domain Tree of Life was established, there was little concern over the fact that technically this tree represented the history of only one gene, even if a universally present and highly conserved one. The 16S RNA was unanimously considered a suitable reference gene to represent the evolution of the respective organisms. Other universal genes, such as 18S RNA ribosomal proteins or RNA polymerase subunits, were thought to be important only to the extent their inclusion could improve the resolution of phylogenetic trees.

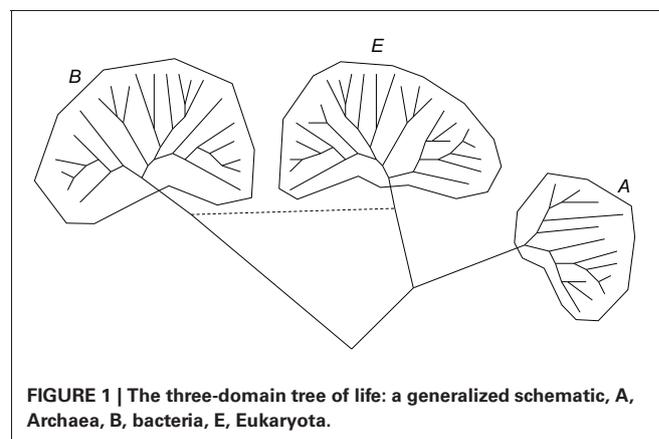


FIGURE 1 | The three-domain tree of life: a generalized schematic, A, Archaea, B, bacteria, E, Eukaryota.

Even long before the advent of the genomic era, microbiologists realized that bacteria had the capacity to exchange genetic information via horizontal gene transfer (HGT), in some cases, producing outcomes of major importance, such as antibiotic resistance (Syvanen and Kado, 2002). Multiple molecular mechanisms of HGT have been described including plasmid exchange, transduction (HGT mediated by bacteriophages), and transformation (Bushman, 2001) [indeed, the phenomenon of transformation was employed by Avery and colleagues to demonstrate the genetic function of DNA in 1944 (Avery et al., 1944a)]. However, despite these discoveries, HGT was generally viewed as a minor phenomenon that was important only under special circumstances and, in any case, did not in any manner jeopardize the Tree of Life that could be reconstructed by phylogenetic analysis of rRNA and other conserved genes.

This comfortable belief was abruptly shattered when the early findings of comparative genomics of bacteria and archaea in the late 1990s have indicated that, at least in some prokaryotic genomes, a substantial fraction of genes were acquired via demonstrable HGT, sometimes across log evolutionary distances. The pathogenicity islands and similar symbiosis islands that comprise over 30% of the genome in many pathogenic and symbiotic bacteria and obviously travel between bacteria via HGT are the prime case in point (Hacker and Kaper, 2000; Perna et al., 2001). Perhaps, more strikingly, comparative analysis of the genomes of hyperthermophilic bacteria and archaea has suggested that in shared habitats even HGT between the two domains of prokaryotes, Archaea and bacteria, can be extensive, with up to 20% of the genes of bacterial hyperthermophiles showing archaeal affinity (Aravind et al., 1998; Nelson et al., 1999; Koonin et al., 2001). Subsequent phylogenomic studies (that is analysis of phylogenies of multiple genes from numerous genomes) have led to a shocking realization: in prokaryotes at least, there seem not to exist two genes with the exact same evolutionary history (Koonin et al., 2001; Gogarten and Townsend, 2005; Gribaldo and Brochier, 2009; Zhaxybayeva, 2009; Boto, 2010; Andam and Gogarten, 2011; Zhaxybayeva and Doolittle, 2011). Apparently, this is so because all genes have experienced HGT at some stage (s) of their evolution. Although some genes, in particular those that encode components of the translation system, show substantial congruency (but not actual identity) between each other and with the standard rRNA tree, the number of such congruent trees is small. In a memorable phrase of Bill Martin and Tal Dagan, the ribosomal tree of a life is at best “a tree of one percent” (of all genes in microbial genomes) (Dagan and Martin, 2006).

Thus, “evolution of prokaryotes and the Tree of Life are two different things” (Baptiste et al., 2009; Martin, 2011). Then, the question arises: is there any substantial tree component in evolution at all and accordingly does it make any sense to speak of HGT? Indeed, horizontal transfer can be defined as such only against some standard of vertical evolution (Baptiste et al., 2005; Doolittle and Baptiste, 2007; Baptiste and Boucher, 2009). As Martin and Dagan wryly notice, if a model (in this case, the Tree of Life model) adequately describes 1% of the data, it might be advisable to abandon it and search for a better one (Dagan and Martin, 2006). Such an alternative indeed has been proposed in the form of a dynamic network of microbial evolution in which

the nodes are bacterial and archaeal genomes, and the edges are the fluxes of genetic information between the genomes (Koonin et al., 2005; Dagan and Martin, 2009; Dagan, 2011; Kloesges et al., 2011). In the extreme, such a network has no vertical, tree-like component whereas the weights of the edges differ depending on the intensity of the gene exchange (Figure 2). Moreover, it has been persuasively argued that “tree thinking in biology” might be a sheer myth, however deeply entrenched in the textbooks and the minds of biologists (Baptiste et al., 2005; Doolittle and Baptiste, 2007; Baptiste and Boucher, 2009). Indeed, there is potential for tree-like patterns to emerge from relationships that have nothing to do with common descent as exemplified by Doolittle and Baptiste by the distribution of human names across the departments of France (Doolittle and Baptiste, 2007).

One could argue, however, that the tree pattern is not at all illusory but, on the contrary, is intrinsic and central to the entire process of biological evolution. The relevance and generality of this pattern plainly follows from the fundamental character of the replication process that underlies the evolution of life (Koonin and Wolf, 2009b). Successive generations of replicating genomes (and accordingly, dividing cells) follows an inherently binary branching pattern that, over generation naturally yields a tree. The tree pattern is predicated on a low rate of intra-genic recombination which is indeed the case for all evolutionary distances large enough to prevent homologous recombination. Accordingly, evolutionary history of individual genes can be adequately represented by trees (the practical problems of accurate phylogeny reconstruction notwithstanding).

A natural, key question to ask then is: are the topologies of the trees for individual genes substantially congruent? In other words, is it possible to identify a statistically significant central trend in the vast “forest” of gene trees? Statistical analysis of thousands of phylogenetic trees for diverse genes of prokaryotes (in fact, all genes with sufficient degree of conservation to obtain a reliable tree topology) has shown that a highly significant central trend is indeed detectable in the phylogenetic forest (Puigbo et al., 2009, 2012; Koonin et al., 2011). Moreover, the consensus topology of the supertree of the (nearly) universal genes (the notorious

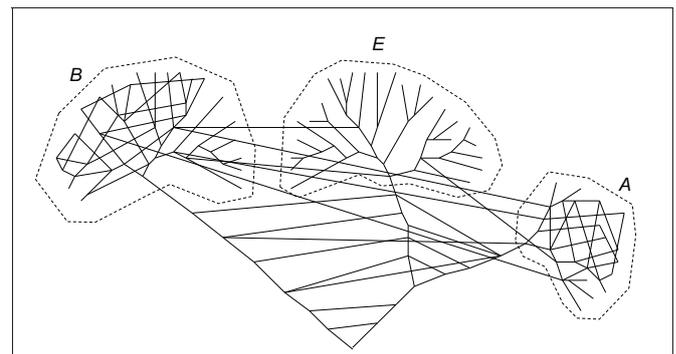


FIGURE 2 | A network representation of the evolutionary process. The network still includes some tree components such that the three domains of cellular life remains distinct but there is also an extensive horizontal component of genetic information flow that in particular dominates the earliest stages of evolution (Koonin and Wolf, 2008).

1%) turned out to be the best approximation of that central trend. Thus, although any phylogenetic tree of a central, conserved component of the cellular information-processing machinery (such as rRNA or the set of universal ribosomal proteins) represents only a minority of the phylogenetic signal across the phylogenetic forest (see details below) and so by no account can be considered an all-encompassing “Tree of Life,” neither is such a phylogeny an arbitrary and irrelevant “tree of 1%.” On the contrary, these trees represent a central evolutionary trend and reflect a “statistical tree of life” (O’Malley and Koonin, 2011).

THE DYNAMIC GENE UNIVERSE

For decades microbiologists knew that bacteria sometimes exchange genes (Low and Porter, 1978; Arber, 1979; Campbell, 1981; Syvanen, 1985, 1994). Moreover, the phenomena of transformation, acquisition of new traits via import of DNA from the environment and integration of the imported molecules into the bacterial genome, and transduction, transfer of genetic markers by bacteriophages, have been studied in considerable detail. In fact, transformation was the basis of the seminal 1944 experiments of Avery and colleagues which demonstrated that the genetic material of bacteria consisted of DNA (Avery et al., 1944b). In addition, microbiologists realized that such HGT could exert well-defined, major biological effects such as conferring pathogenicity (as in Avery’s experiments) or antibiotic resistance on the recipients of horizontally transferred genes. However, all this knowledge notwithstanding, in the pregenomic era, HGT was considered a highly specialized genetic pathway rather than the mainstream of microbial evolution.

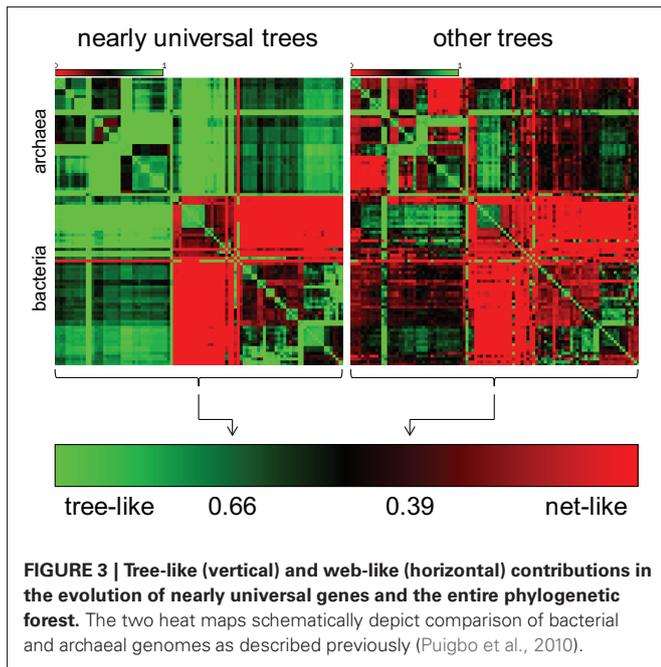
Comparative genomics brought the shocking realization that bacterial and archaeal genomes were literally shaped by HGT. This was clearly demonstrated by early analyses of the genomes of bacterial hyperthermophiles that were shown contain about 20% of genes of obvious archaeal origin (Aravind et al., 1998; Nelson et al., 1999; Koonin et al., 2001); conversely, genomes of mesophilic Archaea, such as *Methanosarcina*, encompass roughly the same proportion of genes clearly derived from bacteria (Deppenmeier et al., 2002; Galagan et al., 2002). These are striking examples of extensive gene exchange between the most distant prokaryotes that is stimulated by cohabitation. Not unexpectedly, the extent of gene exchange is far greater between more closely related organisms, even if often more difficult to detect (Abby et al., 2012). Nevertheless, phylogenomic analysis of a variety of bacteria and archaea clearly reveals their mosaic origins: different genes affiliate with homologs from different organisms (Koonin et al., 2001; Sicheritz-Ponten and Andersson, 2001; Koonin, 2003; Esser et al., 2007; Koonin and Wolf, 2008; Kloesges et al., 2011). These findings have been encapsulated in the concept of the Rhizome of Life under which the history of any given genome can be represented as a rhizome, with diverse sources and evolutionary histories for different genes (Raoult, 2010; Merhej et al., 2011). Recent, detailed studies indicate that at least in tight microbial communities, such as for instance the human gut microbiota, gene exchange is constant and rampant (Smillie et al., 2011).

In the face of the increasingly apparent genomic promiscuity, one cannot help asking whether “horizontal gene transfer” is a viable concept at all: indeed, for any extended span of evolution,

HGT will be identifiable if and only if there is some objectively definable “vertical” standard to compare against. Otherwise, all genetic exchanges would be equal, and the only adequate depiction of evolution would be an undirected network graph. Thus, the validity of the tree representation of evolution and the very existence of HGT are inextricably linked. The results of exhaustive comparison of the individual gene trees in the “phylogenetic forest” discussed in the preceding section reveal the existence of substantial coherence of phylogenetic tree topologies, especially among highly conserved, (nearly) ubiquitous genes that encode components of the translation system (Puigbo et al., 2009). There are many exceptions to this generalization including extensive HGT of genes coding for aminoacyl-tRNA synthetases (Wolf et al., 1999; Woese et al., 2000) and even multiple cases of HGT of genes encoding ribosomal proteins (Brochier et al., 2000; Makarova et al., 2001; Yutin et al., 2012). Nevertheless, these genes appear to comprise a single, co-evolving ensemble, in at least general agreement with the so-called complexity hypothesis (Jain et al., 1999; Wellner et al., 2007; Abby et al., 2012). Under the complexity hypothesis, HGT of genes encoding subunits of macromolecular complexes is largely suppressed because of the deleterious effect caused by disruption of interactions refined by a long time of co-evolution. Indeed, a recent analysis has shown that it is the involvement in complex formation that shows a strong negative correlation with the rate of HGT, rather than any specific biological function (Cohen et al., 2011). Thus, genes encoding many translation system components probably co-evolve and accordingly are rarely horizontally transferred because they are preferentially involved in large complexes (above all, the ribosome itself) rather than owing to their special biological importance or any other peculiarities of their biological function. Other genes show a much weaker but also significant phylogenetic coherence with the nearly universal genes for translation system components, perhaps also reflecting the involvement in complex formation.

The same series of phylogenomic studies that demonstrated the validity of the statistical tree of life quantified the contributions of tree-like (vertical) and web-like (horizontal) gene transmission to the relationships between bacterial and archaeal genomes (Puigbo et al., 2010, 2012). The results came out remarkably different for the ~100 nearly universal trees and the rest of the trees in the phylogenetic forest. The evolution of the nearly universal trees is dominated by the tree-like trend which contributes approximately 2/3 of the evolutionary information whereas in the rest of the forest, the ratio is the opposite, with about 2/3 of the signal coming from horizontal gene exchange (Figure 3).

The extensive HGT that permeates the prokaryote world is the source of gene gain by bacterial and archaeal genomes. Perhaps, the best characterized case of massive gene gain is the emergence of pathogenic bacterial strains that often evolve by acquiring the so-called pathogenicity islands that sometimes comprise over 30% of the pathogen’s genome as first revealed by the comparison of the genomes of laboratory and wild strains of *E. coli* (Perna et al., 2001; Zhang et al., 2007; Eppinger et al., 2011). The opposite trend, gene loss, is at least as prominent as gene gain via HGT (Snel et al., 2002; Mirkin et al., 2003). A prime example



is evolution of intracellular parasites and symbionts, for example, *Buchnera*, a close relative of *E. coli* that lost about 90% of the ancestral genes (Perez-Brocail et al., 2006); several other intracellular bacterial parasites and symbionts show even more drastic genome reduction (Klasson and Andersson, 2004; Perez-Brocail et al., 2006; McCutcheon and Moran, 2012).

The balance between gene gain and gene loss translates into a distinct shape of the distribution of gene occurrence in prokaryote pangenomes at all levels, from closely related bacteria (e.g., those of Enterobacteria) to the entirety of sequenced bacterial and archaeal genomes (Koonin and Wolf, 2008; O’Malley and Koonin, 2011). This universal distribution has an asymmetric U-shape and can be approximated by three exponential functions (Figure 4). The first of these corresponds to a small, highly

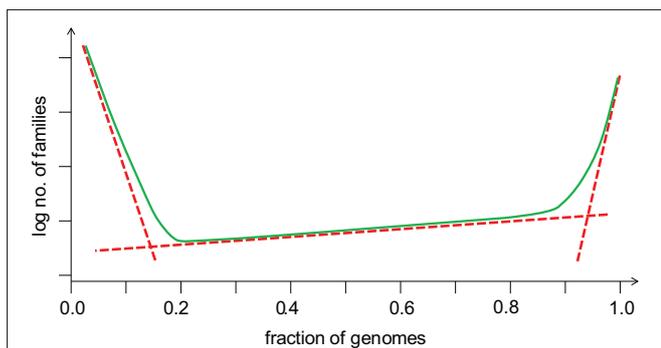


FIGURE 4 | The universal distribution of gene commonality in the microbial genomic universe: a generalized schematic. The three broken lines represent three exponential functions that fit the core (on the right), the shell (in the middle) and the cloud (on the left) of prokaryotic genes (O’Malley and Koonin, 2011).

conserved core (the nearly universal genes discussed above); the second exponent describes the much larger “shell” of genes with limited conservation; and the third one delineates the vast “cloud” of rare, poorly conserved genes. Thus, the gene universe is dominated by rare, sparsely distributed genes most of which are not covered by the limited available sampling of genomes and still remain to be discovered although in each particular genome the moderately conserved “shell” genes comprise the majority (Figure 5). The dynamic, fluid character of the prokaryote genomes yields a distinct, fractal-like structure of the gene universe (O’Malley and Koonin, 2011).

ARE THERE SPECIES IN PROKARYOTES?

The title of Darwin’s seminal book “The Origin of Species” is deeply steeped in traditions of eighteenth and nineteenth century biology that tended to view animal and plant species as key units of biological organization. Darwin himself actually saw species more as an arbitrary category in the continuum of varying life forms than a fundamental unit of life. In the twentieth century the species concept received its biological interpretation, primarily in the work of Ernst Mayr who famously defined a species as a system of panmictic populations that are genetically isolated from other such systems (Mayr, 1944). This concept indeed captures a key feature of the biology of organisms with regular, obligatory sexual reproduction such as, above all, animals and to a lesser extent plants.

Most of the prokaryotes do not engage in regular sex but instead exchange genes via HGT with diverse other microbes that they happen to cohabitate with. In general, in the prokaryote world, there are indeed no discrete, genetically isolated systems of panmictic populations but rather complex webs of gene exchange (Dagan et al., 2008; Koonin and Wolf, 2008). Thus, the very notion of species as a distinct biological category does not apply even though traditionally bacteria and archaea are still denoted by Linnaean species names (e.g., *Escherichia coli* or *Haloferax volcanii*) (Konstantinidis et al., 2006; Cohan and Perry, 2007; Doolittle and Zhaxybayeva, 2009; Fraser et al.,

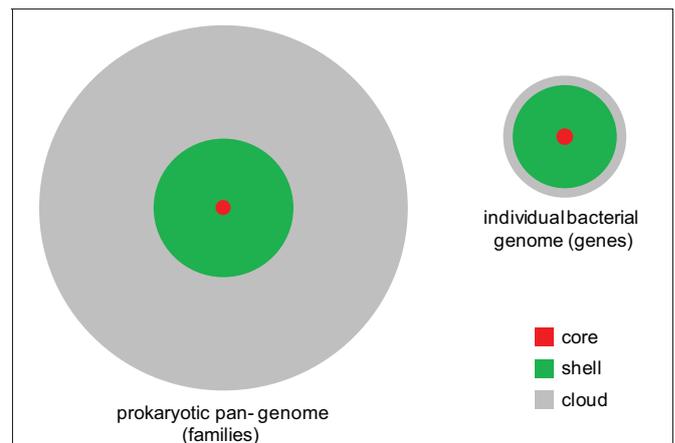


FIGURE 5 | The core, shell, and cloud of microbial genes. A generalized schematic showing the approximate contributions of the core, shell, and cloud to the pangenomes of prokaryotes and individual genomes.

2009). However, the modes of evolution substantially differ across the diversity of prokaryotes, spanning the entire continuum from fully sexual to fully clonal populations (Smith et al., 1993; Doolittle and Zhaxybayeva, 2009). Some bacteria, especially parasites such as for example *Neisseria gonorrhoeae*, have been shown to form largely isolated communities that engage in regular conjugation, the bacterial equivalent of sex, resulting in extensive homologous recombination. For these distinct organisms but not for the majority of bacteria and archaea, Mayr's biological definition of species might be a relevant concept.

The irrelevance of the (traditional) species concept for most prokaryotes by no means implies non-existence of structure in the genome space. Indeed, bacteria and archaea that share common origin in phylogenetic trees of marker genes, such as rRNA, typically also possess similar gene content. The "genome-trees" constructed on the basis of the (dis)similarity of gene content are generally congruent with phylogenetic trees of highly conserved marker genes although interesting deviations that reflect similarities in life style and/or extensive gene exchange have been detected as well (Snel et al., 1999, 2005; Wolf et al., 2002).

Thus, although the bacterial and archaeal "species" are not species in the regular sense, they are "galaxies" in the gene universe that form distinct, hierarchical clusters. Interestingly, it has been shown that, among the processes that lead to the divergence of gene content between evolving lineages of prokaryotes, gene loss appears to occur stochastically and generally follows the divergence of marker genes whereas gene gain (primarily, via HGT) is more episodic (Snel et al., 2002; Novichkov et al., 2004).

DOES EVOLUTION ADVANCE COMPLEXITY?

The idea of a general evolutionary trend toward increasing complexity is extremely popular among both lay public and scientists and certainly was shared by Darwin who wrote, for example, in famous quote: "as natural selection works solely by and for the good of each being, all corporeal and mental endowments will tend to progress toward perfection" (Darwin, 1859). This view does not imply any mysterious strive for perfection as imagined by some pre-Darwinian biologists including Lamarck (1809) or teleology of any kind. Nevertheless, Darwin's position does suggest a trend of evolution from simple to complex forms which is indeed a highly intuitive notion that has some obvious support in well known facts of the history of life on earth. For example, the most organizationally complex organisms with the largest genomes, animals, and plants, appear only at relatively late stages of evolution. Even more generally, at the earliest stages in the evolution of life, origin of complex structures, such as the cell itself, "from so simple a beginning" (Darwin, 1859) appears inevitable. Thus, notwithstanding the numerous cases of reductive evolution, in particular among parasites and symbionts, the belief in a general complexification trend in the evolution of life appears to be common.

However, is complexification the prevailing modality of evolution? Phylogenomic reconstruction, at least for bacteria and Archaea, suggests otherwise. It is not surprising that differential gene loss dominates the evolution of commensal bacteria, such as Lactobacilli, from a complex free-living ancestor (Makarova

et al., 2006). However, a qualitatively similar pattern was detected in evolutionary reconstructions for all bacteria and archaea (Snel et al., 2002; Mirkin et al., 2003; Makarova et al., 2007). Strikingly, more recent reconstructions that were performed using larger genome sets and more sophisticated computational methods confidently indicate that the genome of the last common ancestor of all extant archaea apparently was at least as large and complex as that of typical modern organisms in this domain of cellular life (Csuros and Miklos, 2009). Fully compatible reconstruction results have been reported for the expanded set of cyanobacterial genomes (Larsson et al., 2011). Thus, counter-intuitively, at least in prokaryotes, genome shrinkage that is sometimes called streamlining (Lynch, 2006) and is attributed to increasing selective pressure in successful, large populations (Lynch, 2006; Koonin, 2009b), appears to be no less and probably more common than genome growth and complexification.

THE WRIGTHEAN-DARWINIAN-LAMARCKIAN CONTINUUM OF EVOLUTIONARY PROCESSES

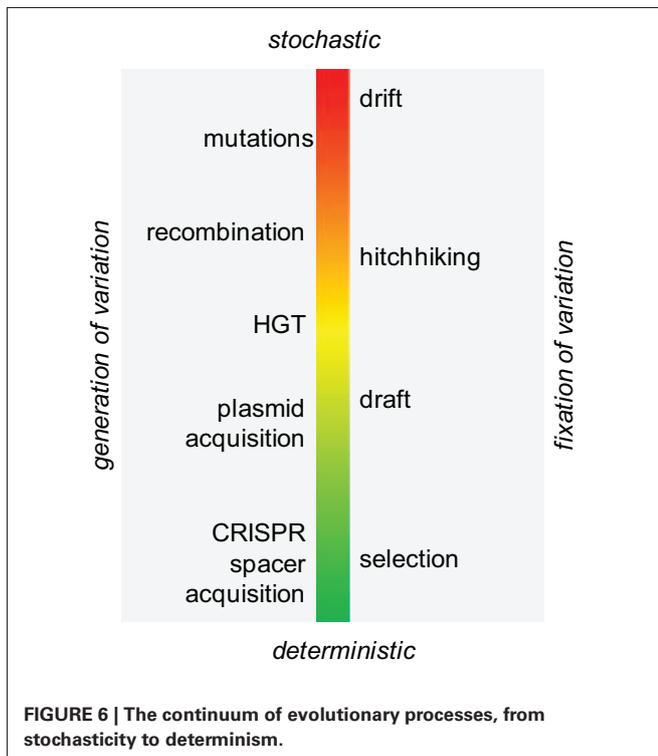
The Modern Synthesis of evolutionary biology emphasizes the randomness of mutations that provide the starting material for selection which engenders survival of the fittest under the given conditions and hence constitutes the adaptive, deterministic component of evolution. The insistence on such strict separation between the stochastic and deterministic aspects of evolution departs from Darwin's view that included the Lamarckian inheritance, with adaptive mutations directly caused by environmental cues, as an important, even if ancillary mechanism of evolution (Darwin, 1872).

Recently, several genetic phenomena with a distinct Lamarckian flavor have been discovered (Koonin and Wolf, 2009a; O'Malley and Koonin, 2011). Probably, the most striking case is the system of adaptive antiviral immunity, known as CRISPR-Cas (Clustered Regularly Interspaced Palindromic Repeats and CRISPR-associated proteins), that is present in most archaea and many bacteria (Koonin and Makarova, 2009; van der Oost et al., 2009; Marraffini and Sontheimer, 2010; Makarova et al., 2011). The CRISPR-Cas system integrates fragments of virus or plasmid DNA into a distinct, repetitive locus in the archaeal or bacterial genome. The transcript of this unique spacer functions as a guide RNA that is incorporated into a specific complex of Cas proteins possessing DNase activity and directs this complex to the cognate alien DNA (or RNA) molecules that are cleaved and accordingly inactivated. The CRISPR-Cas system is amazingly efficient, with only about 10^{-5} failure rate (Deveau et al., 2008). This mechanism qualifies CRISPR-Cas as an adaptive immunity system, i.e., immunity system that adapts to a specific infectious agent, a novelty in prokaryotes (Koonin and Makarova, 2009; Bikard and Marraffini, 2012). Furthermore, the Lamarckian principle of inheritance and evolution is apparent in the mechanism of CRISPR-Cas function. Indeed, this system directly responds to an environmental cue (in this case, foreign DNA) by introducing a genetic change into the genome that is immediately adaptive with respect to that particular cue.

The discovery of the CRISPR-Cas immune system that functions on the Lamarckian principle drew attention to other phenomena that also seem to contain a Lamarckian component

(Koonin and Wolf, 2009a; O'Malley and Koonin, 2011). Some of the common, central evolutionary processes such as HGT and stress-induced mutagenesis show a “quasi-Lamarckian” character. Indeed, even if HGT cannot be viewed as being directly caused by a specific environmental factor, it certainly is the case that the repertoire of the acquired genes depends on the environment. Genes common in a given environment will be acquired often and are likely to possess adaptive value. Stress-induced mutagenesis is triggered directly by environmental stress factors, e.g., desiccation or radiation, and produces variation that is required to develop resistant phenotype (Rosenberg and Hastings, 2003; Ponder et al., 2005; Galhardo et al., 2007; Galhardo and Rosenberg, 2009). The mutations are not specific to the biologically relevant loci but the activity of the molecular machineries of stress-induced mutagenesis [the best characterized of which is the SOS repair-mutagenesis system in bacteria (Sutton et al., 2000)] generates clusters of mutations, thus locally amplifying variability and so increasing the chance of adaptation once a single mutation appears in a relevant gene (Galhardo et al., 2007).

More generally, recent empirical and theoretical studies of diverse processes of stochastic and deterministic change in genomes make it clear that evolution is not limited to the basic Darwinian scheme of random variation that is subject to selection. Evolution can be more adequately depicted as a continuum of processes from completely random ones, under the Wrightian modality defined by random variation and random fixation of changes via genetic drift; to the Darwinian modality with random changes fixed by the deterministic process of selection; to the Lamarckian mode in which both variation and fixation are deterministic (Figure 6) (Koonin and Wolf, 2009a; O'Malley and Koonin, 2011).



EVOLUTION OF EVolvABILITY: DEDICATED MECHANISMS FOR EVOLUTION

All organisms possess a certain degree of evolvability, i.e., the ability to evolve. At the most basic level, evolvability stems from the theoretical impossibility of error-free replication. Genomic variation in evolving organisms is created by a combination of intrinsic replication errors, recombination and mutations induced external agents (mutagens). An intriguing, fundamental question in evolutionary biology is whether or not evolvability itself can evolve under selection, or put another way, whether there are dedicated mechanisms of evolution (Kirschner and Gerhart, 1998; Poole et al., 2003; Pigliucci, 2008; Brookfield, 2009). The prevailing wisdom among biologists seems to be that evolvability is not selectable but is simply maintained at a sufficient level by inevitable errors at all levels of biological information processing. Under this view, selection is always directed at minimization of the error rate but the ability to attain perfection is limited by genetic drift resulting in sufficient evolvability (Lynch, 2011). Evolutionary biologists are usually suspicious of the evolution of evolvability, generally under the old adage, “evolution has no forecast.”

Nevertheless, evidence in support of “evolvability of evolvability” is mounting. The very existence of complex molecular systems for stress-induced mutagenesis (error-prone repair) the activity of which is exquisitely regulated in response to stress implies that mechanisms enhancing variation when variation is needed for survival have evolved (Galhardo et al., 2007). Another remarkable mechanism that appears to have specifically evolved to generate variation involves the Diversity Generating Retroelements (DGR) (Medhekar and Miller, 2007). Strikingly, the DGR are found both in bacteriophages where they generate diversity in cell attachment surface proteins via reverse transcription-mediated mutagenesis, resulting in host tropism switching (Doulatov et al., 2004; Guo et al., 2008), and in bacteria themselves where they produce receptor variation leading to bacteriophage resistance (Bikard and Marraffini, 2012). The analogy between the activity of DGR and hypermutagenesis in animal immune systems is obvious except that the variation generated by the DGR is inherited.

Many bacteria and some archaea possess the natural transformation ability (that was used in the Avery experiment) that requires specialized, complex pumps (recently denoted transformosomes) that internalize DNA from the environment (Claverys et al., 2009; Johnsborg and Havarstein, 2009; Kruger and Stingl, 2011). The transformation machinery potentially could be viewed as a device that evolved under selective pressure to enhance HGT (Johnsborg and Havarstein, 2009). However, one could argue that the enhancement of HGT is only a side effect of the evolution of the transformation system, its actual *raison d'être* being the utilization of DNA as a rich source of replication substrates (or simply food). This argument hardly can hold with regard to the type 4 secretion systems (T4SS) that specialize in secretion of DNA from bacterial cells (Hamilton et al., 2005; Hamilton and Dillard, 2006). The recently discovered Gene Transfer Agents (GTAs) are even more striking devices for DNA donation (Paul, 2008; McDaniel et al., 2010; Lang et al., 2012). The GTAs are a distinct type of defective bacteriophages that

package in the capsid not the phage genome (which remains integrated in the host chromosome) but rather apparently random pieces of the host chromosome. The GTAs have been discovered in diverse bacteria and archaea and have been shown to infect and transfer their genetic content to a broad range of cohabitating prokaryotes (McDaniel et al., 2010). It does not seem conceivable that GTAs are anything but dedicated HGT vehicles. An additional notable aspect of T4SS and GTAs is that these devices mediate donation rather than consumption of DNA, i.e., apparently can directly benefit other microbes (recipients) rather than the donor. This seemingly altruistic behavior can be explained in terms of group selection whereby the object of selection is an ensemble of organisms that jointly benefit from adaptive mutations rather than a single organism. Group selection is a controversial subject in evolutionary biology (Maynard Smith, 1998; Borrello, 2005; Leigh, 2010) but the existence of dedicated devices for DNA donation appears to be a strong argument in its favor.

The discovery of T4SS and GTAs may be the most clear-cut pieces of evidence supporting evolution of evolvability just as the CRISPR-Cas system is the showcase for Lamarckian evolution. However, the case for the evolution of mechanisms for evolution seems to be much more general (O'Malley and Koonin, 2011). Population genetic theory holds that under a broad range of conditions a clonal population is generally doomed to collapse through the action of Muller's ratchet, the irreversible accumulation of deleterious mutations leading to gradual decline in fitness (Leigh, 2010; Bachtrog and Gordo, 2004). The effect of Muller's ratchet that has been directly demonstrated in controlled evolutionary experiments on RNA viruses (Chao, 1990; Duarte et al., 1992) and on bacteria (Andersson and Hughes, 1996). The principal way to escape Muller's ratchet is to enhance recombination via sex (in the form of meiotic crossing over in eukaryotes and in the form of conjugation in prokaryotes) or HGT. Just as sex is generally viewed as a mechanism that evolved to counteract the ratchet, HGT may be best understood as a more general variation-generating process that is supported by various evolved mechanisms. At the risk of being provocative, sex indeed can be legitimately regarded as a specialized form of HGT. Clearly, evolution maintains HGT within the optimal range rather than at the maximum possible level because the latter would eliminate genome stability and wreak havoc into selected high-fitness ensembles of genes (O'Malley and Koonin, 2011). Mechanisms that counter HGT also have evolved: these are the same that provide resistance against virus infection including CRISPR-Cas and restriction-modification (Marraffini and Sontheimer, 2008; Gardner and Olson, 2012).

At a different level, an apparent mechanism of evolution involves unusual, stable phenotype modifications that are widespread in bacteria and lead to coexistence of two distinct phenotypes in a clonal population, the so-called bistability regimes (Dubnau and Losick, 2006; Veening et al., 2008a; Piggot, 2010). For instance, under limited nutrient supply, *Bacillus subtilis* will form two subpopulations of which only the smaller one has the capacity to sporulate and thus yields the only survivors when the conditions become incompatible with cell growth and division

(Veening et al., 2008a,b; Lopez et al., 2009). The coexistence is epigenetically inherited across many bacterial generations, hence this phenomenon has become known as bistability. In theoretical and experimental models bistability is rationalized as "bet hedging": for organisms that live in often and unpredictably changing environments, it is beneficial to maintain a small subpopulation of likely survivors even when their fitness is comparatively low under normal conditions (Veening et al., 2008a; de Jong et al., 2011; Libby and Rainey, 2011; Rainey et al., 2011). The cost of maintaining this subpopulation is more than compensated by the benefit of survival under adverse conditions. Thus, the evolution of the regulatory circuitry that supports bistability appears to be not just a case of evolution of an evolutionary mechanism but more specifically evolution of a kin selection mechanism or evolution of altruism in bacteria. The evolution of kin selection demonstrated by bet hedging is paralleled by the mechanism of altruistic suicide that virus-infected bacteria and archaea commit using the toxin-antitoxin or abortive infection defense systems (Makarova et al., 2009; Van Melderen and Saavedra De Bast, 2009; Hayes and Van Melderen, 2011). In this case, by killing themselves early, before the virus has a chance to replicate, the microbes save their kin from infection. The reality of kin selection, just as that of group selection, is often hotly debated by evolutionary biologists (Nowak et al., 2010; Bourke, 2011; Ferriere and Michod, 2011; Strassmann et al., 2011) but the bistability/bet-hedging phenomena and altruistic suicide in bacteria and archaea seem to plainly demonstrate not only the existence but the evolvability of this form of selection.

In parallel with experimental studies, several theoretical models have been developed that characterize evolvability as a selectable trait in fluctuating environments (Earl and Deem, 2004; Jones et al., 2007; Draghi and Wagner, 2008). Thus, on the whole, and general theoretical doubts notwithstanding, evolution of evolvability appears to be an intrinsic and fundamental, if still poorly understood, aspect of the evolutionary process.

THE VAST, ANCIENT WORLD OF VIRUSES

Viruses are no part of the modern synthesis or more generally the traditional narrative of evolutionary biology. Until very recently, viruses have been viewed primarily as pathogens of animals, plants, and bacteria. Several lines of recent discovery have radically changed this view and promoted viruses to a central position on the stage of evolution. This change in the evolutionary status of viruses and related selfish genetic elements has been discussed in detail elsewhere (Claverie, 2006; Koonin et al., 2006, 2011; Raoult and Forterre, 2008). Here we quickly recapitulate several key points, with a focus on the importance of viruses for evolutionary biology in general. Metagenomic and ecological genomics studies have shown that, astonishingly, viruses are the most common biological entities on earth (Edwards and Rohwer, 2005; Suttle, 2005, 2007). Viruses and/or virus-like mobile elements are present in all cellular life forms. Strikingly, in mammals sequences derived from mobile elements and endogenous viruses account for at least 50% of the genome whereas in plants this fraction can reach 90% (Feschotte et al., 2002; Kazazian, 2004; Devos et al., 2005; Hedges and Batzer, 2005). Even the genomes of some unicellular eukaryotes, such as *Trichomonas vaginalis*, consist mostly

of inactivated transposons (Carlton et al., 2007; Pritham et al., 2007). Recruitment of mobile element sequences for transcription regulation and other cellular functions such as microRNA formation is a common phenomenon the full extent of which is not yet fully appreciated (Jordan et al., 2003; Piriyaopongsa et al., 2007; Lisch and Bennetzen, 2011). Although genomes of prokaryotes are not so overwhelmed by mobile elements, due to the intense purifying selection, nearly all of them encompass multiple prophages and mobile elements. Notably, deletion of all prophages leads to a substantial drop of fitness in *E. coli* (Wang et al., 2010).

In at least some common environments such as ocean water and soil, the number of virus particles exceeds the number of cells by factors of 10–100 (Edwards and Rohwer, 2005; Suttle, 2007; Srinivasiah et al., 2008; Breitbart, 2012). Similarly, the genetic diversity of viruses, measured as the number of distinct genes, substantially exceeds the genetic diversity of cellular life forms. Furthermore, viruses, in particular bacteriophages, are major biogeochemical agents. Periodical killing of microbes, in particular cyanobacteria, has been identified as a major contributor to sediment formation and major contributors to the nutrient cycles in the biosphere (Suttle, 2007; Rohwer and Thurber, 2009). The same process obviously is a key determinant of the population dynamics of the hosts that shapes the selection-drift balance throughout the course of evolution (Weinbauer and Rassoulzadegan, 2004).

The very fact that viruses greatly outnumber bacteria in the environment implies that antiviral defense systems are central to the evolution of bacteria and archaea. This is indeed the case as made evident by the remarkable proliferation of diverse antiviral systems including CRISPR-Cas discussed above as well as multiple restriction-modification, abortive infection, toxin-antitoxin and other, still poorly characterized defense systems that in different combinations and with different abundances are present in most prokaryotes (Juhas et al., 2009; Labrie et al., 2010; Makarova et al., 2011; Martinez-Borra et al., 2012). Taken together, these findings and theoretical considerations strongly support the view that the virus-host arms race is one of the principal processes in all evolution (Forterre and Prangishvili, 2009; Stern and Sorek, 2011).

With regard to the classification of life forms, the only defensible position appears to be that viruses (and related mobile elements) and cells are the two principal categories of biological organization (Figure 7) (Raoult and Forterre, 2008; Koonin, 2010; O'Malley and Koonin, 2011); this view is independent of the semantic issue of viruses being “alive” or not (Koonin et al., 2009; Moreira and Lopez-Garcia, 2009; Raoult, 2009). These two categories of biological entities can be characterized as informational (genetic) parasites, i.e., viruses and other selfish elements, and genetically self-sustained organisms, i.e., cellular life forms. Mathematical modeling indicates that genetic parasites inevitably emerge in any replicator system (Szathmari and Maynard Smith, 1997; Takeuchi and Hogeweg, 2012). This conclusion is certainly intuitively plausible: one expects that cheaters will appear in any system with limited resources—in particular, in any system of replicators, such parasites will attempt to utilize the replication machinery without making it (Koonin and

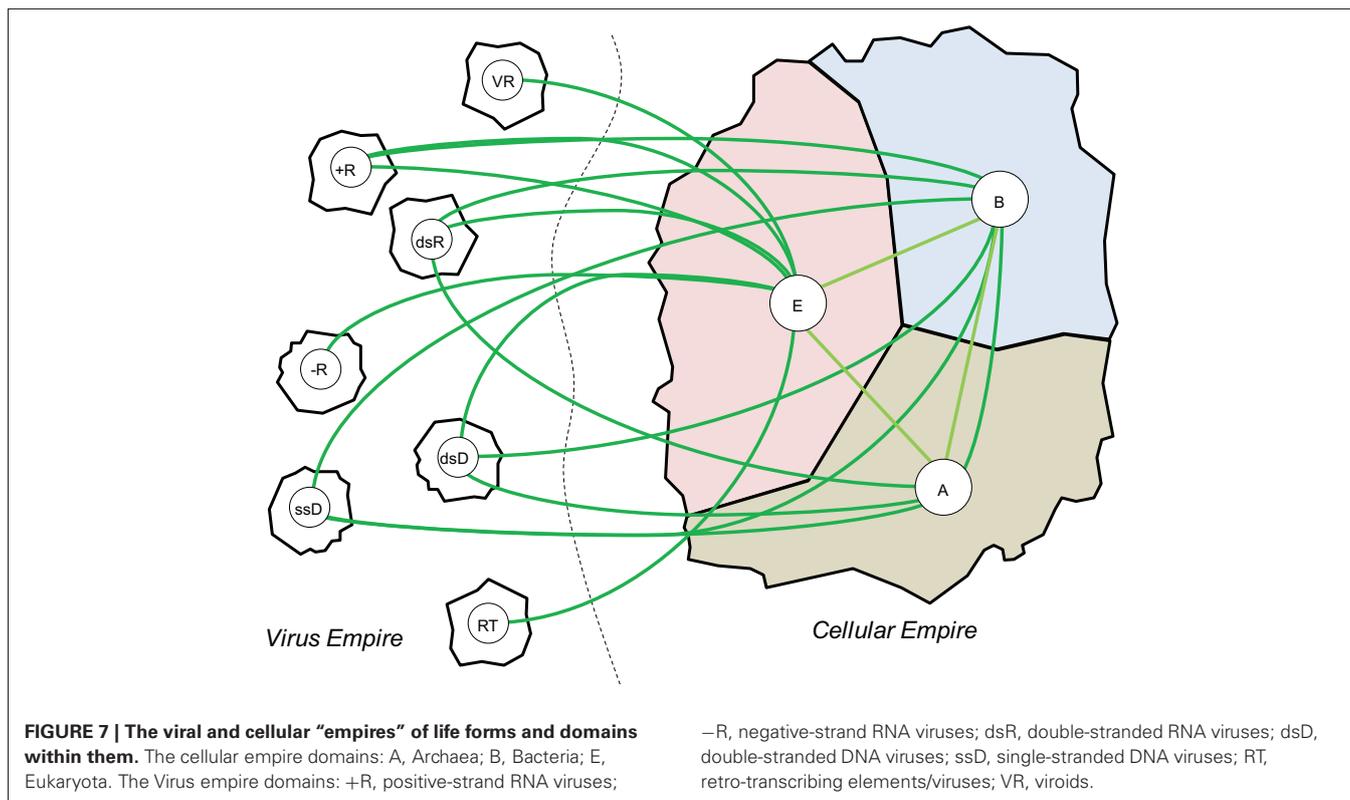
Martin, 2005). Also, the notion that virus-like selfish elements are an intrinsic part of life since its inception [which can be reasonably considered to coincide with the origin of replication (O'Malley and Koonin, 2011)] is compatible with the ubiquity of these elements in nature. In mathematical modeling, the outcome of the virus-host interaction depends on the specific parameters of the adapted model. In homogeneous models, virus-like parasites tend to cause collapse of the entire systems but in models with compartmentalization, which are most relevant for the actual evolution of life, stable host-parasite coexistence is possible (Takeuchi and Hogeweg, 2009). Moreover, the destructive effect of genetic parasites on the host is mitigated when a dedicated genetic information storage medium evolves, which could be one of the driving forces behind the evolution of DNA in the primordial RNA world (Takeuchi et al., 2011).

Further support for the classification of viruses as one of the two “empires” of life is the diversity of the replication-expression cycles that is found among viruses and related elements. Indeed, while cellular life forms all use a uniform replication-expression strategy based on double-stranded (ds)DNA replication, transcription of genes into mRNA or non-coding RNA, and translation of mRNA into protein, viral genome can be represented by all known forms of nucleic acids, and alternative replication processes such as RNA replication and reverse transcription are widely used (Figure 7) (Koonin et al., 2006). Finally, although viral genomes are generally small compared to the genomes of cellular life forms (viruses being the ultimate genetic parasites), the range of genomic complexity is remarkable, from only about 300 nucleotides and no genes in the simplest virus-like parasites, the viroids, to over a megabase and more than 1000 genes (genomes that are more complex than those of many bacterial parasites and symbionts) in the giant mimiviruses (Raoult et al., 2004; Colson et al., 2012). Overall, the conclusion is inescapable that the entire history of life is a story of perennial interplay between genetic parasites and their hosts that is a major driver of evolution for both biological empires.

EVOLUTION OF MICROBES AND VIRUSES: A NEW EVOLUTIONARY PARADIGM?

Prokaryotes (bacteria and archaea) and viruses entered the realm of evolution with the advent of genomics. Has the comparative study of these relatively simple (compared to eukaryotes) organisms radically changed the core tenets of evolutionary biology that were first envisaged by Darwin and were augmented with the genetic foundation in the Modern Synthesis? In terms of Kuhn's concept of the development of science (Kuhn, 1962), did the study of microbial evolution engender a paradigm shift?

It is not easy to answer this question definitively, possibly because the paradigm shift model does not adequately describe the evolution of biology (regardless of whether or not it fits the evolution of physics). Probably, a more appropriate epistemological framework is that of integration, i.e., a relatively smooth incorporation of the classic concepts into the new, more general and versatile theoretical constructs. This model of the evolution of science was recognized by Kuhn himself in his later work (Kuhn, 2002) and was recently examined by O'Malley in the context of biology (O'Malley, 2012; O'Malley and Soyer, 2012).



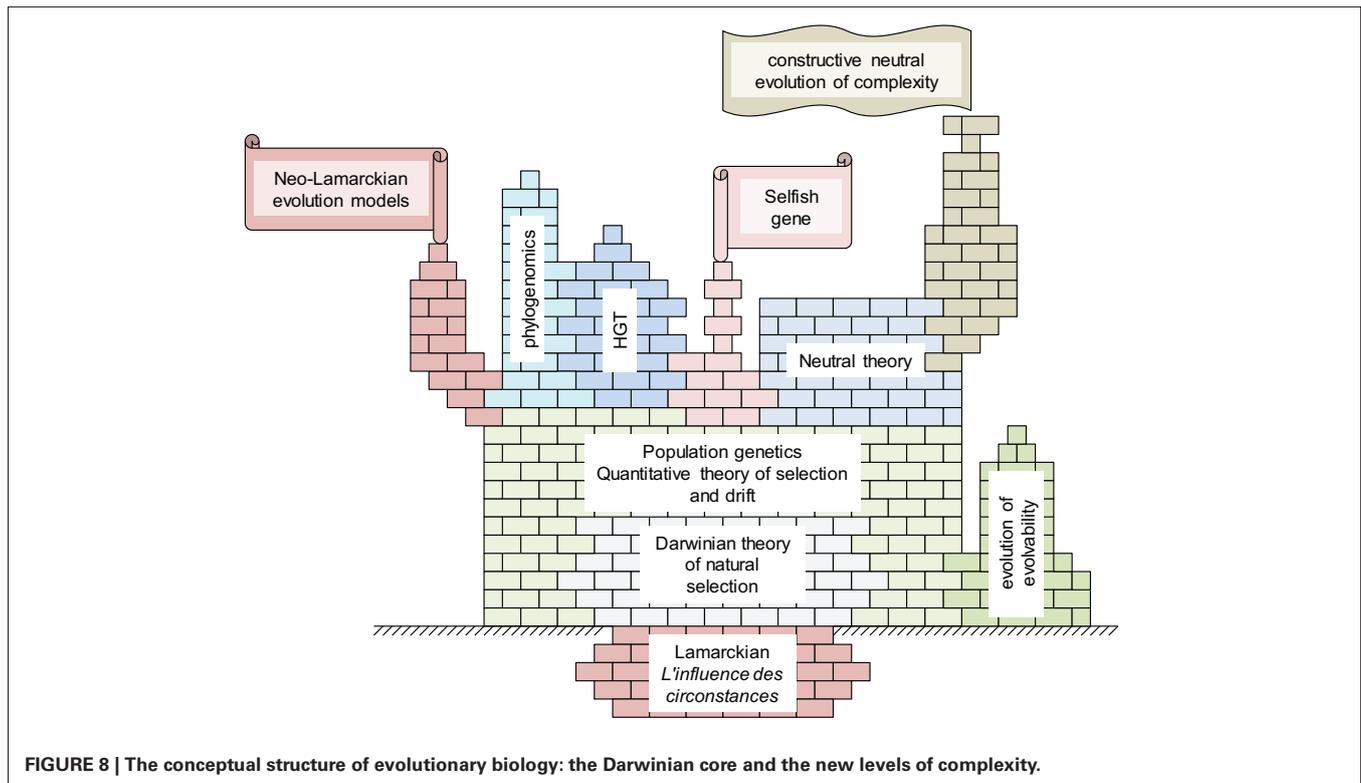
The phylogenomic study of microbes and viruses uncovered new biological realms which Darwin and even the authors of the Modern Synthesis could not possibly fathom. The modes of evolution of these relatively simple organisms that, as we now realize, have dominated the biosphere since its beginning about 4 billion years ago to this day (and into any conceivable future) are different from the evolutionary regimes of animals and plants, the traditional objects of (evolutionary) biology. The study of microbial evolution has shattered the classic idea of a single, all-encompassing tree of life by demonstrating that the evolutionary histories of individual genes are generally different. Remarkably, however, these developments have not rendered trees irrelevant as a key metaphor of evolution (O’Malley and Koonin, 2011). Rather, they have shown that the bona fide unit of tree-like evolution is an individual gene not a genome, and a “tree of life” can only be conceived as a statistical trend in the “forest” of gene trees (Koonin and Wolf, 2009b). Tree-like evolution is a fundamental implication of the binary replication of the genetic material, so it served Darwin well to use a tree as the single illustration of his book. Without, obviously, knowing anything of DNA replication, Darwin grasped the central principle of the evolution of life, descent with modification, and the tree pattern followed naturally.

Microbiology yielded the first clear-cut case of Lamarckian evolution, the CRISPR-Cas system, and subsequent re-examination of other evolutionary phenomena (in both prokaryotes and eukaryotes) has strongly suggested that the (quasi)Lamarckian modality is common and important in all evolving organisms, completing the range of evolutionary

phenomena from purely stochastic (drift, Wrightian evolution) to deterministic (Lamarckian evolution). Again, these findings not so much overturned but rather expanded the vision of Darwin who seriously considered Lamarckian mechanisms as being ancillary to natural selection (only the Modern synthesis banished Lamarck).

Crucially, the study of microbial evolution presented apparently undeniable cases of evolution of evolvability such as the GTAs and the DGRs. Moreover, the discovery of bet-hedging strategies and altruistic suicide in bacteria shows that kin selection (a subject of considerable controversy in evolutionary biology) is evolvable as well. Again, as in the case of Lamarckian mechanisms, these discoveries force one to re-examine many more phenomena and realize that evolution is not limited to fixation of random variation and survival of the fittest but rather is an active process with multiple feedback loops, and that dedicated mechanisms of evolution exist and themselves evolve. This is a major generalization that substantially adds to the overall structure of evolutionary biology but one has to realize that the principle of descent with modification remains at the core of all these complex evolutionary phenomena.

We now realize that evolution of life is to a large extent shaped by the interaction (arms race but also cooperation) between genetic parasites (viruses and other selfish elements) and their cellular hosts. Viruses and related elements, with their distinctive life strategy, informational parasitism, actually dominate the biosphere both physically and genetically, and represent one of the two principal forms of life that as intrinsic to the history of the biosphere as cells are. This new dimension of evolution simply



could not be perceived by Darwin or even the creators of the Modern Synthesis due to the lack of relevant data.

Thus, we are inclined to view the change in evolutionary biology brought about by phylogenomics of microbes and viruses as a case of integration rather than an abrupt departure from the paradigm of the Modern Synthesis (Figure 8). Darwin realized the importance of descent with modification and the tree pattern of evolution it implies whereas Fisher, Wright, and Haldane derived the laws of population genetics that still constitute the core of our understanding of evolution. However, recent advances, in particular those of microbial phylogenomics, added

multiple, new and interconnected layers of complexity (Figure 8) such that the conceptual core is but a small part of the current big picture of evolutionary biology.

AUTHOR CONTRIBUTIONS

Eugene V. Koonin and Yuri I. Wolf wrote the manuscript.

ACKNOWLEDGMENTS

The authors' research is supported by intramural funds of the US Department of Health and Human Services (to National Library of Medicine, NIH).

REFERENCES

- Abby, S. S., Tannier, E., Gouy, M., and Daubin, V. (2012). Lateral gene transfer as a support for the tree of life. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4962–4967.
- Andam, C. P., and Gogarten, J. P. (2011). Biased gene transfer in microbial evolution. *Nat. Rev. Microbiol.* 9, 543–555.
- Andersson, D. I., and Hughes, D. (1996). Muller's ratchet decreases fitness of a DNA-based microbe. *Proc. Natl. Acad. Sci. U.S.A.* 93, 906–907.
- Aravind, L., Tatusov, R. L., Wolf, Y. I., Walker, D. R., and Koonin, E. V. (1998). Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles. *Trends Genet.* 14, 442–444.
- Arber, W. (1979). Promotion and limitation of genetic exchange. *Experientia* 35, 287–293.
- Avery, O. T., MacLeod, C. M., and McCarty, M. (1944a). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* 79, 137–158.
- Avery, O. T., MacLeod, C. M., and McCarty, M. (1944b). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Inductions of transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III. *J. Exp. Med.* 149, 297–326.
- Bachtrog, D., and Gordo, I. (2004). Adaptive evolution of asexual populations under Muller's ratchet. *Evolution* 58, 1403–1413.
- Bapteste, E., and Boucher, Y. (2009). Epistemological impacts of horizontal gene transfer on classification in microbiology. *Methods Mol. Biol.* 532, 55–72.
- Bapteste, E., O'Malley, M. A., Beiko, R. G., Ereshefsky, M., Gogarten, J. P., Franklin-Hall, L., Lapointe, F. J., Dupre, J., Dagan, T., Boucher, Y., and Martin, W. (2009). Prokaryotic evolution and the tree of life are two different things. *Biol. Direct* 4, 34.
- Bapteste, E., Susko, E., Leigh, J., MacLeod, D., Charlebois, R. L., and Doolittle, W. F. (2005). Do orthologous gene phylogenies really support tree-thinking? *BMC Evol. Biol.* 5, 33.
- Bikard, D., and Marraffini, L. A. (2012). Innate and adaptive immunity in bacteria: mechanisms of programmed genetic variation to fight bacteriophages. *Curr. Opin. Immunol.* 24, 15–20.
- Borrello, M. E. (2005). The rise, fall and resurrection of group selection. *Endeavour* 29, 43–47.
- Boto, L. (2010). Horizontal gene transfer in evolution: facts and

- challenges. *Proc. Biol. Sci.* 277, 819–827.
- Bourcier, A. F. (2011). The validity and value of inclusive fitness theory. *Proc. Biol. Sci.* 278, 3313–3320.
- Breitbart, M. (2012). Marine viruses: truth or dare. *Ann. Rev. Mar. Sci.* 4, 425–448.
- Brochier, C., Philippe, H., and Moreira, D. (2000). The evolutionary history of ribosomal protein RpS14, horizontal gene transfer at the heart of the ribosome. *Trends Genet.* 16, 529–533.
- Brookfield, J. F. (2009). Evolution and evolvability: celebrating Darwin 200. *Biol. Lett.* 5, 44–46.
- Browne, J. (2008). Birthdays to remember. *Nature* 456, 324–325.
- Bushman, F. (2001). *Lateral DNA Transfer: Mechanisms and Consequences*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Campbell, A. (1981). Evolutionary significance of accessory DNA elements in bacteria. *Annu. Rev. Microbiol.* 35, 55–83.
- Carlton, J. M., Hirt, R. P., Silva, J. C., Delcher, A. L., Schatz, M., Zhao, Q., Wortman, J. R., Bidwell, S. L., Alsmark, U. C., Besteiro, S., Sichert-Ponten, T., Noel, C. J., Dacks, J. B., Foster, P. G., Simillion, C., Van de Peer, Y., Miranda-Saavedra, D., Barton, G. J., Westrop, G. D., Muller, S., Dessi, D., Fiori, P. L., Ren, Q., Paulsen, I., Zhang, H., Bastida-Corcuera, F. D., Simoes-Barbosa, A., Brown, M. T., Hayes, R. D., Mukherjee, M., Okumura, C. Y., Schneider, R., Smith, A. J., Vanacova, S., Villalvazo, M., Haas, B. J., Perlea, M., Feldblyum, T. V., Utterback, T. R., Shu, C. L., Osoegawa, K., de Jong, P. J., Hrdy, I., Horvathova, L., Zubacova, Z., Dolezal, P., Malik, S. B., Logsdon, J. M. Jr., Henze, K., Gupta, A., Wang, C. C., Dunne, R. L., Upcroft, J. A., Upcroft, P., White, O., Salzberg, S. L., Tang, P., Chiu, C. H., Lee, Y. S., Embley, T. M., Coombs, G. H., Mottram, J. C., Tachezy, J., Fraser-Liggett, C. M., and Johnson, P. J. (2007). Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* 315, 207–212.
- Carus, T. L. (2011). *De Rerum Natura*. New York, NY: Nabu Press.
- Chao, L. (1990). Fitness of RNA virus decreased by Muller's ratchet. *Nature* 348, 454–455.
- Claverie, J. M. (2006). Viruses take center stage in cellular evolution. *Genome Biol.* 7, 110.
- Claverys, J. P., Martin, B., and Polard, P. (2009). The genetic transformation machinery: composition, localization, and mechanism. *FEMS Microbiol. Rev.* 33, 643–656.
- Cohan, F. M., and Perry, E. B. (2007). A systematics for discovering the fundamental units of bacterial diversity. *Curr. Biol.* 17, R373–R386.
- Cohen, O., Gophna, U., and Pupko, T. (2011). The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. *Mol. Biol. Evol.* 28, 1481–1489.
- Colson, P., de Lamballerie, X., Fournous, G., and Raoult, D. (2012). Reclassification of giant viruses composing a fourth domain of life in the new order Megavirales. *Intervirology* 55, 321–332.
- Csuros, M., and Miklos, I. (2009). Streamlining and large ancestral genomes in Archaea inferred with a phylogenetic birth-and-death model. *Mol. Biol. Evol.* 26, 2087–2095.
- Dagan, T. (2011). Phylogenomic networks. *Trends Microbiol.* 19, 483–491.
- Dagan, T., Artzy-Randrup, Y., and Martin, W. (2008). Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc. Natl. Acad. Sci. U.S.A.* 105, 10039–10044.
- Dagan, T., and Martin, W. (2006). The tree of one percent. *Genome Biol.* 7, 118.
- Dagan, T., and Martin, W. (2009). Getting a better picture of microbial evolution en route to a network of genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 2187–2196.
- Darwin, C. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. I. Extract from an unpublished work on species, II. Abstract of a letter from Darwin, C., esq., to Prof. Asa Gray. *J. Proc. Linn. Soc. Lond.* 3, 45–53.
- Darwin, C. (1859). *On the Origin of Species*. London: John Murray.
- Darwin, C. (1872). *Origin of Species*. New York, NY: The Modern Library.
- de Jong, I. G., Haccou, P., and Kuipers, O. P. (2011). Bet hedging or not? A guide to proper classification of microbial survival strategies. *Bioessays* 33, 215–223.
- Deppenmeier, U., Johann, A., Hartsch, T., Merkl, R., Schmitz, R. A., Martinez-Arias, R., Henne, A., Wiezer, A., Baumer, S., Jacobi, C., Bruggemann, H., Lienard, T., Christmann, A., Bomeke, M., Steckel, S., Bhattacharyya, A., Lykidis, A., Overbeek, R., Klenk, H. P., Gunsalus, R. P., Fritz, H. J., and Gottschalk, G. (2002). The genome of *Methanosarcina mazei*: evidence for lateral gene transfer between bacteria and archaea. *J. Mol. Microbiol. Biotechnol.* 4, 453–461.
- Deveau, H., Barrangou, R., Garneau, J. E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D. A., Horvath, P., and Moineau, S. (2008). Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400.
- Devos, K. M., Ma, J., Pontaroli, A. C., Pratt, L. H., and Bennetzen, J. L. (2005). Analysis and mapping of randomly chosen bacterial artificial chromosome clones from hexaploid bread wheat. *Proc. Natl. Acad. Sci. U.S.A.* 102, 19243–19248.
- Dobzhansky, T. (1937). *Genetics and the Origin of Species*. New York, NY: Columbia University Press.
- Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution. *Am. Biol. Teach.* 35, 125–129.
- 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.
- Doolittle, W. F., and Zhaxybayeva, O. (2009). On the origin of prokaryotic species. *Genome Res.* 19, 744–756.
- Doulatov, S., Hodes, A., Dai, L., Mandhana, N., Liu, M., Deora, R., Simons, R. W., Zimmerly, S., and Miller, J. F. (2004). Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. *Nature* 431, 476–481.
- Draghi, J., and Wagner, G. P. (2008). Evolution of evolvability in a developmental model. *Evolution* 62, 301–315.
- Duarte, E., Clarke, D., Moya, A., Domingo, E., and Holland, J. (1992). Rapid fitness losses in mammalian RNA virus clones due to Muller's ratchet. *Proc. Natl. Acad. Sci. U.S.A.* 89, 6015–6019.
- Dubnau, D., and Losick, R. (2006). Bistability in bacteria. *Mol. Microbiol.* 61, 564–572.
- Earl, D. J., and Deem, M. W. (2004). Evolvability is a selectable trait. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11531–11536.
- Edwards, R. A., and Rohwer, F. (2005). Viral metagenomics. *Nat. Rev. Microbiol.* 3, 504–510.
- Eppinger, M., Mammel, M. K., Leclerc, J. E., Ravel, J., and Cebula, T. A. (2011). Genomic anatomy of *Escherichia coli* O157, H7 outbreaks. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20142–20147.
- Esser, C., Martin, W., and Dagan, T. (2007). The origin of mitochondria in light of a fluid prokaryotic chromosome model. *Biol. Lett.* 3, 180–184.
- Ferriere, R., and Michod, R. E. (2011). Inclusive fitness in evolution. *Nature* 471, E6–E8. author reply E9–E10.
- Feschotte, C., Jiang, N., and Wessler, S. R. (2002). Plant transposable elements: where genetics meets genomics. *Nat. Rev. Genet.* 3, 329–341.
- Fisher, R. A. (1930). *The Genetical Theory of Natural Selection*. Oxford: The Clarendon Press.
- Forster, P., and Prangishvili, D. (2009). The great billion-year war between ribosome- and capsid-encoding organisms (cells and viruses) as the major source of evolutionary novelties. *Ann. N.Y. Acad. Sci.* 1178, 65–77.
- Fraser, C., Alm, E. J., Polz, M. F., Spratt, B. G., and Hanage, W. P. (2009). The bacterial species challenge: making sense of genetic and ecological diversity. *Science* 323, 741–746.
- Galagan, J. E., Nusbaum, C., Roy, A., Endrizzi, M. G., Macdonald, P., FitzHugh, W., Calvo, S., Engels, R., Smirnov, S., Atnoor, D., Brown, A., Allen, N., Naylor, J., Stange-Thomann, N., DeArellano, K., Johnson, R., Linton, L., McEwan, P., McKernan, K., Talamas, J., Tirrell, A., Ye, W., Zimmer, A., Barber, R. D., Cann, I., Graham, D. E., Grahame, D. A., Guss, A. M., Hedderich, R., Ingram-Smith, C., Kuettner, H. C., Krzycki, J. A., Leigh, J. A., Li, W., Liu, J., Mukhopadhyay, B., Reeve, J. N., Smith, K., Springer, T. A., Umayam, L. A., White, O., White, R. H., Conway de Macario, E., Ferry, J. G., Jarrell, K. F., Jing, H., Macario, A. J., Paulsen, I., Pritchett, M., Sowers, K. R., Swanson, R. V., Zinder, S. H., Lander, E., Metcalf, W. W., and Birren, B. (2002). The genome of *M. acetivorans* reveals extensive metabolic and physiological diversity. *Genome Res.* 12, 532–542.
- Galhardo, R. S., Hastings, P. J., and Rosenberg, S. M. (2007). Mutation as a stress response and the regulation of evolvability. *Crit. Rev. Biochem. Mol. Biol.* 42, 399–435.
- Galhardo, R. S., and Rosenberg, S. M. (2009). Extreme genome repair. *Cell* 136, 998–1000.
- Gardner, S. P., and Olson, J. W. (2012). Barriers to horizontal gene transfer in *Campylobacter jejuni*. *Adv. Appl. Microbiol.* 79, 19–42.
- Gogarten, J. P., and Townsend, J. P. (2005). Horizontal gene transfer,

- genome innovation and evolution. *Nat. Rev. Microbiol.* 3, 679–687.
- Gordo, I., and Charlesworth, B. (2000). The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* 154, 1379–1387.
- Gribaldo, S., and Brochier, C. (2009). Phylogeny of prokaryotes: does it exist and why should we care? *Res. Microbiol.* 160, 513–521.
- Guo, H., Tse, L. V., Barbalat, R., Sivaamnuaihorn, S., Xu, M., Doulatov, S., and Miller, J. F. (2008). Diversity-generating retroelement homing regenerates target sequences for repeated rounds of codon rewriting and protein diversification. *Mol. Cell* 31, 813–823.
- Hacker, J., and Kaper, J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annu. Rev. Microbiol.* 54, 641–679.
- Haeckel, E. (1997). *The Wonders of Life: A Popular Study of Biological Philosophy*. London: General Books, LLC.
- Haldane, J. B. S. (1932). *The Causes of Evolution*. London, New York, Toronto: Longmans, Green and Co.
- Hamilton, H. L., and Dillard, J. P. (2006). Natural transformation of *Neisseria gonorrhoeae*: from DNA donation to homologous recombination. *Mol. Microbiol.* 59, 376–385.
- Hamilton, H. L., Dominguez, N. M., Schwartz, K. J., Hackett, K. T., and Dillard, J. P. (2005). *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel type IV secretion system. *Mol. Microbiol.* 55, 1704–1721.
- Hayes, F., and Van Melder, L. (2011). Toxins-antitoxins: diversity, evolution and function. *Crit. Rev. Biochem. Mol. Biol.* 46, 386–408.
- Hedges, D. J., and Batzer, M. A. (2005). From the margins of the genome: mobile elements shape primate evolution. *Bioessays* 27, 785–794.
- Jain, R., Rivera, M. C., and Lake, J. A. (1999). Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 96, 3801–3806.
- Johnsborg, O., and Havarstein, L. S. (2009). Regulation of natural genetic transformation and acquisition of transforming DNA in *Streptococcus pneumoniae*. *FEMS Microbiol. Rev.* 33, 627–642.
- Jones, A. G., Arnold, S. J., and Burger, R. (2007). The mutation matrix and the evolution of evolvability. *Evolution* 61, 727–745.
- Jordan, I. K., Rogozin, I. B., Glazko, G. V., and Koonin, E. V. (2003). Origin of a substantial fraction of human regulatory sequences from transposable elements. *Trends Genet.* 19, 68–72.
- Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., and Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS Microbiol. Rev.* 33, 376–393.
- Kazazian, H. H. Jr. (2004). Mobile elements: drivers of genome evolution. *Science* 303, 1626–1632.
- Kirschner, M., and Gerhart, J. (1998). Evolvability. *Proc. Natl. Acad. Sci. U.S.A.* 95, 8420–8427.
- Klasson, L., and Andersson, S. G. (2004). Evolution of minimal-genomes in host-dependent bacteria. *Trends Microbiol.* 12, 37–43.
- Kloesges, T., Popa, O., Martin, W., and Dagan, T. (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.
- Konstantinidis, K. T., Ramette, A., and Tiedje, J. M. (2006). The bacterial species definition in the genomic era. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 361, 1929–1940.
- Koonin, E. V. (2003). Horizontal gene transfer: the path to maturity. *Mol. Microbiol.* 50, 725–727.
- Koonin, E. V. (2009a). Darwinian evolution in the light of genomics. *Nucleic Acids Res.* 37, 1011–1034.
- Koonin, E. V. (2009b). Evolution of genome architecture. *Int. J. Biochem. Cell Biol.* 41, 298–306.
- Koonin, E. V. (2010). The two empires and three domains of life in the postgenomic age. *Nat. Educ.* 3, 27.
- Koonin, E. V. (2011). *The Logic of Chance: The Nature and Origin of Biological Evolution*. Upper Saddle River, NJ: FT Press.
- Koonin, E. V., Makarova, K. S., and Aravind, L. (2001). Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* 55, 709–742.
- Koonin, E. V., and Makarova, K. S. (2009). CRISPR-Cas: an adaptive immunity system in prokaryotes. *F1000 Biol. Rep.* 1, 95.
- Koonin, E. V., and Martin, W. (2005). On the origin of genomes and cells within inorganic compartments. *Trends Genet.* 21, 647–654.
- Koonin, E. V., Puigbo, P., and Wolf, Y. I. (2011). Comparison of phylogenetic trees and search for a central trend in the “forest of life”. *J. Comput. Biol.* 18, 917–924.
- Koonin, E. V., Senkevich, T. G., and Dolja, V. V. (2006). The ancient Virus World and evolution of cells. *Biol. Direct* 1, 29.
- Koonin, E. V., Senkevich, T. G., and Dolja, V. V. (2009). Compelling reasons why viruses are relevant for the origin of cells. *Nat. Rev. Microbiol.* 7, 615.
- Koonin, E. V., and Wolf, Y. I. (2008). Genomics of bacteria and archaea: the emerging dynamic view of the prokaryotic world. *Nucleic Acids Res.* 36, 6688–6719.
- Koonin, E. V., and Wolf, Y. I. (2009a). Is evolution Darwinian or/and Lamarckian? *Biol. Direct* 4, 42.
- Koonin, E. V., and Wolf, Y. I. (2009b). The fundamental units, processes and patterns of evolution, and the Tree of Life conundrum. *Biol. Direct* 4, 33.
- Kruger, N. J., and Stinkl, K. (2011). Two steps away from novelty—principles of bacterial DNA uptake. *Mol. Microbiol.* 80, 860–867.
- Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.
- Kuhn, T. S. (2002). *The Road since Structure*. Chicago, IL: University of Chicago Press.
- Kunin, V., Goldovsky, L., Darzentas, N., and Ouzounis, C. A. (2005). The net of life: reconstructing the microbial phylogenetic network. *Genome Res.* 15, 954–959.
- Labrie, S. J., Samson, J. E., and Moineau, S. (2010). Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* 8, 317–327.
- Lamarck, J.-B. (1809). *Philosophie Zoologique, ou Exposition des Considérations Relatives à l'histoire Naturelle des Animaux*. Paris: Dentu.
- Lang, A. S., Zhaxybayeva, O., and Beatty, J. T. (2012). Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* 10, 472–482.
- Larsson, J., Nylander, J. A., and Bergman, B. (2011). Genome fluctuations in cyanobacteria reflect evolutionary, developmental and adaptive traits. *BMC Evol. Biol.* 11, 187.
- Leigh, E. G. Jr. (2010). The group selection controversy. *J. Evol. Biol.* 23, 6–19.
- Libby, E., and Rainey, P. B. (2011). Exclusion rules, bottlenecks and the evolution of stochastic phenotype switching. *Proc. Biol. Sci.* 278, 3574–3583.
- Lisch, D., and Bennetzen, J. L. (2011). Transposable element origins of epigenetic gene regulation. *Curr. Opin. Plant Biol.* 14, 156–161.
- Lopez, D., Vlamakis, H., and Kolter, R. (2009). Generation of multiple cell types in *Bacillus subtilis*. *FEMS Microbiol. Rev.* 33, 152–163.
- Low, K. B., and Porter, D. D. (1978). Modes of gene transfer and recombination in bacteria. *Annu. Rev. Genet.* 12, 249–287.
- Lynch, M. (2006). Streamlining and simplification of microbial genome architecture. *Annu. Rev. Microbiol.* 60, 327–349.
- Lynch, M. (2011). The lower bound to the evolution of mutation rates. *Genome Biol. Evol.* 3, 1107–1118.
- Makarova, K. S., Haft, D. H., Barrangou, R., Brouns, S. J., Charpentier, E., Horvath, P., Moineau, S., Mojica, F. J., Wolf, Y. I., Yakunin, A. F., van der Oost, J., and Koonin, E. V. (2011). Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* 9, 467–477.
- Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., Shakhova, V., Grigoriev, I., Lou, Y., Rohksar, D., Lucas, S., Huang, K., Goodstein, D. M., Hawkins, T., Plengvidhya, V., Welker, D., Hughes, J., Goh, Y., Benson, A., Baldwin, K., Lee, J. H., Diaz-Muniz, I., Dosti, B., Smeianov, V., Wechter, W., Barabote, R., Lorca, G., Altermann, E., Barrangou, R., Ganesan, B., Xie, Y., Rawsthorne, H., Tamir, D., Parker, C., Breidt, F., Broadbent, J., Hutkins, R., O'Sullivan, D., Steele, J., Unlu, G., Saier, M., Klaenhammer, T., Richardson, P., Kozyavkin, S., Weimer, B., and Mills, D. (2006). Comparative genomics of the lactic acid bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15611–15616.
- Makarova, K. S., Ponomarev, V. A., and Koonin, E. V. (2001). Two C or not two C: recurrent disruption of Zn-ribbons, gene duplication, lineage-specific gene loss, and horizontal gene transfer in evolution of bacterial ribosomal proteins. *Genome Biol.* 2, RESEARCH0033.
- Makarova, K. S., Sorokin, A. V., Novichkov, P. S., Wolf, Y. I., and Koonin, E. V. (2007). Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol. Direct* 2, 33.
- Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2009). Comprehensive comparative-genomic analysis of type 2 toxin-antitoxin systems and related mobile stress response systems in prokaryotes. *Biol. Direct* 4, 19.
- Makarova, K. S., Wolf, Y. I., Snir, S., and Koonin, E. V. (2011). Defense islands in bacterial and archaeal genomes and prediction of novel

- defense systems. *J. Bacteriol.* 193, 6039–6056.
- Marraffini, L. A., and Sontheimer, E. J. (2008). CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845.
- Marraffini, L. A., and Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11, 181–190.
- Martinez-Borra, J., Gonzalez, S., and Lopez-Larrea, C. (2012). The origin of the bacterial immune response. *Adv. Exp. Med. Biol.* 738, 1–13.
- Martin, W. F. (2011). Early evolution without a tree of life. *Biol. Direct* 6, 36.
- Maynard Smith, J. (1998). The units of selection. *Novartis Found. Symp.* 213, 203–211. discussion: 211–217.
- Mayr, E. (1944). *Systematics and the Origin of Species*. New York, NY: Columbia University Press.
- McCutcheon, J. P., and Moran, N. A. (2012). Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* 10, 13–26.
- McDaniel, L. D., Young, E., Delaney, J., Ruhnau, F., Ritchie, K. B., and Paul, J. H. (2010). High frequency of horizontal gene transfer in the oceans. *Science* 330, 50.
- Medhekar, B., and Miller, J. F. (2007). Diversity-generating retroelements. *Curr. Opin. Microbiol.* 10, 388–395.
- Merhej, V., Notredame, C., Royer-Carenzi, M., Pontarotti, P., and Raoult, D. (2011). The rhizome of life: the sympatric *Rickettsia felis* paradigm demonstrates the random transfer of DNA sequences. *Mol. Biol. Evol.* 28, 3213–3223.
- Mirkin, B. G., Fenner, T. I., Galperin, M. Y., and Koonin, E. V. (2003). Algorithms for computing parsimonious evolutionary scenarios for genome evolution, the last universal common ancestor and dominance of horizontal gene transfer in the evolution of prokaryotes. *BMC Evol. Biol.* 3, 2.
- Moreira, D., and Lopez-Garcia, P. (2009). Ten reasons to exclude viruses from the tree of life. *Nat. Rev. Microbiol.* 7, 306–311.
- Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C., and Fraser, C. M. (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* 399, 323–329.
- Novichkov, P. S., Omelchenko, M. V., Gelfand, M. S., Mironov, A. A., Wolf, Y. I., and Koonin, E. V. (2004). Genome-wide molecular clock and horizontal gene transfer in bacterial evolution. *J. Bacteriol.* 186, 6575–6585.
- Nowak, M. A., Tarnita, C. E., and Wilson, E. O. (2010). The evolution of eusociality. *Nature* 466, 1057–1062.
- O'Malley, M. A. (2012). Evolutionary systems biology: historical and philosophical perspectives on an emerging synthesis. *Adv. Exp. Med. Biol.* 751, 1–28.
- O'Malley, M. A., and Koonin, E. V. (2011). How stands the Tree of Life a century and a half after The Origin? *Biol. Direct* 6, 32.
- O'Malley, M. A., and Soyer, O. S. (2012). The roles of integration in molecular systems biology. *Stud. Hist. Philos. Biol. Biomed. Sci.* 43, 58–68.
- Pace, N. R. (1997). A molecular view of microbial diversity and the biosphere. *Science* 276, 734–740.
- Pace, N. R. (2006). Time for a change. *Nature* 441, 289.
- Pace, N. R. (2009). Mapping the tree of life: progress and prospects. *Microbiol. Mol. Biol. Rev.* 73, 565–576.
- Paul, J. H. (2008). Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas? *ISME J.* 2, 579–589.
- Perez-Brocal, V., Gil, R., Ramos, S., Lamelas, A., Postigo, M., Michelena, J. M., Silva, F. J., Moya, A., and Latorre, A. (2006). A small microbial genome: the end of a long symbiotic relationship? *Science* 314, 312–313.
- Perna, N. T., Plunkett, G. 3rd., Burland, V., Mau, B., Glasner, J. D., Rose, D. J., Mayhew, G. F., Evans, P. S., Gregor, J., Kirkpatrick, H. A., Posfai, G., Hackett, J., Klink, S., Boutin, A., Shao, Y., Miller, L., Grotbeck, E. J., Davis, N. W., Lim, A., Dimalanta, E. T., Potamou, K. D., Apodaca, J., Anantharaman, T. S., Lin, J., Yen, G., Schwartz, D. C., Welch, R. A., and Blattner, F. R. (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157, H7. *Nature* 409, 529–533.
- Piggott, P. (2010). Epigenetic switching: bacteria hedge bets about staying or moving. *Curr. Biol.* 20, R480–R482.
- Pigliucci, M. (2008). Is evolvability evolvable? *Nat. Rev. Genet.* 9, 75–82.
- Piriyapongsa, J., Marino-Ramirez, L., and Jordan, I. K. (2007). Origin and evolution of human microRNAs from transposable elements. *Genetics* 176, 1323–1337.
- Ponder, R. G., Fonville, N. C., and Rosenberg, S. M. (2005). A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation. *Mol. Cell* 19, 791–804.
- Poole, A. M., Phillips, M. J., and Penny, D. (2003). Prokaryote and eukaryote evolvability. *Biosystems* 69, 163–185.
- Pritham, E. J., Putliwala, T., and Feschotte, C. (2007). Mavericks, a novel class of giant transposable elements widespread in eukaryotes and related to DNA viruses. *Gene* 390, 3–17.
- Puigbo, P., Wolf, Y. I., and Koonin, E. V. (2009). Search for a Tree of Life in the thicket of the phylogenetic forest. *J. Biol.* 8, 59.
- Puigbo, P., Wolf, Y. I., and Koonin, E. V. (2010). The tree and net components of prokaryote evolution. *Genome Biol. Evol.* 2, 745–756.
- Puigbo, P., Wolf, Y. I., and Koonin, E. V. (2012). Genome-wide comparative analysis of phylogenetic trees: the prokaryotic forest of life. *Methods Mol. Biol.* 856, 53–79.
- Rainey, P. B., Beaumont, H. J., Ferguson, G. C., Gallie, J., Kost, C., Libby, E., and Zhang, X. X. (2011). The evolutionary emergence of stochastic phenotype switching in bacteria. *Microb. Cell Fact.* 10(Suppl. 1), S14.
- Raoult, D. (2009). There is no such thing as a tree of life (and of course viruses are out!). *Nat. Rev. Microbiol.* 7, 615.
- Raoult, D. (2010). The post-Darwinist rhizome of life. *Lancet* 375, 104–105.
- Raoult, D., Audic, S., Robert, C., Abergel, C., Renesto, P., Ogata, H., La Scola, B., Suzan, M., and Claverie, J. M. (2004). The 1.2-megabase genome sequence of Mimivirus. *Science* 306, 1344–1350.
- Raoult, D., and Forterre, P. (2008). Redefining viruses: lessons from Mimivirus. *Nat. Rev. Microbiol.* 6, 315–319.
- Rohwer, F., and Thurber, R. V. (2009). Viruses manipulate the marine environment. *Nature* 459, 207–212.
- Rose, M. R., and Oakley, T. H. (2007). The new biology: beyond the Modern Synthesis. *Biol. Direct* 2, 30.
- Rosenberg, S. M., and Hastings, P. J. (2003). Microbiology and evolution. Modulating mutation rates in the wild. *Science* 300, 1382–1383.
- Sicheritz-Ponten, T., and Andersson, S. G. (2001). A phylogenomic approach to microbial evolution. *Nucleic Acids Res.* 29, 545–552.
- Simpson, G. G. (1944). *Tempo and Mode in Evolution*. New York, NY: Columbia University Press.
- Smillie, C. S., Smith, M. B., Friedman, J., Cordero, O. X., David, L. A., and Alm, E. J. (2011). Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* 480, 241–244.
- Smith, J. M., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993). How clonal are bacteria? *Proc. Natl. Acad. Sci. U.S.A.* 90, 4384–4388.
- Snel, B., Bork, P., and Huynen, M. A. (1999). Genome phylogeny based on gene content. *Nat. Genet.* 21, 108–110.
- Snel, B., Bork, P., and Huynen, M. A. (2002). Genomes in flux: the evolution of archaeal and proteobacterial gene content. *Genome Res.* 12, 17–25.
- Snel, B., Huynen, M. A., and Dutilh, B. E. (2005). Genome trees and the nature of genome evolution. *Annu. Rev. Microbiol.* 59, 191–209.
- Srinivasiah, S., Bhavsar, J., Thapar, K., Liles, M., Schoenfeld, T., and Wommack, K. E. (2008). Phages across the biosphere: contrasts of viruses in soil and aquatic environments. *Res. Microbiol.* 159, 349–357.
- Stanier, R. Y., and Van Niel, C. B. (1962). The concept of a bacterium. *Arch. Mikrobiol.* 42, 17–35.
- Stern, A., and Sorek, R. (2011). The phage-host arms race: shaping the evolution of microbes. *Bioessays* 33, 43–51.
- Strassmann, J. E., Page, R. E. Jr., Robinson, G. E., and Seeley, T. D. (2011). Kin selection and eusociality. *Nature* 471, E5–E6. author reply E9–E10.
- Suttle, C. A. (2005). Viruses in the sea. *Nature* 437, 356–361.
- Suttle, C. A. (2007). Marine viruses—major players in the global ecosystem. *Nat. Rev. Microbiol.* 5, 801–812.
- Sutton, M. D., Smith, B. T., Godoy, V. G., and Walker, G. C. (2000). The SOS response: recent insights into umuDC-dependent mutagenesis and DNA damage tolerance. *Annu. Rev. Genet.* 34, 479–497.
- Syvanen, M. (1985). Cross-species gene transfer; implications for a new theory of evolution. *J. Theor. Biol.* 112, 333–343.
- Syvanen, M. (1994). Horizontal gene transfer: evidence and possible consequences. *Annu. Rev. Genet.* 28, 237–261.

- Syvanen, M., and Kado, C. I. (2002). *Horizontal Gene Transfer*. San Diego, CA: Academic Press.
- Szathmary, E., and Maynard Smith, J. (1997). From replicators to reproducers: the first major transitions leading to life. *J. Theor. Biol.* 187, 555–571.
- Takeuchi, N., and Hogeweg, P. (2009). Multilevel selection in models of prebiotic evolution II: a direct comparison of compartmentalization and spatial self-organization. *PLoS Comput. Biol.* 5:e1000542. doi: 10.1371/journal.pcbi.1000542
- Takeuchi, N., and Hogeweg, P. (2012). Evolutionary dynamics of RNA-like replicator systems: a bioinformatic approach to the origin of life. *Phys. Life Rev.* doi: 10.1016/j.plrev.2012.06.001. [Epub ahead of print].
- Takeuchi, N., Hogeweg, P., and Koonin, E. V. (2011). On the origin of DNA genomes: evolution of the division of labor between template and catalyst in model replicator systems. *PLoS Comput. Biol.* 7:e1002024. doi: 10.1371/journal.pcbi.1002024
- Tax, S., and Callender, C. (1960). *Evolution after Darwin; the University of Chicago Centennial*. Chicago, IL: University of Chicago Press.
- van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., and Brouns, S. J. (2009). CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* 34, 401–407.
- Van Melderen, L., and Saavedra De Bast, M. (2009). Bacterial toxin-antitoxin systems: more than selfish entities? *PLoS Genet.* 5:e1000437. doi: 10.1371/journal.pgen.1000437
- Veening, J. W., Smits, W. K., and Kuipers, O. P. (2008a). Bistability, epigenetics, and bet-hedging in bacteria. *Annu. Rev. Microbiol.* 62, 193–210.
- Veening, J. W., Stewart, E. J., Berngruber, T. W., Taddei, F., Kuipers, O. P., and Hamoen, L. W. (2008b). Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4393–4398.
- Wallace, A. R. (1858). On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural mean of selection. III. On the tendency of varieties to depart indefinitely from the original type. *J. Proc. Linn. Soc. London* 3, 53–62.
- Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturino, J. M., and Wood, T. K. (2010). Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1, 147.
- Weinbauer, M. G., and Rassoulzadegan, F. (2004). Are viruses driving microbial diversification and diversity? *Environ. Microbiol.* 6, 1–11.
- Wellner, A., Lurie, M. N., and Gophna, U. (2007). Complexity, connectivity, and duplicability as barriers to lateral gene transfer. *Genome Biol.* 8, R156.
- Woese, C. R. (1987). Bacterial evolution. *Microbiol. Rev.* 51, 221–271.
- Woese, C. R. (2004). The Archaeal Concept and the World it Lives in: A Retrospective. *Photosynth. Res.* 80, 361–372.
- Woese, C. R., and Fox, G. E. (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5088–5090.
- Woese, C. R., and Goldenfeld, N. (2009). How the microbial world saved evolution from the scylla of molecular biology and the charybdis of the modern synthesis. *Microbiol. Mol. Biol. Rev.* 73, 14–21.
- Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci. U.S.A.* 87, 4576–4579.
- Woese, C. R., Olsen, G. J., Ibba, M., and Soll, D. (2000). Aminoacyl-tRNA synthetases, RNA, the genetic code, and the evolutionary process. *Microbiol. Mol. Biol. Rev.* 64, 202–236.
- Wolf, Y. I., Aravind, L., Grishin, N. V., and Koonin, E. V. (1999). Evolution of aminoacyl-tRNA synthetases—analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* 9, 689–710.
- Wolf, Y. I., Rogozin, I. B., Grishin, N. V., and Koonin, E. V. (2002). Genome trees and the tree of life. *Trends Genet.* 18, 472–479.
- Yutin, N., Puigbo, P., Koonin, E. V., and Wolf, Y. I. (2012). Phylogenomics of prokaryotic ribosomal proteins. *PLoS ONE* 7:e36972. doi: 10.1371/journal.pone.0036972
- Zhang, Y., Laing, C., Steele, M., Ziebell, K., Johnson, R., Benson, A. K., Taboada, E., and Gannon, V. P. (2007). Genome evolution in major *Escherichia coli* O157, H7 lineages. *BMC Genomics* 8, 121.
- Zhaxybayeva, O. (2009). Detection and quantitative assessment of horizontal gene transfer. *Methods Mol. Biol.* 532, 195–213.
- Zhaxybayeva, O., and Doolittle, W. F. (2011). Lateral gene transfer. *Curr. Biol.* 21, R242–R246.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 August 2012; accepted: 27 August 2012; published online: 13 September 2012.

Citation: Koonin EV and Wolf YI (2012) Evolution of microbes and viruses: a paradigm shift in evolutionary biology? *Front. Cell. Inf. Microbio.* 2:119. doi: 10.3389/fcimb.2012.00119

Copyright © 2012 Koonin and Wolf. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.