Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes

Chuan Ku[a,1], Shijulal Nelson-Sathi[a,1], Mayo Roettger[a], Sriram Garg[a], Einat Hazkani-Covo[b], and William F. Martin[a,2]

[a]Institute of Molecular Evolution, Heinrich Heine University, 40225 Düsseldorf, Germany; and [b]Department of Natural and Life Sciences, The Open University of Israel, Ra'anana 43107, Israel

Edited by John P. McCutcheon, University of Montana, Missoula, MT, and accepted by the Editorial Board February 5, 2015 (received for review December 12, 2014)

Endosymbiotic theory in eukaryotic-cell evolution rests upon a foundation of three cornerstone partners—the plastid (a cyanobacterium), the mitochondrion (a proteobacterium), and its host (an archaeon)—and carries a corollary that, over time, the majority of genes once present in the organelle genomes were relinquished to the chromosomes of the host (endosymbiotic gene transfer). However, notwithstanding eukaryote-specific gene inventions, single-gene phylogenies have never traced eukaryotic genes to three single prokaryotic sources, an issue that hinges crucially upon factors influencing phylogenetic inference. In the age of genomes, single-gene trees, once used to test the predictions of endosymbiotic theory, now spawn new theories that stand to eventually replace endosymbiotic theory with descriptive, gene tree-based variants featuring supernumerary symbionts: prokaryotic partners distinct from the cornerstone trio and whose existence is inferred solely from single-gene trees. We reason that the endosymbiotic ancestors of mitochondria and chloroplasts brought into the eukaryotic—and plant and algal—lineage a genome-sized sample of genes from the proteobacterial and cyanobacterial pangenomes of their respective day and that, even if molecular phylogeny were artifact-free, sampling prokaryotic pangenomes through endosymbiotic gene transfer would lead to inherited chimerism. Recombination in prokaryotes (transduction, conjugation, transformation) differs from recombination in eukaryotes (sex). Prokaryotic recombination leads to pangenomes, and eukaryotic recombination leads to vertical inheritance. Viewed from the perspective of endosymbiotic theory, the critical transition at the eukaryote origin that allowed escape from Muller's ratchet—the origin of eukaryotic recombination, or sex—might have required surprisingly little evolutionary innovation.

endosymbiosis | evolution | mitochondria | lateral gene transfer | plastids

The origin of eukaryotes was one of life's major evolutionary transitions (1, 2). Despite much progress in recent years, the issue is far from being resolved to everyone's satisfaction. There is broad agreement that the last eukaryotic common ancestor (LECA) possessed numerous features that are lacking in prokaryotes, including a mitochondrion, a nucleus, an extensive endomembrane traffic system, meiosis, sex, spliceosomal introns, a eukaryotic flagellum, a cytoskeleton, and the like (2, 3). The order of events that gave rise to those attributes is still debated (3–5), as are issues concerning (*i*) the number and nature of prokaryotic partners that were involved in eukaryotic symbioses, (*ii*) the role of gene transfers from the ancestral mitochondrion, and (*iii*) the possible role of lateral gene transfer (LGT) from donors that were distinct from the mitochondrial (or plastid) endosymbiont, or its host.

Three recent developments have shed new light on the problem of eukaryote origins. The first is the insight that the host for the origin of eukaryotes is now best understood as a garden-variety archaeon, one that branches within the diversity of known archaeal lineages (4, 6–9). An origin of the host from within the TACK superphylum (4, 7, 9) is the position most widely discussed at present, but the TACK superphylum was itself only recently recognized through the discovery of new archaeal lineages (7). It is possible that, as new archaeal lineages become discovered, the phylogenetic arrangement of eukaryotes and archaea might undergo further adjustments still (10).

A second development is the recognition that the origin of the roughly 2,000 gene families that underpinned the origin of eukaryotic-specific traits in the eukaryote ancestor required the biochemical power of internalized bioenergetic membranes that mitochondria provided (3). Mitochondria, not oxygen, made the energetic difference that separates eukaryotes from prokaryotes. That is because anaerobic mitochondria generate about five ATP per glucose and fermentations in eukaryotes generate two to four ATP per glucose (11), such that the meager 5- to 10-fold increase in ATP yield per glucose conferred by oxygen respiration is dwarfed by the 10^4 to 10^5 increase in ATP yield per gene manifest in cells with mitochondria (3). The key to the orders of magnitude increase in energy available for evolutionary invention that mitochondria conferred is the eukaryotic configuration of internal, compartmentalized bioenergetic membranes relative to genes (3, 5). After all, had oxygen been the key to eukaryote complexity, *Escherichia coli* would have become eukaryotic for the same reason. Furthermore, eukaryotic aerobes and anaerobes interleave across eukaryote phylogeny (11), and bioenergetics point to a mitochondrion ancestor with a facultatively anaerobic lifestyle (12). Only those cells became complex that experienced the increased energy per gene afforded by mitochondria, and the long puzzling lack of true intermediates in the prokaryote–eukaryote transition has a bioenergetic cause (3).

A third, and more involved, development is the recognition of genomic chimerism in eukaryotes (13), an issue that has been brewing for some time (13–22). Genome analyses showed that genes of bacterial origin outnumber genes of archaeal origin in yeast (21) and other eukaryotic genomes (23, 24) by a factor of about 3:1 and that roughly 15–20% of the nuclear genes in photosynthetic eukaryotes are acquisitions attributable to the endosymbiotic origin of plastids from cyanobacteria (25–27).

However, many of the gene acquisitions in photosynthetic eukaryotes do not trace, in gene trees, directly to a cyanobacterial, and thus obviously plastid, origin. Fewer still among the threefold

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Symbioses Becoming Permanent: The Origins and Evolutionary Trajectories of Organelles," held October 15–17, 2014, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/Symbioses.

Author contributions: C.K., S.N.-S., M.R., and W.F.M. analyzed data; and C.K., S.N.-S., M.R., S.G., E.H.-C., and W.F.M. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. J.P.M. is a guest editor invited by the Editorial Board.

[1]C.K. and S.N.-S. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: bill@hhu.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1421385112/-/DCSupplemental.

excess of bacterial genes over archaeal genes in eukaryote genomes trace directly via gene trees to proteobacteria. The excess of bacterial genes in eukaryotes continues to generate new thoughts, new explanations, and debate. There are several different schools of thought on the issue of how the excess of bacterial genes in eukaryotes is best explained. Eukaryotic gene acquisitions from resident organelles (plastids and mitochondria), lateral gene transfers from casual bacterial acquaintances, and pitfalls of inferring eukaryotic gene origins from gene trees alone stand in the foreground.

Unexpected Bacterial Genes in Eukaryotic Genomes

Efforts to explain bacterial genes in eukaryotes that have unexpected branching patterns often involve "supernumerary symbionts," hypothetical cellular partners that are distinct from the mitochondrion or its host but that donated genes to eukaryotes as the only remnant of their ephemeral existence. This idea probably goes back to Zillig et al. (28), who found genes of bacterial origin in *Giardia* long before anyone suspected that it possessed reduced mitochondria (29). Zillig et al. suggested that such genes betray the existence of a bacterial symbiont *incertae sedis* that preceded the origin of mitochondria and that brought extra bacterial genes into the eukaryotic lineage. Gupta and Golding (17) reasoned similarly, as did others (30, 31), who favored the view that the nucleus was an archaeal endosymbiont, which the extra bacterium engulfed, and which became the nucleus. Supernumerary symbionts were thus allied with endosymbiotic theory, but with an important twist that all of the genes that branched "unexpectedly" were attributed to the same supernumerary donor, whereby the expectations were too seldom spelled out (19).

Another school invokes gene acquisition from "food bacteria" (32): that is, the ancestral eukaryote was a phagotroph (33) that fed on bacteria and occasionally incorporated genes so ingested. A different suggestion has it that eukaryotes and archaea are directly descended from actinobacteria, but that the cause of higher sequence similarity in eukaryote–bacterial comparisons stems from cataclysmic elevation of the substitution rate in archaea, which are however suggested to have arisen about 800 My ago (33), despite evidence that archaea are far more ancient (34). De Duve argued that the host for the origin of mitochondria was a bacterium, the archaeal genes (and ribosomes) of eukaryotes having been acquired via LGT from archaea (35). More recent is Gray's "premitochondrial hypothesis" (36), which posits that mitochondrial proteins that do not branch with alphaproteobacterial homologues are relicts from a premitochondrion that existed in the host, although no suggestion is offered for why the host had bacterial genes to begin with (they are just "there"), nor is the existence or origin of bacterial proteins in the eukaryotic cytosol addressed.

Similar to the situation for the eukaryote common ancestor, the plant lineage was also found to harbor many nuclear genes whose gene distributions—shared only by plants and prokaryotes—strongly suggest that they are acquisitions via endosymbiotic gene transfer from the plastid ancestor even though they do not all branch with cyanobacteria in phylogenetic trees (25, 37). Other suggestions have appeared in the literature to address the excess plant-specific bacterial genes. The shopping bag model (38) was introduced to explain the observation that plant nuclear genes acquired from plastids do not all branch with the same cyanobacterial donor (25). In a nutshell, the shopping bag model invokes a different donor bacterium for every gene that does not branch as expected although the expectation is not explicitly formulated. In that respect it is similar to Doolittle's food bacteria theory (32) for eukaryotic heterotrophs. At the same time, it entails a distinctly gradualist view of endosymbiotic theory: that is, the gradual accumulation of genes in preparation for obtaining a plastid, such that the actual acquisition of a plastid was a small final step in a long process preparing the host for its endosymbiont, an element that is also contained in Gray's premitochondrion theory (36). A problem with the shopping bag model is that acquired nuclear genes for plastid functions are quite useless for a host that has neither a plastid nor a TIC/TOC protein routing machinery to direct nuclear encoded gene products to the plastid should it finally acquire one, such that gene acquisitions before the acquisition of the plastid itself would hardly have a selectable function and would thus be more likely to be lost than be fixed.

Inherited Chimerism: Cutting Trees a Bit of Slack

As an alternative to supernumerary symbionts, perhaps the too many bacterial genes in eukaryotes are acquisitions, by an archaeal host, via gene transfer from the mitochondrion itself (39), whereby the excess of bacterial genes that do not tend to branch with any bacterial group in particular, including alphaproteobacteria, is best explained as gene acquisitions from the mitochondrion followed by LGT among prokaryotes, in addition to the many technical shortcomings of deep phylogeny (40). In that view, the localization of bacterial proteins in the cytosol of nonphotosynthetic eukaryotes comes mainly from endosymbiotic gene transfer out of the mitochondrion to the host before the origin of a mitochondrial protein import apparatus, giving rise to bacterially related cytosolic proteins encoded by nuclear genes of mitochondrial origin (19, 39, 41). With the advent of the mitochondrial protein import machinery, and some gene tinkering in the nucleus, the same transfer mechanism could also give rise to nuclear encoded mitochondrial proteins. That view, termed here "inherited chimerism," has stressed two main aspects: (*i*) we cannot take single-gene phylogenies that span over a billion years back to the origin of mitochondria (and plastids) at face value; we need to be skeptical of their topologies, especially at the deepest branches; and (*ii*) LGT among prokaryotes complicates things in a manner too seldom appreciated, in that genes acquired via the mitochondrion and the plastid were sequestered in the eukaryotic lineage whereas their homologues in prokaryotes were free to continue undergoing recombination, within and across taxon boundaries (21, 40, 42–44). Pangenomes, which arise from the mechanisms of inheritance in prokaryotes, play an underappreciated role in this issue, as the following brief consideration of recombination in prokaryotes and eukaryotes illustrates.

Prokaryotes vs. Eukaryotes, Pangenomes vs. Lineages

Differences in the mechanisms of inheritance across the prokaryote–eukaryote divide generate, over long time frames, different patterns of variation. In both prokaryotes and eukaryotes, there are clonally propagating species that seem never to undergo recombination. Because mutation is inevitable (45), prokaryotic or eukaryotic species that never undergo recombination will continuously accumulate sublethal mutations, which they cannot purge from their genomes. This process continuously increases genetic load, for which reason they will eventually go extinct, a process known as Muller's ratchet (46–49). Recombination has an important role in evolution in that it rescues genomes from Muller's ratchet.

In prokaryotes, three main mechanisms of recombination introduce new genes or alleles into the genome to counteract Muller's ratchet: conjugation, transduction, and transformation (50), in addition to other mechanisms that are restricted to only some lineages, such as gene-transfer agents (51). Over evolutionary timescales, these mechanisms are superimposed upon the clonal patterns of variation that prokaryotic cell division produces (52), leading to a continuous increase in genome size that eventually must be counterbalanced by gene losses and results in clonally descended clusters of sequences that differ substantially in gene content (Fig. 1*A*). The genes shared by all members of the group are called the core genome, those differentially present across the genomes in question are called the dispensable or

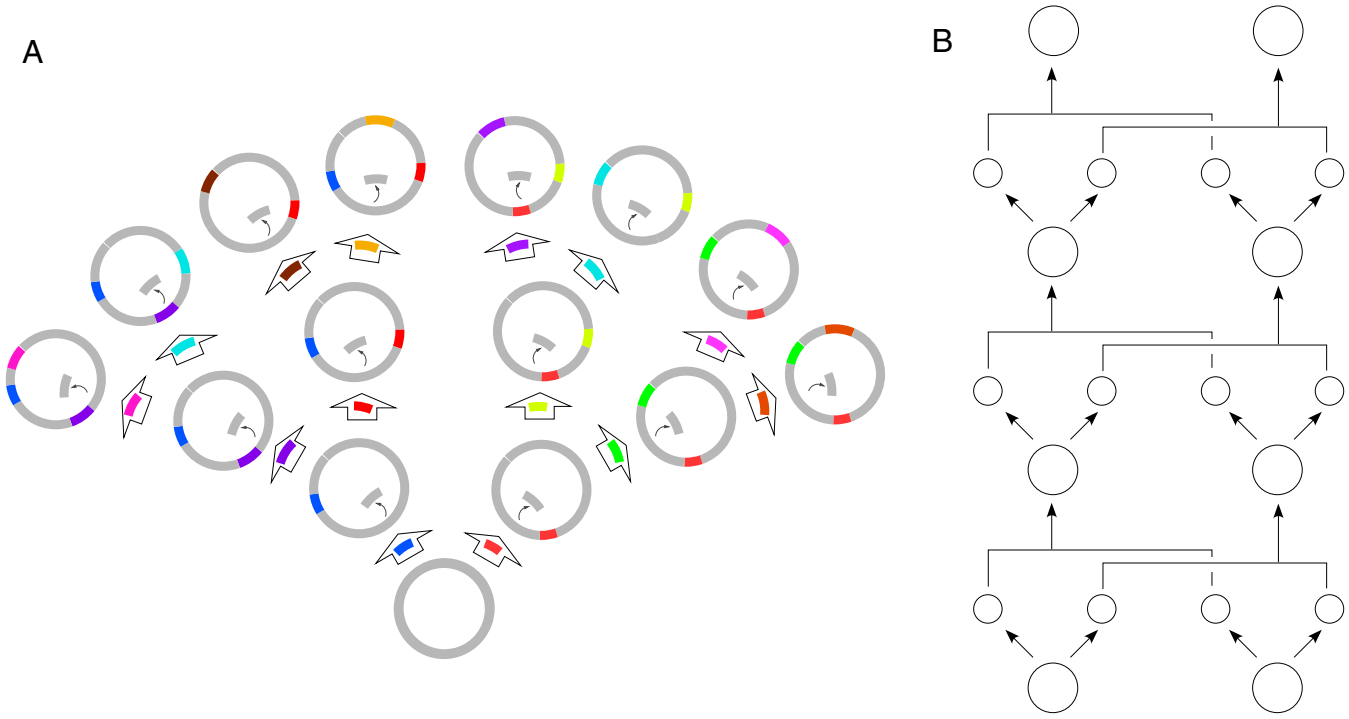

Fig. 1. Recombination and inheritance in prokaryotes and eukaryotes. (*A*) Gene transfer in prokaryotes leads to new genes in different clonally propagating lines. Gene gain (colored segments) is counterbalanced by differential loss. (*B*) Recombination and gamete fusion in eukaryotes (highly schematic) lead to vertically evolving lineages.

accessory genome, and the sum of these components is called the pangenome (53, 54). Importantly, recombination in prokaryotes is not reciprocal, but unidirectional from donor to acceptor, even in archaea that fuse (55). Furthermore, the donor DNA need not come from individuals of the same species; rather, it can come from any taxon or it can even come from dead cells (the environment) (49).

In eukaryotes, the mechanism that counteracts Muller's ratchet is sex. Although there are many variations on the theme (56–59), the underlying principle is that gametes containing different combinations of genes from the same species fuse to produce individuals containing two sets of chromosomes harboring variants (alleles) of the same genes. Meiotic recombination generates new assortments of alleles in the next generation of gametes. Notwithstanding the occasional hybridization, allopolyploidization, or introgression events among closely related species, the process of recombination in eukaryotes produces lineages and patterns that reflect, over geological timescales, vertical descent and new combinations of alleles from within the same gene set (Fig. 1*B*).

It is noteworthy that the mechanisms of recombination in prokaryotes are simultaneously the mechanisms of LGT. Their operation upon clonal lineages over time produces pangenomes whereas the mechanisms of recombination in eukaryotes produce lineages with vertical inheritance. LGT in prokaryotes is just natural variation in action, and microbiologists have always known that there was something like a pangenome out there for prokaryotes because they built 70% DNA–DNA hybridization into the species definition (60, 61), fully aware that 70% hybridization meant 70% shared DNA sequences, not 30% sequence divergence (62).

What Do Pangenomes Look Like?

Pangenomes are collections of genes within the species (or within any taxon) that are or are not uniformly or universally distributed across individual genomes (53), as shown in Fig. 2, where we display the distribution of genes for 54 *E. coli* genomes (Fig. 2*A*), 44 cyanobacterial genomes (Fig. 2*B*), and 208 alphaproteobacterial genomes (Fig. 2*C*). Note that the basic nature of the gene distribution is the same at the species and at the phylum or class level, except for larger numbers of genes at the higher levels, which result from the mechanism in Fig. 1*A* working for greater amounts of time.

Fig. 2 shows only how the genes are shared within the taxa whereas Fig. 3 shows how the genes are distributed across taxa, which is also relevant for the issue of inherited chimerism. This effect is seen for cyanobacteria in Fig. 3*A* and for alphaproteobacteria in Fig. 3*B*. The vast majority of genes found either in this sample of cyanobacteria or in this sample of alphaproteobacteria are not specific to the taxonomic group. Rather, they are shared with other groups. However, they are not shared with all other groups because only about 33 protein-coding genes are universal to all genomes (67), the rest being distributed in some manner. How specifically they are distributed goes beyond the scope of this paper, but it is clear that the distributions mainly entail network-like patterns of sharing (68–70), not tree-like patterns of inclusive hierarchy. The point is this: Were we to reenact endosymbiosis today and allow one of the cyanobacteria in Fig. 3*A* to become the plastid, we would be selecting and sequestering a genome-sized sample of the cyanobacterial pangenome. By putting it into the eukaryotic lineage, we would not affect the ability of the genes shared by the new plastid ancestor and other taxa to undergo LGT and reassortment among the free-living species. If we allow many genes to be relocated to the nucleus while the free-living prokaryotes undergo recombination for the next 1.5 billion years (roughly the age of plastid origin) (71), we might end up with the situation we observe for plants today: Many or most genes that came in with our new plastid will not branch with homologs from a particular cyanobacterial lineage, even if our gene phylogeny is artifact-free. We repeat the experiment for one of the alphaproteobacteria in Fig. 3*B*, which becomes our new mitochondrion, but this time we wait for ~1.8 Ga (roughly

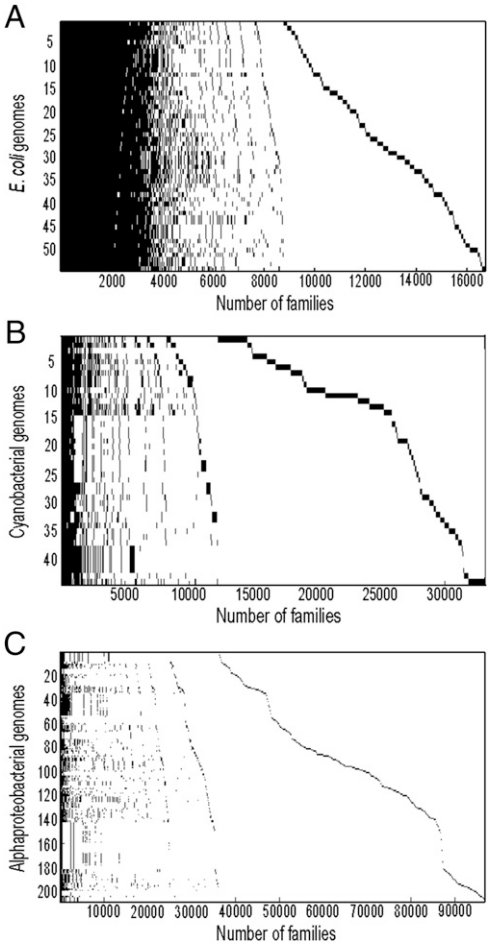

Fig. 2. Bacterial pangenome distribution. The bidirectional best BLAST hit approach (63) was performed on protein sequences of 1,981 complete prokaryotic genomes [see Nelson-Sathi et al. (64) for the full list] that had hits with ≥25% local identity and e-value $<10^{-10}$ in BLAST (65) search. Grouping into protein families was performed using the Markov Chain clustering procedure (66). Patterns of presence (black) and absence (white) of all protein-coding genes are shown. Each genome is represented by a row and gene families by columns. Gene families are sorted in decreasing order (left to right) of their presence in the total genomes. The core genes are shown on the left side and genome-specific genes to the right. (*A*) Distribution of 16,725 genes in 54 *E. coli* genomes, with 8,776 (52.5%) clusters present in at least two genomes and 7,949 (47.5%) unique to individual strains (singletons). Among the singletons, 1,132 genes have at least one homolog in non-*E. coli* species. (*B*) Distribution of 33,118 genes in 44 cyanobacterial genomes, including 12,236 found in at least two genomes and 20,882 singletons. (*C*) A total of 96,916 genes are present in 208 alphaproteobacterial genomes, including 36,176 in at least 2 genomes and 60,740 singletons.

the age of LECA) (71): Many, or even most genes that came in with our new mitochondrion will not branch with a particular alphaproteobacterial lineage, even if our gene trees are free of phylogeny-reconstruction artifacts.

Supernumerary Symbionts or Inherited Chimerism?

Directly from the forest of trees for the excess bacterial genes in eukaryotic genomes, a different category of supernumerary symbionts has emerged that might be called supernumerary phylobionts because their existence is inferred exclusively from phylogenetic trees—trees in which the nearest neighbor of a eukaryotic gene is inferred as the donor. Phylobionts arise directly from observations in gene trees, without independent evidence, and as such their existence and nature are subject to all of the vagaries of phylogenetic methods and lineage sampling. Examples of supernumerary phylobionts include the idea of a supernumerary chlamydiae symbiont that has been repeatedly claimed to have helped the cyanobacterial ancestor of plastids to make the transition from endosymbiont to organelle (72), or various gene-donating bacteria that supposedly helped plants conquer the land (73). The chlamydiae helper symbiont (72, 74) and other hypotheses that summon supernumerary phylobionts from trees are problematic (75, 76)—if we think things through in full, supernumerary phylobionts entail the inference of an additional supernumerary partner for every eukaryotic nuclear gene with prokaryotic homologs, of which there are thousands in eukaryotic genomes (24, 25, 27). As our sample of prokaryotic genomes grows, and as phylogenetic methodologies evolve, it is already evident that, for every eukaryotic gene family, there will eventually be a new and different sister group in phylogenetic trees, and each tree could give rise to some story. In the framework of supernumerary phylobionts, this reasoning will lead to thousands of individual gene donors to the eukaryotic ancestor and the archaeplastidan ancestor. That proposition is untenable. How so? An example illustrates.

What would happen if we were to use the same methodology—single-gene trees—as people have been using to infer the origins of eukaryotic nuclear genes to infer the origin of genes that are still present in the mitochondrion or the plastid? To see, we constructed alignments and single-gene maximum likelihood trees (see *SI Text* for the detailed methods) for those 51 (out of 67) protein-coding genes from the *Reclinomonas americana* mitochondrial genome (77) that are sufficiently well-conserved to make trees and the best conserved 183 out of 209 protein-coding genes in the *Porphyra purpurea* plastid genome (78) in the context of 1,981 prokaryotic genomes (64). The results (Dataset S1 and Figs. S2 and S3) show that, for *Reclinomonas*, 43 different sister groups were obtained, and, in 20 cases, the mitochondrial sister group differs in trees based on the forward and reverse alignments (79) using the same algorithm (Fig. S2). For the *Porphyra* plastid proteins, 124 different sister groups were obtained, and, in 52 cases, the plastid sister group is different in the reverse-alignment trees (Fig. S3).

Using the logic germane to supernumerary phylobiont inference, the findings in Dataset S1 and Figs. S2 and S3 would be interpreted as evidence that neither the mitochondrion nor the plastid arose via endosymbiosis; rather, each would be the product of 43 and 124 independent gene transfers, respectively, from different donors, thus one at a time, to the eukaryotic ancestor and the archaeplastidan ancestor, but the transfers would have to be directed to some kind of preexisting compartment, not dissimilar to Gray's premitochondrion, where rRNA operons and tRNAs also became donated, enabling the result of such transfer to morph into a bioenergetic organelle, but only mimicking a bona fide endosymbiotic origin, the real mechanism being LGT: So say the single-gene trees. We say: That scenario cannot possibly be true. However, why can it not be true? It cannot be true because exactly the same kinds of transfers—one at a time and from independent donors—for exactly the right kinds of genes to support the function of the bioenergetic membrane in mitochondria and the bioenergetic membrane in plastids (in addition to the other biochemical and physiological functions of the organelles) would have to be going on to the nucleus as well, the crux being that, until the whole organelle is assembled through such imaginary LGT, none of the transferred genes have a selectable function. Without selection for function, they would all become pseudogenes, and no organelle would emerge at all. A free-living prokaryote brings along the complete and selectable functional unit, which can then be transferred a chunk at a time to the host, but from a continuously selected and replicating functional source. There is something very wrong with the supernumerary phylobiont stories, and the core of the problem is rooted in trees.

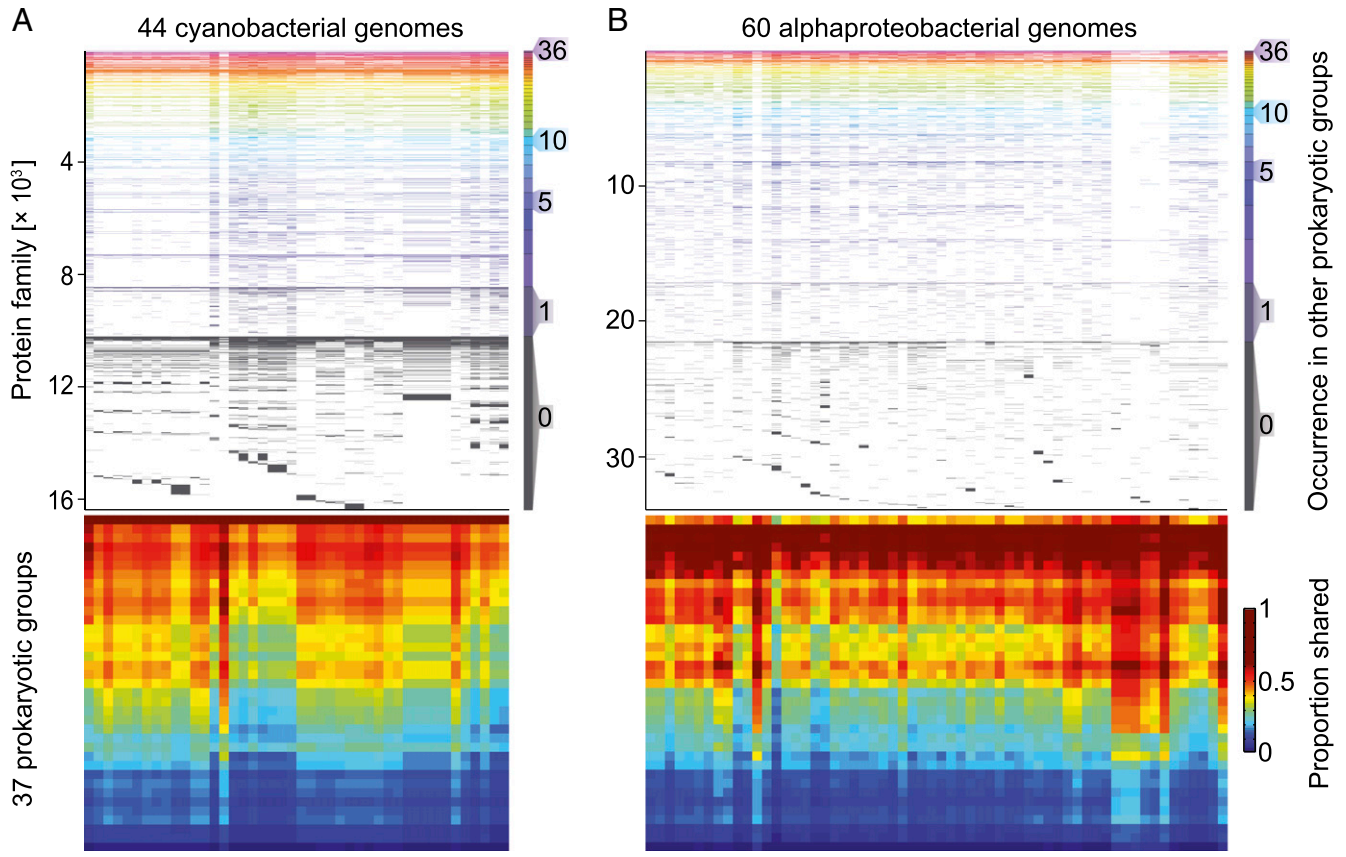

Fig. 3. Gene sharing among prokaryotes. (*A*, *Upper*) Presence/absence of protein families in cyanobacterial genomes, sorted according to the number of taxonomic groups sharing the corresponding gene. (*Lower*) Proportion of genes in each cyanobacterial genome shared with other taxonomic groups. *k*-means clustering was applied to sort taxonomic groups according to pattern similarity. (*B*) Showing the same patterns as in *A* for alphaproteobacteria. (*Lower*) Taxonomic groups are sorted as in *A*. This figure is based on the nonsingleton clusters from those described in Fig. 2. For the complete figure with taxon labels, see Fig. S1.

We have to relax our expectations regarding the ability of single-gene trees to provide a crayon with which we can draw eukaryotic genome history. If we take gene trees at face value, we would have to reject the proposition that plastids and mitochondria descend via endosymbiosis from free-living prokaryotes in favor of a biochemically untenable view of single-gene assembly based on LGTs inferred from gene trees. Endosymbiosis is clearly the better supported alternative, whereby inherited chimerism is a corollary whose function is to help explain odd branches in gene trees so that we do not throw out the baby (endosymbiotic theory) with the bathwater (gene trees). Concatenation is the answer, some might say, but concatenation for prokaryotic genes is very problematic (80), and, if we do combine the eukaryotic trees into categories justifiable by tree-independent methods, what we find is evidence for a plastid, a mitochondrion, and an archaeal host (81).

Where Can We Go Wrong with Trees and Where Will It Lead Us?

Asking all genes that came into the eukaryotic lineage via the mitochondrial and plastid symbiosis, respectively, to branch with homologs from one and the same present-day proteobacterial and one and the same present-day cyanobacterial genome is simply asking too much. If we adopt a vertical, static view of prokaryotic genome evolution, where genes in a prokaryotic lineage can be passed down only within the lineage, then a tree of an endosymbiotically acquired gene would always show a prokaryotic sister group to eukaryotes that consists of only taxa from the lineage to which the organelle belongs, the true donor lineage (Fig. 4*A*). Because of LGT among prokaryotes, however, the prokaryotic homologs of eukaryotic genes almost never show the prokaryotic groups to be monophyletic (Fig. 4*B* and table 1 from ref. 24). Add to that gene loss [which has to be as common in gene evolution as LGT; otherwise genomes would constantly be expanding (82, 83)] and incomplete prokaryote genome sampling, which results in a tree where the prokaryotic sister group is sparsely populated by (Fig. 4*C*) or sometimes even without any representative from the true sister group (Fig. 4*D*). The gene donors we infer from trees today are thus ephemeral (Fig. 4 *B*–*E*). For example, the first plant–chlamydiae gene connection was the plastid ATP/ADP translocase (84), which, until 2007 (72), was found only in Rickettsiales, and sparked heated debate on its origin (85). As of November 2014, the *Arabidopsis* plastid ATP/ADP translocase (NP_173003) detects homologs in alphaproteobacteria outside Rickettsiales, in beta-, gamma- and deltaproteobacteria, and in bacteroidetes (Table S1). In 2021, there will be more. These factors (Fig. 4) are sufficient to generate patterns of apparent transfer from prokaryotes to eukaryotes, not to mention tree-building artifacts (38) that can also produce trees showing apparent gene transfer (86). By ignoring such factors, and by naively believing trees at face value, a view is emerging that LGT, not endosymbiosis, is the main mechanism behind the origin of plastids (87, 88). Should we believe that?

Gene Transfer from Organelles to the Nucleus: At Least It's Real

If LGT from prokaryotes to eukaryotes were really as common in genome evolution as such studies would have us believe, then eukaryotic chromosomes should be replete with recently acquired

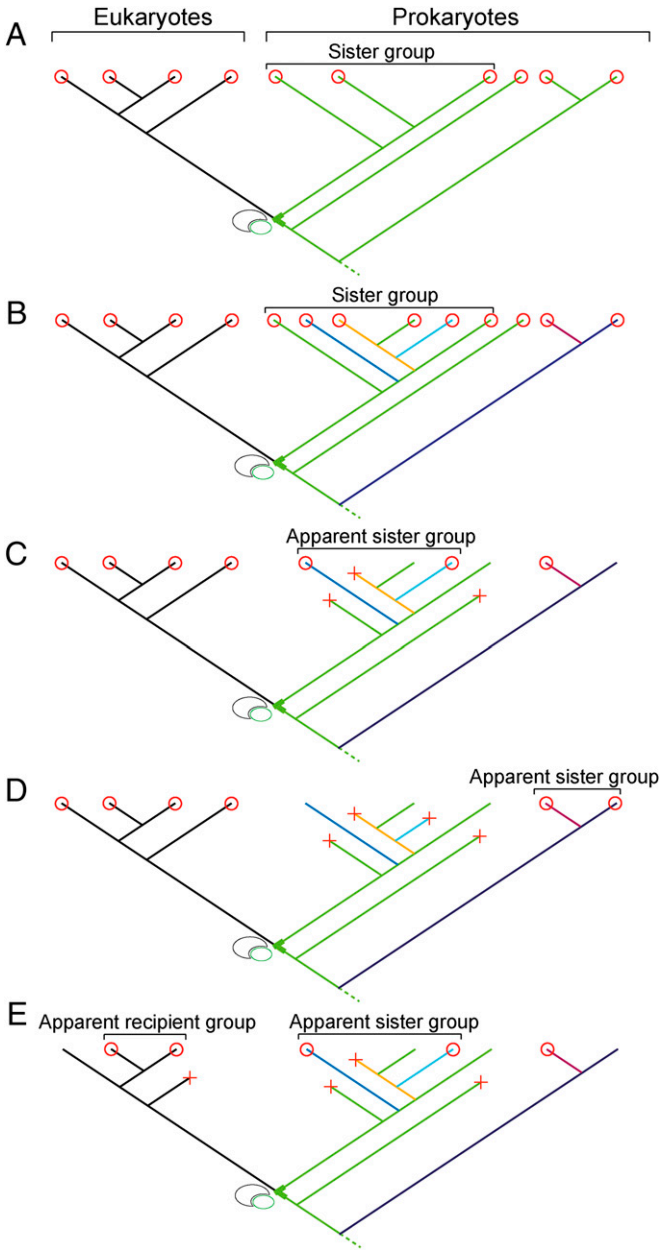

Fig. 4. Histories hidden behind trees. (*A*) In an ideal tree of a gene acquired endosymbiotically from a donor prokaryotic lineage (green), eukaryotes (black) should be nested within present-day representatives of that lineage. (*B*) LGT among prokaryotes results in a prokaryotic sister group consisting of homologs from both donor and nondonor lineages (nongreen colors). Further complicated by gene loss (crosses) and incomplete sampling (only circled homologs are sampled and used for phylogenetic analyses), the sister group observed in the tree is an apparent one that is a subsample of the complete sister group (*C*) or does not contain any representative from the true sister group (*D*). (*E*) The same factors also influence sampling of eukaryotic homologs, resulting in an apparent acquisition of the gene by a subgroup of the eukaryotic clade involved in the endosymbiosis event.

bacterial DNA. However, bona fide recent bacterial gene acquisitions are very rare, and most—but not all—of the bacterial sequences that are reported in genome-sequencing projects are ultimately removed from the databases because they are contaminations from the genome-sequencing process. Important exceptions are the genomes of phloem-feeding insects, which are regularly found to harbor insertions of bacterial DNA that stems from the obligate bacterial endosymbionts that grow in the bacteriome, a specialized organ that houses the symbionts, which provide essential functions to their host, most commonly amino acid biosynthesis. Genome sequences of pea aphids (89), mealybugs (90), psyllids (91), and invertebrates infected by *Wolbachia* (92) have revealed DNA segments that have been integrated from endosymbionts. However, such recent DNA transfers from bacteria are generally quite rare in eukaryotes, which is probably why they get so much attention when such verified cases are reported.

By comparison, the transfer of DNA from organelles to the nuclear genome is ubiquitous among eukaryotic genomes. DNA transfer from organelles to the nucleus occurs in all eukaryote genomes studied to date (93). *Numts*, for nuclear mitochondrial DNA copies (and *nupts* for the plastid) (94), are typical components of eukaryotic genomes (93–95) whereas segments of bacterial chromosomes are not. For example, our genomes harbor 53 *numts* that are specific to the human lineage (96), with 12 *numts* that are polymorphic in human populations (93), and more *numts* continuously being found in the human 1,000 Genomes data (97). Five human *numts* are associated with disease (93), one of which involves a 72-bp numt insertion into exon 14 of the GLI3 gene, causing a premature stop codon, in a rare case of Pallister–Hall syndrome stemming from the Chernobyl incident (98). No human genomes are (yet) known to be polymorphic for recent bacterial DNA insertions.

The mechanism of gene transfer from organelles to the nucleus entails the incorporation of bulk organelle DNA into nuclear chromosomes. Very large copies can be inserted, as the 262-kb mtDNA of *Arabidopsis* (99.91% identical) and the 131-kb complete rice chloroplast genome (99.77% identical) attest (99), suggesting that, during the early phases of organelle origins, large segments or even whole chromosomes were also being transferred, followed by the normal DNA dynamics of mutation, recombination, fixation, and deletion. *Numts* and *nupts* are inserted into double-strand breaks by the nonhomologous end-joining machinery (100, 101) and enter the genome in open chromatin regions (101, 102). *Numts* can be integrated into chromosomes with a short microhomology of 1–7 bp, implicating a submechanism of nonhomologous end joining known as microhomology-mediated repair (103), but insertion can also occur without microhomology—a process known as blunt-end repair.

Analysis of 90 recent *numt* insertions in human and chimpanzee suggests that 35% of the fusion points involve microhomology of at least 2 bp; thus, it seems that repair involving microhomology plays some role in *numt* integration but is not strictly required (103). No analyses of recent insertions of bacterial DNA into the human and chimpanzee lineages have been reported. Notwithstanding the cases of plant-feeding insects and their tightly associated bacteria, why we do not observe recent bacterial transfers, as we do for *numts* and *nupts*? And if all of the prokaryote-to-eukaryote LGT reports are real, then, at some point, we need to see evidence for its long-term effects in terms of different lineages of eukaryotes harboring fundamentally different collections of genes, as we see in prokaryotes (64). However, except for photosynthetic eukaryotes, which acquired the plastid and many genes with it, different eukaryotic lineages tend to possess the very same collections of genes having prokaryotic homologs, which is not true for prokaryotes (Fig. 1). We are saying that prokaryotes recombine via LGT but that eukaryotes have remained genetically isolated from prokaryotes (except at the origins of organelles) because they recombine via sex. Our critics will thus ask: Where did sex come from?

Did Sex Rescue the Ancestral Eukaryote from Muller's Ratchet?

Like eukaryotes, the origin of sex also counts as one of the major evolutionary transitions (1) and remains one of evolutionary biology's toughest problems. Existing theories seek the origin of sex

in a haploid cell with fully fledged eukaryotic mitosis (104), but it is more likely that mitosis and sex arose in a cell that had a mitochondrion (3, 5). During the prokaryote-to-eukaryote transition, eukaryotes seem to have lost the standard mechanisms that prokaryotes use to escape Muller's ratchet—transduction, transformation, and conjugation—because they are lacking in all eukaryotic groups. Had eukaryotes retained one or all three of those mechanisms, it seems unlikely that they would have evolved sex on top of them, and, indeed, cells that never had mitochondria (prokaryotes) never evolved sex. The machinery involved in eukaryotic recombination was surely present at the time of mitochondrial symbiosis because the main enzymes involved are homologous to their prokaryotic counterparts: Spo11, Mre11, Dmc1, Rad51, Mlh1, and Pms1 (105, 106). Did a simple form of eukaryotic recombination, catalyzed by enzymes that are homologous to the enzymes of prokaryotic recombination, rescue nascent eukaryotes from Muller's ratchet? The basic machinery required might have been a property of the host. It is a curiously underpublicized observation that various archaea can fuse their cells (55, 107) and that, in some haloarchaea, fusion is accompanied by recombination (108) whereas, in others, only recombination is observed (109). One needs to be careful not to (over-)state that "archaea have sex," but, in some rare documented examples, they do undergo outright cell fusion (an otherwise curious property of gametes) and, in some rarer cases, recombination and fusion are observed (108).

Thus, it could be that the essentials of the machinery required for sex—fusion of cells from the same species and ability to generate recombinants in fused cells—was present in the host lineage that acquired the mitochondrion. Without such a capability, extinction would have been the alternative. That suggestion would help to ease one more evolutionary transition in the origin of eukaryotes (sex), which would go a long way toward explaining the differences between inheritance in prokaryotes and eukaryotes (Fig. 1), without solving the problem in full or explaining (*i*) how mitosis and meiosis are related to one another, (*ii*) where the cell cycle comes from, or (*iii*) why eukaryotes, in contrast to all prokaryotes, shut down their gene expression at cell division. Such longstanding questions concerning the major evolutionary transition at eukaryote origin (1), are arguably more tractable than ever before, given progress concerning the archaeal nature of the host that acquired mitochondria (4, 7, 81).

Conclusion

Inherited chimerism is an alternative to the problematic practice of conjuring up additional, gene-donating symbionts at organelle origins to explain gene trees. It merely requires a selective force to associate the symbiont (either plastid or mitochondrion) to its host so that the endosymbiosis (one cell living within another) can be established and gene transfer from the symbiont can commence. It places no constraints on the collections of genes that the plastid and the mitochondrial symbionts possessed, other than that it needs to be a genome-sized collection, not tens of thousands of genes, and it allows freely for LGT among prokaryotes before the endosymbionts become organelles and afterward. LGT among prokaryotes has received much attention in the past decades. Inherited chimerism incorporates LGT among prokaryotes into endosymbiotic theory.

ACKNOWLEDGMENTS. We thank Eric Bapteste, Nick Lane, and Eörs Szathmáry for input and constructive comments on the text, and we especially thank James O. McInerney for pointing out papers on archaeal recombination and fusion. This work was supported by grants from the European Research Council (Grant 232975) (to W.F.M.) and from the Open University of Israel Research Fund (to E.H.-C.). C.K. is grateful to the Deutscher Akademischer Austauschdienst for a PhD stipend.

1. Szathmáry E, Smith JM (1995) The major evolutionary transitions. *Nature* 374(6519): 227–232.
2. Lane N (2009) *Life Ascending: The Ten Great Inventions of Evolution* (Norton, New York).
3. Lane N, Martin W (2010) The energetics of genome complexity. *Nature* 467(7318): 929–934.
4. Williams TA, Foster PG, Cox CJ, Embley TM (2013) An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* 504(7479):231–236.
5. Lane N (2014) Bioenergetic constraints on the evolution of complex life. *Cold Spring Harb Perspect Biol* 6(5):a015982.
6. Cox CJ, Foster PG, Hirt RP, Harris SR, Embley TM (2008) The archaebacterial origin of eukaryotes. *Proc Natl Acad Sci USA* 105(51):20356–20361.
7. Guy L, Ettema TJG (2011) The archaeal 'TACK' superphylum and the origin of eukaryotes. *Trends Microbiol* 19(12):580–587.
8. Kelly S, Wickstead B, Gull K (2011) Archaeal phylogenomics provides evidence in support of a methanogenic origin of the Archaea and a thaumarchaeal origin for the eukaryotes. *Proc Biol Sci* 278(1708):1009–1018.
9. Koonin EV, Yutin N (2014) The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harb Perspect Biol* 6(4):a016188.
10. Williams TA, Embley TM (2014) Archaeal "dark matter" and the origin of eukaryotes. *Genome Biol Evol* 6(3):474–481.
11. Müller M, et al. (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev* 76(2):444–495.
12. Degli Esposti M (2014) Bioenergetic evolution in proteobacteria and mitochondria. *Genome Biol Evol* 6(12):3238–3251.
13. McInerney JO, O'Connell MJ, Pisani D (2014) The hybrid nature of the Eukaryota and a consilient view of life on Earth. *Nat Rev Microbiol* 12(6):449–455.
14. Martin W, Cerff R (1986) Prokaryotic features of a nucleus-encoded enzyme: cDNA sequences for chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenases from mustard (Sinapis alba). *Eur J Biochem* 159(2):323–331.
15. Martin W, Brinkmann H, Savonna C, Cerff R (1993) Evidence for a chimeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc Natl Acad Sci USA* 90(18):8692–8696.
16. Brinkmann H, Martin W (1996) Higher-plant chloroplast and cytosolic 3-phosphoglycerate kinases: A case of endosymbiotic gene replacement. *Plant Mol Biol* 30(1): 65–75.
17. Gupta RS, Golding GB (1996) The origin of the eukaryotic cell. *Trends Biochem Sci* 21(5):166–171.
18. Brown JR, Doolittle WF (1997) Archaea and the prokaryote-to-eukaryote transition. *Microbiol Mol Biol Rev* 61(4):456–502.
19. Martin W, Schnarrenberger C (1997) The evolution of the Calvin cycle from prokaryotic to eukaryotic chromosomes: A case study of functional redundancy in ancient pathways through endosymbiosis. *Curr Genet* 32(1):1–18.
20. Rivera MC, Jain R, Moore JE, Lake JA (1998) Genomic evidence for two functionally distinct gene classes. *Proc Natl Acad Sci USA* 95(11):6239–6244.
21. Esser C, et al. (2004) A genome phylogeny for mitochondria among α-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol Biol Evol* 21(9): 1643–1660.
22. Lane CE, Archibald JM (2008) The eukaryotic tree of life: Endosymbiosis takes its TOL. *Trends Ecol Evol* 23(5):268–275.
23. Cotton JA, McInerney JO (2010) Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. *Proc Natl Acad Sci USA* 107(40):17252–17255.
24. Thiergart T, Landan G, Schenk M, Dagan T, Martin WF (2012) An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. *Genome Biol Evol* 4(4):466–485.
25. Martin W, et al. (2002) Evolutionary analysis of *Arabidopsis*, cyanobacterial, and chloroplast genomes reveals plastid phylogeny and thousands of cyanobacterial genes in the nucleus. *Proc Natl Acad Sci USA* 99(19):12246–12251.
26. Deusch O, et al. (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol* 25(4):748–761.
27. Dagan T, et al. (2013) Genomes of Stigonematalean cyanobacteria (subsection V) and the evolution of oxygenic photosynthesis from prokaryotes to plastids. *Genome Biol Evol* 5(1):31–44.
28. Zillig W, et al. (1989) Did eukaryotes originate by a fusion event. *Endocyt Cell Res* 6(1):1–25.
29. Tovar J, et al. (2003) Mitochondrial remnant organelles of *Giardia* function in iron-sulphur protein maturation. *Nature* 426(6963):172–176.
30. Horiike T, Hamada K, Kanaya S, Shinozawa T (2001) Origin of eukaryotic cell nuclei by symbiosis of Archaea in Bacteria is revealed by homology-hit analysis. *Nat Cell Biol* 3(2):210–214.
31. Forterre P, Philippe H (1999) Where is the root of the universal tree of life? *BioEssays* 21(10):871–879.
32. Doolittle WF (1998) You are what you eat: A gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet* 14(8):307–311.
33. Cavalier-Smith T (2002) The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int J Syst Evol Microbiol* 52(Pt 1):7–76.
34. Liu Y, Beer LL, Whitman WB (2012) Methanogens: A window into ancient sulfur metabolism. *Trends Microbiol* 20(5):251–258.

35. de Duve C (2007) The origin of eukaryotes: A reappraisal. *Nat Rev Genet* 8(5): 395–403.
36. Gray MW (2014) The pre-endosymbiont hypothesis: A new perspective on the origin and evolution of mitochondria. *Cold Spring Harb Perspect Biol* 6(3):a016097.
37. Rujan T, Martin W (2001) How many genes in *Arabidopsis* come from cyanobacteria? An estimate from 386 protein phylogenies. *Trends Genet* 17(3):113–120.
38. Larkum AWD, Lockhart PJ, Howe CJ (2007) Shopping for plastids. *Trends Plant Sci* 12(5):189–195.
39. Martin W, Müller M (1998) The hydrogen hypothesis for the first eukaryote. *Nature* 392(6671):37–41.
40. Martin W (1999) Mosaic bacterial chromosomes: A challenge en route to a tree of genomes. *BioEssays* 21(2):99–104.
41. Martin W (2010) Evolutionary origins of metabolic compartmentalization in eukaryotes. *Philos Trans R Soc Lond B Biol Sci* 365(1541):847–855.
42. Martin WF (1996) Is something wrong with the tree of life? *BioEssays* 18(7):523–527.
43. Martin W, Herrmann RG (1998) Gene transfer from organelles to the nucleus: How much, what happens, and why? *Plant Physiol* 118(1):9–17.
44. Schnarrenberger C, Martin W (2002) Evolution of the enzymes of the citric acid cycle and the glyoxylate cycle of higher plants: A case study of endosymbiotic gene transfer. *Eur J Biochem* 269(3):868–883.
45. Nei M (2013) *Mutation-Driven Evolution* (Oxford Univ Press, Oxford).
46. Muller HJ (1964) The relation of recombination to mutational advance. *Mutat Res* 106(1):2–9.
47. Felsenstein J (1974) The evolutionary advantage of recombination. *Genetics* 78(2): 737–756.
48. Lynch M, Bürger R, Butcher D, Gabriel W (1993) The mutational meltdown in asexual populations. *J Hered* 84(5):339–344.
49. Takeuchi N, Kaneko K, Koonin EV (2014) Horizontal gene transfer can rescue prokaryotes from Muller's ratchet: benefit of DNA from dead cells and population subdivision. *G3 (Bethesda)* 4(2):325–339.
50. Popa O, Dagan T (2011) Trends and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol* 14(5):615–623.
51. Lang AS, Zhaxybayeva O, Beatty JT (2012) Gene transfer agents: Phage-like elements of genetic exchange. *Nat Rev Microbiol* 10(7):472–482.
52. Milkman R (1997) Recombination and population structure in *Escherichia coli*. *Genetics* 146(3):745–750.
53. Medini D, Donati C, Tettelin H, Masignani V, Rappuoli R (2005) The microbial pangenome. *Curr Opin Genet Dev* 15(6):589–594.
54. Young JPW, et al. (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* 7(4):R34.
55. Naor A, Lapierre P, Mevarech M, Papke RT, Gophna U (2012) Low species barriers in halophilic archaea and the formation of recombinant hybrids. *Curr Biol* 22(15): 1444–1448.
56. Bell G (1988) *Sex and Death in Protozoa* (Cambridge Univ Press, Cambridge, UK).
57. Lee SC, Ni M, Li W, Shertz C, Heitman J (2010) The evolution of sex: A perspective from the fungal kingdom. *Microbiol Mol Biol Rev* 74(2):298–340.
58. Parfrey LW, Katz LA (2010) Dynamic genomes of eukaryotes and the maintenance of genomic integrity. *Microbe* 5(4):156–163.
59. Roach KC, Heitman J (2014) Unisexual reproduction reverses Muller's ratchet. *Genetics* 198(3):1059–1069.
60. Goris J, et al. (2007) DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *Int J Syst Evol Microbiol* 57(Pt 1):81–91.
61. Stackebrandt E, Goebel BM (1994) A place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *Int J Syst Bacteriol* 44(4):846–849.
62. Beltz GA, Jacobs KA, Eickbush TH, Cherbas PT, Kafatos FC (1983) Isolation of multigene families and determination of homologies by filter hybridization methods. *Methods Enzymol* 100:266–285.
63. Tatusov RL, Koonin EV, Lipman DJ (1997) A genomic perspective on protein families. *Science* 278(5338):631–637.
64. Nelson-Sathi S, et al. (2015) Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* 517(7532):77–80.
65. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402.
66. Enright AJ, Van Dongen S, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30(7):1575–1584.
67. Hansmann S, Martin W (2000) Phylogeny of 33 ribosomal and six other proteins encoded in an ancient gene cluster that is conserved across prokaryotic genomes: Influence of excluding poorly alignable sites from analysis. *Int J Syst Evol Microbiol* 50(Pt 4):1655–1663.
68. Dagan T, Artzy-Randrup Y, Martin W (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc Natl Acad Sci USA* 105(29):10039–10044.
69. Bapteste E, et al. (2013) Networks: Expanding evolutionary thinking. *Trends Genet* 29(8):439–441.
70. Alvarez-Ponce D, Lopez P, Bapteste E, McInerney JO (2013) Gene similarity networks provide tools for understanding eukaryote origins and evolution. *Proc Natl Acad Sci USA* 110(17):E1594–E1603.
71. Parfrey LW, Lahr DJG, Knoll AH, Katz LA (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci USA* 108(33):13624–13629.
72. Huang J, Gogarten JP (2007) Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol* 8(6):R99.
73. Yue J, Hu X, Sun H, Yang Y, Huang J (2012) Widespread impact of horizontal gene transfer on plant colonization of land. *Nat Commun* 3:1152.
74. Ball SG, et al. (2013) Metabolic effectors secreted by bacterial pathogens: Essential facilitators of plastid endosymbiosis? *Plant Cell* 25(1):7–21.
75. Zimorski V, Ku C, Martin WF, Gould SB (2014) Endosymbiotic theory for organelle origins. *Curr Opin Microbiol* 22C:38–48.
76. Deschamps P (2014) Primary endosymbiosis: Have cyanobacteria and Chlamydiae ever been roommates? *Acta Soc Bot Pol* 83:291–302.
77. Lang BF, et al. (1997) An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. *Nature* 387(6632):493–497.
78. Reith M, Munholland J (1995) Complete nucleotide sequence of the *Porphyra purpurea* chloroplast genome. *Plant Mol Biol Rep* 13(4):333–335.
79. Landan G, Graur D (2007) Heads or tails: A simple reliability check for multiple sequence alignments. *Mol Biol Evol* 24(6):1380–1383.
80. Bapteste E, et al. (2008) Alternative methods for concatenation of core genes indicate a lack of resolution in deep nodes of the prokaryotic phylogeny. *Mol Biol Evol* 25(1):83–91.
81. Pisani D, Cotton JA, McInerney JO (2007) Supertrees disentangle the chimerical origin of eukaryotic genomes. *Mol Biol Evol* 24(8):1752–1760.
82. Wolf YI, Koonin EV (2013) Genome reduction as the dominant mode of evolution. *BioEssays* 35(9):829–837.
83. Dagan T, Martin W (2007) Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proc Natl Acad Sci USA* 104(3): 870–875.
84. Stephens RS, et al. (1998) Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* 282(5389):754–759.
85. Schmitz-Esser S, et al. (2004) ATP/ADP translocases: A common feature of obligate intracellular amoebal symbionts related to Chlamydiae and Rickettsiae. *J Bacteriol* 186(3):683–691.
86. Moreira D, Deschamps P (2014) What was the real contribution of endosymbionts to the eukaryotic nucleus? Insights from photosynthetic eukaryotes. *Cold Spring Harb Perspect Biol* 6(7):a016014.
87. Reyes-Prieto A, Moustafa A (2012) Plastid-localized amino acid biosynthetic pathways of Plantae are predominantly composed of non-cyanobacterial enzymes. *Sci Rep* 2:955.
88. Qiu H, et al. (2013) Assessing the bacterial contribution to the plastid proteome. *Trends Plant Sci* 18(12):680–687.
89. Richards S, et al.; International Aphid Genomics Consortium (2010) Genome sequence of the pea aphid *Acyrthosiphon pisum*. *PLoS Biol* 8(2):e1000313.
90. Husnik F, et al. (2013) Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* 153(7):1567–1578.
91. Sloan DB, et al. (2014) Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol Biol Evol* 31(4):857–871.
92. Dunning Hotopp JC, et al. (2007) Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* 317(5845):1753–1756.
93. Hazkani-Covo E, Zeller RM, Martin W (2010) Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet* 6(2):e1000834.
94. Kleine T, Maier UG, Leister D (2009) DNA transfer from organelles to the nucleus: The idiosyncratic genetics of endosymbiosis. *Annu Rev Plant Biol* 60:115–138.
95. Timmis JN, Ayliffe MA, Huang CY, Martin W (2004) Endosymbiotic gene transfer: Organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet* 5(2):123–135.
96. Lang M, et al. (2012) Polymorphic NumtS trace human population relationships. *Hum Genet* 131(5):757–771.
97. Dayama G, Emery SB, Kidd JM, Mills RE (2014) The genomic landscape of polymorphic human nuclear mitochondrial insertions. *Nucleic Acids Res* 42(20): 12640–12649.
98. Turner C, et al. (2003) Human genetic disease caused by de novo mitochondrial-nuclear DNA transfer. *Hum Genet* 112(3):303–309.
99. Huang CY, Grünheit N, Ahmadinejad N, Timmis JN, Martin W (2005) Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm nuclear chromosomes. *Plant Physiol* 138(3):1723–1733.
100. Ricchetti M, Tekaia F, Dujon B (2004) Continued colonization of the human genome by mitochondrial DNA. *PLoS Biol* 2(9):E273.
101. Lloyd AH, Timmis JN (2011) The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. *Mol Biol Evol* 28(7):2019–2028.
102. Wang D, Timmis JN (2013) Cytoplasmic organelle DNA preferentially inserts into open chromatin. *Genome Biol Evol* 5(6):1060–1064.
103. Hazkani-Covo E, Covo S (2008) Numt-mediated double-strand break repair mitigates deletions during primate genome evolution. *PLoS Genet* 4(10):e1000237.
104. Wilkins AS, Holliday R (2009) The evolution of meiosis from mitosis. *Genetics* 181(1): 3–12.
105. Malik SB, Pightling AW, Stefaniak LM, Schurko AM, Logsdon JM, Jr (2008) An expanded inventory of conserved meiotic genes provides evidence for sex in *Trichomonas vaginalis*. *PLoS ONE* 3(8):e2879.
106. Hörandl E, Hadacek F (2013) The oxidative damage initiation hypothesis for meiosis. *Plant Reprod* 26(4):351–367.
107. Kuwabara T, et al. (2005) *Thermococcus coalescens* sp. nov., a cell-fusing hyperthermophilic archaeon from Suiyo Seamount. *Int J Syst Evol Microbiol* 55(Pt 6): 2507–2514.
108. Naor A, Gophna U (2013) Cell fusion and hybrids in Archaea: Prospects for genome shuffling and accelerated strain development for biotechnology. *Bioengineered* 4(3):126–129.
109. Papke RT, Koenig JE, Rodríguez-Valera F, Doolittle WF (2004) Frequent recombination in a saltern population of *Halorubrum*. *Science* 306(5703):1928–1929.