Modern endosymbiotic theory: Getting lateral gene transfer into the equation

William Martin[1]*, Mayo Roettger[1], Thorsten Kloesges[1,2], Thorsten Thiergart[1], Christian Woehle[1], Sven Gould[1] and Tal Dagan[3]

[1]Institute of Molecular Evolution, Universität Düsseldorf, 40225 Düsseldorf, Germany; [2]current address: Institute of Population Genetics, Universität Düsseldorf, 40225 Düsseldorf, Germany; [3]Institute of Genomic Microbiology, Universität Düsseldorf, 40225 Düsseldorf, Germany;
*correspondence to: William Martin, e-mail: bill@hhu.de

Modern views of eukaryote origins always entail endosymbiosis, but the number and nature of endosymbiotic events is still debated. It is generally agreed among all theories that the mitochondrion arose from proteobacteria and that the plastid arose from cyanobactera. It is also agreed that those were singlular and rare events in the history of life. The remaining debate concerns the origin of the host that acquired the mitochondrion and whether it was a prokaryote or a eukaryote and if the latter, whether additional endosymbioses were involved in its origin. Bioenergetic considerations indicate that the host that acquired the mitochondrion was a prokaryote, because the energy per gene that mitochondria provide was necessary to support the origin of the numerous evolutionary novelties that distinguish eukaryotes from prokaryotes. Yet other theories holding that the eukaryotic state arose before the acquisition of mitochondria remain popular. These include the notion that the nucleus arose as an endosymbiont or the view that eukaryotes are directly descended from planctomycetes. Theories for the origin of eukaryotes are, in principle, testable with genome data and inference tools of molecular phylogeny. Comparisons of genes present in the eukaryote common ancestor to genes in prokaryote genomes have uncovered evidence for the roles of an archaebacterium as the source of the genetic apparatus or informational genes, and a eubacterium as the source of energy metabolic functions, but for no other significant contributors. At the same time, the way in which microbial evolution, and lateral gene transfer (LGT) in particular, figure into endosymbiotic theory is gradually becoming apparent.

Reconstructing the ancestors of bioenergetic organelles

Endosymbiotic theories for the origin of organelles have been around for a hundred years. Mereschkowsky (1910) and Wallin (1926) were important early proponents of the theory for chloroplasts and mitochondria respectively. Popular in their early days (Geus and Höxtermann 2006), endosymbiosis theories were scorned by Wilson in the 1920s (Martin and Kowallik 1999), and stayed scorned until the 1960s, when the late Lynn Margulis revived them (Sagan 1967). There is still no universal agreement among cognoscenti as to how many symbioses involving prokaryotic partners occurred during eukaryotic evolution. The minimum number of partners at the origin of eukaryotes is two: the mitochondrial endosymbiont and its host (Martin and Müller 1998; Embley and Martin 2006; Martin and Koonin 2006; Lane and Martin 2010), the latter of which under the simplest models was itself a prokaryote, by inference an archaebacterium, because of the archaebacterial nature of ribosomes in the eukaryotic cytosol. Indeed, the eukaryotic translation apparatus reflects well their endosymbiotic origin, with archaebacterial ribosomes in the cytosol and eubacterial ribosomes in the mitochondrion. In "prokaryote host" versions of endosymbitic theory, as in all modern endosymbiotic theories, the mitochondrial endosymbiont is seen as a proteobacterium; the host is seen as either a methanogen or as a *Thermoplasma*-like archaebacterium (Embley and Martin 2006).

Models involving more than two endosymbiotic partners at the origin of eukaryotes (">2 theories") are also popular. Zillig's model (Zillig et al. 1989) suggested a "fusion" event between an archaebacetrium and a eubacterium at the origin of eukaryotes. This suggestion came from the observation that some genes in eukaryotes were eubacterial in ancestry, but whose origin could not be specifically ascribed to the mitochondrion at that time. The idea that the nucleus arose from an endosymbiosis of an archaebacterium in a eubacterial host has its supporters. One version of that theory has it that the host was a deltaproteobacterium and the nucleus-generating endosymbiont was a methanogen (a euryarchaeote) (Lopez-Garcia and Moriera 1999; Lopez-Garcia and Moriera 2006), another version has it that the host was a planctomycete and the endosymbiont was a mesophilic crenarchaeote (Forterre 2011). Another ">2" theory posits that in addition to mitochondrion and archaebacterial host, an endosymbiotic spirochaete gave rise to the flagellum and the nucleus (Margulis et al. 2001). These theories generate predictions about the patterns of similarity that one might find when comparing eukaryotic nuclear genes to their prokaryotic homologues. Accordingly, genomes are a useful source of data when it comes to testing evolutionary theories.

What do genomes say about symbiotic eukartyote origins?

Several recent studies have addressed the origin of mitochondria, but have focused on sequences residing in mito-

chondrial DNA (mtDNA) only (Thrash et al. 2011; Brindefalk et al. 2011; Georgiades and Raoult 2011). Those studies delivered widely conflicting results because of the small sample of mitochondrion-encoded protein available, about 55 of which at most can be used to generate trees (Esser et al. 2004). Furthermore, the phylogenetic biases introduced by the rapid evolutionary rate and AT richness of mtDNA can cause mtDNA-encoded proteins to artefactually group together with homologues from rapidly evolving and AT-rich bacterial lineages such as Rickettsiales (Thrash et al. 2011; Rodríguez-Ezpeleta and Embley 2012). Nuclear-encoded proteins should, in principle, be less subject to AT bias and the elevated substitution rate of mitochondrially encoded proteins. Investigations of mitochondrial origin using nuclear genes of mitochondrial origin are still scarce. Trees for pyruvate dehydrogenase subunits pointed to Rickettsia-like ancestors (Kurland and Andersson 2000). Trees for Krebs cycle and glyoxylate cycle enzymes (Schnarrenberger and Martin 2002) as well as trees for 220 nuclear-encoded mitochondrial proteins from Chlamydomonas point more frequently to origins from generalist, facultatively anaerobic α-proteobacteria (Atteia et al. 2009), than to Rickettsia-like ancestors, whereby many proteins indicated a eubacterial, but not a specifically α-proteobacterial ancestry. Recent analysis of 86 yeast nuclear-encoded mitochondrial proteins produced a similar result: some point to Rickettsia-like ancestors and some point to facultatively anaerobic Rhodobacter-like ancestors (Abhishek et al. 2011). The consideration of anaerobic energy metabolism in eukaryotes, especially its manifestation in mitochondria (Müller et al. 2012) and investigation of eukaryotic nuclear genes (Thiergart et al. 2012) support the view that the ancestor of mitochondria was a gerenalist, facultatively anaerobic proteobacterium.

The origin of plastids has also come under investigation. Deusch et al. (2008) investigated the evolutionary relationship of all individual genes encoded in nine cyanobacterial genomes to all nuclear genes among four sequenced eukaryotic genomes from the plant lineage. Those results indicated that among the cyanobacterial genomes then sampled, representatives of section IV in Stanier's traditional classification — filamentous, but non-branching, heterocyst-forming cyanobacteria — harbor collections of genes that far more closely reflect those found in plant nuclear genomes than representatives from non-nitrogen-fixing cyanobacterial groups. That finding is consistent with the view pioneered by Maier and colleagues (Kneip et al. 2007, 2008) that nitrogen fixation might have played a critical role in the establishment of the ancestral cyanobacteria endosymbiont that gave rise to plastids at the origin of the plant (Archaeplastida) lineage. Criscuolo and Gribaldo (2011) recently studied concatenated gene phylogenies and found evidence for a deep branching position of plastids among cyanobacteria, but that work, like previous studies was lacking sequences from subsection V. Indeed, among the ca. 60 cyanobacterial genomes currently available for comparison, no genome sequences have been determined from the other main heterocyst-forming lineage, subection V — filamentous, branching, heterocyst-forming cyanobacteria. It would be desirable to have sequences for representatives from subsection V, transformation methods for which now exist (Stucken et al. 2012), but that will have to await future work. As in the case of mitochondria (Thrash et al. 2011), the very long branches exhibited by plastid-encoded and plastid derived sequences causes them to branch deeply and topologically attract to

other long branch lineages in concatenated pyhlogenies (e.g. Criscuolo and Gribaldo 2011) rendering the assignment of sister group relationships problematic. The ways in which LGT among prokaryotes shuffles microbial chromosomes (Bapteste et al. 2009) further complicates phylogenetic reconstructions of the organellar ancestors (Esser et al. 2007; Richards and Archibald 2011).

Fluid prokaryotic chromosomes

To test the predictions of competing and mutually exclusive hypotheses for the origin of eukaryotes, Thiergart et al. (2012) studied a sample of 27 sequenced eukaryotic and 994 sequenced prokaryotic genomes and found 571 genes that were present in the eukaryote common ancestor and that have readily identifiable homologs among eubacterial and archaebacterial genomes. They reconstructed maximum likelihood trees to identify the prokaryotic genomes that most frequently contained genes branching as the sister to the eukaryotic nuclear homologs. The resulting distribution showed that among the archaebacteria, euryarchaeote representatives appeared most frequently as the sister to the eukaryotic nuclear gene. Among eubacteria, the α-proteobacteria were most frequently represented within the sister group. Homologs from α-proteobacterial genomes that branched as the sister to nuclear genes were found more frequently in genomes of facultatively anaerobic members the rhiozobiales and rhodospirilliales than in obligate intracellular rickettsial parasites. Following α-proteobacteria in line, the most frequent eubacterial sister lineages were γ-proteobacteria, δ-proteobacteria and firmicutes. Do we need to add additional endosymbionts at the origin of eukaryotes in order to account for those sequence similarity patterns? Thiergart et al. (2012) did not consider that possibility, rather they pointed out that because prokaryotic "lineages" have been laterally acquiring and disseminating genes for more than 1.5 billion years since eukaryote origins, it is quite unlikely that any modern prokaryote would harbor the same collection of genes as the ancestral mitochondrion. Their findings underscored the archaebacterial (host) nature of the eukaryotic informational genes and the eubacterial (mitochondrial) nature of eukaryotic energy metabolism (Thiergart et al. 2012).

Some of us have been saying for a long time that LGT among prokaryotes needs to be taken better into account in endosymbiotic theory. Martin (1999, p. 102) for example, wrote:

"Clearly, the origin of mitochondria involved a straightforward sampling process of a genome-sized aliquot of eubacterial genes by the host's genome. But ancient transfer events between eubacteria complicate matters surrounding the biological source (hence the biological context of acquisition) of eubacterial genes in eukaryotic chromosomes [...]. This is not a trivial issue, because there are many more genes of eubacterial origin in eukaryotic chromosomes than can be accounted for by traditional formulations of the endosymbiont hypothesis."

Studying the enzymes of the citric acid cycle, Schnarrenberger and Martin (2002, p. 1) surmised:

"...about half of the proteins involved in this eukaryotic pathway are most similar to their α-proteobacterial homologues, whereas the remain-

der are most similar to eubacterial, but not specifically α-proteobacterial, homologues. A consideration of i) the process of lateral gene transfer among free-living prokaryotes and ii) the mechanistics of endosymbiotic (symbiont-to-host) gene transfer reveals that it is unrealistic to expect all nuclear genes that were acquired from the α-proteobacterial ancestor of mitochondria to branch specifically with their homologues encoded in the genomes of contemporary α-proteobacteria. Rather, even if molecular phylogenetics were to work perfectly (which it doesn't), then some nuclear-encoded proteins that were acquired from the α-proteobacterial ancestor of mitochondria should, in phylogenetic trees, branch with homologues that are no longer found in most α-proteobacterial genomes, and some should reside on long branches that reveal affinity to eubacterial rather than archaebacterial homologues, but no particular affinity to any specific eubacterial donor."

Studying genes that plants acquired from cyanobacteria, Martin et al. (2002, p. 12248) concluded:

"...we found a surprisingly large fraction of *Arabidopsis* proteins that branch with their homologues from Grampositive (G+ve) bacteria. For example, more *Arabidopsis* proteins branched with their homolgues from *Mycobacterium* (148 proteins) than did with either *Prochlorococcus* (102) or *Synechocystis* (82) (Fig. 1). Naively, this might be interpreted as suggesting that the *Arabidopsis* lineage acquired genes specifically from a G+ve donor subsequent to its divergence from the yeast lineage. But by that same measure, the data in Fig. 1 would suggest at face value that the *Arabidopsis* lineage acquired genes from all organisms sampled in this study. Such interpretations can hardly be true and are at odds with the finding that the data in Fig. 1 would suggest at face value the *Arabidopsis* lineage to have acquired genes not from one cyanobacterium, but from all three sampled [even at a bootstrap probability (BP) = 0.95, whereby that view contradicts independent evidence suggesting a single origin of plastids from one cyanobacterium (42, 43), not three or more in the *Arabidopsis* lineage. The G+ve signal in the *Arabidopsis* data most likely reflects an overall similarity of many proteins in G+ve genomes to homologues in cyanobacteria. [...] In our view, the G+ve signal in the *Arabidopsis* data are most easily attributed to genes that entered the plant lineage through the ancestors of plastids, even though the gene trees recover a G+ve branch, either because of shared ancestry or lateral transfer of G+ve and cyanobacterial genes."

The same reasoning applied by Martin et al. (2002) to sequences that branch with Gram positive homologs applies directly to chlamydia related genes in plant genomes that Martin et al. (2002) also observed, and which are currently a hot topic (Moustafa et al. 2008; Price et al. 2012). Looking specifically at the phylogenetic purity of a dozen sequenced α-proteobacterial genomes, Esser et al. (2007, p. 4) wrote:

"Our findings indicate that modern α-proteobacterial genomes represent transient collections of genes that stem from diverse sources. By infer-

ence, the ancestor of mitochondria had a mosaic genome as well; hence, a criterion that is often used to infer whether a eukaryotic nuclear gene of eubacterial origin stems from the mitochondrion or not—namely branching with an α-proteobacterial gene—is probably too strict, because it tacitly assumes a static model of bacterial chromosome evolution in which LGT and gene loss do not exist, either now or in the past. Incorporating a fluid bacterial chromosome model into endosymbiotic theory generates the prediction that nuclear genes acquired by eukaryotes from the ancestor of mitochondria should tend to reflect a single common eubacterial ancestry—provided that molecular phylogeny can accurately recover events that occurred more than 1.5 billion years ago—but that they should not necessarily belong to the known set of contemporary α-proteobacterial genes, regardless of how one were to define it."

Those passages lean quite heavily on the view that LGT among prokaryotes is indeed very prevalent in nature. Is that really true?

Microbial genome evolution really does impact endosymbiotic theory

No matter what model for the evolution of organelles ones agrees with, it is widely accepted that eukaryotic organelles evolved from prokaryotic endosymbionts. Accordingly, if we want to reconstruct the origin of organelles from genomic comparisons with present-day microbes, one has to take the lateral realities of microbial genome evolution into account (Esser et al. 2007; Richards and Archibald 2011). This is more challenging than it might seem at first sight. Prokaryotic genomes evolve not only through vertical inheritance but also through DNA acquisition via lateral gene transfer. In an LGT event, a recipient genome acquires genetic material from a donor genome. The acquired DNA becomes an integral part of the recipient genome and is inherited by its descendants (Babic et al. 2008). LGT is a major mechanism for natural variation in prokaryotes, and several well-characterized mechanisms for DNA acquisition have evolved, including transformation (Chen and Dubnau 2004), transduction (Thomas and Nielsen 2005), conjugation (Chen et al. 2005), and gene transfer agents (Lang et al. 2012).

Among prokaryotes, LGT serves as a significant source of new gene families (Treangen and Rocha 2011) and a source of genetic novelty (Bapteste et al. 2009). Evolutionary speaking, gene acquisition in prokaryotes fulfills the functions that both sex and gene (and genome) duplications fulfill in eukaryotes. But in marked contrast to meiotic recombination in eukaryotes, LGT in prokaryotes is never reciprocal, it is always unidirectional, from donor to recipient. Currently there are no known mechanisms for LGT in eukaryotes except in the case of endosymbiosis where donors and recipients are found in intimate relations making gene transfer feasible and sometimes also beneficial to the holobiont (Nikoh et al. 2008; Gilbert et al. 2010; Dupuy et al. 2011). Eukaryotes however have undergone extensive LGT from the endosymbiotic ancestors of cholorplasts and mitochondria to the host nuclear genome. This process is called endosymbiotic gene transfer (Martin et al. 1993; Timmis et al. 2004; Kleine et al. 2009; Hazcani Covo et al. 2010).

While evolution by gene transfer during eukaryotic evolution is either ancient or rare, LGT during prokaryotic evolution is frequent and abundant (Doolittle 1999; Ochman 2000). The frequency of protein families affected by LGT during microbial evolution as inferred from gene phylogenies is estimated to range between 60% (Kunin et al. 2005; Dagan and Martin 2007) and 90% (Mirkin et al. 2003). Other authors reported much lower frequencies: Ge et al. estimated that merely 2% of the protein families evolved by LGT (Ge et al. 2005), and Beiko et al. detected LGT in only 14% of the protein families (Beiko et al. 2005). However, these low estimates should be contrasted with the experimental assessment of LGT frequency calculated *in vitro* from gene acquisition rate in *Escherichia coli* (Sorek et al. 2007). Out of 246,045 LGTs from 79 different donor species via a plasmid (similar to LGT by transformation or conjugation), only 1,402 instances failed to integrate into the *E. coli* genome. In remaining 99.4% of the transfers the gene was transferred successfully (Sorek et al. 2007). Genes that were identified as resistant to lateral transfer are common among proteins involved in complex biological mechanisms, such as the ribosome, where both sequence conservation and gene copy number confer major selective constraints on protein function (Sorek et al. 2007). Cellular pathways posing a barrier to LGT of their member proteins are most common in information processing pathways (Jain et al. 1999) but are found in metabolism and cellular processes pathways as well (Cohen et al. 2011).

So it turns out that LGT is real in the prokaryotic world (Popa and Dagan 2011; Popa et al. 2011) and that it can be studied using network methods (Dagan et al. 2008; Kloesges et al. 2010; Dagan 2011). When it comes to the reconstruction of eukaryotic ancestry, the widespread occurrence of LGT during microbial evolution leads to the scrambling of the genetic record. That makes testing of endosymbiotic hypotheses for the origins of eukaryotes and their organelles a bit tougher, but it can still be done. We just have to look at all the data openly, and think things through in full. After all, nothing in the biology of bioenergetic organelles makes sense except in the light of endosymbiosis.

Acknowledgement

References

Abhishek A, Bavishi A, Choudhary M. (2011) Bacterial genome chimaerism and the origin of mitochondria. *Can J Microbiol*. 57:49-61.

Atteia A, Adrait A, Brugière S, van Lis R, Tardif M, Deusch O, et al. (2009) A proteomic survey of *Chlamydomonas reinhardtii* mitochondria sheds new light on the metabolic plasticity of the organelle and on the nature of the α-proteobacterial mitochondrial ancestor. *Mol Biol Evol*. 29:1533-1548.

Babic A, Lindner AB, Vulic M, Stewart EJ, Radman M. (2008) Direct visualization of horizontal gene transfer. *Science*. 319:1533-1536.

Bapteste E, O'Malley M, Beiko RG, Ereshefsky M, Gogarten JP, Franklin-Hall L, et al. (2009) Prokaryotic evolution and the tree of life are two different things. *Biol Direct*. 4:34.

Beiko RG, Harlow TJ, Ragan MA. (2005) Highways of gene sharing in prokaryotes. *Proc Natl Acad Sci USA*. 102:14332-14337.

Brindefalk B, Ettema TJG, Viklund J, Thollesson M, Andersson SGE. (2011) A phylometagenomic exploration of oceanic alphaproteobacteria reveals mitochondrial relatives unrelated to the SAR11 clade. *PLoS One*. 6: e24457.

Chen I, Christie PJ, Dubnau D. (2005) The ins and outs of DNA transfer in bacteria. *Science*. 310:1456-1460.

Chen I, Dubnau D. (2004) DNA uptake during bacterial transformation. *Nat Rev Microbiol*. 2:241-249.

Cohen O, Gophna U, Pupko T. (2011) The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. *Mol Biol Evol*. 28:1481-1489.

Criscuolo A, Gribaldo S. (2011) Large-scale phylogenomic analyses indicate a deep origin of primary plastids within cyanobacteria. *Mol Biol Evol*. 28:3019-3032.

Dagan T. (2011) Phylogenomic networks. *Trends Microbiol*. 19:483-491.

Dagan T, Artzy-Randrup Y, Martin WF. (2008) Modular networks and cumulative impact of lateral transfer in prokaryote genome evolution. *Proc Natl Acad Sci USA*. 105:10039-10044.

Dagan T, Martin WF. (2007) Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. *Proc Natl Acad Sci USA*. 104:870-875.

Deusch O, Landan G, Roettger M, Gruenheit N, Kowallik KV, Allen JF, et al. (2008) Genes of cyanobacterial origin in plant nuclear genomes point to a heterocyst-forming plastid ancestor. *Mol Biol Evol*. 25:748–761.

Doolittle WF. (1999) Phylogenetic classification and the universal tree. *Science*. 284:2124-2128.

Dupuy C, Periquet G, Serbielle C, Bézier A, Louis F, Drezen JM. (2011) Transfer of a chromosomal Maverick to endogenous bracovirus in a parasitoid wasp. *Genetica*. 139:489-496.

Embley TM, Martin WF. (2006) Eukaryote evolution: changes and challenges. *Nature*. 440:623-630.

Esser C, Martin W, Dagan T. (2007) The origin of mitochondria in light of a fluid prokaryotic chromosome model. *Biol Lett*. 3:180-184.

Esser C, Ahmadinejad N, Wiegand C, Rotte C, Sebastaini F, Gelius-Dietrich G, et al. (2004) A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol Biol Evol*. 21:1643-1660.

Forterre P. (2011) A new fusion hypothesis for the origin of Eukarya: better than previous ones, but probably also wrong. *Res Microbiol*. 162:77-91.

Ge F, Wang LS, Kim J. (2005) The cobweb of life revealed by genome-scale estimates of horizontal gene transfer. *PLoS Biol*. 3:e316.

Geus A, Höxtermann E. (2006) Evolution durch Kooperation und Integration. Basilikum-Presse, Marburg, pp. 751.

Georgiades K, Raoult D. (2011) The rhizome of *Reclinomonas americana*, *Homo sapiens*, *Pediculus humanus* and *Saccharomyces cerevisiae* mitochondria. *Biol Direct*. 6:55.

Gilbert C, Schaack S, Pace JK, Brindley PJ, Feschotte C. (2010) A role for host-parasite interactions in the horizontal transfer of transposons across phyla. *Nature*. 464:1347-1350.

Hazkani-Covo E, Zeller RM, Martin WF. (2010) Molecular poltergeists: mitochondrial DNA copies (*numts*) in sequenced nuclear genomes. *PLoS Genet*. 6:e1000834.

Jain R, Rivera MC, Lake JA. (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *Proc Natl Acad Sci USA*. 96:3801-3806.

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.

Kloesges T, Popa O, Martin WF, 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.

Kneip C, Voss C, Lockhart PJ, Maier UG. (2008) The cyanobacterial endosymbiont of the unicellular algae *Rhopalodia gibba* shows reductive genome evolution. *BMC Evol Biol*. 8:30.

Kneip C, Lockhart P, Voss C, Maier UG. (2007) Nitrogen fixation in eukaryotes—new models for symbiosis. *BMC Evol Biol*. 7:55.

Kunin V, Goldovsky L, Darzentas N, Ouzounis CA. (2005) The net of life: reconstructing the microbial phylogenetic network. *Genome Res.* 15:954-959.

Kurland CG, Andersson SG. (2000) Origin and evolution of the mitochondrial proteome. *Microbiol Mol Biol Rev.* 64:786-820.

Lane N, Martin WF. (2010) The energetics of genome complexity. *Nature.* 467:929–934.

Lang AS, Zhaxybayeva O, Beatty JT. (2012) Gene transfer agents: phage-like elements of genetic exchange. *Nat Rev Micro.* 10:472-482.

Lopez-Garcia P, Moreira D. (2006) Selective forces for the origin of the eukaryotic nucleus. *Bioessays.* 28:525-533.

Lopez-Garcia P, Moreira D. (1999) Metabolic symbiosis at the origin of eukaryotes. *Trends Biochem Sci.* 24:88-93.

Margulis L, Dolan MF, Guerrero R. (2000) The chimeric eukaryote: Origin of the nucleus from the karyomastigont in a mitochondriate protists. *Proc Natl Acad Sci USA.* 97:6954-6959.

Martin W. (1999) Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. *Bioessays.* 21:99-104.

Martin W, Brinkmann H, Savona C, Cerff R. (1993) Evidence for a chimaeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. *Proc Natl Acad Sci USA.* 90:8692-8696.

Martin W, Kowallik KV. (1999) Annotated English translation of Mereschkowsky's 1905 paper "Über Natur und Ursprung der Chromatophoren im Pflanzenreiche" *Eur J Phycol.* 34:287-295.

Martin W, Koonin EV. (2006) Introns and the origin of nucleus-cytosol compartmentation. *Nature.* 440:41-45.

Martin W, Müller M. (1998) The hydrogen hypothesis for the first eukaryote. *Nature.* 392:37-41.

Martin W, Rujan T, Richly E, Hansen A, Cornelsen S, Lins T, 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.* 99:12246-12251.

Martin W, Stoebe B, Goremykin V, Hansmann S, Hasegawa M, Kowallik KV. (1998) Gene transfer to the nucleus and the evolution of chloroplasts. *Nature.* 393:162-165.

Mereschkowsky C. (1910) Theorie der zwei Plasmaarten als Grundlage der Symbiogenesis, einer neuen Lehre von der Entstehung der Organismen. *Biol Centralbl.* 30:353-367.

Mirkin BG, Fenner TI, Galperin MY, Koonin EV. (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.

Moustafa A, Reyes-Prieto A, Bhattacharya D. (2008) Chlamydiae has contributed at least 55 genes to plantae with predominantly plastid functions. *PLoS One.* 3:e2205.

Müller M, Mentel M, van Hellemond J, Henze K, Woehle C, Gould SB, et al. (2012) Biochemistry and evolution of anaerobic energy metabolism in eukaryotes. *Microbiol Mol Biol Rev.* 76:444-495.

Nikoh N, Tanaka K, Shibata F, Kondo N, Hizume M, Shimada M, et al. (2008) *Wolbachia* genome integrated in an insect chromosome: evolution and fate of laterally transferred endosymbiont genes. *Genome Res.* 18:272-280.

Ochman H, Lawrence JG, Groisman EA. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature.* 405:299-304.

Popa O, Dagan T. (2011) Trend and barriers to lateral gene transfer in prokaryotes. *Curr Opin Microbiol.* 14:615-623.

Popa O, Hazkani-Covo E, Landan G, Martin WF, Dagan T. (2011) Directed networks reveal barriers and bypasses to lateral gene transfer in prokaryotes. *Genome Res.* 21:599-609.

Price DC, Chan CX, Yoon HS, Yang EC, Qiu H, Weber APM, et al. (2012) Cyanophora paradoxa genome elucidates origin of photosynthesis in algae and plants. *Science.* 335:843-847.

Richards TA, Archibald JM. (2011) Gene transfer agents and the origin of mitochondria. *Curr Biol.* 21:R112-R114.

Rodríguez-Ezpeleta N, Embley TM. (2012) The SAR11 group of alpha-proteobacteria is not related to the origin of mitochondria. *PLoS One.* 7:e30520.

Sagan L. (1967) On the origin of mitosing cells. *J Theoret Biol.* 14:225-274.

Schnarrenberger C, Martin WF. (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.* 269:1-15.

Sorek R, Zhu Y, Creevey CJ, Francino MP, Bork P, Rubin EM. (2007) Genome-wide experimental determination of barriers to horizontal gene transfer. *Science.* 318:1449-1452.

Stucken K, Ilhan J, Roettger M, Dagan T, Martin WF. (2012) Transformation and conjugal transfer of foreign genes into the filamentous multicellular cyanobacteria (subsection V) *Fischerella* and *Chlorogloeopsis. Curr Microbiol.* (in press)

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.* 4:466-485.

Thomas CM, Nielsen KM. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol.* 3:711-721.

Thrash JC, Boyd A, Huggett MJ, Grote J, Carini P, Yoder RJ, et al. (2011) Phylogenomic evidence for a common ancestor of mitochondria and the SAR11 clade. *Sci Rep.* 1:e13.

Timmis JN, Ayliffe MA, Huang CY, Martin WF. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat Rev Genet.* 5:123-135.

Treangen TJ, Rocha EPC. (2011) Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PloS Genetics.* 7:e1001284.

Wallin IE. (1927) Symbionticism and the origin of species. In: Bailliere, Tindall and Cox, London, pp. 171.

Zillig W, Klenk H-P, Palm P, Leffers H, Pühler G, Gropp F, et al. (1989) Did eukaryotes originate by a fusion event? *Endocytobiosis Cell Res.* 6:1-25.