Journal of Endocytobiosis and Cell Research (2012) 16-20 | International Society of Endocytobiology
zs.thulb.uni-jena.de/content/main/journals/ecb/info.xml

Journal of Endocytobiosis and Cell Research

Gene transfers from organelles to the nucleus: How much, what happens, and why none in *Elysia*?

William Martin[1]*, Einat Hazkani-Covo[2], Liat Shavit-Greivink[1,3], Valerie Schmitt[1], Katharina Händeler[1], Sven B. Gould[1], Giddy Landan[4], Dan Graur[5] and Tal Dagan[4]

[1]Institute of Molecular Evolution, Universität Düsseldorf, 40225 Düsseldorf, Germany; [2]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA; [3]present address: The Edmond and Lily Safra Center for Brain Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel; [4]Institute of Genomic Microbiology, Universität Düsseldorf, 40225 Düsseldorf, Germany; [5]Department of Biology and Biochemistry, University of Houston, Houston, TX 77204-5001, USA; *correspondence to: William Martin, e-mail: bill@hhu.de

Gene transfers from organelles to the nucleus are important in evolution and they are quantitatively great. Studies from higher plants and algae have shown that a substantial proportion of plant genomes, more than 15% in some species, consists of genes that were ultimately acquired from the cyanobacterial ancestor of plastids. Looking back into early eukaryote evolution, among heterotrophic eukaryotes, well over half of the sequences that have prokaryotic homologues are derived from eubacterial rather than archaebacterial genes. It can be argued that these ultimately stem from the proteobacterial ancestor of mitochondria. Reflecting more recent genome evolutionary processes, all sequenced eukaryotic genomes from species that have DNA in their mitochondria contain fragments of mitochondrial or plastid DNA in their nuclear chromosomes. The haploid human genome contains about 260,000 bases of sequences recently derived from insertions of mitochondrial DNA (mtDNA), 12 human loci are polymporphic for mtDNA insertions and five human mtDNA insertions cause disease (Hazkani-Covo et al. 2010). The process of gene transfer into the chromosomes is well characterized and entails nonhomologous end joining mechanism, not cDNA intermediates as was once thought. Gene transfers from organelles to the nucleus have become a very commonplace phenomenon in biology. Because endosymbiotic gene transfer is so well documented for plastid origins, there has long been an expectation that gene transfer underpins the biology of the stolen chloroplasts (kleptoplasts) in the photosynthetic sacoglossan slugs from the genus *Elysia*. However, genome wide expression data show that no expressed genes in these photosynthetic slugs have been transferred from algae.

Journal of Endocytobiosis and Cell Research (2012) 16-20
Keywords: Endosymbiotic gene transfer, plastids, mitochondria, sacoglossans, photosynthetic slugs

Endosymbiotic gene transfer

Endosymbiosis is integral to eukaryote evolution, and gene transfers from organelles to the nucleus — a process known as endosymbiotic gene transfer (Martin et al. 1993) — is integral to endosymbiotic origin of organelles. Gene transfers from organelles were an important mechanism of genetic variation that participated centrally in the prokaryote-to-eukaryote transition (Timmis et al. 2004; Gould et al. 2008; Kleine et al. 2009). Though DNA can be experimentally relocated from organelles to the nucleus in the laboratory (Huang et al. 2003; Ricchetti et al. 1999), the more far-reaching experiment is the one ongoing in nature over evolutionary time. All genome sequences from eukaryotes that possess mitochondrial DNA — for exceptions lacking mtDNA see van der Giezen (2009) and Müller et al. (2012) — harbour evidence for the ongoing process of organelle-to-nuclear DNA transfer in the form of nuclear copies of mitochondrial and, in the case of plants, chloroplast DNA (Leister 2005). Genome sequences from those eukaryotes that have lost their mitochondrial DNA altogether still harbour evidence for gene transfers from the mitochondrion during the early phases of eukaryote history (Tovar et al. 2003; Müller et al. 2012).

During plastid evolution, the process of gene transfer continued during the primary diversification of the archaeplastida lineage into glaucocystophytes, red algae, and green algae (Martin et al. 1998, 2002). It continues also today, where it can be observed as *nupt*s — nuclear sequences of plastid origin — showing ecotype-specific polymorphisms among different *Arabidopsis* ecotypes (Leister 2003). In genetically engineered tobacco, for example, gene transfer from plastids to the nucleus occurs at a rate of about 1/16000 gametes (Huang et al. 2003), and the rate can be increased by stress conditions (Wang et al. 2012). Similarly, *numt*s — nuclear sequences of mitochondrial origin (Lopez et al. 1994) — are very abundant in nuclear genomes and are frequently transferred (Bensasson et al. 2001; Hazcani-Covo et al. 2003). The largest nuclear insertions of organelle DNA known are the complete, 131 kb chloroplast DNA insertion in the rice genome and the complete 376 kb mitochondrial DNA insertion in the *Arabidopsis* genome, and both insertions are recent, having occurred less than about 200,000 years ago (Huang et al. 2005).

The mechanism of integrating organelle DNA in the nucleus involves double strand break repair and nonhomologous end joining (Ricchetti et al. 1999; Hazkani-Covo and Covo 2008; Hazcani-Covo et al. 2010). It was once thought that cDNA intermediates (Nugent and Palmer 1991) were involved in organelle to nucleus gene transfer, but genome sequences revealed that integration of organelle chromosomes into nuclear DNA directly is the mechanism by which genes are transferred (Henze and Martin

2001). As a process, gene transfer to the nucleus can be broken down into four distinct and independent phases (Allen 1993; Martin and Herrmann 1998; Timmis et al. 2004; Hazckani-Covo et al. 2010): i) genes are first copied to the nucleus, that is, the original copy remains functional in the organelle (Allen 1993), ii) the nuclear copy accumulates mutations, both point mutations at the nuclear rate, which is higher than the organelle rate in plants, and mutations involving recombination, insertion and deletion, yet under strongly reduced functional constraints, because the functional plastid copy remains under selection, iii) the nuclear copy either undergoes pseudogenization, or its encoded (but mutated) protein product becomes expressed and functional (= selected and fixed), either in the organelle (a fate only possible once the protein translocon complexes for protein import into the plastid had evolved), or in other cell compartments involving the acquisition of such targeting signals as exist at the time of transfer in the given lineage, followed by iv) loss of the organelle copy, either subsequent to its functional replacement by the nuclear copy or following the loss of the corresponding gene function in the given lineage (for example phycobilisomes in green algae).

The ancestors of mitochondria and plastids donated many genes to the host nucleus via endosymbiotic gene transfer. But more than protein coding genes entered the nucleus thus. Cech (1986) proposed that eukaryotic introns and their cognate spliceosomal small nuclear RNAs (snRNAs) originated from disarticulate group II introns, Cavalier-Smith (1991) extended that idea by suggesting that they evolved specifically from group II introns that invaded the ancestrally intron-less eukaryotic genome through the mitochondrial endosymbiont. Thus group II introns also entered the eukaryotic lineage, most likely via gene transfer from the mitochondrial endosymbiont. This has been implicated as the source of the specific selective pressures that precipitated the origin of the nuclear compartment (Koonin and Martin 2006) in a mitochondrion-bearing cell.

Ancient phylogenies mean methodological problems

When it comes to trying to pinpoint the evolutionary source or sources of eukaryotic nuclear genes, it is important to note that immediately following gene transfer to the nucleus, a phase commences in which the nuclear gene copy accumulates mutations under strongly reduced functional constraints until its product becomes expressed, targeted and selected. In molecular evolutionary and phylogenetic terms, this should correspond to elevated numbers of radical or unique substitutions relative to copies present in the plastid and in cyanobacteria. This aberrant phase of mutational accumulation should, in principle, therefore lead to "long branches" in phylogenies, which has severe consequences for identifying genes that were acquired by plant genomes from cyanobacteria (plastids) or proteobacteria (mitochondria), because it leads to high rates of misidentification of gene sources using phylogenetic methods.

When one is investigating ancient events, such as the origin of plastids, amino acid sequence conservation is low, such that all aspects of phylogenomics, including alignment, become problematic. Current estimates for the cyanobacterial fraction of plant genomes (~15% on average) are based on all genes that can be aligned E-value ≤ 10^{-10} and ≥ 25% amino acid identity in the BLAST pairwise alignments of the query (the plant protein) to the search set comprising 223 prokaryotes and 13 eukaryotes (Deusch et al. 2008). At 20% amino acid identity or below, one is clearly in the "twilight zone" of sequence similarity, and no phylogenetic methods perform reliably when sequences are on average less that 50% identical (Nei 1996). However, very few sequences are that well conserved (50% amino acid sequence identity) in eukaryote-prokaryote comparisons. Worse, there is no generally recognized "cut-off" for amino acid sequence identity in the field of phylogenetics or phylogenomics below which it is seen as "impermissible" to interpret results, such that some phylogenomic studies include proteins that are as little as 10% identical, with the unsurprising result that they infer lots of strange branches that are (generally) interpreted as lateral gene transfers (Gophna et al. 2006).

Where is the limit at which we have to stop making alignments and trees? Landan and Graur (2007) have developed an outstanding method to approach this problem that is as simple as it is elegant, it is called the Heads or Tails (HoT) method. In a nutshell, the HoT method starts by comparing two alignments of the same sequences, whereby both alignments are generated using the *same programm* (such as MUSCLE; Edgar 2004) and using the *same parameters*, but one alignment ("heads") is generated by reading the sequences in N-terminus to C-terminus, while the other is generated by reading the sequences in C-terminus to N-terminus ("tails"). Trees can be constructed from both alignments and the proportion of identical branches in both trees for a given protein can be compared. For example, among the 11,569 trees that were investigated by Deusch et al. (2008) only ~40% of actually produced identical topologies in the heads and tails alignments. This is because the heads and tails alignments differ, because there are two equally optimal solutions to any pairwise aligments and in multiple alignments (the high road and the low road). One or the other must be chosen by convention so that a multiple alignment results (Landan and Graur 2007). But the site patterns (columns of amino acids) that are generated in the two alignments differ. Hence, the trees that are inferred from the computational analysis of the site patterns often differ too. If the same data gives two different trees dependent upon the random variable of whether the sequences were read into the alignment algorithm left-to-right (heads) or right-to-left (tails) then something is wrong with at least one of the trees, or more likely with both of them.

A convenient way to see where the problem lies is to examine the column score (CS), which is simply the proportion of columns (site patterns) that are constructed identically in the heads and tails alignments. If the CS is 1.0, all site patterns are independent of the random variable (heads or tails), if the CS is 0.1, only 10% of the columns are independent of the random variable. In real data from real genomes, as we have analyzed in this project, with an E-value threshold of ≤10^{-10} and at least 25% amino acid iden-

tity, the average column score for some 11,000 alignments is 0.6, which means that our average fraction of cyanobacterial genes per genome (~15%) is too low for reasons of false negatives. The higher the column score, the higher the proportion of genes is inferred to be of cyanobacterial origin (Deusch et al. 2008). If one looks at only the most reliable alignments (CS > 0.8), the proportion of cyanobacterial genes in each plant genome is ~20%, as opposed to ~15% for the whole data, which contains many false negatives (Deusch et al. 2008). Clearly, the origin of organelles had a very substantial quantitative impact upon the content and nature of genes in the genome of the founding eukaryotic (and plant) lineages.

While various laboratories are striving to make estimates for the fraction of genes that eukaryotes acquired from organelles (Lane and Archibald 2008), or to infer the lineages from which mitochondria (Abhishek et al. 2011; Thrash et al. 2011; Brindefalk et al. 2011; Georgiades and Raoult 2011) and chloroplasts descend (Criscuolo and Gribaldo 2011), often using very highly parameterized phylogenetic methods, there is at least as much cause, in our view, to worry about more fundamental parameters, such as alignment reliability (Penn et al. 2010).

Elysia: Gene transfer that isn't

Elysia is a genus belonging to the sacoglossan sea slugs, which are unique in the animal kingdom in that they sequester and maintain active plastids that they acquire from the siphonaceous algae upon which they feed, making the animals photosynthetic. Four species from the sacoglossan family Plakobranchidae retain their stolen plastids (kleptoplasts) in a photosynthetically active state on timescales of weeks to months (the other roughly 30 species just digest the plastids as food). Perhaps because lots of genes had been found to be transferred during the origin of plastids in the plant kingdom, there have been many discussions of whether or not genes for photosynthetic functions transferred had been transferred to the nuclei of sacoglossan slugs (Rumpho et al. 2000). These discussions reached a crescendo recently with the publication that one sacoglossan, *Elysia chlorotica*, which feeds on the siphoaceous xanthophyte *Vaucheria*, had acquired the gene for a component of photosystem II, psbO, and that the slug provides its plastids with that protein (Rumpho et al. 2008).

There were various reasons to doubt whether *Elysia chlorotica* had really acquired the psbO gene. But the most distressing aspect was that Rumpho et al. (2008) described in detail a bipartite targeting sequence in the putatively transferred psbO gene of the slug that would, in their interpretations, direct the product to the plastid. The plastids of *Vaucheria* are indeed surrounded by four membranes in the alga (Gould et al. 2008) and do require such complex targeting signals to traverse the four membranes when the plastids reside in the *Vaucheria* cytosol. But the *Vaucheria* plastids sequestered in *Elysia* are only surrounded by two membranes (Rumpho et al. 2001), that is, the outer two membranes are stripped off in the animal. Rumpho et al. (2011) pointed out that this circumstance "has potential implications for protein targeting, but will not be discussed here" (p. 306). What they did not discuss is that if the bipartite targeting signals were indeed present, the psbO protein would exit the slug cell via the secretory pathway (Gould et al. 2008; Wägele et al. 2011), and not be targeted to the plastid.

Wägele et al. (2011) reasoned that if there are transferred genes for photosynthetic functions in the nuclei of sacoglossan slugs, those genes should be expressed as mRNA, and thus they should readily be detectable in a deep sequencing EST experiment using mRNA extracted from photosynthesizing slugs. Wägele et al. (2011) performed EST analyses of two sacoglossan species that retain photosynthticall yactive plastids over periods of months: *Plakobranchus ocellatus* and *Elysia timida* with 77.000 expressed sequence contigs for *P. ocellatus* and 25.000 contigs for *E. timida* (a total of 1.5 million reads and 64 Mb of nonredundant sequence data). The mRNA sample was extracted from animals that were demonstrably photosynthetic at the time of harvesting and that had sequestered their plastids for at least three weeks. They compared the extensive *Arabidopsis* EST data, where a wealth of information on nuclear encoded genes for chloroplast biogenesis exists, to the slug ESTs and to the limited EST data then available for *Acetabularia*, the food alga of *E. timida* as a control. The comparison to *Acetabularia* reveals whether, using *Arabidopsis* query sequences, one would be able to detect expressed algal genes in *Elysia* by sequence comparison if they were there. The comparison to the slugs reveals which homologues of nuclear encoded *Arabidopsis* genes for chloroplast proteins, if any, are expressed as mRNA in the animals. The results were striking and clearly ruled out the possibility that any genes of algal origin are expressed by the animals. Since well over 500 nuclear encoded genes are required to make a functional plastid (Leister 2003; Wägele et al. 2011) and since none are expressed by the slugs, one can firmly conclude that no genes were transferred in those two species, and by inference, in any of the other sacoglossans. The slugs clearly have no need for transferred genes to support their photosynthetic plastids, that they just sequester long-lived plastids — which is possibly even more interesting.

But gene transfer stories are hard to stop once they get going, even with strong data of the type the Wägele et al. (2011) presented indicating that no genes were transferred. Pelletreau et al. (2012) repeated that EST experiment and analysed *Elysia chlorotica*, reporting 148 Mb of sequence data. They also found no evidence for expression of transferred genes, although they found a very low frequency of about 20 non-*Elysia* sequences (possibly contaminants), none of which indicate an involvement in photosynthesis. However, they nonetheless concluded that "multiple" lines of evidence that indicate that nuclear algal genes have been transferred.

More recently, Pierce et al. (2012) reported 98,000,000 reads and 8.8 billion base pairs of next generation transcriptome sequencing data for *Elysia chlorotica.* Undoubtedly, a substantial sequencing effort. Among their 98 million reads, they only found 101 that match *Vaucheria* nuclear sequence data better than animal sequence data, corresponding to 52 transcripts. However, from that Pierce

et al. (2012, p. 1545) concluded "*that a variety of functional algal genes have been transferred into the slug genome*", even though evidence for expression of the collection of well over 500 nuclear encoded genes that are required for functional plastids is altogether lacking. Arguably, the main finding of their study is that 98 million reads from *Elysia chlorotica* nucleic acid preparations produce 100 reads with undeniable similarity to *Vaucheria* nuclear genes; that finding suggests to us, and should have suggested to Pierce et al. (2012) that their nucleic acid preparations used for sequencing were 99.9999% free of contamination, because that is the frequency (one part per million) at which they find *Elysia* sequences. A purity of 99.9999% is outstanding by any measure, but the case for gene transfer to *Elysia* is exactly that: one in a million.

Why no gene transfer in *Elysia*?

The title of this contribution promissed an answer to the question of "why" no gene transfer in *Elysia*. Sacoglossans acquire their plastids from their food source, every generation anew, in daimetric opposition to the ground-floor-and-basement foundation basis of endosymbiotic theory for the origin of organelles. When Schimper (1883, p. 112-113) was kicking the idea around of the endosymbiotic origin of plastids, he did so in a footnote, and his salient observation was the continuity of plastids "*Sollte es sich definitiv bestätigen, dass die Plastiden in den Eizellen nicht neu gebildet werden, so würde ihre Beziehung zu dem sie enthaltenden Organismus einigermassen an eine Symbiose erinnern. Möglicherweise verdanken die grüinen Pflanzen wirklich einer Vereinigung eines farblosen Organismus mit einem von Chlorophyll gleichmäissig tingierten ihren Ursprung.*" [If it can be conclusively confirmed that plastids do not arise *de novo* in egg cells, the relationship between plastids and the organisms within which they are contained would then be somewhat reminiscent of a symbiosis. Green plants may in fact owe their origin to the unification of a colorless organism with one uniformly tinged with chlorophyll" translation from Martin et al. (1992)]. That is basically all Schimper said about endosymbiosis. Mereschkowsky (1905) of course fully devloped the idea, and in that paper, which set forth the foundation of modern endosymbiotic theory, he listed several reasons why plastids (which he called chromatophores) are derived from cyanobacteria. The first of those reasons was "Die Kontinuität der Chromatophoren" (Mereschkowsky 1905, p. 596).

The sacoglossan situation can be understood in the context of classical endosymbiotic theory. If the slugs had established a large repertiore of algal genes to maintain their plastids, why would they bother to acquire new ones every generation? The plastids of sacoglossans are not continuous organelles in the sense that they have to be re-established every generation. The sacoglossan plastids are not organelles, they are not even endosymbionts — the difference being protein import from the cytosol (Theissen and Martin 2006) — they are just undigested food. And while there is plenty of good theoretical ground to expect gene transfer from organelles to the nucleus in animals, there is no good theory to expect gene transfer from food to the nucleus. Indeed, in every newly sequenced animal genome we see hundreds and thousands of recently transferred DNA fragments from organelle genomes (Hazkani-Covo et al. 2010), but no DNA fragments that stem from what we or our ancestors recently had for lunch. Humans have existed since at least Sunday the 23rd of October, 4004 BCE (Graur and Martin 2004), ample time to integrate many numts into our geonome, but not enough time to incorporate genes from our food. That is something to keep in mind.

Acknowledgement

This work was kindly funded by the DFG in SFB-Tr1, project A10. GL is grateful to the rectorate of the University of Düsseldorf for a postdoctoral stipend, LS-G is grateful to the AvH foundation for a postdoctoral stipend, DG is grateful to the AvH foundation for a research award. We thank Heike Wägele, Bonn, for fruitful cooperation and unsightful discussion.

References

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

Allen JF. (1993) Control of gene expression by redox potential and the requirement for chloroplast and mitochondrial genomes. *J Theor Biol*. 165:609-631.

Bensasson D, Zhang D, Hartl DL, Hewitt GM. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol Evol*. 16:314-321.

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.

Cavalier-Smith T. (1991) Intron phylogeny: A new hypothesis. *Trends Genet*. 7:145-148.

Cech TR. (1986) The generality of self-splicing RNA: Relationship to nuclear mRNA splicing. *Cell*. 44:207-210.

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

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.

Edgar RC. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 32:1792-1797.

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

Gophna U, et al. (2006) Ancient lateral gene transfer in the evolution of *Bdellovibrio bacteriovorus*. *Trends Microbiol*. 14:64-69.

Gould SB, Waller RF, McFadden GI. (2008) Plastid evolution. *Annu Rev Plant Biol*. 59:491–517.

Graur D, Martin W. (2004) Reading the entrails of chickens: Molecular timescales of evolution and the illusion of precision. *Trends Genet*. 20:80-86.

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

Hazkani-Covo E, Covo S. (2008) *Numt*-mediated double strand break repair mitigates deletions during primate genome evolution. *PLoS Genetics*. 4:e1000237.

Hazkani-Covo E, Sorek R, Graur D. (2003) Evolutionary dynamics of large numts in the human genome: rarity of independent in-

sertions and abundance of post-insertion duplications. *J Mol Evol.* 56:169-174.

Henze K, Martin WF. (2001) How do mitochondrial genes get into the nucleus? *Trends Genet.* 17:383-387.

Huang CY, Grünheit N, Ahmadinejad N, Timmis JN, Martin WF. (2005) Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm nuclear chromosomes. *Plant Physiol.* 138:1723-1733.

Huang CY, Ayliffe MA, Timmis JN. (2003) Direct measurement of the transfer rate of chloroplast DNA into the nucleus. *Nature.* 422:72-76.

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.

Landan G, Graur D. (2007) Heads or tails: A simple reliability check for multiple sequence alignments. *Mol Biol Evol.* 24:1380-1383.

Lane CE, Archibald JM. (2008) The eukaryotic tree of life: endosymbiosis takes its TOL. *Trends Ecol Evol.* 23:268-275.

Leister D. (2005) Origin, evolution and genetic effects of nuclear insertions of organelle DNA. *Trends Genet.* 21:655-663.

Leister D. (2003) Chloroplast research in the genomics age. *Trends Genet.* 19:47-56.

Lopez JV, Yuhki N, Masuda R, Modi W, O'Brien SJ. (1994) Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat. *J Mol Evol*. 39:174-190.

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

Martin WF, 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 WF, Herrmann RG. (1998) Gene transfer from organelles to the nucleus: How much, what happens and why? *Plant Physiol.* 118:9-17.

Martin WF, 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 WF, Somerville CC, Loiseaux-de Goer S. (1992) Molecular phylogenies of plastid origins and algal evolution. *J Mol Evol*. 35:385-403.

Mereschkowsky C. (1905) Über Natur und Ursprung der Chromatophoren im Pflanzenreiche. *Biol Centralbl.* 25:593-604. [English translation: In: Martin W, Kowallik KV. (1999) *Eur J Phycol.* 34:287-295.]

Müller M, Mentel M, van Hellemond JJ, Henze K, Wöhle C, Gould SB, et al. (2012) Biochemistry of evolution of anaerobic energy metabolism in eukaryotes. *MMBR.* 76:444-495.

Nei M. (1996) Phylogenetic analysis in molecular evolutionary genetics. *Annu Rev Genet.* 30:371-403.

Nugent JM, Palmer JD. (1991) RNA-mediated transfer of the gene coxII from the mitochondrion to the nucleus during flowering plant evolution. *Cell.* 66:473-481.

Pelletreau KN, Bhattacharya D, Price DC, Worful JM, Moustafa A, Rumpho ME. (2011) Sea slug kleptoplasty and plastid maintenance in a metazoan. *Plant Physiol*. 155:1561-1565.

Penn O, Privman E, Ashkenazy H, Landan G, Graur D, Pupko T. (2010) GUIDANCE: a web server for assessing alignment confidence scores. *Nucleic Acids Res*. 38(Web Server issue):W23-W28.

Pierce SK, Fang X, Schwartz JA, Jiang X, Zhao W, Curtis NE, et al. (2012) Transcriptomic evidence for the expression of horizontally transferred algal nuclear genes in the photosynthetic sea slug, *Elysia chlorotica*. *Mol Biol Evol.* 29:1545-1556.

Rumpho ME, Pelletreau KN, Moustafa A, Bhattacharya D. (2011) The making of a photosynthetic animal. *J Exp Biol*. 214:303-311.

Rumpho ME, Worful JM, Lee J, Kannan K, Tyler MS, Bhattacharya D, et al. (2008) Horizontal gene transfer of the algal nuclear gene psbO to the photosynthetic sea slug *Elysia chlorotica*. *Proc Natl Acad Sci USA.* 105:17867-17871.

Rumpho ME, Summer EJ, Green BJ, Fox TC, Manhart JR. (2001) Mollusc/algal chloroplast symbiosis: how can isolated chloroplasts continue to function for months in the cytosol of a sea slug in the absence of an algal nucleus? *Zoology*. 104:303-312.

Rumpho ME, Summer EJ, Manhart JR. (2000) Solar-powered sea slugs. Mollusc/algal chloroplast symbiosis. *Plant Physiol.* 132:29-38.

Ricchetti M, Fairhead C, Dujon B. (1999) Mitochondrial DNA repairs doublestrand breaks in yeast chromosomes. *Nature.* 402:96-100.

Schimper AFW. (1883) Über die Entwicklung der Chlorophyllkörner und Farbkörper. *Bot Ztg*. 41:105-114.

Theissen U, Martin W. (2006) The difference between endosymbionts and organelles. *Curr Biol.* 16:R1016-R1017.

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

Tovar J, Leon-Avila G, Sanchez LB, Sutak R, Tachezy J, van der Giezen M, et al. (2003) Mitochondrial remnant organelles of Giardia function in iron-sulphur protein maturation. *Nature.* 426:172-176.

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.

van der Giezen M. (2009) Hydrogenosomes and mitosomes: conservation and evolution of functions. *J Eukaryot Microbiol.* 56:221-231.

Waegele H, Deusch O, Haendeler K, Martin R, Schmitt V, Christa G, et al. (2011) Transcriptomic evidence that longevity of acquired plastids in the photosynthetic slugs *Elysia timida* and *Plakobranchus ocellatus* does not entail lateral transfer of algal nuclear genes. *Mol Biol Evol.* 28:699-706.

Wang D, Lloyd AH, Timmis JN. (2012) Environmental stress increases the entry of cytoplasmic organellar DNA into the nucleus in plants. *Proc Natl Acad Sci USA*. 109:2444-2448.