

Review

Lateral gene transfer in eukaryotes

J. O. Andersson

Institute of Cell and Molecular Biology, Uppsala University, Biomedical Center, Box 596, 751 24 Uppsala (Sweden),
Fax: +46 18 530396, e-mail: Jan.Andersson@icm.uu.se

Received 3 December 2004; received after revision 24 January 2005; accepted 1 February 2005
Available online 09 March 2005

Abstract. Lateral gene transfer – the transfer of genetic material between species – has been acknowledged as a major mechanism in prokaryotic genome evolution for some time. Recently accumulating data indicate that the process also occurs in the evolution of eukaryotic genomes. However, there are large rate variations between groups of eukaryotes; animals and fungi seem to be largely unaffected, with a few exceptions, while lateral gene transfer frequently occurs in protists with phagotrophic lifestyles, possibly with rates comparable to

prokaryotic organisms. Gene transfers often facilitate the acquisition of functions encoded in prokaryotic genomes by eukaryotic organisms, which may enable them to colonize new environments. Transfers between eukaryotes also occur, mainly into larger phagotrophic eukaryotes that ingest eukaryotic cells, but also between plant lineages. These findings have implications for eukaryotic genomic research in general, and studies of the origin and phylogeny of eukaryotes in particular.

Key words. Horizontal gene transfer; lateral gene transfer; phylogeny; origin of eukaryotes; phagotrophy, endosymbiotic gene transfer; eukaryote phylogeny.

Introduction

Lateral, or horizontal, gene transfer is the process of exchange of genetic material between distantly related species. In prokaryotes, comparative genomics of whole genome data has led to the suggestion that the process of lateral gene transfer (LGT) is a more influential evolutionary mechanism than the 20th-century microbiologists ever thought [1–3], although more conservative interpretations exist [4, 5]. The research on LGT has focused on prokaryotic rather than eukaryotic organisms, mainly for two reasons; there has been a high amount of sequence data from diverse prokaryotic lineages available for some time, and the process has been assumed to be of limited significance in eukaryotes. However, our understanding of the phenomenon is changing as more and more genomic sequences from diverse eukaryotes become available, and recent reports involving many different lineages (tables 1 and 2) have indicated LGT as a potentially im-

portant evolutionary mechanism also in eukaryotic organisms. There are two distinct types of gene transfer in eukaryotes: the transfer of genes from the organelles with an endosymbiotic origin (the mitochondrion and plastid) to the nucleus of the eukaryotic cell (endosymbiotic gene transfer) and LGT between unrelated species (fig. 1). Endosymbiotic gene transfer is widely accepted as an important source of genetic material in eukaryotic lineages [6]. The occurrence of LGT in eukaryotes, on the other hand, is much more controversial. The claim that over 100 genes were recently transferred from prokaryotes to the human genome may have been the weakest point in an otherwise landmark paper [7] – the reported cases did not hold up for reanalysis using phylogenetic methods [8, 9]. In fact, transfer of genes from prokaryotes to animals with sequestered germ lines appear to be extremely rare, although it cannot be formally excluded [9, 10]. Still, these results do not preclude the possibility that LGT is important for non-human eukaryotes (tables 1 and 2) –

Table 1. Interdomain gene transfers.

Recipient lineage	Protein	Phagotrophic	Reference
Apicomplexa (<i>Cryptosporidium</i>)	5-monophosphate-dehydrogenase	–	[36]
Apicomplexa	a number of proteins	–	[34]
Apicomplexa (<i>Cryptosporidium</i>)	thymidine kinase	–	[37]
Apicomplexa (<i>Cryptosporidium</i>)	24 proteins	–	[35]
Chlorarachniophytes	a number of plastid-targeted proteins	+	[54]
Ciliates (<i>Entodinium</i>)	glutamate dehydrogenase	+	[44]
Diplomonads	3-hydroxy-3-methylglutaryl coenzyme a reductase	+	[45]
Diplomonads	a dozen proteins	+	[77]
Diplomonads	NADH oxidase	+	[108]
Diplomonads	phosphoenolpyruvate carboxykinase	+	[109]
Diplomonads	sulfide dehydrogenase	+	[41]
<i>Entamoeba</i>	8 proteins	+	[77]
<i>Entamoeba</i>	IscS and IscU	+	[110]
<i>Entamoeba</i>	malic enzyme, acetyl-CoA synthetase and alcohol dehydrogenase	+	[111]
<i>Entamoeba</i>	NADH oxidase	+	[108]
Euglenozoa	glyceraldehyd-3-phosphate dehydrogenase	+	[55, 56, 58]
Euglenozoa (diplonemids)	glyceraldehyd-3-phosphate dehydrogenase	+	[58]
Euglenozoa (<i>Trypanosoma</i>)	glutamate dehydrogenase	+	[44]
Fungi (filamentous fungi)*	pea pathogenicity genes	–	[87]
Fungi (rumen fungi)	glycosyl hydrolases	–	[86]
Fungi (yeast)	flavoheemoglobin	–	[77]
Metazoa	deoxyribose-phosphate aldolase and threonine dehydratase	–	[77]
Metazoa	DNA polymerase gamma subunit 2	–	[9]
Metazoa (<i>Callosobruchus</i>)	genome fragment of endosymbiont	–	[91]
Metazoa (<i>Ciona</i>)	cellulose synthase	–	[88, 89]
Metazoa (<i>Meloidogyne</i>)	12 putatively plant-parasitic genes	–	[90]
Parabasalids	alcohol dehydrogenase	+	[77]
Parabasalids	glyceraldehyde-3-phosphate dehydrogenase	+	[56, 58]
Parabasalids	N-acetylneuraminase lyase	+	[112]
Parabasalids	potential surface protein	+	[42]
Parabasalids and diplomonads	alanyl-tRNA and prolyl-tRNA synthetase	+	[65]
Parabasalids and diplomonads	glucokinase and glucosephosphate isomerase	+	[113]
Plants (<i>Nicotiana</i>)	<i>Agrobacterium</i> genes	–	[92]
<i>Prostherobacter</i> **	tubulins		[75]
Proteobacteria**	deoxyribose-phosphate aldolase		[77]

* Unknown source.

** Eukaryotic donor lineage.

the vast majority of the diversity of the eukaryotic life is indeed unicellular with lifestyle that resemble that of prokaryotes in many aspects. This review will focus on transfer of genes independent of organelles; transfer of mobile elements, such as transposons and introns, have been covered previously [11, 12] and will not be discussed here.

Mechanisms

Gene transfers are rare evolutionary events that are detected by their consequences – similar genes are present in distantly related organisms – rather than at the time when they are happening. Therefore, it is often problematic to deduce the exact mechanism by which a gene transfer has occurred. Nonetheless, to gain insight into the process an understanding of the steps involved in a successful gene transfer between two lineages is needed.

First, the foreign genetic material must enter the cell, either as naked DNA, or together with the cell that harbors the gene. Once inside, the gene has to be incorporated in the host nucleus and expressed into a functional protein. For the foreign genetic material to be maintained, the protein must provide a function which is selected in the population. It may indeed intuitively seem very unlikely that a prokaryotic gene could be successfully transferred to a eukaryote – if a single step fails, no transfer will occur. Incorporation of foreign genetic material into the nucleus of eukaryotes does indeed occur at a substantial rate; a large amount of genetic material from the mitochondria has been incorporated into the eukaryotic genomes, some of which has been shown to be functional [6, 13]. It is reasonable to assume that foreign genetic material from sources other than the endosymbionts may also be functionally incorporated into the nucleus via the same mechanisms [13], if it is present within the eukaryotic cell. Consequently, the critical step for LGT into eukaryotes is

Table 2. Intra-domain gene transfers.

Recipient lineage	Protein	Phagotrophic	Reference
Animals (<i>Hydra</i>) or parabasalids	<i>flp</i> gene	+/-	[43]
Chlorarachniophytes	a number of plastid-targeted proteins	+	[54]
Ciliates	alanyl-tRNA synthetase	+	[65]
Dinoflagellates	enolase	+	[60, 61]
Dinoflagellates	glyceraldehyde-3-phosphate dehydrogenase	+	[57]
Diplomonads or mycetozoa	threonine dehydratase	+	[77]
<i>Entamoeba</i>	alanyl-tRNA synthetase	+	[65]
Many lineages	EF-like protein – replacement of EF-1 α	+/-	[63]
Plants	mitochondrial genes	-	[70]
Plants	<i>Nad1B-C</i>	-	[71]
Plants	part of mitochondrial <i>nad1</i> gene	-	[69]

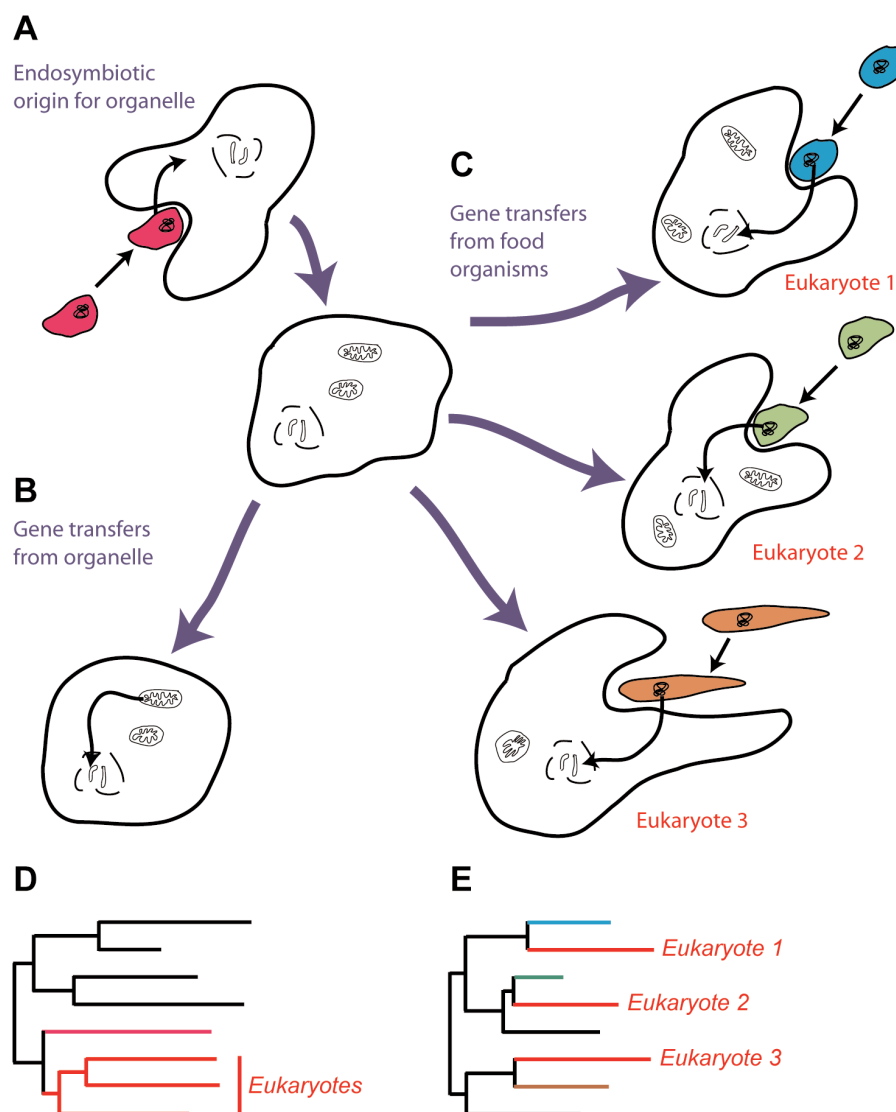


Figure 1. Schematic drawings of processes that may introduce prokaryotic genes into the eukaryotic nucleus. The endosymbiotic origin of organelles (i.e. mitochondria and plastids) (A) may lead to transfer of genes from the organelle to the nucleus (B). Ingestion of prokaryotic cells may lead to transfer of genes from the food organism to the nucleus (C) – the ‘You are what you eat’ hypothesis [14, 15]. The two pathways for the introduction of prokaryotic genes into eukaryotes have different predicted phylogenetic outcomes; transfer from an organelle is expected to give trees with monophyletic eukaryotes (D), while independent transfer from food organisms should give trees with polyphyletic eukaryotes (E), unless the transfers occurred before the extant eukaryotes diverged.

likely the uptake of DNA into the cell. For prokaryotes, the processes of transformation, transduction and conjugation are usually invoked as potential mechanisms to explain LGT events. These processes are probably less wide-spread in eukaryotes, which make the mechanisms of the uptake of genetic material for gene transfers elusive in many putative cases.

Nevertheless, the lifestyles of the inferred donor and recipient lineages often provide some clues about how and where the transfers may have taken place. For example, the phagotrophic lifestyle of many unicellular eukaryotes and the presence of genes of eubacterial origin in these lineages prompted Doolittle to propose a ratchet-like mechanism by which prokaryotic genes from food or symbiont bacteria could replace ancient eukaryotic genes over evolutionary time: the 'You are what you eat' hypothesis [14, 15] (fig. 1). Many of the species in table 1 are indeed phagotrophs which utilize prokaryotic cells as food – indicating that this mechanism of transfer operates over evolutionary time. In practice, the mechanism should not be limited to interdomain transfers – many lineages phagocytose eukaryotic cells, which would provide a mechanism for intradomain transfer (table 2). Additional mechanisms for gene transfers must exist – several lineages in tables 1 and 2 are not phagotrophic. Transfection of a virus between unrelated species and physical contact of symbiotic or host-parasite relationships have been suggested as possibilities for DNA of unrelated species to be introduced into eukaryotic cells [16].

Genetic and/or biochemical studies are needed to verify whether a gene of foreign origin has been successfully incorporated into the nucleus – its presence does not necessarily indicate that it is expressed and that the protein is functional. However, unsuccessful incorporation of DNA will most likely either turn into pseudogenes, or be lost altogether, since genes which are not under purifying selection will accumulate deleterious mutations [17]. Successful transfers, on the other hand, may be maintained in the population. Thus, if the same gene is present in related eukaryotes, the pattern of sequence variation of the gene may be used to deduce whether the amino acid sequence is under purifying selection. The recipient lineages are represented by a natural group of eukaryotes in some, but not all, cases of gene transfer listed in tables 1 and 2, which suggests that the genes are functionally maintained. Nevertheless, further studies are needed to determine the frequency of failed LGT in eukaryotic genomes, and also to understand how the expressed proteins of foreign origins function within the cell.

Methods for detecting of LGT

Several different methods have been used to test the occurrence of LGT in prokaryotes [15, 18–23], most of

which have been applied to eukaryotes, with different levels of success. Since all methods have their own limitations, several independent methods should ideally be used before concluding that a gene has been acquired by LGT. A combination of phylogenetic methods, analyses of the distribution pattern and identification of a potential overlap of life habitats between the inferred donor and recipient lineages has been particularly useful in pinpointing LGT events affecting eukaryotes, as exemplified below.

The easiest and probably most widely used method to detect cases of gene transfers is similarity searches, such as BLAST. A result where the closest homolog of a sequence is a gene from a distantly related organism is often taken as evidence for a gene transfer event. In some cases, this seems reasonable; the high fraction of genes with best hits to archaeal genes in the genome of the hyperthermophilic eubacterium *Thermotoga maritima* did indeed hold up for subsequent phylogenetic analyses [24, 25]. However, BLAST searches are only very crude indicators of phylogenetic relationships [26] and should not be invoked as the only evidence for gene transfer events [27]. Luckily, apart from the initial analysis of the human genome, claims of LGT affecting eukaryotes based solely on unexpected results from similarity searches have been few [23].

Another method for detecting LGTs in prokaryotes has been to analyze the codon usage pattern between genes in genomes [2, 28, 29]. Although potentially very powerful for detecting recent transfers into genomes, the method has been shown to be inconsistent with other methods [21]. Furthermore, an extensive knowledge of the patterns of codon usage within the genome is needed to reliably detect recent gene transfers using the composition variations between gene sequences, which usually is lacking in eukaryotes. Therefore, these methods are currently of limited value for eukaryotes.

Unexpected phylogenetic relationships

The vast majority of the reported cases of LGT in eukaryotes are inferred from the observation that the phylogenetic relationship based on single gene sequences differs from the expected organismal relationships (tables 1 and 2). For example, a eukaryotic sequence may be found nested within prokaryotic sequences (fig. 1E). Such results are indeed regarded as the most reliable indicators of gene transfers, although any topology obtained in phylogenetic studies that supports an LGT event may also be explained by gene duplication and lineage-specific gene loss events [22, 30], if no limit is set for the number of such events compared to LGT events. As a matter of fact, the reverse is equally true; any gene phylogeny that resembles the organismal phylogeny may be explained

solely by LGT events, if gene transfer is assumed to be very frequent [31]. Thus, to distinguish between LGT scenarios and gene duplication and loss scenarios, the likelihoods of these evolutionary events have to be judged against each other. This is problematic, especially since the frequencies of such events are expected to vary between lineages as well as throughout evolution. Nevertheless, gene duplication and differential gene loss scenarios very often become extremely complicated for real datasets with a large number of taxa that disagree with the expected organismal phylogenies, since these scenarios usually involve maintenance of multiple copies of genes in genomes over extended evolutionary timescales followed by independent parallel recent losses in many lineages, and/or rooting of the phylogenetic trees in improbable places.

The phylogenetic tree of alcohol dehydrogenase is an example of a real dataset which indicates many LGT events; the eukaryotic sequences are found in at least six clearly separated classes, all with different prokaryotic affinities (fig. 2). To explain such a pattern with a scenario only including ancient duplications and subsequent differential gene losses, the ancestral eukaryote would be required to encode at least six copies of the *adh* gene, and these paralogs must have been retained for a long evolutionary time and only recently been lost in most eukaryotes in parallel evolutionary events. For example, the last common excavate (taxa labeled brown in fig. 2) ancestor must have encoded five divergent copies of the alcohol dehydrogenase gene, three of which were lost in each of the lineages leading to *Trichomonas vaginalis* and *Giardia lamblia*. Since it is rather unlikely that this ancestor had a much larger coding potential than the current species, a scenario where the alcohol dehydrogenase genes were introduced into Excavata from various lineages over evolutionary time; the *Naegleria gruberi* sequence probably from a eubacteria, one of the *Trichomonas* and one of the *Giardia* sequences from different low G+C Gram positives, the other *Giardia* sequences from a prokaryote, and the other *T. vaginalis* possibly from a prokaryote (fig. 2), appears much more likely. Such a scenario also readily explains the specific and unexpected relationships between prokaryotic and eukaryotic lineages, such as the branching of the *Trichomonas* sequence as a sister to low G+C Gram positives (fig. 2). A duplication and loss scenario, on the other hand, would require an extreme number of additional duplications and losses to account also for these relationships. For instance, the specific relationship between the animal+fungi cluster and the *Thermoplasma* sequence would indicate that this paralog was lost in all other lineages of eukaryotes as well as all other archaeal species. At the same time, the specific relationships between the *Giardia* and *Trichomonas* sequences with different lineages of low G+C Gram positives would indicate that these paralogs were lost not only in all other

eukaryotes and archaea, but also in all eubacteria except some lineages of low G+C Gram positives. Still, this would not explain the observation that the *Giardia* sequence is nested within low G+C Gram positives (fig. 2). Phylogenetic methods are imperfect and may falsely lead to unexpected trees. However, given the large sequence distances between the six separated groups of eukaryotic alcohol dehydrogenases compared to the distance to the closest prokaryotic sequence for each group as well as the high bootstrap support values separating the groups, it seems rather unlikely the phylogenetic artefacts are responsible for this overall topology. The detailed topology, on the other hand, may be influenced by phylogenetic artefacts; the grouping of an *Entamoeba* sequence with the *T. vaginalis* sequences is difficult to evaluate; at face value it indicates a prokaryote-to-eukaryote LGT followed by an intradomain transfer (*Entamoeba* and *Trichomonas* are only distantly related eukaryotes). However, both the *Entamoeba* and the *T. vaginalis* sequences are long, and the grouping may be artefactual due to long-branch attraction, which would argue for distinct prokaryotic origins of alcohol dehydrogenase in the two lineages. Likewise, it is unclear whether *Naegleria*, *Trypanosoma* and the three fungal sequences make up a eukaryotic clade, which makes the number and direction of LGT events problematic to deduce for these eukaryotic sequences (fig. 2). To summarize, the phylogenetic tree of alcohol dehydrogenase indicates a very complex evolutionary history of this gene, which includes a number of LGT events involving both prokaryotes and eukaryotes. Unfortunately, as for prokaryotes [15, 18–23], there are no strict objective criteria against which it can be tested whether a eukaryotic gene originated via LGT, which makes it very problematic to deduce the exact number of transfer events from a phylogenetic tree. For example, radically different results may be reached from the same phylogeny if strict vertical inheritance is used as the null hypothesis (all sampled homologous genes are orthologs originating from a common organismal ancestor), and gene transfers are inferred only when there is strong support for such events, compared to the other extreme, if gene transfer is used as the null hypothesis (all sample homologous genes are xenologs that originated from LGT events). In practice, this may not be such a huge problem on a gene-to-gene scale as it might seem; different genes have distinct rates of transfer, suggesting that different null hypotheses should be used (fig. 3). If several of the resolved nodes do indicate LGT events, the rate of transfer of the gene is relatively high and gene transfer should be the null hypothesis; suggesting that some transfers most likely have occurred in the unresolved part of the tree (fig. 3B). On the other hand, if all resolved nodes indicate the expected organismal phylogeny, the gene is rarely transferred, and vertical transmission should be the null hypothesis. In such a case few, if any, transfers can be

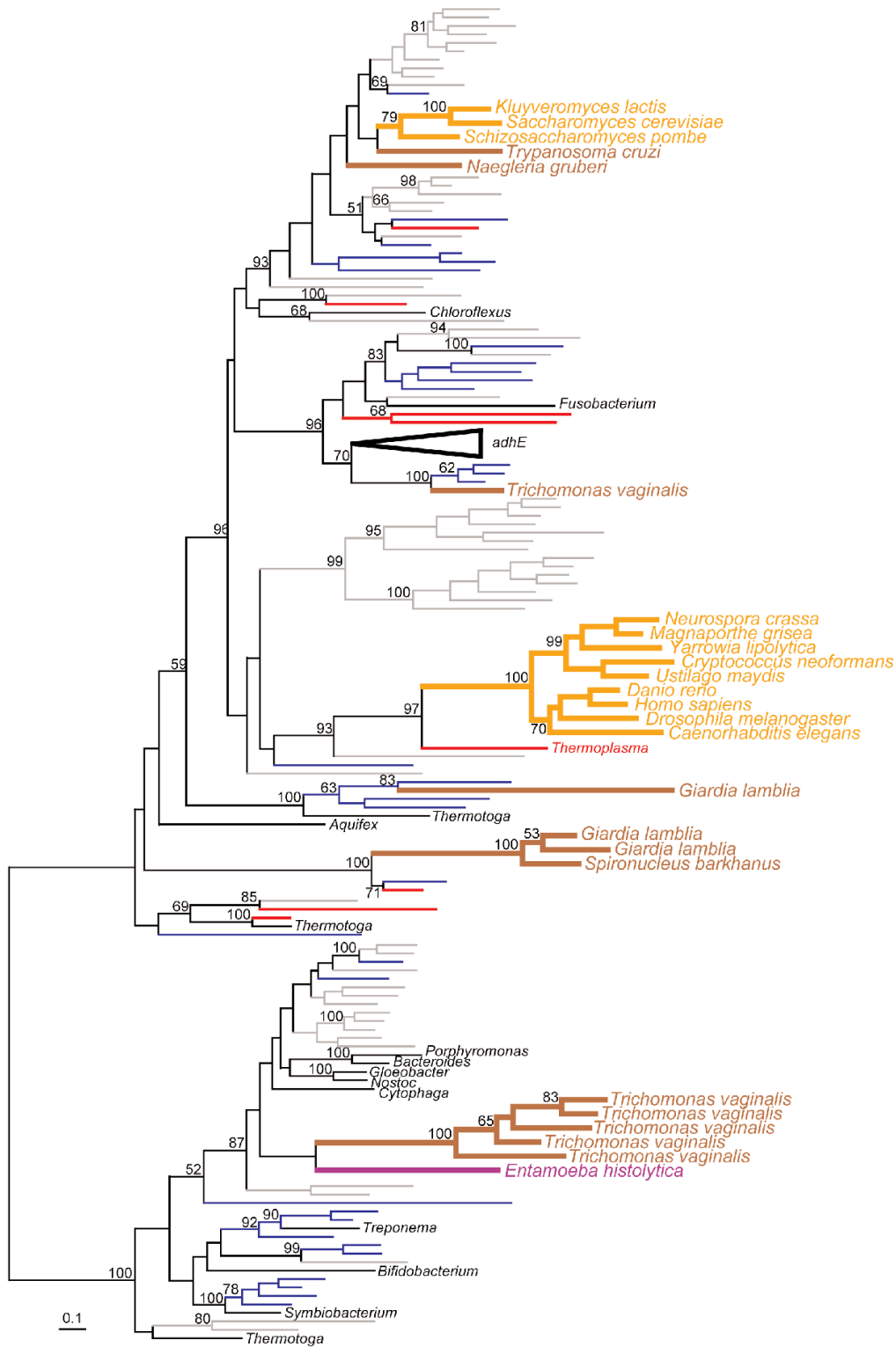


Figure 2. Phylogenetic tree of alcohol dehydrogenase. Protein maximum likelihood tree inferred using the program PHYML, version 2.1b [114], on 253 unambiguously aligned amino acid positions. Close homologs and sequences that failed the χ^2 tests for deviation of amino acid frequencies implemented in TREE-PUZZLE, version 5.1 [115], were excluded, as previously described [77]. Bootstrap support values above 50% based on 100 replicates are shown. Proteobacteria, low G+C Gram positives and archaea are indicated by gray, blue and red branches, respectively, without any names. Other prokaryotic species are shown with genus names. The complete names of eukaryotic sequences are shown in large fonts and bold branches labeled according to their classification into 'supergroups' [50, 79, 80]: opisthokonts (orange), amoebozoa (purple) and excavates (brown). The wedge indicates the position of the alcohol dehydrogenase E gene family, which has been distributed via LGT affecting both prokaryotes and eukaryotes [77].

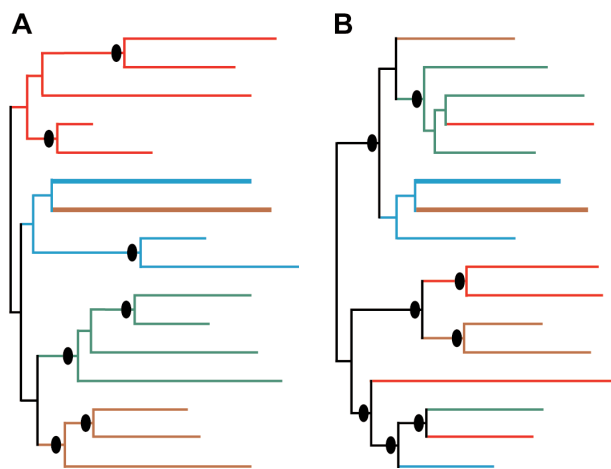


Figure 3. Hypothetical trees that illustrate the detection of LGT. Two hypothetical phylogenetic trees based on two imaginary genes present in 16 species belonging to four accepted organismal groups (indicated by different colors). Filled ovals indicate nodes strongly supported by statistical analysis (i.e. bootstrap analysis). The two genes have different frequencies of LGT. (A) The absence of indication of gene transfer among the supported bipartitions suggests that the gene is rarely transferred. Therefore, the grouping of the brown and blue branches is unlikely due to LGT. (B) A hypothetical tree in the presence of LGT. Several putative cases of LGT are supported by statistical analyses, indicating that the gene is frequently transferred between organismal groups. For that reason, the grouping of the brown and blue branches presumably is due to a gene transfer event, even though the phylogenetic analysis only weakly support this hypothesis.

expected to have occurred in the unresolved part of the tree (fig. 3A). In the case of alcohol dehydrogenase, some of the nodes with high bootstrap support unite known organismal groups, such as animals and fungi, while others indicate transfer events, such as the grouping of *T. vaginalis* with low G+C Gram positives (fig. 2). Thus, for this gene, neither strict vertical inheritance nor gene transfer only could be used as the null hypothesis. This suggests that neither process should be assumed for any of the nodes of the unresolved part of the tree; some are expected to have originated via LGT, and some via vertical inheritance – indicating that the number of transfer events with strong support should be treated as the minimum, rather than the maximum.

The weakest point in using phylogenetic inferences to detect cases of lateral gene transfer may be that the phylogenetic methods are not always powerful enough to estimate reliable trees based on single gene alignment; artefactual trees may falsely be interpreted as indicators of gene transfer [30]. There are several reasons for the limitations of the power of the phylogenetic analyses; the models for the evolution of gene sequences may not be realistic, and the information content in single gene sequences may be insufficient to resolve all relationships in the tree. Indeed, concatenated gene sequences are needed to recover deep branches in the eukaryotic tree [32, 33].

Since LGT events probably most often occur on the single gene scale, concatenation of gene sequences is not a feasible approach to identify transfer events. For these reasons, phylogenetic studies should be combined with other sources of information, like gene distribution patterns, to make up a more complete picture of the contribution of LGT in the evolution of individual genes and gene families. Nevertheless, phylogenetic analyses are a powerful tool in identifying clear cases of transfer, such as some of the cases mentioned in the alcohol dehydrogenase phylogenetic tree (fig. 2), although such analyses are not expected to reveal every case of LGT for any gene.

Obvious cases of LGT should also be possible to detect with large-scale phylogenetic methods. Indeed, automated phylogenetic approaches using data from apicomplexan genomes revealed a large number of genes with unexpected phylogenetic relationships [34, 35]. In a preliminary analysis, over 100 sequences from each genome showed affinity to eubacteria, although inspection of individual datasets revealed that less than 50% of these trees had sufficient resolution or bootstrap support to infer transfer events, illustrating the difficulties with automated phylogenetic approaches [34]. In a more thorough large-scale phylogenetic analysis, 24 *Cryptosporidium parvum* genes were shown to be of eubacterial origin [35], and together with more detailed studies of individual genes [36, 37], these large-scale phylogenetic analyses [34, 35] indicate that LGT has been an important evolutionary mechanism in Apicomplexa.

Patchy phyletic distribution

Maybe the most convincing evidence for the importance of gene transfers in prokaryotes comes from the comparison of gene content of closely related strains and species; only 39% of the genes found in any of the three sequenced strains of *Escherichia coli* are present in all three [38], and only 51% of the genes found in any of the two sequenced strains of the marine cyanobacterium *Prochlorococcus* are found in both [39]. Although explanations other than gene transfer surely accounts for some of these differences, they are very strong indications of the importance of LGT in prokaryotes. For protists, no pairs of genome sequences from closely related species or strains are available yet, except for the case of *Plasmodium*. The comparison of draft genome sequences of a rodent and a human malaria parasite, *P. yoelii yoelii* and *P. falciparum*, indicated the presence of lineage-specific genes, although the possibility of LGT was not suggested [40]. However, given that phylogenetic studies of apicomplexan genomes in combination with analyses of the phyletic distribution of genes have indicated numerous gene transfers in these parasites [34–37], some of the lin-

eage-specific genes in the malaria parasites could be of LGT origin.

Comparative genomics approaches have been very useful in studying the occurrence of LGT in eukaryotes, despite the shortage of whole-genome sequences for most eukaryotic groups. For example, gene families that have a very patchy distribution in the tree of life, i.e. only present in a few distantly related lineages, are likely to have been distributed by gene transfer events. Sulfide dehydrogenase, which was only found in some unrelated prokaryotic lineages and a single eukaryotic lineage, the diplomonads, is a good example of such a gene family [41]. Two evolutionary scenarios could account for such a distribution; either the gene family was present in the last universal ancestor and subsequently lost from all sampled eukaryotic lineages, except diplomonads, and most prokaryotic lineages, or the gene was distributed between lineages via LGT. The latter alternative seems more likely since all lineages that do have the genes are found in oxygen-poor environments, which would indicate that the organisms possessing the gene could have been in physical proximity. The phylogenetic analysis indeed recovered the diplomonad sequences within a prokaryotic cluster [41]. A similar example is the presence of a potential surface protein with a leucine-rich repeat in the parasite *Trichomonas vaginalis*, which is shared with divergent eubacterial and archaeal lineages [42]. Gene transfer appears to be the most likely explanation for the discontinuous appearance of this gene within the domains of life, since all species that encode the gene are either known to be parasites or commensals of mammalian mucosa, or have been isolated from such an environment [42]. The fact that the leucine-rich repeats are too divergent for meaningful phylogenetic analyses neither invalidates the suggestion nor makes it stronger. The unique presence of a gene in a metazoan lineage and *T. vaginalis* is another unexpected phyletic distribution involving parabasalids [43]. An expressed sequence tag (EST) clone from *Hydra* – a member the metazoan phylum Cnidaria – showed high sequence identity to three *flp* genes identified in *T. vaginalis*. The absence of homologs of the gene in any other studied species, including several protist lineages, prompted the authors to suggest that LGT is the most likely explanation for the distribution of the gene in these two distantly related eukaryotic lineages, although an origin in the last common ancestor of *Trichomonas* and *Hydra*, followed by differential loss in the lineages leading to organisms for which the genome sequences are available, could not be formally excluded [43].

However, gene families that are present in most lineages can also provide information regarding LGT events, for example if there are several easily recognized classes within the family. One such example is the four classes of glutamate dehydrogenase. No genome encodes more than

two of these four classes, yet the phylogenetic distribution of the members of the four classes mirror the organismal relationships very poorly [44]. If it is assumed that ancient genomes also encoded two classes at most, a number of LGT events affecting both prokaryotes and eukaryotes have to be inferred to explain the distribution, which indeed is supported by phylogenies of the individual glutamate dehydrogenase classes [44]. The pattern of replacements with genes with a similar function, but different origin has also been observed for whole metabolic pathways. Two alternative pathways can synthesize isopentenyl diphosphate, a universal precursor of isoprenoids; the 1-deoxy-D-xylulose 5-phosphate pathway and the mevalonate pathways. Boucher and Doolittle surveyed the distribution of these two pathways in the tree of life and found a scattered distribution which is incompatible with vertical transmission of the two pathways, and argued that LGT more parsimoniously explains the observed pattern [45]. Phylogenetic analysis of the first enzyme in the mevalonate pathway further supported that conclusion; the *Giardia lamblia* homolog belongs to a different class of the enzyme together with mostly eubacteria, to the exclusion of all other included eukaryotes – a strong case for interdomain gene transfer [45].

Endosymbiotic gene transfer from mitochondria and primary plastids vs. LGT

The endosymbionts that gave rise to the mitochondria and chloroplast of eukaryotic cells are large sources of genetic material in the eukaryotic nucleus (fig. 1AB); 18% of the *Arabidopsis* genome may have originated from the cyanobacterial ancestor of the chloroplast [46], and it has been suggested that the majority of the genes in the nucleus may have originated from the α -proteobacterial ancestor of the mitochondria [47]. Nevertheless, some of the genes that appear to be of organellar origin may indeed have been acquired via LGT. It may seem very difficult to distinguish between endosymbiotic gene transfers and eubacteria-to-eukaryote LGT events using phylogenetic methods, given that frequent gene transfer events among prokaryotes may obscure the phylogenetic relationship between nuclear genes derived from the endosymbionts and their prokaryotic ancestors [48]. However, that mainly applies for gene transfers that occurred before the extant eukaryotic lineages diverged. More recent LGT events are expected to produce rather different patterns if the taxonomic sampling is large enough and transfer events are frequent enough; endosymbiotic gene transfer events should give a single eukaryotic clade, ideally grouping with cyanobacteria or α -proteobacteria (fig. 1D), while LGT events should give polyphyletic eukaryotes, ideally nested within recognized prokaryotic groups (fig. 1E).

A recent paper that compared the yeast genome to a large number of prokaryotic genomes using a method based on similarity searches found that about 75% of the genes that showed detectable similarities with prokaryotic genes had greater amino acid identities to eubacterial than archaeal homologs [47]. The authors argued that endosymbiotic gene transfer explained the vast majority of these similarities – LGT from eubacteria was dismissed for three reasons: yeast is not phagotrophic; only six genes (0.7% of the genes with prokaryotic homologs) were absent from other eukaryotic species; and the study did not reveal any ‘recent’ LGT events – while recent mitochondrial transfers are observed frequently [49]. One problem with the detection of prokaryote-to-eukaryote gene transfers is that the donor lineage is not expected to be present in the sequence databases since only a tiny fraction of the prokaryotic diversity has been sequenced. Therefore, LGT events appear more ancient than they are, in contrast to recent transfers from the mitochondrion to the nucleus where the donor genome almost always is sequenced [49]. Also, the authors failed to test whether the eubacterial yeast genes present in other genomes showed greater sequence identity to other eukaryotes than to any eubacteria, as they should if they originated from the mitochondria rather than from diverse eubacterial lineages (fig. 1). If not, the presence of the gene in other eukaryotic lineages does not make the argument for endosymbiotic gene transfer any stronger, rather the opposite, although phylogenetic analyses are needed to draw any firm conclusions. Finally, the non-phagotrophic lifestyle of yeast is possibly a derived feature [50] – they may have had phagotrophic ancestors that acquired genes from prokaryotes via LGT. Thus, the conclusion that only a very tiny fraction of the yeast genes with eubacterial origin could be explained by LGT [47] is premature; further analyses are needed to identify the origin of these eukaryotic genes.

Gene transfers between eukaryotes independent of secondary endosymbionts?

For gene transfer events within eukaryotes it is more complicated to distinguish between LGT and endosymbiotic gene transfer events mediated by plastids, mainly because the number and nature of the secondary endosymbioses that have occurred during algal evolution are not completely understood [51–53]. However, organisms that are capable of initiate secondary endosymbiosis should also be able to ingest eukaryotic cells for other reasons, which may lead to LGT events. Therefore, the presence of a gene with foreign eukaryotic origin in a lineage harboring secondary plastids should not naively be assumed to have originated from the endosymbiont. An instructive example of this comes from a study of the algal group

Chlorachniophytes, which are amoeboflagellates that have recruited their plastids secondarily by engulfing a green alga [54]. Phylogenetic analyses of 78 plastid-targeted proteins in the *Bigeloviella natans* (a chlorachniophyte) indicated that ~20% of these were not derived from the green algal endosymbiont, as expected, but recruited from various lineages via LGT. Most of these were of non-green algal origin, but some were recruited from eubacteria [54]. Interestingly, the homologous genes of the chlorophyte *Chlamydomonas* do not show any indications of LGT. These differences in the rates of gene transfer were hypothesized to be due to the different lifestyles of the chlorophyte and the chlorachniophyte; the former is a strict autotroph, while the latter is a mixotroph which is able to be photosynthetic as well as phagotrophic – an important trait for gene transfer in eukaryotes [54].

There are additional hints that genes may be transferred between eukaryotes independent of endosymbioses. Phylogenetic analyses of the glyceraldehyde-3-phosphate (GAPDH) genes have revealed a very complex pattern of relationships which only can be explained by frequent gene transfers among and between prokaryotes and eukaryotes [55–58]. Interestingly, different lineages of dinoflagellates have acquired genes from euglenozoan lineages two, or maybe three, times independently [57]. It is currently unknown whether these intradomain transfers occurred from an endosymbiotic euglenozoan, or a euglenozoan ingested as food by the dinoflagellate [57, 59]. Enolase is another metabolic enzyme with a very complex evolutionary history – shared insertion sequences between apicomplexan parasites, ciliates and distantly related plants have been inferred as indication of an LGT event [60, 61]. Since the phylogeny of the whole gene disagreed with the distribution of the insertions, the authors suggested that only a part of the enolase gene was transferred in this event and recombined in the recipient lineage [60]. Subsequent more taxon-rich phylogenetic analyses of enolase complicated the pattern – a single LGT event in eukaryotes was clearly insufficient to explain the topology [61]. The pattern of the insertion sequences is very complex, with the species possessing the insertion showing up in five positions in the tree, indicating that LGT is unlikely to solely explain the distribution of the insertion. At any rate, the phylogenetic analyses of enolase strongly suggested a couple of eukaryote intradomain transfers. For example, the three dinoflagellate enolase paralogs show up in distinct regions, indicating separate origins via LGT from different eukaryotic lineages [61], which suggest LGT events – dinoflagellates are known to graze on eukaryotes [62]. Overall, detailed phylogenetic analyses of the metabolic enzymes GAPDH and enolase have revealed very complex evolutionary histories which indicate that many evolutionary forces have been at work, including several clear cases of in-

tradomain gene transfer, which may or may not involve secondary endosymbiosis [57, 59–61]. A better understanding of endosymbiotic events within eukaryotes is clearly needed before firm conclusions can be drawn from these kinds of data.

Gene transfer between eukaryotes that do not harbor secondary endosymbionts

Fortunately, for non-algal eukaryotes it is somewhat easier to identify intradomain transfers. For example, the finding that there are two phylogenetically distinct, but functional interchangeable, proteins involved in the eukaryotic translational machinery – the canonical EF-1 α and an EF-like (EFL) protein – is very intriguing [63]; key components in the translational machinery are thought to be highly resistant to gene transfer [5, 64]. The EFL protein is found in diverse lineages of eukaryotes, and its presence is almost always coupled with the absence of a canonical EF-1 α . Although ancient duplication and differential loss of the two classes, in principle, could explain the distribution, a scenario with frequent intradomain gene transfers of EFL coupled with losses of EF-1 α appears much more consistent with the observation [63], which suggests that proteins with complex interactions may also be replaced by functional homologs in eukaryotes. Since the putative LGT events involve both algal and non-algal eukaryotes, a scenario where only endosymbiotic events distributed the gene seems very unlikely. Another example of intradomain gene transfer of proteins involved in the translational machinery is the acquisitions of alanyl-transfer RNA (tRNA) synthetase genes in ciliates and *Entamoeba* from the parabasalid lineage [65]. Endosymbiotic relationships are unknown between the inferred donor and recipient lineages, which argues against the possibility that these genes were spread via endosymbiotic events. Rather, both ciliates and *Entamoeba* are known to be able to ingest eukaryotic cells [66–68], which suggests a mode of transfer via food.

The reports of gene transfer of mitochondrial genes between unrelated species of plants were surprising [69–71]. In all three cases strongly supported phylogenetic relationships of genes that disagree with organismal phylogenies were inferred as LGT events. The mechanism in the first two cases is unknown, although viruses have been suggested [69, 70]. The third reported case involved an LGT to a parasitic plant from the host, where the direct and prolonged contact make suggesting any intermediate vector for the transfer unnecessary [71]. Although these three reports have focused on transfer of mitochondrial genes, there is no reason to believe that the mechanism of gene transfer is not operating on the plant nuclear genome [70].

The direction of gene transfer

The vast majority of the reported interdomain transfers between prokaryotes and eukaryotes are in the direction to eukaryotes (table 1), although there have been reports that prokaryotes acquired eukaryotic genes at a significant rate [23, 72]. However, it is unclear how the direction of transfer was deduced from the similarity search-based method applied [23], and the large-scale phylogenetic analyses are problematic to evaluate since only distance-based methods were applied and detailed bootstrap support values were omitted [72]. Furthermore, rigorous phylogenetic analyses of mycobacterial genes proposed to be of eukaryotic origin [73] showed that alternative explanation not involving LGT in most cases more easily explained the distribution of the genes [74]. Thus, there seem to be very few cases of putative eukaryote-to-prokaryote gene transfer that hold up for careful analyses. The tubulin genes found in *Prostecobacter* might be an exception [75]; these eukaryotic genes are only found in this single prokaryotic lineage which live in close proximity to eukaryotes as ectosymbionts of ciliates [76]. The divergent nature of the *Prostecobacter* tubulin gene sequences revealed in the phylogenetic trees [75] most likely is due to relaxed functional constraints in the prokaryotic cell, rather than a sign of an ancient presence in prokaryotes. Another example might be the deoxyribose-phosphate aldolase gene in some proteobacteria, which appears to have a metazoan origin [77].

The skewed distribution of genome sequences in favor of prokaryotes may, in principle, bias the observations towards prokaryote-to-eukaryote transfers. However, there are several likely biological reasons for the bias. The presence of introns in many eukaryotic lineages is a potential barrier against gene transfer since an intron-containing eukaryotic gene cannot be expressed in a prokaryote. In theory, a gene transfer event could involve a spliced messenger RNA (mRNA) intermediate, although this would require a direct physical contact between the donor and recipient organisms, since RNA molecules are expected to be degraded quickly in the environment. Maybe a stronger argument for the dominance of transfers in the direction to eukaryotes is that there are many more possibilities in this direction. The last common eukaryotic ancestor most probably was a mitochondria-containing organism [78] which had two genomes – the nuclear and the mitochondrial, with a defined number of genes. LGT from prokaryotes may have been an evolutionary mechanism accelerating the diversification of eukaryotes from this common ancestor. The prokaryotic diversity was probably already huge with large biochemical diversity, and acquisitions of prokaryotic genes may have allowed eukaryotes to inhabit diverse environments, which were inaccessible for the ancestral eukaryote.

Indeed, many of the observed interdomain gene transfers involve metabolic genes that confer a new function on the eukaryotic recipients (table 1). For example, half of the genes found to have originated via LGT in a survey of diplomonad genomes turned out to be involved in anaerobic processes, suggesting that diplomonads have adapted to an anaerobic lifestyle secondarily from an aerobic ancestor [77]. Strikingly, *Entamoeba histolytica*, which have adapted to a similar lifestyle in oxygen-poor environments, acquired most of the anaerobic genes in separate LGT events [77], and several other genes have been acquired by both diplomonads and *Entamoeba* from anaerobic prokaryotes (table 1). In fact, any gene function that is present in a prokaryote and could be advantageous for a eukaryote may, in principle, be acquired by a eukaryote. Classes of genes that could be expected to be hot spots for such transfers are genes responsible for biochemistry not normally found in eukaryotes and genes that provide access to new ecological niches; LGT may turn out to be an evolutionary mechanism that is able to change the ecological and pathogenic character of microbial eukaryotes, as it has been suggested for prokaryotes [2]. In contrast, the eukaryotic genes that would be favorable in a prokaryotic organism are most likely many fewer, with the putative transfer of tubulin genes – which are central to the eukaryotic cell – as an intriguing exception [75].

Inter- or intradomain transfers?

The majority of the reported cases of gene transfers affecting eukaryotes are transfers from prokaryotes (interdomain transfers) (tables 1 and 2). Since the expression of a functional gene product is necessary for a laterally transferred gene to be maintained in the population and the gene expression machineries are different in prokaryotes and eukaryotes, the probability for a gene acquired from a prokaryote to be successfully expressed in a eukaryote should be lower than for a gene acquired from another eukaryote. Therefore, the observed bias towards interdomain transfers is counterintuitive, and may indeed be a false reflection of the underlying biological pattern. Interdomain transfers are relatively easy to detect with phylogenetic methods since the taxonomic sampling of prokaryotic genome sequences is becoming reasonably good, and eukaryotic sequences nested within prokaryotic sequences is a strong signal in favor of a transfer event. On the other hand, the taxonomic sampling of eukaryotic genome sequences is still very sparse, with only a minority of major eukaryotic groups represented, making transfers between eukaryotes difficult to detect – the probability that a close relative to both the eukaryotic donor and recipient lineage is sampled should be very small. Also, the knowledge of eukaryote phylogeny is rel-

atively poor, and only recently have a handful of ‘super-groups’ been identified [50, 79, 80]. Thus, in the absence of good knowledge of organismal phylogeny, it has been difficult to identify unexpected phylogenetic relationships between eukaryotic lineages in gene trees that would indicate LGT events.

However, the observed bias towards inter- compared with intradomain transfers may also be due to a higher exposure of prokaryotic than eukaryotic DNA in many eukaryotic lineages. If ingestion of cells via food is indeed an important mechanism for gene transfer in eukaryotes [14, 15], a higher fraction of prokaryotes than eukaryotes in the food for phagotrophic eukaryotes could explain the dominance of interdomain transfer. Indeed, many interdomain transfers have been reported for diplomonads – small flagellates that mainly ingest prokaryotes [81] – as the recipient lineages (table 1). On the other hand, larger phagotrophic species, such as dinoflagellates and ciliates, which ingest eukaryotes as well as prokaryotes [62], appear more often to be the recipient lineages in intradomain LGT events (table 2). Comparative studies of the frequency and patterns of LGT in phagotrophic species will be able to test this hypothesis – the number of events is currently very small. If it turns out that there exists a connection between feeding habits and patterns of gene transfer, it is a strong indication that the phagotrophy indeed is most relevant as a mechanism for gene transfers [14, 15].

LGT is frequent in phagotrophic eukaryotes ...

The cases listed in tables 1 and 2 indicate that LGT is an evolutionary mechanism that operates in many eukaryotic lineages, although there seem to be large variations in frequency between lineages. Generally, phagotrophs appear to be affected at a substantial rate; the number of reports in some lineages, such as diplomonads, parabasalids and *Entamoeba*, hints at an important role for LGT in these groups (table 1), and several individual cases have been reported from other phagotrophic lineages, for example dinoflagellates and ciliates (tables 1 and 2). In an attempt to estimate the frequencies of gene transfer in the glutamate dehydrogenase gene families, comparable rates were found for eukaryotes and prokaryotes; three transfers identified between eubacteria and eukaryotes, two transfers between the two prokaryotic domains and seven transfers within eubacteria [44]. Similarly, analyses of prolyl- and alanyl-tRNA synthetase identified comparable numbers of LGT event affecting prokaryotes and phagotrophic eukaryotes [65]. Two intradomain and a putative interdomain prokaryotic LGT event were inferred, while two archaea-to-eukaryote transfers, two eukaryote-to-eukaryote and two putative eubacteria-to-eukaryote transfers were inferred from the

phylogenetic analyses, all affecting phagotrophic eukaryotes [65]. Similar conclusions may be drawn from inspections of phylogenetic analyses of metabolic enzymes, such as GAPDH [55–58] and alcohol dehydrogenase (fig. 2). Thus, rates of gene transfer in phagotrophic eukaryotes may be comparable to rates of free-living prokaryotes, although a much larger number of phylogenetic analyses including homologous genes from diverse organisms of all three domains of life are needed to test this hypothesis.

... but much rarer in fungi and animals

There seems to be a high variation of the frequency of LGT within eukaryotes; animals and fungi are much less affected than phagotrophic protists. Two recent comparative genomics studies have failed to identify gene transfer as an important mechanism in the evolution of yeast genomes [47, 82]. Only a few lineage-specific yeast proteins that showed similarities to prokaryotic homologs were detected, indicating a low rate of gene transfer in yeasts. Still, clusters of genes encoding secondary metabolite enzymes in fungi have been taken as an indication that LGT is important in their evolution [11, 83] – gene transfer has been suggested to maintain gene clusters in prokaryotes (‘selfish operons’) [84], although phylogenetic analyses of one of the families of secondary metabolite enzymes – polyketide synthetases – indicated that gene duplications and losses explained the data equally well, suggesting only a minor role of LGT [85]. Nevertheless, gene transfer events are not totally absent from the genome evolution of fungi; the yeast flavohemoglobin has been identified to have bacterial origin [77], and most enzymes that play important roles in the degradation of cellulose in rumen fungi have been shown to have been acquired via LGT from rumen bacteria, indicating that the process of gene transfer was probably the key event for the colonization of the rumen by fungi [86]. Additionally, discontinuous phylogenetic distribution of pea pathogenicity gene clusters in two filamentous fungi suggests that transfer of genes could, at least occasionally, be involved in the evolution of fungal pathogenicity [87]. There have been few reports of LGT affecting animals that have held up for rigorous analyses, despite the availability of a large amount of genome sequence data. There are indeed a few notable exceptions; ascidian urochordates appear to have acquired a cellulose synthase gene from eubacteria [88, 89], 12 genes involved in the plant-parasitic lifestyle of the nematode *Meloidogyne* are strong candidates to have been acquired via LGT, mostly from rhizobial eubacterial species [90], phylogenetic analyses identified eubacterial sources for two metazoan metabolic genes [77] and a fragment of the genome of the *Wolbachia* endosymbiont of the beetle *Callosobruchus*

chinensis apparently has been transferred to the host nucleus [91]. Still, gene transfer is likely a very rare event in the evolution of animals. Similarly, the lack of examples of LGT from non-plant sources into plants – the transfer of *Agrobacterium* genes to the *Nicotiana* lineage is an exception [92] – indicates that the frequency of such transfers is very low. Nonetheless, the recent findings that gene transfer is a widespread phenomenon among the plant mitochondrial genomes [69–71] suggest that gene transfer between plant nuclear genomes may also be common [70]. Further studies of animal mitochondrial, and nuclear, genes are needed to examine whether animal-to-animal LGT occurs, as in the cases of plants.

Large variations of the rates of LGT in eukaryotes

The differences in the frequencies of gene transfers between eukaryotic lineages are likely to be due to a number of factors. Exposure to foreign DNA of the germ line is probably a main factor; it is easy to imagine that the genome of a phagotrophic protist experiences a larger amount of foreign genes than the sequestered germ lines of animals. The observation that ciliates do participate in gene transfer events (tables 1 and 2), although they have sequestered germ lines, is likely explained by the fact that they are unicellular and the germ lines possibly are exposed to foreign genetic material since they are phagotrophic. Other factors affecting the lineage-specific rates of LGT might include the efficiency of incorporation of DNA into the genome, and the flexibility of gene expression mechanisms (i.e. the same prokaryotic gene may have widely different probabilities to be expressed in different eukaryotes). In addition, factors such as the strength of selection for the incorporated gene, and population sizes, vary between lineages and affect the probability of a successful LGT. Obviously, understanding the process of gene transfer in eukaryotes is currently limited; phagotrophy likely explains a number of transfers (tables 1 and 2) [14, 15], but the mechanisms behind some of the examples remain enigmatic. For example, the lifestyle of Apicomplexa – they are non-phagotrophic intracellular parasites – may suggest that they are resistance to gene transfer, yet there are many published examples from this group [34–37]. Although transfer in a free-living phagotrophic ancestor [93] is a putative explanation for some of these, it is not satisfactory for all, since many of the transfers are unique to a subset of the available apicomplexan genomes [34–37].

LGT and eukaryotic origin

The occurrence of LGT makes the evolutionary history of eukaryotic genomes much more complex than previously

thought, which has important implication for the various 'fusion' hypotheses for the origin of the eukaryotic cells [47, 94–101] that have been proposed to explain the observation that many eukaryotic genes are more closely related to Eubacteria than Archaea [47, 95, 99–101]. The origins of all eukaryotic members of a gene family may not necessarily be the same, since some lineages may have acquired the gene by LGT from prokaryotes after divergence of extant lineages. Therefore, the fusion hypothesis cannot be based on prokaryotic phylogenetic affinities of genes in a single eukaryotic genome, such as yeast [95, 98]. Rather, the sampling of eukaryotes should be as broad as possible to identify genes that truly were present in the last common eukaryotic ancestor. However, even such genes could have been acquired from prokaryotes via LGT after the eukaryotic lineage diverged from Archaea and Eubacteria. The small subunit glutamate synthase may be one such example; this gene was probably transferred from a low G+C Gram-positive eubacteria to an ancestor of animals and plants, which indicates that it was already present in the last common ancestor of all extant eukaryotes [41].

Obviously, the support for the various 'fusion' hypotheses needs to be reanalyzed. The signal from proposed fusion events may be detectable behind the noise introduced by subsequent LGT events. Perhaps more likely, phylogenetic support may fall apart when analyzed with more realistic assumptions. If so, we are left with one of the largest unsolved problems in evolutionary biology: how the eukaryote cell originated from prokaryotic cells. Clearly, the possibility that continuous transfer from LGT from various prokaryotic lineages played an important role in the origin and evolution of eukaryotic organisms should be taken into account in any future hypotheses about the origin of the eukaryotic cell [14, 15, 77, 102, 103].

LGT and eukaryote phylogeny

The effects of LGT on the possibilities of reconstructing an organismal tree for prokaryotes has been widely debated [1, 4, 5, 31]. Some authors have questioned the meaning of an organismal phylogeny for prokaryotes in the presence of frequent gene transfer [1, 31]. Similarly, the possibility to resolve the eukaryote phylogeny is dependent on the frequency of LGT. Intradomain gene replacements – which appear to be relatively frequent (table 2) – are especially problematic; these are difficult to detect since phylogenetic signals from such events are easily mistaken as signals from organismal relationships. For example, the recent finding that EF-1 α probably was replaced by a functional similar EF-like homolog in many eukaryotic lineages suggests that EF-1 α may also undergo gene transfer [63], indicating that phylogenies

based on this marker, which already previously has been identified as problematic [104], should be treated cautiously. Thus, LGT will probably turn out to be a major challenge for efforts to reconstruct eukaryote phylogenies.

Nevertheless, important advances in the field of eukaryote phylogeny have occurred recently [50, 79, 80], although the phylogeny of eukaryotes is far from resolved. In addition to datasets of concatenated protein sequences [32, 33], rare evolutionary events such as insertions in gene sequences and gene fusions have been used to pinpoint phylogenetic relationships between eukaryotic groups [105, 106]. These markers have been shown to be very useful complements to methods of phylogenetic inference based on gene sequences to resolve deep relationships within eukaryotes, since such relationships have been notoriously difficult to resolve with traditional molecular methods. Similarly, interdomain gene transfer events shared between eukaryotic groups may be very useful in untangling eukaryote phylogeny. For example, two interdomain transfer events from the archaeal phylum Nanoarchaeota shared between diplomonads and parabasalids unite these two eukaryotic groups, which both have been very difficult to place in the tree of life [65]. Clearly, gene transfer events can be both advantageous and disadvantageous for efforts to reconstruct eukaryote phylogeny.

Concluding remarks

The occurrence of LGT in eukaryotes has been neglected for a long time. Recent findings indicate that the process does occur in most lineages of eukaryotes to some degree, and probably is an important mechanism in many. Therefore, the possibility of gene transfer should not be dismissed in eukaryotic comparative genomics studies; rather it should be integrated as an additional mechanism by which eukaryotic organisms may evolve. If incorporated, the process of LGT will probably explain aspects of eukaryote genome evolution that have been problematic to understand with previously acknowledged processes such as gene duplication and divergence, and gene transfer from organelles. This is obviously important to consider for a deeper understanding of the origin and phylogeny of eukaryotes, but an awareness of LGT is also needed for everybody using 'lower' eukaryotes as model organisms to study basic eukaryotic functions. Strict vertical transmission of these functions should not naively be assumed; the presence of a function in one eukaryotic lineage does not necessarily indicate its presence in all lineages. These kind of studies could also have practical implications – identification of prokaryotic genes in human parasites may be used to identify suitable drug targets. Also, a better understanding of the mechanisms and pat-

terns of gene transfer in eukaryotes would help risk assessments for release of genetically modified organisms in the field [107]. Future studies of LGT will most likely contribute new insights into many different aspects of eukaryotic genomics in unforeseen ways.

Acknowledgements. The author is supported by a Swedish Research Council (VR) Grant.

- 1 Doolittle W. F. (1999) Phylogenetic classification and the universal tree. *Science* **284**: 2124–2129
- 2 Ochman H., Lawrence J. G. and Groisman E. A. (2000) Lateral gene transfer and the nature of bacterial innovation. *Nature* **405**: 299–304
- 3 Boucher Y., Douady C. J., Papke R. T., Walsh D. A., Boudreau M. E. R., Nesbø C. L. et al. (2003) Lateral gene transfer and the origins of prokaryotic groups. *Annu. Rev. Genet.* **37**: 283–328
- 4 Kurland C. G., Canback B. and Berg O. G. (2003) Horizontal gene transfer: a critical view. *Proc. Natl. Acad. Sci. USA* **100**: 9658–9662
- 5 Daubin V., Moran N. A. and Ochman H. (2003) Phylogenetics and the cohesion of bacterial genomes. *Science* **301**: 829–832
- 6 Timmis J. N., Ayliffe M. A., Huang C. Y. and Martin W. (2004) Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5**: 123–135
- 7 The International Human Genome Sequencing Consortium (2001) Initial sequencing and analysis of the human genome. *Nature* **409**: 860–921
- 8 Stanhope M. J., Lupas A., Italia M. J., Koretke K. K., Volker C. and Brown J. R. (2001) Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. *Nature* **411**: 940–944
- 9 Geneux D. P. and Logsdon J. M. Jr (2003) Much ado about bacteria-to-vertebrate lateral gene transfer. *Trends Genet.* **19**: 191–195
- 10 Andersson J. O., Doolittle W. F. and Nesbø C. L. (2001) Are there bugs in our genome? *Science* **292**: 1848–1850
- 11 Rosewich U. L. and Kistler H. C. (2000) Role of horizontal gene transfer in the evolution of fungi. *Annu. Rev. Phytopathol.* **38**: 325–363
- 12 Kidwell M. G. (1993) Lateral transfer in natural populations of eukaryotes. *Annu. Rev. Genet.* **27**: 235–256
- 13 Adams K. L. and Palmer J. D. (2003) Evolution of mitochondrial gene content: gene loss and transfer to the nucleus. *Mol. Phylogenet. Evol.* **29**: 380–395
- 14 Doolittle W. F. (1998) You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* **14**: 307–311
- 15 Doolittle W. F., Boucher Y., Nesbø C. L., Douady C. J., Andersson J. O. and Roger A. J. (2003) How big is the iceberg of which organellar genes in nuclear genomes are but the tip? *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **358**: 39–58
- 16 Gogarten J. P. (2003) Gene transfer: gene swapping craze reaches eukaryotes. *Curr. Biol.* **13**: R53–R54
- 17 Andersson J. O. and Andersson S. G. E. (1999) Insights into the evolutionary process of genome degradation. *Curr. Opin. Genet. Dev.* **9**: 664–671
- 18 Lawrence J. G. and Ochman H. (2002) Reconciling the many faces of lateral gene transfer. *Trends Microbiol.* **10**: 1–4
- 19 Brown J. R. (2003) Ancient horizontal gene transfer. *Nat. Rev. Genet.* **4**: 121–132
- 20 Lawrence J. G. and Hendrickson H. (2003) Lateral gene transfer: when will adolescence end? *Mol. Microbiol.* **50**: 739–749
- 21 Ragan M. A. (2001) Detection of lateral gene transfer among microbial genomes. *Curr. Opin. Genet. Dev.* **11**: 620–626
- 22 Koonin E. V. (2003) Horizontal gene transfer: the path to maturity. *Mol. Microbiol.* **50**: 725–727
- 23 Koonin E. V., Makarova K. S. and Aravind L. (2001) Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55**: 709–742
- 24 Nesbø C. L., L'Haridon S., Stetter K. O. and Doolittle W. F. (2001) Phylogenetic analyses of two 'archaeal' genes in *Thermotoga maritima* reveal multiple transfers between Archaea and Bacteria. *Mol. Biol. Evol.* **18**: 362–375
- 25 Nelson K. E., Clayton R. A., Gill S. R., Gwinn M. L., Dodson R. J., Haft D. H. et al. (1999) Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**: 323–329
- 26 Koski L. B. and Golding G. B. (2001) The closest BLAST hit is often not the nearest neighbor. *J. Mol. Evol.* **52**: 540–542
- 27 Andersson J. O. (2003) Bacterial DNA in the human genome. In: *The Encyclopedia of the Human Genome*, Cooper D. N. (ed.), Nature Publishing Group, London
- 28 Lawrence J. G. and Ochman H. (1997) Amelioration of bacterial genomes: rates of change and exchange. *J. Mol. Evol.* **44**: 383–397
- 29 Lawrence J. G. and Ochman H. (1998) Molecular archaeology of the *Escherichia coli* genome. *Proc. Natl. Acad. Sci. USA* **95**: 9413–9417
- 30 Richards T. A., Hirt R. P., Williams B. A. and Embley T. M. (2003) Horizontal gene transfer and the evolution of parasitic protozoa. *Protist* **154**: 17–32
- 31 Gogarten J. P., Doolittle W. F. and Lawrence J. G. (2002) Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* **19**: 2226–2238
- 32 Baldauf S. L., Roger A. J., Wenk-Siefert I. and Doolittle W. F. (2000) A kingdom-level phylogeny of eukaryotes based on combined protein data. *Science* **290**: 972–977
- 33 Baptiste E., Brinkmann H., Lee J. A., Moore D. V., Sensen C. W., Gordon P. et al. (2002) The analysis of 100 genes supports the grouping of three highly divergent amoebae: *Dictyostelium*, *Entamoeba* and *Mastigamoeba*. *Proc. Natl. Acad. Sci. USA* **99**: 1414–1419
- 34 Huang J., Mullapudi N., Sicheritz-Pontén T. and Kissinger J. C. (2004) A first glimpse into the pattern and scale of gene transfer in Apicomplexa. *Int. J. Parasitol.* **34**: 265–274
- 35 Huang J., Mullapudi N., Lancto C. A., Scott M., Abrahamsen M. S. and Kissinger J. C. (2004) Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol.* **5**: R88
- 36 Striepen B., White M. W., Li C., Guerini M. N., Malik S.-B., Logsdon J. M. Jr et al. (2002) Genetic complementation in apicomplexan parasites. *Proc. Natl. Acad. Sci. USA* **99**: 6304–6309
- 37 Striepen B., Pruijssers A. J., Huang J., Li C., Gubbels M. J., Umejiego N. N. et al. (2004) Gene transfer in the evolution of parasite nucleotide biosynthesis. *Proc. Natl. Acad. Sci. USA* **101**: 3154–3159
- 38 Welch R. A., Burland V., Plunkett G. 3rd, Redford P., Roesch P., Rasko D. et al. (2002) Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc. Natl. Acad. Sci. USA* **99**: 17020–17024
- 39 Rocap G., Larimer F. W., Lamerdin J., Malfatti S., Chain P., Ahlgren N. A. et al. (2003) Genome divergence in two *Prochlorococcus* ecotypes reflects oceanic niche differentiation. *Nature* **424**: 1042–1047
- 40 Carlton J. M., Angiuoli S. V., Suh B. B., Kooij T. W., Pertea M., Silva J. C. et al. (2002) Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. *Nature* **419**: 512–519
- 41 Andersson J. O. and Roger A. J. (2002) Evolutionary analyses of the small subunit of glutamate synthase: gene order con-

- servation, gene fusions and prokaryote-to-eukaryote lateral gene transfers. *Eukaryot. Cell* **1**: 304–310
- 42 Hirt R. P., Harriman N., Kajava A. V. and Embley T. M. (2002) A novel potential surface protein in *Trichomonas vaginalis* contains a leucine-rich repeat shared by micro-organisms from all three domains of life. *Mol. Biochem. Parasitol.* **125**: 195–199
 - 43 Steele R. E., Hampson S. E., Stover N. A., Kibler D. F. and Bode H. R. (2004) Probable horizontal transfer of a gene between a protist and a cnidarian. *Curr. Biol.* **14**: R298–299
 - 44 Andersson J. O. and Roger A. J. (2003) Evolution of glutamate dehydrogenase genes: evidence for lateral gene transfer within and between prokaryotes and eukaryotes. *BMC Evol. Biol.* **3**: 14
 - 45 Boucher Y. and Doolittle W. F. (2000) The role of lateral gene transfer in the evolution of isoprenoid biosynthesis pathways. *Mol. Microbiol.* **37**: 703–716
 - 46 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
 - 47 Esser C., Ahmadinejad N., Wiegand C., Rotte C., Sebastiani F., Gelius-Dietrich G. et al. (2004) A genome phylogeny for mitochondria among α -proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* **21**: 1643–1660
 - 48 Martin W. (1999) Mosaic bacterial chromosomes: a challenge en route to a tree of genomes. *Bioessays* **21**: 99–104
 - 49 Bensasson D., Zhang D., Hartl D. L. and Hewitt G. M. (2001) Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends Ecol. Evol.* **16**: 314–321
 - 50 Cavalier-Smith T. (2002) The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. *Int. J. Syst. Evol. Microbiol.* **52**: 297–354
 - 51 Archibald J. M. and Keeling P. J. (2002) Recycled plastids: a 'green movement' in eukaryotic evolution. *Trends Genet.* **18**: 577–584
 - 52 Cavalier-Smith T. (2002) Chloroplast evolution: secondary symbiogenesis and multiple losses. *Curr. Biol.* **12**: R62–64
 - 53 Bhattacharya D., Yoon H. S. and Hackett J. D. (2004) Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *Bioessays* **26**: 50–60
 - 54 Archibald J. M., Rogers M. B., Toop M., Ishida K. and Keeling P. J. (2003) Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga *Bigelowiella natans*. *Proc. Natl. Acad. Sci. USA* **100**: 7678–7683
 - 55 Figge R. M., Schubert M., Brinkmann H. and Cerff R. (1999) Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra- and inter-kingdom gene transfer. *Mol. Biol. Evol.* **16**: 429–440
 - 56 Figge R. M. and Cerff R. (2001) GAPDH gene diversity in spirochetes: a paradigm for genetic promiscuity. *Mol. Biol. Evol.* **18**: 2240–2249
 - 57 Takishita K., Ishida K. and Maruyama T. (2003) An enigmatic GAPDH gene in the symbiotic dinoflagellate genus *Symbiodinium* and its related species (the order Süssiales): possible lateral gene transfer between two eukaryotic algae, dinoflagellate and euglenophyte. *Protist* **154**: 443–454
 - 58 Qian Q. and Keeling P. J. (2001) Diplonemid glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and prokaryote-to-eukaryote lateral gene transfer. *Protist* **152**: 193–201
 - 59 Fagan T. F. and Woodland Hastings J. (2002) Phylogenetic analysis indicates multiple origins of chloroplast glyceraldehyde-3-phosphate dehydrogenase genes in dinoflagellates. *Mol. Biol. Evol.* **19**: 1203–1207
 - 60 Keeling P. J. and Palmer J. D. (2001) Lateral transfer at the gene and subgenomic levels in the evolution of eukaryotic enolase. *Proc. Natl. Acad. Sci. USA* **98**: 10745–10750
 - 61 Harper J. T. and Keeling P. J. (2004) Lateral gene transfer and the complex distribution of insertions in eukaryotic enolase. *Gene* **340**: 227–235
 - 62 Tillmann U. (2004) Interactions between planktonic microalgae and protozoan grazers. *J. Eukaryot. Microbiol.* **51**: 156–168
 - 63 Keeling P. J. and Inagaki Y. (2004) A class of eukaryotic GT-Pase with a punctate distribution suggesting multiple functional replacements of translation elongation factor 1 α . *Proc. Natl. Acad. Sci. USA* **101**: 15380–15385
 - 64 Jain R., Rivera M. C. and Lake J. A. (1999) Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. USA* **96**: 3801–3806
 - 65 Andersson J. O., Sarchfield S. W. and Roger A. J. (2004) Gene transfers from Nanoarchaeota to an ancestor of diplomonads and parabasalids. *Mol. Biol. Evol.* **22**: 85–90
 - 66 Pierce R. W. and Turner J. P. (1992) Ecology of planktonic ciliates in marine food webs. *Rev. Aquat. Sci.* **6**: 139–181
 - 67 Clark C. G. and Diamond L. S. (2002) Methods for cultivation of luminal parasitic protists of clinical importance. *Clin. Microbiol. Rev.* **15**: 329–341
 - 68 Diamond L. S. (1968) Improved method for the monoxenic cultivation of *Entamoeba histolytica* Schaudinn, 1903 and *E. histolytica*-like amebae with trypanosomatids. *J. Parasitol.* **54**: 715–719
 - 69 Won H. and Renner S. S. (2003) Horizontal gene transfer from flowering plants to Gnetum. *Proc. Natl. Acad. Sci. USA* **100**: 10824–10829
 - 70 Bergthorsson U., Adams K. L., Thomason B. and Palmer J. D. (2003) Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature* **424**: 197–201
 - 71 Davis C. C. and Wurdack K. J. (2004) Host-to-parasite gene transfer in flowering plants: phylogenetic evidence from Malpighiales. *Science* **305**: 676–678
 - 72 Wolf Y. I., Aravind L., Grishin N. V. and Koonin E. V. (1999) Evolution of aminoacyl-tRNA synthetases-analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res.* **9**: 689–710
 - 73 Gamielidien J., Ptitsyn A. and Hide W. (2002) Eukaryotic genes in *Mycobacterium tuberculosis* could have a role in pathogenesis and immunomodulation. *Trends Genet.* **18**: 5–8
 - 74 Kinsella R. J. and McInerney J. O. (2003) Eukaryotic genes in *Mycobacterium tuberculosis*? Possible alternative explanations. *Trends Genet.* **19**: 687–689
 - 75 Jenkins C., Samudrala R., Anderson I., Hedlund B. P., Petroni G., Michailova N. et al. (2002) Genes for the cytoskeletal protein tubulin in the bacterial genus *Prosthecoacter*. *Proc. Natl. Acad. Sci. USA* **99**: 17049–17054
 - 76 Petroni G., Spring S., Schleifer K. H., Verni F. and Rosati G. (2000) Defensive extrusive ectosymbionts of *Euplotidium* (Ciliophora) that contain microtubule-like structures are bacteria related to *Verrucomicrobia*. *Proc. Natl. Acad. Sci. USA* **97**: 1813–1817
 - 77 Andersson J. O., Sjögren Å. M., Davis L. A. M., Embley T. M. and Roger A. J. (2003) Phylogenetic analyses of diplomonad genes reveal frequent lateral gene transfers affecting eukaryotes. *Curr. Biol.* **13**: 94–104
 - 78 Roger A. J. and Silberman J. D. (2002) Cell evolution: mitochondria in hiding. *Nature* **418**: 827–829
 - 79 Baldauf S. L. (2003) The deep roots of eukaryotes. *Science* **300**: 1703–1706
 - 80 Simpson A. G. and Roger A. J. (2004) The real 'kingdoms' of eukaryotes. *Curr. Biol.* **14**: R693–696
 - 81 Brugerolle G. and Lee J. J. (2002) Order Diplomonadida. In: *An Illustrated Guide to the Protozoa*, 2nd edn, vol. pp. 1125–1135, Lee J. J., Leedale G. F., Bradbury P. (eds.), Society of Protozoologists, Lawrence, Kansas

- 82 Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Lafontaine I. et al. (2004) Genome evolution in yeasts. *Nature* **430**: 35–44
- 83 Walton J. D. (2000) Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: an hypothesis. *Fungal Genet. Biol.* **30**: 167–171
- 84 Lawrence J. G. and Roth J. R. (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**: 1843–1860
- 85 Kroken S., Glass N. L., Taylor J. W., Yoder O. C. and Turgeon B. G. (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proc. Natl. Acad. Sci. USA* **100**: 15670–15675
- 86 Garcia-Vallve S., Romeu A. and Palau J. (2000) Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. *Mol. Biol. Evol.* **17**: 352–361
- 87 Temporini E. D. and VanEtten H. D. (2004) An analysis of the phylogenetic distribution of the pea pathogenicity genes of *Nectria haematococca* MPVI supports the hypothesis of their origin by horizontal transfer and uncovers a potentially new pathogen of garden pea: *Neocosmospora boniensis*. *Curr. Genet.* **46**: 29–36
- 88 Nakashima K., Yamada L., Satou Y., Azuma J. and Satoh N. (2004) The evolutionary origin of animal cellulose synthase. *Dev. Genes Evol.* **214**: 81–88
- 89 Matthyse A. G., Deschet K., Williams M., Marry M., White A. R. and Smith W. C. (2004) A functional cellulose synthase from ascidian epidermis. *Proc. Natl. Acad. Sci. USA* **101**: 986–991
- 90 Scholl E. H., Thorne J. L., McCarter J. P. and Bird D. M. (2003) Horizontally transferred genes in plant-parasitic nematodes: a high-throughput genomic approach. *Genome Biol.* **4**: R39
- 91 Kondo N., Nikoh N., Ijichi N., Shimada M. and Fukatsu T. (2002) Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect. *Proc. Natl. Acad. Sci. USA* **99**: 14280–14285
- 92 Intriери M. C. and Buiatti M. (2001) The horizontal transfer of *Agrobacterium rhizogenes* genes and the evolution of the genus *Nicotiana*. *Mol. Phylogenet. Evol.* **20**: 100–110
- 93 Leander B. S., Kuvardina O. N., Aleshin V. V., Mylnikov A. P. and Keeling P. J. (2003) Molecular phylogeny and surface morphology of *Colpodella edax* (Alveolata): insights into the phagotrophic ancestry of apicomplexans. *J. Eukaryot. Microbiol.* **50**: 334–340
- 94 Gupta R. S. and Golding G. B. (1996) The origin of the eukaryotic cell. *Trends Biochem. Sci.* **21**: 166–171
- 95 Horiike T., Hamada K., Kanaya S. and Shinozawa T. (2001) Origin of eukaryotic cell nuclei by symbiosis of Archaea in Bacteria is revealed by homology-hit analysis. *Nat. Cell. Biol.* **3**: 210–214
- 96 Hartman H. and Fedorov A. (2002) The origin of the eukaryotic cell: a genomic investigation. *Proc. Natl. Acad. Sci. USA* **99**: 1420–1425.
- 97 Gupta R. S. (1998) Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria and eukaryotes. *Microbiol. Mol. Biol. Rev.* **62**: 1435–1491
- 98 Rivera M. C. and Lake J. A. (2004) The ring of life provides evidence for a genome fusion origin of eukaryotes. *Nature* **431**: 152–155
- 99 Ribeiro S. and Golding G. B. (1998) The mosaic nature of the eukaryotic nucleus. *Mol. Biol. Evol.* **15**: 779–788
- 100 Rivera M. C., Jain R., Moore J. E. and Lake J. A. (1998) Genomic evidence for two functionally distinct gene classes. *Proc. Natl. Acad. Sci. USA* **95**: 6239–6244
- 101 Brown J. R. and Doolittle W. F. (1997) Archaea and the prokaryote-to-eukaryote transition. *Microbiol. Mol. Biol. Rev.* **61**: 456–502
- 102 Katz L. A. (1999) The tangled web: gene genealogies and the origin of eukaryotes. *Am. Nat.* **154**: S137–S145
- 103 Katz L. A. (2002) Lateral gene transfers and the evolution of eukaryotes: theories and data. *Int. J. Syst. Evol. Microbiol.* **52**: 1893–1900
- 104 Roger A. J., Sandblom O., Doolittle W. F. and Philippe H. (1999) An evaluation of elongation factor 1 alpha as a phylogenetic marker for eukaryotes. *Mol. Biol. Evol.* **16**: 218–233
- 105 Stechmann A. and Cavalier-Smith T. (2002) Rooting the eukaryote tree by using a derived gene fusion. *Science* **297**: 89–91
- 106 Archibald J. M., Longet D., Pawlowski J. and Keeling P. J. (2003) A novel polyubiquitin structure in cercozoa and foraminifera: evidence for a new eukaryotic supergroup. *Mol. Biol. Evol.* **20**: 62–66
- 107 Dröge M., Pühler A. and Selbitschka W. (1998) Horizontal gene transfer as a biosafety issue: a natural phenomenon of public concern. *J. Biotechnol.* **64**: 75–90
- 108 Nixon J. E. J., Wang A., Field J., Morrison H. G., McArthur A. G., Sogin M. L. et al. (2002) Evidence for lateral transfer of genes encoding ferredoxins, nitroreductases, NADH oxidase, and alcohol dehydrogenase 3 from anaerobic prokaryotes to *Giardia lamblia* and *Entamoeba histolytica*. *Eukaryot. Cell* **1**: 181–190
- 109 Suguri S., Henze K., Sánchez L. B., Moore D. V. and Müller M. (2001) Archaeobacterial relationships of the phosphoenolpyruvate carboxykinase gene reveal mosaicism of *Giardia intestinalis* core metabolism. *J. Eukaryot. Microbiol.* **48**: 493–497
- 110 van der Giezen M., Cox S. and Tovar J. (2004) The iron-sulfur cluster assembly genes *iscS* and *iscU* of *Entamoeba histolytica* were acquired by horizontal gene transfer. *BMC Evol. Biol.* **4**: 7
- 111 Field J., Rosenthal B. and Samuelson J. (2000) Early lateral transfer of genes encoding malic enzyme, acetyl-CoA synthetase and alcohol dehydrogenases from anaerobic prokaryotes to *Entamoeba histolytica*. *Mol. Microbiol.* **38**: 446–455
- 112 de Koning A. P., Brinkman F. S. L., Jones S. J. M. and Keeling P. J. (2000) Lateral gene transfer and metabolic adaptation in the human parasite *Trichomonas vaginalis*. *Mol. Biol. Evol.* **17**: 1769–1773
- 113 Henze K., Horner D. S., Suguri S., Moore D. V., Sánchez L. B., Müller M. et al. (2001) Unique phylogenetic relationship of glucokinase and glucosephosphate isomerase of the amitochondriate eukaryotes *Giardia intestinalis*, *Spironucleus barkhanus* and *Trichomonas vaginalis*. *Gene* **281**: 123–131
- 114 Guindon S. and Gascuel O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**: 696–704
- 115 Strimmer K. and von Haeseler A. (1996) Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**: 964–969