Functional horizontal gene transfer from bacteria to eukaryotes

Filip Husnik[1,2] and John P. McCutcheon[2]

1. Faculty of Science, University of South Bohemia and Institute of Parasitology, Biology Centre ASCR, Ceske Budejovice 370 05, Czech Republic.

2. Division of Biological Sciences, University of Montana, Missoula, MT 59812, United States of America.

Abstract

The antiquity of bacteria and archaea in part explains why they, along with viruses, encode most of the genetic and biochemical diversity on Earth. Eukaryotic life evolved into a world teeming with prokaryotes, and so bacteria (especially) have inevitably affected eukaryotic biology as parasitic, commensal, or beneficial symbionts. But along with these important organismal interactions, the ubiquity and diversity of bacteria have also made them frequent sources of horizontally transferred DNA into eukaryotic genomes. Here we survey the role of bacterial genes throughout the eukaryotic lineage. We review what steps horizontal gene transfers (HGTs) take in becoming functional, what bacterial groups these HGTs come from, and what functions these HGTs typically bestow on their eukaryotic recipient. We classify HGTs into two broad types: those that maintain preexisting functions and those that add new functionality to the recipient. We find that genes involved in host nutrition, protection, and adaptation to extreme environments are the most common HGTs from bacterial to eukaryotic genomes.

Key points

- Genome fragments are sometimes transferred from bacteria to eukaryotes (Horizontal Gene Transfers, or HGTs)
- When these DNA fragments contain genes, these genes can retain their functionality in some cases
- If these bacterial HGTs are maintained for long periods of time, they can acquire eukaryotic features such as introns
- If the eukaryotic recipient retains a stable bacterial endosymbiont, these HGTs can compensate for genome reduction in the endosymbiont
- These HGTs can also allow the eukaryotic recipient to protect itself from other organisms, survive in new environments, and use new food sources
- Further study of neglected eukaryotic groups will help clarify the frequency of bacteria-eukaryote HGT

Glossary terms:

Horizontal Gene Transfer (HGT): movement of genetic material between organisms (also called lateral gene transfer, LGT) via non-vertical (not parent to offspring) transmission.

Nuclear mitochondrial/plastid transfer (numt/nupt): transfer of mitochondrial/plastid DNA into the nuclear genome of its eukaryotic host, which often becomes non-functional.

Nuclear *Wolbachia* transfer (nuwt): transfer of *Wolbachia* DNA into the nuclear genome of its eukaryotic host.

Primary symbiosis: a process in which an archaeon/eukaryote acquires a bacterium, such as in the origin of mitochondria and plastids.

Non-homologous end joining (NHEJ): a pathway for direct repair of double-strand DNA breaks without a homologous template.

Peptidoglycan: a structural matrix in bacterial cell walls formed by alternating N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues, where peptide chains of up to five amino acids link NAM to other NAM-connected peptides.

Reproductive manipulator: a bacterium such as *Wolbachia* that is transmitted in the egg cytoplasm of arthropods and shifts the sex ratio of the host population.

Pore-forming toxins (PFTs): proteins used by bacteria to make holes in membranes of target cells.

Nucleomorph: a remnant nucleus found between two sets of membranes in some complex plastids of cryptomonads and chlorarachniophytes that suggests that the plastid originated from a eukaryote (alga) and was acquired by another eukaryote.

Nuclear nucleomorph-derived transfer (nunm): transfer of nucleomorph DNA into the nuclear genome of its eukaryotic host.

Glycoside hydrolases (GHs): enzymes that assist in hydrolysis of glycosidic bonds of complex sugars, for example plant cell wall (cellulose, hemicellulose, and starch) degrading enzymes, fungal cell wall and animal exoskeleton (chitin) degrading enzymes (chitinases), and bacterial cell wall (peptidoglycan) degrading enzymes (lysozymes)

Main text

Introduction. An organism's genome is usually passed vertically through parent-to-offspring relationships. As such, in the simplest case, the evolutionary history of a genome should reflect the evolutionary history of the organism. But genomes are dynamic in content, size, and rates of evolution. Genes can be gained through both duplication within genomes and acquisition from foreign sources (horizontal gene transfer, or HGT), genes can be lost, non-coding or selfish genomic regions can expand or contract over relatively short time scales, and different loci can evolve at different rates because of unequal selective pressures. Processes such as incomplete lineage sorting and introgression can make inferring organismal relationships from gene trees difficult for (relatively) closely related lineages. HGT can in principle occur between any two DNA-based organisms, and so can lead to different genes on the same genome possessing wildly different evolutionary histories and can confound phylogenies for even distantly related organisms. But HGT is not equally likely between all branches of life.

HGT is now understood to be a major driver of genome evolution in bacteria and archaea [1–9]. So common, in fact, that our ability to infer organismal relationships through gene trees has been questioned for prokaryotes [10,11]. The frequency and significance of HGT between prokaryotes and eukaryotes is less clear and remains somewhat controversial [9,12–16]. Clarity is especially difficult to find for animals [17–19], at least partly due to the large size and complexity of their genomes and to the almost inevitable problem of microbial contamination in genome projects. These problems have resulted in a fraught history of misattributions of HGT in animal genomes such as human [20–23] or, more recently, tardigrade [24–27] **(Box 1—methods and problems in calling HGTs)**.

Despite these problems, many cases of *bona fide* functional inter-Kingdom HGT have been documented, including several recent examples of bacteria-to-animal transfers [28–37] **(Figure 1, Table 1)**. Here we review these recent findings, focusing on how HGTs become functional and on what kinds of functions are typically endowed to eukaryotic lineages by genes from bacteria. We have limited

ourselves to transfers from bacteria to eukaryotes because they seem to be the most common—or at least the most commonly reported [38]—but we note that HGT between nearly all branches of life have been described **(Table 1, Supplementary table 1)**.

Which eukaryotes tend to gain genes by HGT, and where do these genes come from? The frequency with which an organism is the recipient or donor of an HGT is dictated by both mechanistic and serendipitous forces. An organism's ability to take up foreign DNA, whether or not its germ line is sequestered, its recombinogenic tendencies, the frequency of the donor DNA in the environment, and the presence or absence of endosymbionts all affect the frequency of gene transfer. Once acquired, foreign DNA has the potential to give the recipient new abilities, but first must be expressed in a way that is visible to selection [39] **(Figure 2)**. If not activated in a selectable way, the potential of the transfer is not realized and the new DNA erodes away in a manner dependent on the host organism. Because of the differences in gene structure between bacteria and eukaryotes, gene inactivation and erosion is likely the most common outcome of HGT between these organisms [40–43]. But transferred DNA does become functional in many instances. Here we use the large number of HGT reports **(Table 1, Supplementary table 1)** to look for commonalities or themes that might allow us to predict when a transfer is more likely to become functional.

Three issues make it difficult to precisely identify the donor organism for many HGTs. The first is that prokaryote genomes can be taxonomically mosaic due to relatively high levels of HGT among them, so the donor DNA may or may not have the same evolutionary history as its bacterial host [44,45]. The second is that the bacterium that transferred the DNA in question may no longer be associated with the recipient when the HGT is discovered. The third is that phylogenetic problems such as long-branch attraction, differential gene loss, and inadequate taxon sampling can often make the provenance of DNA difficult to infer [11,46,47]. Additionally, in multicellular eukaryotes, foreign DNA needs to be exposed to the germ line to become heritable in future generations. If one assumes that the more common the source organism is in the environment of the gene recipient, the more likely transfer is

between organisms, it is not surprising that bacterivorous or parasitic single-cell eukaryotes—where every cell is a germ line—are among the most frequent recipients of gene acquisition from bacteria [14,48,49]. Many other examples also suggest that proximity matters: soil-dwelling nematodes tend to acquire genes from putative soil bacteria [50], thermoacidophilic algae acquire genes from putative thermoacidophilic bacteria [51,52], plant pathogenic oomycetes acquire genes from plant-associated bacteria [53], stramenopile pathogens of human digestive tracts acquire genes from relatives of common gut microbiome inhabitants [38], and many invertebrates acquire genes from common reproductive manipulators such as *Wolbachia* [28,35,37,54]. However, in some HGT-rich organisms such as rotifers [43,55] or fungi [56], there is no clear environmental source of transferred DNA. These organisms seem to collect genes from multiple donors as they explore and adapt to diverse environments.

Various models have been put forward to explain these patterns [13,57,58]. In our view, the model which best summarizes the HGT process in eukaryotes is called the weak-link model [59]. This model expands previous hypotheses [13,57] by suggesting that genes primarily enter cells of their recipient organisms at 'unprotected' stages of their lifecycle in natural environments. For unicellular eukaryotes, every cell is a weak link because any DNA incorporated into the genome is also incorporated into the germ line [59]. For asexual multicellular eukaryotes with no germ line, every clonal cell is similar to a unicellular eukaryote. Consequently, bdelloid rotifers, microscopic freshwater animals that reproduce asexually for millions of years, are the most HGT-rich animals reported to date [43]. For multicellular eukaryotes, early developmental stages fully exposed to the environment of plants or aquatic animals (e.g. spores, zygotes, or embryos) are the weak links [59]. In cases where early developmental stages are not exposed to the environment, the weak link is often the infection of germ line cells by parasites, symbionts or pathogens, which use this cell tropism as a vehicle to ensure vertical transmission by their hosts but also expose the host to HGT which can be incorporated and maintained in the genome [31,59,60].

How does foreign DNA get acquired and become functional in eukaryotes? The short answer is that there is more than one way to do it, and that clear rules are hard to find since so much of the

process seems dependent on chance. Mechanisms of gene transfer include conjugation [61], transduction [62], transformation [63], introgression [64], cell fusion [65], gene transfer agents [66], and intracellular gene transfer from endosymbionts [67]. Non-homologous end joining (NHEJ) seems to be the major mechanism of incorporation of foreign DNA in eukaryotes [68], although recombination with a homologue already present in the genome can also occur, e.g. in plant mitochondria [69]. As these mechanisms have been thoroughly discussed elsewhere [5,13], we focus mainly on the steps that allow genes in foreign DNA to become visible to selection through expression, and thus provide the host with new functions.

Whether a newly arrived foreign sequence attains functionality probably depends on its features (length of the acquired DNA, GC content, codon usage, genetic code, epigenetic marks, etc.) and where it lands in the genome (heterochromatin vs. euchromatin). The size of the transferred DNA seems to be a major factor in whether or not an HGT becomes functional **(Figure 2)**. Several examples from diverse eukaryotes show that while large HGT events may be dramatic, they are more likely to become junk than small fragments [42,67,70]. For example, transfers of huge fragments of *Wolbachia* DNA into arthropod and nematode genomes are fairly common [40,70–73]. These transfers usually originate as large fragments or even entire genomes, but in most cases undergo rapid nonfunctionalization and have very low transcription levels [41,74]. Many functional bacteria-to-animal HGTs are also from *Wolbachia* (discussed below), but in these cases the transfers seem either to be much smaller in size or from larger transfers that were somehow quickly reduced in size. A fascinating exception to this rule comes from the pillbug *Armadillidium vulgare*, where incorporation of an ~1.5 Mb section of *Wolbachia* DNA into the pillbug genome (which subsequently duplicated to a final size of 3 Mb) seems to have become functional by creating a new female sex chromosome [73]. The transcriptional profile of this transfer is currently unknown, but this example serves as a cautionary note to our labeling of large bacteria-to-eukaryote transfers as non-functional simply because of their size.

Eukaryotic cells are often surrounded by 'dead' DNA fragments, and these fragments sometimes get

incorporated into the host chromosomes but their fate is almost always rapid nonfunctionalization or loss. When these fragments come from organelle genomes, the transfers are called *numts* and *nupts* (nuclear mitochondrial and plastid transfers, usually pronounced as “new-mights” and “new-peats”) or sometimes also organellar gene transfers (OGTs) [59]. These transfers were originally found when searching for mitochondrial genomes in eukaryotic genome assemblies, and have been studied intensely due to their possible role in several human diseases [75]. Several studies have shown that numts and nupts are inserted via non-homologous end joining at double-strand breaks [76–78] and microhomology-mediated or blunt-end repair are involved in the DNA incorporation [79,80]. The incorporation is enriched at open chromatin regions [78], and can occur in different genome regions in different lineages. For example, numts are enriched inside introns in the human genome [77], but not in the intron-poor brewer's yeast genome [81]. Flanking regions are rich in retrotransposons and the insertion often occurs immediately adjacent to AT oligomers in mammals [82]. Both ancient and recent examples of all sizes from only several base pairs to several hundred kbp [83] are known, and numpts can be further amplified after acquisition and give rise to tandem repeats [84]. The frequency of non-functional HGT events from *Wolbachia* has motivated a similar name, *nuwts* (nuclear *Wolbachi*a transfers) to highlight the similarity to organelle transfers [60]. These *nuwts* have been found in hosts both with and without extant *Wolbachia* symbionts [40,71,72].

In principle, foreign DNA could be incorporated anywhere on a recipient genome that does not disrupt an existing genomic element that is under selection for function. Several examples suggest that, like numts, HGTs often land in genome regions rich in DNA transposons and retrotransposons [50,85–88] (**Figure 2**). In some genomes, older HGTs are found in gene-rich regions (**Figure 2**), while more recent HGTs occupy less conserved and gene-dense locations such as telomeric regions or within or around transposable elements [31,88]. For example, HGTs acquired by mealybugs tens of million years ago are on gene-rich, likely euchromatic, scaffolds (**Figure 3**), while HGTs that are more recent are on less gene-rich and more poorly assembled scaffolds [54]. Overall, however, a relative scarcity of data and the poor

assembly quality of many eukaryotic genomes make drawing firm conclusions on the importance of HGT landing position difficult.

Gene duplication also seems to sometimes play an important role in helping a transferred gene become functional, possibly because it provides raw material for evolutionary experimentation (**Figure 2**). Tandem duplications of HGTs are frequently observed [14,28,31,36,53,73,87,89]. For example, 38 out of 48 HGTs present in oomycete genomes were found in at least two copies [53]. It thus seems likely that the dynamism of some genomic regions might also be useful as generators of HGT innovation through gene duplication (**Figure 2**).

Finally, intron gain can be an important part of a prokaryotic gene becoming more eukaryotic in nature (**Figure 2, 3**), although this probably depends on the intron density of the recipient genome. That is, HGTs that gain introns in genomes with high densities of introns in native genes might be more likely to become expressed and subject to selection, while intron gain might not be as important in genomes with low intron densities such as some fungi. Introns are commonly found in functional HGTs in many eukaryotes [15,29,31,32,50,56,88,90–92] (**Figure 3**), although these introns are sometimes found in 5' or 3' UTRs of the gene [31]. Intron gain might be an important step in increasing the expression of a transferred gene, as the presence of introns have been shown to increase gene expression in many eukaryotes [93]. After intron gain, the new gene can then evolve as any eukaryotic gene would: its GC content and codon usage can gradually adjust to its host, and it can acquire cell compartment/tissue specific expression, and other eukaryote-specific features of gene expression such as alternative splicing can evolve. In summary, successful HGT seems to often (but not always, of course) involve transfers of relatively short foreign DNA fragments, which more often than not originally land in gene-poor dynamic parts of the genome, undergo duplication, gain introns, and eventually move to more stable and gene-rich parts of the genome over time (**Figure 2**).

What kinds of horizontally transferred genes become functional in eukaryotes? In this review,

we tried to take a broad look at the kinds of genes and functions that are gained in eukaryotes by HGT, and to classify these transfers in meaningful (and hopefully useful) ways. Classification in biology is always problematic—some might argue with our categories, and in a few cases we found it difficult to neatly classify some gene families. We have broken down functional HGTs into two broad categories: those that bring a new function not present in the recipient organism (which we call 'innovative' transfers) and those that replace or maintain a functional loss in the recipient organism (which we call 'maintenance' transfers). Innovative transfers are commonly used by recipients to protect themselves from attacks by other organisms, feed on nutritionally poor or toxic diets, parasitize other eukaryotes, or to survive in cold, hot, acidic, anaerobic or toxic environments **(Table 1)**. Maintenance transfers are most common in recipients that also house a bacterial endosymbiont that is required for normal host function, and can often be explained as a mechanism to maintain the function initially encoded on a degenerating symbiont or organelle genome **(Table 1)**.

Maintenance transfers: enabling the loss of endosymbiont function. All eukaryotes rely on at least one bacterial endosymbiont, now called the mitochondrion, while some others also rely on a photosynthetic bacterial endosymbiont now called the plastid or chloroplast. Modern mitochondrial and plastid genomes are tiny and encode few genes, and organelle function therefore requires extensive participation from the host. Many of these host-encoded genes have been transferred from bacteria, the proteins products of which are transported back into the organelle. As such, maintenance HGT from bacteria has shaped the content of all eukaryote genomes [13,44,94,95].

But the sources of these endosymbiont-related HGTs are difficult to establish. While it is not disputed that the endosymbiont that became the mitochondrion was an alphaproteobacterium (and the plastid a cyanobacterium), controversy arises because the bacterial HGTs that are targeted to these organelles affiliate with numerous other bacterial groups in phylogenetic trees [44,96]. This taxonomic diversity has motivated several scenarios to explain the data [45,97–106], which fall roughly into two camps. The first camp hypothesizes that the endosymbionts that eventually became organelles had highly mosaic

genomes due to bacteria-bacteria HGT [44,45]. The second camp hypothesizes that the cell that became host to the organelle already had previously acquired genes by HGT before the endosymbiont got fixed [107–110]. These (not mutually exclusive) hypotheses will likely always be difficult to differentiate due to the antiquity of organellogenesis. Might more recent endosymbioses provide some insight into the possible general mechanisms used to build mosaic metabolic pathways? The advantage of more recent endosymbioses is that inferring the origin of transferred genes, and how these HGTs interact with the host and its endosymbionts, is much more straightforward. The disadvantage—of course—is that the cell biological and genetic context of these symbioses are different than those encountered at organellogenesis.

Paulinella chromatophora is an amoeboid protist that acquired an organelle-like cyanobacterial endosymbiont (relatively) recently, roughly 100 million years ago [111,112]. In this case, only about 25% of its 229 nuclear genes of bacterial origin seem to result from transfer from the cyanobacterial endosymbiont [48]. The remaining 75% taxonomically affiliate with other bacterial groups in phylogenetic tress. Similarly, the incomplete nutritional pathways in some trypanosomatid bacterial endosymbionts are complemented by horizontal gene transfer from diverse bacteria to the protist genome [113,114]. Finally, analyses of insect endosymbioses in aphids, psyllids, whiteflies and mealybugs also show similar patterns of mosaic pathways built from multiple HGT events [28,35,37,54]. Many of these HGTs seem to compensate for genome reduction in their highly degraded nutritional endosymbionts, but—as in the protist examples above—most of the the HGT is from bacteria that are not the existing endosymbionts. Importantly, these insect transfers come from reproductive manipulators such as *Wolbachia* and *Cardinium,* bacteria that are extremely common in insect germ cells. This makes the provenance of these insect transfers clear: they have not come from the endosymbiont that is degenerating, they have come from other endosymbiotic bacteria with germ line cell tropisms. Taken together, these results show that gene transfer from degenerate (or degenerating) extant endosymbionts to the host is not necessarily needed, or perhaps even common, and in some hosts HGT from other sources can compensate for gene loss by the endosymbiont [31]. We argue that these data

support the idea that the taxonomic diversity of HGT in eukaryotic genomes could have resulted from previously existing HGT from non-organelle sources [48,57,115,116,96,117], perhaps in combination with the 'inherited chimerism' of organelle progenitor resulting from bacteria-bacteria HGT that predated endosymbiosis [44,45].

Innovative transfers: eukaryotes fight bacteria using bacterial weapons. Genes of bacterial origin that target the bacterial cell envelope are commonly found as functional HGTs in eukaryote genomes. [The phenomenon of eukaryotes using bacterial genes to defend themselves against other bacteria has been called *The Eukaryotes Strike Back* [118]]. Bacterial cell envelopes can be disrupted by several mechanisms, but peptidoglycan is a common target. Peptidoglycan is a structural matrix in bacterial cell walls formed by alternating N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) residues, where peptide chains of up to five amino acids link NAM to other NAM-connected peptides (**Figure 4**).

A dramatic example of eukaryotes acquiring bacterial genes to protect themselves from other bacteria involves a family of amide bond-breaking genes called *tae* (**t**ype IV secretion system **a**midase **e**ffectors) [33]. These Tae proteins all cleave the amide linkages of peptidoglycan, and were domesticated in at least three eukaryotic supergroups **(Table 1, Figure 4)**. Some of these transfers date to more than 800 million years ago. Importantly, the functional role of these HGTs was experimentally verified, and it was demonstrated that they limit *Borrelia burgdorferi* proliferation in the deer tick *Ixodes scapularis* [33]. Other examples show that bacterial genes encoding lysozymes, which cleave peptidoglycan between NAG and NAM moieties, have also been acquired by eukaryotes numerous times independently [28,34] **(Figure 4)**. In some cases, the antibacterial effect of the transferred lysozyme genes was also verified experimentally *in vitro* [34]. Perhaps the most astounding toolkit for bacterial cell envelope disruption acquired through HGT was reported from the marine protist *Trepanomas* sp. PC1 (Diplomonadida) [49]. This species clusters deeply within a clade of parasites, but seems to have secondarily transitioned to a free-living and bacterivorous lifestyle. *Trepanomas*'s lifestyle transition is

hypothesized to have been enabled by the acquisition of 40 bactericidal permeability-increasing proteins (which bind to lipopolysaccharides), 20 cell wall hydrolases, four *N*-acetylmuramoyl-*L*-alanine amidases, and five lysozymes.

Innovative or maintenance transfers? The confusing case of peptidoglycan. Bacteria-to-eukaryote HGTs include not only genes for peptidoglycan degradation, but also genes for peptidoglycan synthesis. Although the functional role of these transfers remains unclear, putative peptidoglycan-producing HGTs have been found to be expressed in bdelloid rotifers [88], aphids [28], and mealybugs [54]. Interestingly, peptidoglycan construction in some archaeplastidal chloroplasts seem to also require numerous genes of HGT origin [119,120]. These genes have been shown to be essential for chloroplast division in both moss [121] and *Arabidopsis* [122]. While a peptidoglycan layer has been found surrounding chloroplasts of moss [120], no detectable peptidoglycan layer seems to exist in *Arabidopsis* [120,122].

A similarly murky situation exists for animals with obligate endosymbionts that have acquired amidase HGTs [28,31,54]. Amidases are related to peptidoglycan recognition proteins (PGRPs), which are normally used by the animal innate immune system as bacterial sensors [123]. Some PGRPs possess amidase activity which can cleave peptidoglycan at positions distinct from tae proteins. Interestingly, PGRPs have been shown to be involved in endosymbiont maintenance rather than pathogen defense [124–127]. For example, native amidase-active PGRPs are used by tsetse flies to shield their nutritional endosymbionts from the host immune system by recycling peptidoglycan from lysed endosymbiont cells [128]. Some sap-feeding insects with obligate bacterial endosymbionts have acquired bacterial amidaes by HGT [28,31,54] **(Figure 4)**, raising the prospect that some of these horizontally acquired amidases are used not as anti-bacterial control measures by the host, but rather for endosymbiont maintenance. In a potentially parallel process, bacterial beta-lactamases have been found to have been transferred into the genomes of mealybugs and the slime mold *Dictyostelium discoideum* [54,129]. Beta-lactamases are enzymes that provide resistance to beta-lactam antibiotics such as penicillin. Why

would eukaryotes need genes targeting bacterial antibiotics? It seems plausible that, similar to PGRPs, these enzymes could be used for symbiont protection rather than defense from pathogens, since both slime molds and mealybugs are tightly associated with beneficial bacteria [54,130].

Innovative transfers: eukaryotes also fight other eukaryotes using bacterial weapons. Genomes from many bacterial pathogens encode proteins called pore-forming toxins (PFTs) to make holes in the membranes of target cells [reviewed in [131]]. One class of these toxins are called aerolysins, first discovered from the bacterium *Aeromonas* [132]. Aerolysin genes have been transferred in several independent acquisitions in numerous eukaryotes [29] **(Figure 4)**. These transferred PFTs have been hypothesized to be involved in disintegration of prey in *Hydra* and sea anemones [29,133], blood cell lysis by ticks and other bloodsucking arthropods [29,134], and water stress regulation in plants [29,135].

Herbivorous eukaryotes must cope with complex mixtures of toxic compounds produced by plants. Both plants and their herbivores have taken advantage of HGT in a process that resembles an evolutionary arms race (reviewed for arthropods in [19]). For example, HGTs for detoxifying cyanide were found in phytophagous mites and various lepidopterans [36]; the beetle *Hypothenemus hampei* can feed exclusively on coffee beans due to a mannanase HGT [86]; and the silkworm *Bombyx mori* [136] has used HGT for overcoming alkaloids in the latex of mulberry plants. A very similar history of HGT likely shaped some plant-parasitic nematode [137] and oomycete genomes [53]. On the plant side of the arms race, defensive HGTs have been hypothesized to be used by some mosses and land plants [119].

Epichloe fungi are intercellular symbionts of grasses, where they protect their hosts from insect herbivores using a cocktail of fungal alkaloids. This anti-insect arsenal of epichloe fungi seems to be supplemented by the *mcf* gene (makes caterpillars floppy), which was acquired from a bacterial HGT to the fungal genome [138]. The authors show that *mcf* encodes a toxin that the endosymbiotic fungi produce to help kill insect larvae that feed on their host grass plants. Interestingly, the *mcf* gene is also found in the genome of the bacterial symbionts of entomopathogenic nematodes, where it has been

shown to be sufficient to kill insects [139]. Thus, HGT from one tripartite interaction (nematodes, bacteria, insects) has possibly altered interactions in another tripartite interaction (grasses, fungi, insects) [138].

Innovative transfers: eukaryotes exploit novel environments using bacterial genes. Plant material represents an enormous amount of Earth's biomass, and so it is no surprise that numerous eukaryotes have evolved adaptations to feed on both living and dead plants. But herbivory comes with the price of not only dealing with plant defenses, but also with the degradation of complex plant carbohydrates. Genes for degrading these complex carbohydrates have been acquired in numerous eukaryotic lineages. HGT of genes involved in carbohydrate metabolism have been found in herbivorous insects [19], rumen ciliates [140], oomycetes and fungi [141], plant-parasitic nematodes [50,92,142], necromenic nematodes from the *Pristionchus* genus [143–145], rotifers [55], and choanoflagellates [146]. Conversely, tunicates can synthesize cellulose for their eponymous protective exoskeleton (the tunic) because of an ancient cellulose synthase HGT [147,148].

HGT of biosynthetic enzymes has allowed many lineages to live in extremely nutrient-poor environments. Amino acid, vitamin, and carbohydrate metabolism genes are perhaps most often involved in HGT (and in cases where the host has an endosymbiont, are probably maintenance transfers [31,35,37,89]). Most of the HGTs found in 13 genomes of unicellular protists involve amino acid and carbohydrate metabolism [14]. For example, in *Cryptosporidium* spp. genomes there are HGTs such as tryptophan synthase, aspartate-ammonia ligase, and glutamine synthetase [149] and trypanosomatids have acquired numerous HGTs involved in arginine, tryptophan, threonine, methionine, cysteine, lysine and vitamin B5 metabolism from diverse donors [113,114]. Marine environments can be notoriously poor in nitrogen and iron, and HGTs have played a role in allowing several distinct eukaryotic lineages to scavenge these growth-limiting compounds from their environments. Iron-binding proteins of bacterial origin were found in diatoms [150,151], an extremophilic red alga *Galdieria sulphuraria* [51], and a soil-dwelling amoeba *Dictyostelium* [129,152]. Nitrogen metabolism was influenced by HGTs at least in

diatoms [91] and green algae from the *Micromonas* genus [153]. In other environments, the limiting nutrient is not that obvious, for example N-acetylneuraminate lyase HGT likely allows *Trichomonas vaginalis* to scavenge sialic acids from its host for nutrition [154]. Transfers of amino acid and vitamin genes from bacteria to animals include those involved in the diaminopimelic acid pathway in the choanoflagellate *Monosiga* [155], haeme biosynthesis pathway (a ferrochelatase HGT) in the nematode *Brugia malayi* [32], and B1, B5, and B6 vitamin pathways in the plant nematode *Heterodera glycines* [90,156].

Finally, HGTs have allowed eukaryotes to live in extremely hot, cold, and otherwise toxic environments. HGTs of bacterial genes encoding ice-binding proteins (IBPs), which can allow its recipients to survive in extremely cold environments, were found in the diatom *Fragilariopsis cylindrus* [157], the haptophyte *Phaeocystis antarctica* [158], and the green algae *Pyramimonas gelidicola* [158] and *Coccomyxa subellipsoidea* [159]. HGTs allowing recipients to survive hot, metal-rich, and acidic environments were detected in the genome of red alga *Galdieria sulphuraria* [51]. HGTs that protect hosts from other sources of environmental stress, such as desiccation and oxidative and osmotic stress, seem to be frequent in rotifer species from desiccating habitats [43]. Likewise, two HGTs seem to protect the microsporidium *Antonospora locustae* from oxidative and UV-induced damage [160,161].

Concluding remarks

Both early and recent missteps in HGT research has made it clear that care is required when hypothesizing that a gene has moved into a genome from an unrelated organism. But numerous papers reporting HGT from bacteria to eukaryotes stand on solid conceptual and methodological ground. While the field of HGT research is vibrant, there is clearly room for improvements both in the methods used to screen for contamination in genome projects and in the methods used to verify putative HGT events. As we show in this review, adaptive HGTs are used both to maintain functionality—often in situations involving an endosymbiont that is subject to genomic erosion—but also to acquire new functions such as nutrient acquisition, protection, and environmental buffering. To better understand the true phylogenetic and functional scope of bacterial genes in eukaryotes, a focus

on phylogenetically important but neglected groups such as protists is warranted since the majority of eukaryotic diversity resides in these understudied organisms.

Display items

Box 1: Methodologies and caveats in HGT identification.

Sample preparation (caveat: contamination): HGT detection tends to start with a genome project. As such, the first step in trying to detect HGTs in any organism should be careful sample and sequencing library preparation. For example, tissues known to house dense communities of bacteria can sometimes be removed, and the outer surface of sequenced individuals can be bleached to reduce contaminating environmental bacteria. Guidelines for reducing reagent and laboratory contamination are well-defined for microbiome studies [162], and we recommend using these approaches when generating data that might include HGT detection. Several software programs are available for high-throughput HGT detection [163–168], but none of these programs deal well with contamination.

Sequence database searches (caveat: database sampling): Numerous BLAST-based approaches have been developed to find candidates of horizontal gene transfer in eukaryotic genomes [37,55,88,169]. Perhaps the most rigorous implementations of this approach use two or more alternative search databases and calculate differences between either E-values (a so-called "Alien index") or bit score (HGT index) of the highest HGT and the highest non-HGT blast hit [55,88,169]. The HGT index might be preferred because bit scores are not dependent on database sizes the way E-values are, but arbitrary cutoffs are always needed to separate HGT candidates from non-HGT genes. However, the performance of any sequence alignment-based method is only as good as the databases used in the searches. It is relatively reliable for taxa with numerous sequenced and well-annotated genomes (e.g. some Bacteria and Metazoa), but HGT candidates found with these approaches should be interpreted with extreme skepticism for non-model organism, for detection of ancient and highly diverged HGTs, or if any related reference genomes are of low quality (or possibly contaminated with bacterial

sequences) themselves.

Phylogenetic evidence (caveats: poor taxon sampling and phylogenetic artifacts): Sequence database searches should primarily be used to narrow-down the total number of HGT candidates to a computationally feasible number of candidates for phylogenetic analysis. Phylogenetic conflict, i.e. incongruence of a single-gene tree with a known 'species' phylogeny is the method of choice for HGT detection. However, a phylogenetic tree is still just a hypothesis. HGTs likely evolve under different selection pressures than native genes, especially at first, and many have long branches. Methods that reduce long-branch attraction [170] can be used on single-gene datasets of HGT candidates, and the HGT-topology can be tested with statistical approaches (e.g. the approximately unbiased test [171] can be used to compare the HGT-topology to a constrained non-HGT topology). Finally, differential gene loss in different lineages can lead to situations that look like HGT but are not [21,45], especially in situations where taxon sampling is limited.

Genomic evidence (caveats: endosymbionts and contamination): The co-assembly of an HGT candidate with one or more eukaryotic genes on a single genomic scaffold is probably the best genomic evidence of gene transfer. This does not mean that a promising HGT candidate is simply present in the same assembly, which can easily be contamination, but rather shares the same contig or scaffold as a high-confidence native gene. Typically this means that the average depth of sequencing coverage of the scaffold with the HGT candidate should be similar to other eukaryotic scaffold. (In some cases, the HGT might be present in several copies that are collapsed into a single region, resulting in higher than average coverage, or might be localized on an under- or over-represented/replicated genome region such as a B- or sex-chromosome that has a completely different copy number from autosomal regions.) The use of PCR-free library methods can help reduce uneveness in sequencing depth of coverage. Using very strict sequencing depth filtering can help to remove contamination, but it can also remove some false negatives. Software such as Blobtools [172] [https://blobtools.readme.io/] which bins and displays assemblies by characteristics such as depth of

coverage, GC content, codon usage, and k-mer frequencies can also be very helpful in discriminating contamination from HGT candidates [25,54]. However, very recent HGTs will sometimes be difficult to distinguish from contamination or symbionts [40,72].

Population, microscopy, and functional evidence (caveats: sample availability, cost): Presence of an HGT candidate in individuals from geographically distinct populations (or even species or higher taxonomic units) can help to distinguish it from most contaminants, especially if the gene is present on the same genome region and its phylogeny reflects the species phylogeny [54]. Localizing the HGT candidate on the host chromosome using FISH with specific fluorescently labeled probes is probably the best piece of evidence available for any HGT candidate. This approach was shown to be essential to corroborate whole *Wolbachia* genome inserts into arthropod genomes because these HGTs lack many of other HGT signatures [70–72]. Preparing good chromosome spreads and doing single-gene hybridizations can be, however, impossible or very laborious with some organisms. On the other hand, generating RNA-Seq data is becoming a must for any genome project and these data are very useful for detection of expressed and likely functional HGTs. Genes of such HGTs tend to have eukaryotic gene structure with canonical introns, untranslated 5' and 3' regions, a polyA tail, and a eukaryotic signal peptide [31,88]. Tissue-specific expression combined with experimental validation of the enzymatic function previously unknown from the target taxon was used as supporting evidence for several HGTs, e.g. in arthropods (reviewed in [19]).

Figures and tables

Fig 1: Functional HGTs mapped on the eukaryotic phylogeny. (A) A phylogenetic tree with with all cases of primary and complex plastids highlighted by differently colored ovals. Selected functional bacteria-eukaryote HGTs are displayed in various colors and highlight parallel nature of HGT across the tree of life. (B) Horizontal gene transfers in insects are shown in greater detail, revealing diverse types of transfers from numerous bacterial groups in many insect orders.

Fig 2: How does foreign DNA become functional in eukaryotes? Here we schematize the steps any foreign bacterial sequence may take when becoming functional in a eukaryotic genome. (A) The genomic players are shown as large and small fragments of bacterial genome (bacterial genes shown in blue) and a section of the recipient eukaryotic genome with a gene-rich region on the left and a gene-poor region on the right (eukaryotic genes are shown in purple; transposable elements are shown in green). (B) Acquisition of a large fragment of bacterial genome by a eukaryote often results in bacterial gene nonfunctionalization (shown in grey) and the formation of junk DNA. (C) Acquisition of a small fragment of bacterial genome is more likely to result in a functional transfer. We have shown the process as one that occurs step-wise, but in reality, functional transfers can enter and exit at any part of this pathway, and can move through the steps in any order.

Fig 3: Two examples of functional HGT from the mealybug *Planococcus citri*. Genome context, gene structure, and patterns of expression are shown for two HGTs, *dapF* and *ribD*. Native eukaryotic genes are shown in red, and the bacterial HGTs are shown in yellow.

Fig 4: Parallel HGTs targeting peptidoglycan bonds. Parallel adaptive HGTs used by various eukaryotes for defense from bacterial pathogens by degradation of their cell envelopes.

Table 1. Functional and putatively functional horizontal gene transfers reported from eukaryotes. We note that some of the categories can overlap with others, and our categorization here reflects a simplification of reality. Here we focus only on HGTs with an assigned functional categorization, so the list is not exhaustive. #The *Nematostella* genome was suggested several times to be highly contaminated by bacterial data, so here we include only this one well-supported HGT event (type VI secretion amidase effector Tae). *ecotins, bacterial inhibitors of animal serine peptidases. Column one designates eukaryotic supergroup, abbreviations: Exc=Excavata, Opi=Opisthokonta, Arch=Archaeplastida, Amo=Amoebozoa, Hac=Hacrobia, SAR=Stramenopila-Alveolata-Rhizaria.

(A) Maintenance transfers

Compensating for gene loss in obligate endosymbionts		
S	**Eukaryotic lineage**	**Ref**
Opi	mealybugs, aphids, psyllids, whiteflies	28,35,37,48,54,114
SAR	*Paulinella chromatophora*	
Exc	*Angomonas* and *Strigomonas* trypanosomatids	

(B) Innovative transfers

Protection from bacteria (lysis of bacterial cells)		
S	**Eukaryotic lineage**	**Ref**
Opi	mites and ticks, *Daphnia, Capitella* and mollusks, lancelet, acorn worm, sea anemone, *Monosiga,* hemipteran insects, *Trepanomas*, bdelloid rotifers, fungi	28,33,34,49,54,88

Arch	*Selaginella*	
SAR	*Oxytricha*	
Exc	*Naegleria*	
Protection from eukaryotes (e.g. pore forming toxins) and their metabolites		
Opi	*Hydra*, ticks, fungi, #*Nematostella, Epichloe* fungi, herbivorous arthropods	19,29,138
Arch	various plants	
Exc	*Leishmannia+*Trypanosoma spp.*	

546

Nutrition (amino acid and nitrogen metabolism)		
S	**Eukaryotic lineage**	**Ref**
Opi	herbivorous arthropods, *Monosiga*, fungi	19,38,91,113,114,149,153,155,174
Amo	*Dictyostelium, Entamoeba*	
Arch	*Micromonas*	
SAR	diatoms, *Apicomplexa* species, Blastocystis	
Exc	Kinetoplastida species, *Trichomonas*, *Giardia*	
Nutrition (co-factors, vitamins, and iron limitation)		
Opi	hemipteran insects, *Brugia malayi*, *Heterodera glycines*, fungi	19,32,51,113,129,150–152,156,174

Amo	*Dictyostelium, Entamoeba*	
Arch	*Galdieria sulphuraria*	
SAR	diatoms	
Exc	Kinetoplastida spp., *Trichomonas*	
Nutrition (plant carbohydrates)		
Opi	herbivorous arthropods, rumen chytrid fungi, plant parasitic nematodes, *Pristionchus* necronemic nematodes, choanoflagellates (likely algal origin), rotifers	19,50,55,92,140–146,173
SAR	rumen ciliates, oomycetes	
Nutrition (carbohydrates)		
Opi	fungi	14,38,174
Amo	*Entamoeba, Dictyostelium*	
SAR	*Cryptosporidium, Blastocystis*	
Exc	*Trichomonas vaginalis,* Kinetoplastida species	

547

Extreme and toxic conditions (cold, hot, acidic, arsenic, telluric, …) and environmental stress (oxidative, osmotic, UV-induced, …)		
S	**Eukaryotic lineage**	**Ref**
Opi	rotifers, *Antonospora locustae* microsporidium,	43,56,129,158–

	fungi	161
Amo	*Dictyostelium*	
Arch	*Coccomyxa subellipsoidea, Galdieria sulphuraria*, various plants	
SAR	diatoms	
Hac	*Phaeocystis antarctica*	
Facultative or obligate anaerobiosis		
Opi	chytrid fungi and yeasts, *Amoebidium parasiticum*	38,140,149,177–188
Amo	*Entamoeba* spp., *Mastigamoeba balamuthi*	
Arch	*Cyanophora paradoxa, Porphyra haitanensis,* Prasinophyte algae, Chlorophyte algae	
SAR	*Thalassiosira pseudonana*, *Cryptosporidium parvum***,** rumen ciliates, *Nyctotherus ovalis, Blastocystis*	
Hac	*Prymnesium parvum*	
Exc	*Giardia lamblia, Trichomonas vaginalis, Trimastix pyriformis*	
Nucleotide metabolism		
Opi	microsporidia	14
Amo	*Entamoeba*	
SAR	*Cryptosporidium, Plasmodium*	
Exc	*Trichomonas,* Kinetoplastida species*, Giardia,*	

	Spironucleus	
RNA and DNA modifications		
Opi	mealybugs, psyllids	14,31,35
Amo	*Dictyostelium*	
SAR	*Plasmodium, Toxoplasma*	
Exc	*Trichomonas,* Kinetoplastida spp.	
Parasitism and pathogenicity		
Opi	various microsporidia	14,38,175,176
Amo	*Entamoeba*	
SAR	*Cryptosporidium, Plasmodium, Toxoplasma* spp., *Blastocystis*	
Exc	*Trichomonas, Giardia, Leishmannia, Trypanosoma*	

Supplementary information

Table S1: HGTs from bacteria to eukaryotes. A comprehensive table with additional references to articles reporting both functional and non-functional HGTs from bacteria to eukaryotes is freely available at the following URL, open for comments from anyone, and for edits upon request: [https://docs.google.com/spreadsheets/d/1F2UKsfTfMawU4N_yv_oep9desviiq7fl7iKk5Qg8rg/edit?usp=sharing].

Competing interests statement

The authors declare that they have no competing interests.

Author biographies

Filip Husnik is an EMBO postdoctoral fellow in the laboratory of Patrick Keeling at the University of British Columbia. He is fascinated by diverse symbiotic interactions between bacteria, protists, and animals and studies how horizontal gene transfer influences ecology and evolution of these organisms.

John McCutcheon is an Associate Professor in the Division of Biological Sciences at the University of Montana in Missoula. He is interested in how symbioses involving microorganisms arise, how they are maintained, and how they sometimes break down.

Acknowledgements

FH was supported by the Fulbright Commission and a fellowship from EMBO (ALTF 1260-2016) while writing this review. JPM was supported by grants from the National Science Foundation (IOS-1256680 and IOS-1553529), the National Aeronautics and Space Administration Astrobiology Institute (NNA15BB04A), and the Gordon and Betty Moore Foundation (5602).

References

1. Koonin, E. V, Makarova, K. S. & Aravind, L. Horizontal gene transfer in prokaryotes: quantification and classification. *Annu. Rev. Microbiol.* **55,** 709–742 (2001).

2. Gogarten, J. P., Doolittle, W. F. & Lawrence, J. G. Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* **19,** 2226–2238 (2002).

3. Gogarten, J. P. & Townsend, J. P. Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **3,** 679–687 (2005).

4. Thomas, C. M. & Nielsen, K. M. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **3,** 711–21 (2005).

5. Soucy, S. M., Huang, J. & Gogarten, J. P. Horizontal gene transfer: building the web of life. *Nat. Rev. Genet.* **16,** 472–482 (2015).

6. Koonin, E. V. Horizontal gene transfer: essentiality and evolvability in prokaryotes, and roles in evolutionary transitions. *F1000Research* **5,** (2016).

7. Wagner, A. *et al.* Mechanisms of gene flow in archaea. *Nat. Rev. Microbiol.* (2017). doi:10.1038/nrmicro.2017.41

8. Williams, T. A. *et al.* Integrative modeling of gene and genome evolution roots the archaeal tree of life. *Proc. Natl. Acad. Sci.* 201618463 (2017). doi:10.1073/PNAS.1618463114

9. Syvanen, M. Evolutionary implications of horizontal gene transfer. *Annu. Rev. Genet.* **46,** 341–58 (2012).

10. Doolittle, W. F. *et al.* What is the tree of life? *PLOS Genet.* **12,** e1005912 (2016).

11. Philippe, H. & Douady, C. J. Horizontal gene transfer and phylogenetics. *Curr. Opin. Microbiol.* **6,** 498–505 (2003).

12. Timmis, J. N., Ayliffe, M. A., Huang, C. Y. & Martin, W. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nat. Rev. Genet.* **5,** 123–135 (2004).

13. Keeling, P. J. & Palmer, J. D. Horizontal gene transfer in eukaryotic evolution. *Nat. Rev. Genet.* **9,**

605–618 (2008).

14. Alsmark, C. *et al.* Patterns of prokaryotic lateral gene transfers affecting parasitic microbial eukaryotes. *Genome Biol.* **14,** R19 (2013). ***Systematical reanalysis of 13 parasitic microbial eukaryotes showing that majority of HGTs in these species are involved in amino acid and sugar metabolism and that putative donors of these genes are significantly enriched in bacterial groups sharing the same habitats, for example human microbiota.***

15. Schönknecht, G., Weber, A. P. M. & Lercher, M. J. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. *BioEssays* **36,** 9–20 (2014).

16. Lacroix, B. & Citovsky, V. Transfer of DNA from bacteria to eukaryotes. *MBio* **7,** e00863-16 (2016).

17. Robinson, K. M., Sieber, K. B. & Dunning Hotopp, J. C. A review of bacteria-animal lateral gene transfer may inform our understanding of diseases like cancer. *PLoS Genet.* **9,** e1003877 (2013).

18. Boto, L. & B, P. R. S. Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc. Biol. Sci.* **281,** 20132450 (2014).

19. Wybouw, N., Pauchet, Y., Heckel, D. G. & Van Leeuwen, T. Horizontal gene transfer contributes to the evolution of arthropod herbivory. *Genome Biol. Evol.* **8,** 1785–1801 (2016).

20. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409,** 860–921 (2001).

21. Salzberg, S. L. *et al.* Microbial genes in the human genome: lateral transfer or gene loss? *Science* **292,** 1903–6 (2001).

22. Stanhope, M. J. *et al.* Phylogenetic analyses do not support horizontal gene transfers from bacteria to vertebrates. *Nature* **411,** 940–944 (2001).

23. Salzberg, S. L. Horizontal gene transfer is not a hallmark of the human genome. *Genome Biol.* **18,** 85 (2017).

24. Boothby, T. C. *et al.* Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. **112,** 15976–15981 (2015).

25. Koutsovoulos, G. *et al.* No evidence for extensive horizontal gene transfer in the genome of the tardigrade *Hypsibius dujardini*. *Proc. Natl. Acad. Sci.* **113,** 5053–5058 (2016).

26. Delmont, T. O. & Eren, A. M. Identifying contamination with advanced visualization and analysis practices: metagenomic approaches for eukaryotic genome assemblies. *PeerJ* **4,** e1839 (2016).

27. Arakawa, K. No evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proc. Natl. Acad. Sci.* **113,** E3057–E3057 (2016).

28. Nikoh, N. *et al.* Bacterial genes in the aphid genome: absence of functional gene transfer from Buchnera to its host. *Plos Genet.* **6,** e1000827 (2010).

29. Moran, Y., Fredman, D., Szczesny, P., Grynberg, M. & Technau, U. Recurrent horizontal transfer of bacterial toxin genes to eukaryotes. *Mol. Biol. Evol.* **29,** 2223–30 (2012). ***Bacteria use pore-forming toxins to lyse cellular membranes, but this article revealed several independent transfers of these toxins to eukaryotic genomes and discusses their possible role in the ecology (defense and predation) of multicellular eukaryotes.***

30. Flot, J. *et al.* Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* **500,** 453–457 (2014).

31. Husnik, F. *et al.* Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell* **153,** 1567–1578 (2013).

32. Wu, B. *et al.* Interdomain lateral gene transfer of an essential ferrochelatase gene in human parasitic nematodes. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 7748–53 (2013).

33. Chou, S. *et al.* Transferred interbacterial antagonism genes augment eukaryotic innate immune function. *Nature* **518,** 98–101 (2014). ***An important study that shows multiple independent horizontal gene transfers of antibacterial amidases across the tree of life and experimentally verifies their function.***

34. Metcalf, J. A., Funkhouser-Jones, L. J., Brileya, K., Reysenbach, A.-L. & Bordenstein, S. R. Antibacterial gene transfer across the tree of life. *Elife* **3,** e04266 (2014). ***Similar to Chou et al. (2014), this article also shows and experimentally verifies horizontal gene transfers for bacterial cell wall degradation enzymes occurring in parallel across the tree of life, but this time the transferred genes are lysozymes.***

35. Sloan, D. B. *et al.* Parallel histories of horizontal gene transfer facilitated extreme reduction of endosymbiont genomes in sap-feeding insects. *Mol. Biol. Evol.* **31,** 857–871 (2014).

36. Wybouw, N. *et al.* A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. *Elife* **3,** e02365 (2014). ***An elegant analysis of a cysteine synthase family HGTs found in several lineages of herbivorous arthropods. Functional expression analysis suggested that the enzyme can be used for both cysteine biosynthesis and cyanide detoxification, but enzyme kinetics favored that its main function is to detoxify cyanide produced by plants.***

37. Luan, J.-B. *et al.* Metabolic coevolution in the bacterial symbiosis of whiteflies and related plant sap-feeding insects. *Genome Biol. Evol.* **7,** 2635–47 (2015).

38. Eme, L., Gentekaki, E., Curtis, B., Archibald, J. M. & Roger, A. J. Lateral gene transfer in the adaptation of the anaerobic parasite *Blastocystis* to the gut. *Curr. Biol.* **27,** 807–820 (2017). ***This rigorous phylogeny-based analysis of HGT in Blastocystis reports 2.5% of its genes have been relatively recently acquired from both prokaryotes and eukaryotes. In particular, genes involved in carbohydrate scavenging and metabolism, anaerobic amino acid and nitrogen metabolism, oxygen-stress resistance, and pH homeostasis were detected. These HGTs are hypothesized to assist Blastocystis in adaptation to the gut environment it inhabits.***

39. Baltrus, D. A. Exploring the costs of horizontal gene transfer. *Trends Ecol. Evol.* **28,** 489–95 (2013).

40. Koutsovoulos, G. *et al.* Palaeosymbiosis revealed by genomic fossils of *Wolbachia* in a

strongyloidean nematode. *PLoS Genet.* **10,** e1004397 (2014).

41. Choi, J. Y., Bubnell, J. E. & Aquadro, C. F. Population genomics of infectious and integrated *Wolbachia pipientis* genomes in *Drosophila ananassae*. *Genome Biol. Evol.* **7,** 2362–2382 (2015).

42. Derelle, E. *et al.* Genome analysis of the smallest free-living eukaryote *Ostreococcus tauri* unveils many unique features. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 11647–52 (2006).

43. Eyres, I. *et al.* Horizontal gene transfer in bdelloid rotifers is ancient, ongoing and more frequent in species from desiccating habitats. *BMC Biol.* **13,** 90 (2015).

44. Ku, C. *et al.* Endosymbiotic gene transfer from prokaryotic pangenomes: Inherited chimerism in eukaryotes. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 10139–10146 (2015).

45. Ku, C. *et al.* Endosymbiotic origin and differential loss of eukaryotic genes. *Nature* **524,** 427–432 (2015).

46. Katz, L. A. Recent events dominate interdomain lateral gene transfers between prokaryotes and eukaryotes and, with the exception of endosymbiotic gene transfers, few ancient transfer events persist. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **370,** 20140324 (2015).

47. Brown, J. R. Ancient horizontal gene transfer. *Nat. Rev. Genet.* **4,** 121–132 (2003).

48. Nowack, E. C. M. *et al.* Gene transfers from diverse bacteria compensate for reductive genome evolution in the chromatophore of *Paulinella chromatophora*. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 12214–12219 (2016)**.** ***Paulinella chromatophora is an amoeba that contains photosynthetic organelles that have originated relatively recently and independently of other plastids. Genomic and transcriptomic data generated from mostly bacteria-free culture allowed the study to untangle how many bacterial genes in the Paulinella genome come from other phylogenetic sources than the current symbiont/organelle and suggest that HGT from diverse bacteria compensates for gene loss of the symbiont/organelle.***

49. Xu, F. *et al.* On the reversibility of parasitism: adaptation to a free-living lifestyle via gene acquisitions in the diplomonad *Trepomonas* sp. PC1. *BMC Biol.* **14,** 62 (2016).

50. Paganini, J. *et al.* Contribution of lateral gene transfers to the genome composition and parasitic ability of root-knot nematodes. *PLoS One* **7,** e50875 (2012).

51. Schönknecht, G. *et al.* Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* **339,** 1207–10 (2013). ***There are various ways how an organism can adapt to extreme environments and this study introduces the significant role HGT from bacteria had in adaptation to hot, toxic, and acidic environments for Galdieria red algae.***

52. Qiu, H. *et al.* Adaptation through horizontal gene transfer in the cryptoendolithic red alga *Galdieria phlegrea*. *Curr. Biol.* **23,** R865–R866 (2013).

53. Savory, F., Leonard, G. & Richards, T. A. The role of horizontal gene transfer in the evolution of the oomycetes. *PLoS Pathog.* **11,** e1004805 (2015).

54. Husnik, F. & McCutcheon, J. P. Repeated replacement of an intrabacterial symbiont in the tripartite nested mealybug symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E5416–E5424 (2016). ***One common argument against HGT from diverse bacteria compensating for gene loss of symbiont/organelle present in the eukaryotic host is that complex pathways/structures are impossible to build by a gene-by-gene fashion unless individual components are continuously kept under selection. In this study, we showed how complex and mosaic pathways can be built sequentially with HGTs from diverse bacteria and that preexisting HGTs can remain stable on genomes in the face of extensive symbiont turnover.***

55. Boschetti, C. *et al.* Biochemical diversification through foreign gene expression in bdelloid rotifers. *PLoS Genet.* **8,** e1003035 (2012).

56. Marcet-Houben, M. & Gabaldón, T. Acquisition of prokaryotic genes by fungal genomes. *Trends Genet.* **26,** 5–8 (2010).

57. Ford Doolittle, W. You are what you eat: A gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* **14,** 307–311 (1998).

58. Gluck-Thaler, E. & Slot, J. C. Dimensions of horizontal gene transfer in eukaryotic microbial pathogens. *PLoS Pathog.* **11,** e1005156 (2015).

59. Huang, J. Horizontal gene transfer in eukaryotes: the weak-link model. *Bioessays* **35,** 868–75 (2013**). *The weak-link model posits that foreign genes enter eukaryotic genomes at unprotected stages of their development and the model thus expands previous models for multicellular sexual eukaryotes.***

60. Dunning Hotopp, J. C. Horizontal gene transfer between bacteria and animals. *Trends Genet.* **27,** 157–163 (2011).

61. Heinemann, J. A. & Sprague, G. F. Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340,** 205–209 (1989).

62. Davison, J. Genetic exchange between bacteria in the environment. *Plasmid* **42,** 73–91 (1999).

63. Lorenz, M. G. & Wackernagel, W. Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol. Rev.* **58,** 563–602 (1994).

64. Ellstrand, N. C., Prentice, H. C. & Hancock, J. F. Gene flow and introgression from domesticated plants into their wild relatives. *Annu. Rev. Ecol. Syst.* **30,** 539–563 (1999).

65. Naor, A., Lapierre, P., Mevarech, M., Papke, R. T. & Gophna, U. Low species barriers in halophilic archaea and the formation of recombinant hybrids. *Curr. Biol.* **22,** 1444–1448 (2012).

66. Lang, A. S., Zhaxybayeva, O. & Beatty, J. T. Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* **10,** 472–82 (2012).

67. Hotopp, J. C. D. *et al.* Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* **317,** 1753–1756 (2007).

68. Lieber, M. R. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. *Annu. Rev. Biochem.* **79,** 181–211 (2010).

69. Hao, W., Richardson, A. O., Zheng, Y. & Palmer, J. D. Gorgeous mosaic of mitochondrial genes

created by horizontal transfer and gene conversion. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 21576–81 (2010).

70. Funkhouser-Jones, L. J. *et al. Wolbachia* co-infection in a hybrid zone: discovery of horizontal gene transfers from two *Wolbachia* supergroups into an animal genome. *PeerJ* **3,** e1479 (2015).

71. Klasson, L. *et al.* Extensive duplication of the *Wolbachia* DNA in chromosome four of *Drosophila ananassae*. *BMC Genomics* **15,** 1097 (2014).

72. Brelsfoard, C. *et al.* Presence of extensive *Wolbachia* symbiont insertions discovered in the genome of its host *Glossina morsitans morsitans*. *PLoS Negl. Trop. Dis.* **8,** e2728 (2014).

73. Leclercq, S. *et al.* Birth of a W sex chromosome by horizontal transfer of *Wolbachia* bacterial symbiont genome. *Proc. Natl. Acad. Sci.* **113,** 15036–15041 (2016).

74. Kumar, N. *et al.* Efficient subtraction of insect rRNA prior to transcriptome analysis of *Wolbachia-Drosophila* lateral gene transfer. *BMC Res. Notes* **5,** 230 (2012).

75. Hazkani-Covo, E., Zeller, R. M. & Martin, W. Molecular poltergeists: mitochondrial DNA copies (numts) in sequenced nuclear genomes. *PLoS Genet.* **6,** e1000834 (2010).

76. Lloyd, A. H. & Timmis, J. N. The origin and characterization of new nuclear genes originating from a cytoplasmic organellar genome. *Mol. Biol. Evol.* **28,** 2019–2028 (2011).

77. Ricchetti, M., Tekaia, F. & Dujon, B. Continued colonization of the human genome by mitochondrial DNA. *PLoS Biol.* **2,** E273 (2004).

78. Wang, D. & Timmis, J. N. Cytoplasmic organelle DNA preferentially inserts into open chromatin. *Genome Biol. Evol.* **5,** 1060–4 (2013).

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

80. Hazkani-Covo, E. Mitochondrial insertions into primate nuclear genomes suggest the use of numts as a tool for phylogeny. *Mol. Biol. Evol.* **26,** 2175–2179 (2009).

81. Ricchetti, M., Fairhead, C. & Dujon, B. Mitochondrial DNA repairs double-strand breaks in yeast chromosomes. *Nature* **402,** 96–100 (1999).

82. Tsuji, J., Frith, M. C., Tomii, K. & Horton, P. Mammalian NUMT insertion is non-random. *Nucleic Acids Res.* **40,** 9073–9088 (2012).

83. Huang, C. Y., Grünheit, N., Ahmadinejad, N., Timmis, J. N. & Martin, W. Mutational decay and age of chloroplast and mitochondrial genomes transferred recently to angiosperm nuclear chromosomes. *Plant Physiol.* **138,** 1723–33 (2005).

84. Tourmen, Y. *et al.* Structure and chromosomal distribution of human mitochondrial pseudogenes. *Genomics* **80,** 71–7 (2002).

85. McNulty, S. N. *et al.* Endosymbiont DNA in endobacteria-free filarial nematodes indicates ancient horizontal genetic transfer. *PLoS One* **5,** e11029 (2010).

86. Acuna, R. *et al.* Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. *Proc. Natl. Acad. Sci. U. S. A.* **109,** 4197–4202 (2012).

87. Pauchet, Y. & Heckel, D. G. The genome of the mustard leaf beetle encodes two active xylanases originally acquired from bacteria through horizontal gene transfer. *Proc. Biol. Sci.* **280,** 20131021 (2013).

88. Gladyshev, E. A., Meselson, M. & Arkhipova, I. R. Massive horizontal gene transfer in bdelloid rotifers. *Science* **320,** 1210–3 (2008). ***One of the first studies to describe massive HGT to an animal genome with details on gene localization, structure, and function.***

89. Sun, B. F. *et al.* Multiple ancient horizontal gene transfers and duplications in lepidopteran species. *Insect Mol. Biol.* **22,** 72–87 (2013).

90. Craig, J. P., Bekal, S., Niblack, T., Domier, L. & Lambert, K. N. Evidence for horizontally transferred genes involved in the biosynthesis of vitamin B(1), B(5), and B(7) in *Heterodera glycines*. *J. Nematol.* **41,** 281–90 (2009).

91. Bowler, C. *et al.* The *Phaeodactylum* genome reveals the evolutionary history of diatom

genomes. *Nature* **456,** 239–44 (2008).

92. Danchin, E. G. J. *et al.* Multiple lateral gene transfers and duplications have promoted plant parasitism ability in nematodes. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 17651–17656 (2010).

93. Le Hir, H., Nott, A. & Moore, M. J. How introns influence and enhance eukaryotic gene expression. *Trends Biochem. Sci.* **28,** 215–220 (2003).

94. Gray, M. W. Mitochondrial evolution. *Cold Spring Harb. Perspect. Biol.* **4,** a011403 (2012).

95. Koonin, E. V & Yutin, N. The dispersed archaeal eukaryome and the complex archaeal ancestor of eukaryotes. *Cold Spring Harb. Perspect. Biol.* **6,** a016188 (2014).

96. Gray, M. W. Mosaic nature of the mitochondrial proteome: Implications for the origin and evolution of mitochondria. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 10133–10138 (2015).

97. Huynen, M. A., Duarte, I. & Szklarczyk, R. Loss, replacement and gain of proteins at the origin of the mitochondria. *Biochim. Biophys. Acta* **1827,** 224–31 (2013).

98. Gabaldón, T. & Huynen, M. A. From endosymbiont to host-controlled organelle: the hijacking of mitochondrial protein synthesis and metabolism. *PLoS Comput. Biol.* **3,** e219 (2007).

99. Gabaldón, T. & Huynen, M. A. Lineage-specific gene loss following mitochondrial endosymbiosis and its potential for function prediction in eukaryotes. *Bioinformatics* **21,** ii144-150 (2005).

100. Esser, C. *et al.* A genome phylogeny for mitochondria among alpha-proteobacteria and a predominantly eubacterial ancestry of yeast nuclear genes. *Mol. Biol. Evol.* **21,** 1643–60 (2004).

101. Thiergart, T., Landan, G., Schenk, M., Dagan, T. & Martin, W. F. An evolutionary network of genes present in the eukaryote common ancestor polls genomes on eukaryotic and mitochondrial origin. *Genome Biol. Evol.* **4,** 466–85 (2012).

102. Cotton, J. A. & McInerney, J. O. Eukaryotic genes of archaebacterial origin are more important than the more numerous eubacterial genes, irrespective of function. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 17252–5 (2010).

103. Reyes-Prieto, A. & Moustafa, A. Plastid-localized amino acid biosynthetic pathways of Plantae are predominantly composed of non-cyanobacterial enzymes. *Sci. Rep.* **2,** 955 (2012).

104. Kurland, C. G. & Andersson, S. G. Origin and evolution of the mitochondrial proteome. *Microbiol. Mol. Biol. Rev.* **64,** 786–820 (2000).

105. Gabaldón, T. & Huynen, M. A. Shaping the mitochondrial proteome. *Biochim. Biophys. Acta - Bioenerg.* **1659,** 212–220 (2004).

106. Suzuki, K. & Miyagishima, S. Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses. **27,** (2010).

107. Zaremba-Niedzwiedzka, K. *et al.* Metagenomic exploration of Asgard arcahea illuminates the origin of eukaryotic cellular complexity. *Nature* **541,** 353–358 (2017).

108. Pittis, A. A. & Gabaldón, T. Late acquisition of mitochondria by a host with chimaeric prokaryotic ancestry. *Nature* **531,** 101–104 (2016).

109. Koonin, E. V. Archaeal ancestors of eukaryotes: not so elusive any more. *BMC Biol.* **13,** 84 (2015).

110. Spang, A. *et al.* Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521,** 173–179 (2015).

111. Nowack, E. C. M. *Paulinella chromatophora* – rethinking the transition from endosymbiont to organelle. *Acta Soc. Bot. Pol.* **83,** 387–397 (2014).

112. Delaye, L., Valadez-Cano, C. & Pérez-Zamorano, B. How really ancient is *Paulinella chromatophora*? *PLoS Curr.* (2016). doi:10.1371/currents.tol.e68a099364bb1a1e129a17b4e06b0c6b

113. Klein, C. C. *et al.* Biosynthesis of vitamins and cofactors in bacterium-harbouring trypanosomatids depends on the symbiotic association as revealed by genomic analyses. *PLoS One* **8,** e79786 (2013).

114. Alves, J. M. P. *et al.* Endosymbiosis in trypanosomatids: the genomic cooperation between bacterium and host in the synthesis of essential amino acids is heavily influenced by multiple horizontal gene transfers. *BMC Evol. Biol.* **13,** 190 (2013).

115. Larkum, A. W. D., Lockhart, P. J. & Howe, C. J. Shopping for plastids. *Trends Plant Sci.* **12,** 189–195 (2007).

116. Gray, M. W. The pre-endosymbiont hypothesis: a new perspective on the origin and evolution of mitochondria. *Cold Spring Harb. Perspect. Biol.* **6,** 1–13 (2014).

117. Qiu, H. *et al.* Assessing the bacterial contribution to the plastid proteome. *Trends Plant Sci.* **18,** 680–7 (2013).

118. Dunning Hotopp, J. C. *et al.* Biology wars: the eukaryotes strike back. *Cell Host Microbe* **16,** 701–703 (2014).

119. Yue, J., Hu, X., Sun, H., Yang, Y. & Huang, J. Widespread impact of horizontal gene transfer on plant colonization of land. *Nat. Commun.* **3,** 1152 (2012).

120. Hirano, T. *et al.* Moss chloroplasts are surrounded by a peptidoglycan wall containing D-amino acids. *Plant Cell* **28,** 1521–32 (2016).

121. Machida, M. *et al.* Genes for the peptidoglycan synthesis pathway are essential for chloroplast division in moss. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 6753–8 (2006).

122. Garcia, M. *et al.* An *Arabidopsis* homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *Plant J.* **53,** 924–34 (2008).

123. Royet, J., Gupta, D. & Dziarski, R. Peptidoglycan recognition proteins: modulators of the microbiome and inflammation. *Nat. Rev. Immunol.* **11,** 837 (2011).

124. McFall-Ngai, M. *et al.* Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 3229–3236 (2013).

125. Troll, J. V. *et al.* Peptidoglycan induces loss of a nuclear peptidoglycan recognition protein

during host tissue development in a beneficial animal-bacterial symbiosis. *Cell. Microbiol.* **11,** 1114–1127 (2009).

126. Koropatnick, T. A. *et al.* Microbial factor-mediated development in a host-bacterial mutualism. *Science* **306,** (2004).

127. Nyholm, S. V & Graf, J. Knowing your friends: invertebrate innate immunity fosters beneficial bacterial symbioses. *Nat. Rev. Microbiol.* **10,** 815–27 (2012).

128. Wang, J. W. & Aksoy, S. PGRP-LB is a maternally transmitted immune milk protein that influences symbiosis and parasitism in tsetse's offspring. *Proc. Natl. Acad. Sci.* **109,** 10552–10557 (2012).

129. Eichinger, L. *et al.* The genome of the social amoeba *Dictyostelium discoideum*. *Nature* **435,** 43–57 (2005).

130. DiSalvo, S. *et al. Burkholderia* bacteria infectiously induce the proto-farming symbiosis of *Dictyostelium* amoebae and food bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **112,** E5029-37 (2015).

131. Peraro, M. D. & van der Goot, F. G. Pore-forming toxins: ancient, but never really out of fashion. *Nat. Rev. Microbiol.* **14,** 77–92 (2015).

132. Bernheimer, A. W. & Avigad, L. S. Partial characterization of aerolysin, a lytic exotoxin from *Aeromonas hydrophila*. *Infect. Immun.* **9,** 1016–21 (1974).

133. Sher, D., Fishman, Y., Melamed-Book, N., Zhang, M. & Zlotkin, E. Osmotically driven prey disintegration in the gastrovascular cavity of the green hydra by a pore-forming protein. *FASEB J.* **22,** 207–214 (2007).

134. Amino, R. *et al.* Trialysin, a novel pore-forming protein from saliva of hematophagous insects activated by limited proteolysis. *J. Biol. Chem.* **277,** 6207–13 (2002).

135. Hoang, Q. T. *et al.* An actinoporin plays a key role in water stress in the moss *Physcomitrella patens*. *New Phytol.* **184,** 502–510 (2009).

136. Daimon, T. *et al.* Beta-Fructofuranosidase genes of the silkworm, *Bombyx mori*: Insights into enzymatic adaptation of B. mori to toxic alkaloids in mulberry latex. *J. Biol. Chem.* **283,** 15271–15279 (2008).

137. Haegeman, A., Jones, J. T. & Danchin, E. G. J. Horizontal gene transfer in nematodes: a catalyst for plant parasitism? *Mol. Plant. Microbe. Interact.* **24,** 879–887 (2011).

138. Ambrose, K. V, Koppenhöfer, A. M. & Belanger, F. C. Horizontal gene transfer of a bacterial insect toxin gene into the *Epichloë* fungal symbionts of grasses. *Sci. Rep.* **4,** 5562 (2014).

139. Daborn, P. J. *et al.* A single *Photorhabdus* gene, makes caterpillars floppy (mcf), allows *Escherichia* coli to persist within and kill insects. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 10742–7 (2002).

140. Ricard, G. *et al.* Horizontal gene transfer from Bacteria to rumen Ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. *BMC Genomics* **7,** 22 (2006).

141. Richards, T. A. *et al.* Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proc. Natl. Acad. Sci. U. S. A.* **108,** 15258–15263 (2011). ***Our review focuses on HGT from bacteria, however, HGT between eukaryotes seems to be more common in at least some eukaryotic lineages. This study demonstrated extensive HGT from fungi that assisted the oomycetes when becoming successful plant parasites.***

142. Danchin, E. G. J. & Rosso, M.-N. Lateral gene transfers have polished animal genomes: lessons from nematodes. *Front. Cell. Infect. Microbiol.* **2,** 27 (2012).

143. Dieterich, C. *et al.* The *Pristionchus pacificus* genome provides a unique perspective on nematode lifestyle and parasitism. *Nat. Genet.* **40,** 1193–8 (2008).

144. Mayer, W. E., Schuster, L. N., Bartelmes, G., Dieterich, C. & Sommer, R. J. Horizontal gene transfer of microbial cellulases into nematode genomes is associated with functional assimilation and gene turnover. *BMC Evol. Biol.* **11,** 13 (2011).

145. Schuster, L. N. & Sommer, R. J. Expressional and functional variation of horizontally acquired

cellulases in the nematode Pristionchus pacificus. *Gene* **506,** 274–82 (2012).

146. Sun, G., Yang, Z., Ishwar, A. & Huang, J. Algal genes in the closest relatives of animals. *Mol. Biol. Evol.* **27,** 2879–89 (2010).

147. Nakashima, K., Yamada, L., Satou, Y., Azuma, J.-I. I. & Satoh, N. The evolutionary origin of animal cellulose synthase. *Dev. Genes Evol.* **214,** 81–88 (2004).

148. Sasakura, Y. *et al.* Transcriptional regulation of a horizontally transferred gene from bacterium to chordate. *Proc. R. Soc. London B Biol. Sci.* **283,** (2016).

149. Huang, J. *et al.* Phylogenomic evidence supports past endosymbiosis, intracellular and horizontal gene transfer in *Cryptosporidium parvum*. *Genome Biol.* **5,** R88 (2004).

150. Marchetti, A. *et al.* Ferritin is used for iron storage in bloom-forming marine pennate diatoms. *Nature* **457,** 467–70 (2009).

151. Allen, A. E. *et al.* Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 10438–43 (2008).

152. Andersson, J. O. Evolution of patchily distributed proteins shared between eukaryotes and prokaryotes: *Dictyostelium* as a case study. *J. Mol. Microbiol. Biotechnol.* **20,** 83–95 (2011).

153. Worden, A. Z. *et al.* Green evolution and dynamic adaptations revealed by genomes of the marine picoeukaryotes Micromonas. *Science* **375,** 268–272 (2009).

154. de Koning, A. P., Brinkman, F. S. L., Jones, S. J. M. & Keeling, P. J. Lateral gene transfer and metabolic adaptation in the human parasite *Trichomonas vaginalis*. *Mol. Biol. Evol.* **17,** 1769–1773 (2000).

155. Sun, G. & Huang, J. Horizontally acquired DAP pathway as a unit of self-regulation. *J. Evol. Biol.* **24,** 587–595 (2011).

156. Craig, J. P. *et al.* Analysis of a horizontally transferred pathway involved in vitamin B6 biosynthesis from the soybean cyst nematode *Heterodera glycines*. *Mol. Biol. Evol.* **25,** 2085–98

(2008).

157. Mock, T. *et al.* Evolutionary genomics of the cold-adapted diatom *Fragilariopsis cylindrus*. *Nature* **541,** 536–540 (2017).

158. Raymond, J. A. *et al.* Possible role of horizontal gene transfer in the colonization of sea ice by algae. *PLoS One* **7,** e35968 (2012).

159. Blanc, G. *et al.* The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol.* **13,** R39 (2012).

160. Slamovits, C. H. & Keeling, P. J. Class II photolyase in a microsporidian intracellular parasite. *J. Mol. Biol.* **341,** 713–21 (2004).

161. Fast, N. M., Law, J. S., Williams, B. A. P. & Keeling, P. J. Bacterial catalase in the microsporidian *Nosema locustae*: implications for microsporidian metabolism and genome evolution. *Eukaryot. Cell* **2,** 1069–75 (2003).

162. Salter, S. J. *et al.* Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12,** 87 (2014).

163. Podell, S. *et al.* DarkHorse: a method for genome-wide prediction of horizontal gene transfer. *Genome Biol.* **8,** R16 (2007).

164. Podell, S. *et al.* A database of phylogenetically atypical genes in archaeal and bacterial genomes, identified using the DarkHorse algorithm. *BMC Bioinformatics* **9,** 419 (2008).

165. Frickey, T., Lupas, A. N. & Tancred Frickey, A. N. L. PhyloGenie: automated phylome generation and analysis. *Nucleic Acids Res.* **32,** 5231–8 (2004).

166. Abby, S. S., Tannier, E., Gouy, M. & Daubin, V. Detecting lateral gene transfers by statistical reconciliation of phylogenetic forests. *BMC Bioinformatics* **11,** 324 (2010).

167. Zhu, Q., Kosoy, M. & Dittmar, K. HGTector: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* **15,** 717 (2014).

168. Nguyen, M., Ekstrom, A., Li, X. & Yin, Y. HGT-Finder: A new tool for horizontal gene transfer finding and application to *Aspergillus* genomes. *Toxins (Basel).* **7,** 4035–4053 (2015).

169. Crisp, A. *et al.* Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol.* **16,** 50 (2015).

170. Bergsten, J. A review of long-branch attraction. *Cladistics* **21,** 163–193 (2005).

171. Shimodaira, H. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51,** 492–508 (2002).

172. Kumar, S., Jones, M., Koutsovoulos, G., Clarke, M. & Blaxter, M. Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Front. Genet.* **4,** 237 (2013).

173. Garcia-Vallve, S., Romeu, A. & Palau, J. Horizontal gene transfer of glycosyl hydrolases of the rumen fungi. *Mol. Biol. Evol.* **17,** 352–361 (2000).

174. Richards, T. A., Leonard, G., Soanes, D. M. & Talbot, N. J. Gene transfer into the fungi. *Fungal Biol. Rev.* **25,** 98–110 (2011).

175. Tsaousis, A. D. *et al.* A novel route for ATP acquisition by the remnant mitochondria of *Encephalitozoon cuniculi*. *Nature* **453,** (2008).

176. Alexander, W. G., Wisecaver, J. H., Rokas, A. & Hittinger, C. T. Horizontally acquired genes in early-diverging pathogenic fungi enable the use of host nucleosides and nucleotides. *Proc. Natl. Acad. Sci.* **113,** 4116–4121 (2016). ***This comprehensive analysis of HGTs in Microsporidia and Cryptomycota fungi detected dozens of novel HGT candidates and especially parallel acquisitions that enable these pathogenic fungi to steal nucleosides and nucleotides from their hosts. One of the HGTs was also suggested as a possible treatment route for microsporidiosis.***

177. Stairs, C. W., Roger, A. J. & Hampl, V. Eukaryotic pyruvate formate lyase and its activating enzyme were acquired laterally from a firmicute. *Mol. Biol. Evol.* **28,** 2087–2099 (2011).

178. Loftus, B. *et al.* The genome of the protist parasite *Entamoeba histolytica*. *Nature* **433,** 865–868 (2005).

179. Carlton, J. M. *et al.* Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. *Science* **315,** 207–212 (2007).

180. Morrison, H. G. *et al.* Genomic minimalism in the early diverging intestinal parasite *Giardia lamblia*. *Science* **317,** 1921–6 (2007).

181. Slamovits, C. H. & Keeling, P. J. Pyruvate-phosphate dikinase of oxymonads and parabasalia and the evolution of pyrophosphate-dependent glycolysis in anaerobic eukaryotes. *Eukaryot. Cell* **5,** 148–54 (2006).

182. Stechmann, A. *et al.* The glycolytic pathway of *Trimastix pyriformis* is an evolutionary mosaic. *BMC Evol. Biol.* **6,** 101 (2006).

183. Boxma, B. *et al.* The [FeFe] hydrogenase of *Nyctotherus ovalis* has a chimeric origin. *BMC Evol. Biol.* **7,** 230 (2007).

184. Müller, M., Lee, J. A., Gordon, P., Gaasterland, T. & Sensen, C. W. Presence of prokaryotic and eukaryotic species in all subgroups of the PP(i)-dependent group II phosphofructokinase protein family. *J. Bacteriol.* **183,** 6714–6 (2001).

185. Barberà, M. J., Ruiz-Trillo, I., Leigh, J., Hug, L. A. & Roger, A. J. in *Origin of Mitochondria and Hydrogenosomes* 239–275 (Springer Berlin Heidelberg, 2007). doi:10.1007/978-3-540-38502-8_10

186. Nixon, J. E. J. *et al.* 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,** 181–90 (2002).

187. Sateriale, A. & Striepen, B. Beg, borrow and steal: three aspects of horizontal gene transfer in the protozoan parasite, *Cryptosporidium parvum*. *PLoS Pathog.* **12,** e1005429 (2016).

188. Gojković, Z. *et al.* Horizontal gene transfer promoted evolution of the ability to propagate under

anaerobic conditions in yeasts. *Mol. Genet. Genomics* **271,** 387–93 (2004).