Horizontal gene transfer: building the web of life

Shannon M. Soucy[1], Jinling Huang[2] and Johann Peter Gogarten[1,3]

Abstract | Horizontal gene transfer (HGT) is the sharing of genetic material between organisms that are not in a parent–offspring relationship. HGT is a widely recognized mechanism for adaptation in bacteria and archaea. Microbial antibiotic resistance and pathogenicity are often associated with HGT, but the scope of HGT extends far beyond disease-causing organisms. In this Review, we describe how HGT has shaped the web of life using examples of HGT among prokaryotes, between prokaryotes and eukaryotes, and even between multicellular eukaryotes. We discuss replacement and additive HGT, the proposed mechanisms of HGT, selective forces that influence HGT, and the evolutionary impact of HGT on ancestral populations and existing populations such as the human microbiome.

Selfish genetic element
A gene or group of genes that enhance their own transmission and reproductive success without making a positive contribution to the host's fitness.

[1]*Department of Molecular and Cell Biology, University of Connecticut, 91 North Eagleville Road, Storrs, Connecticut 06269–3125, USA.*
[2]*Department of Biology, East Carolina University, Greenville, North Carolina 27858, USA.*
[3]*Institute for Systems Genomics, University of Connecticut, Connecticut 06269–3125, USA.*
Correspondence to J.P.G.
e-mail: gogarten@uconn.edu
doi:10.1038/nrg3962

Horizontal gene transfer (HGT) was first described in microorganisms in the late 1940s[1], and around 20 years later it was speculated to have a role in the adaptation of multicellular eukaryotes — specifically plants[2]. Since then, methods to detect HGT have improved, and these have revealed the surprising extent and relevance of HGT to the variation of viral, prokaryotic and eukaryotic gene content. Many apparent gene duplications, for example, are now known to be the result of HGT, not autochthonous gene duplication, resulting in a 'web of life' rather than in a steadily bifurcating tree[3,4].

For a transferred gene to survive in the recipient lineage for long periods of time, the gene usually needs to provide a selective advantage either to itself (in the case of a selfish genetic element) or to the recipient, and research on HGT initially focused on such genes. However, it is now known that many of the genes that have been identified as transferred through comparative genomics between close relatives have neutral or nearly neutral effects in the recipient in both prokaryotic and eukaryotic organisms[5]. One rule for transferred genes seems to be 'first do no harm' — genes that are successfully integrated into a recipient are often expressed at low levels and encode functions at the periphery of metabolism[6]. These neutral acquisitions, however, can later provide novel combinations of genetic material for selection to act on — in some cases, the transferred material becomes domesticated over time and produces a beneficial phenotype. In other cases, when the imported genes remain neutral and there is no obvious benefit associated with their retention, the genes are likely to be lost over time.

HGT has long been recognized as an important force in the evolution of bacteria and archaea. However, the exchange of genetic information between prokaryotic symbionts and their eukaryotic hosts, and even between eukaryotes, signifies that HGT in eukaryotes occurs more frequently than previously thought[7,8]. Often these transfers involve gene donations to unicellular eukaryotes[9] and are frequently associated with bacterial endosymbionts[10] (known as endosymbiotic gene transfer (EGT) or intracellular gene transfer (IGT)). However, bacterial genes can also be transferred to multicellular eukaryotes[8]. Recent interest in the human microbiome has reinvigorated the search for HGTs from symbionts into the human genome. Although transfers of bacterial genes into the human germ line[11,12] have not be confirmed, evidence is accumulating of HGT from bacteria to human somatic cells[13]. These findings demonstrate the enduring influence of HGT on the evolution of all parts of the web of life, eukaryotes included.

In this Review, we present an overview of how HGT has contributed to innovation throughout the web of life by providing novel combinations of gene sequences for selection to act upon, thus shaping the evolution of species ranging from single-celled microorganisms to multicellular eukaryotes. Advances in the understanding of mechanisms of HGT, methods of identifying HGT events and the growth of genome databases have facilitated these insights.

Mechanisms of HGT

The three most recognized mechanisms of HGT in prokaryotes are conjugation, transformation and

transduction (FIG. 1). Conjugation requires physical contact between a donor and a recipient cell via a conjugation pilus, through which genetic material is transferred. Conjugation is canonically restricted to bacterial cells as the donor and recipient, however, *Agrobacterium* spp. is an exception and uses its conjugation machinery for HGT into plant cells[14,15]. Transformation is the uptake of exogenous DNA from the environment and has been reported in both archaea and bacteria[16,17]. Transduction is the delivery of genetic material through phage predation owing to the integration of exogenous host genetic material into a phage genome, and this phenomenon has been observed in both bacteria and archaea. There are two types of transduction: generalized, in which a random piece of the host DNA is incorporated during cell lysis; and specialized, in which a prophage imprecisely excises itself from a host genome and incorporates some of the flanking host DNAs.

Other mechanisms of gene transfer, such as gene transfer agents (GTAs) and cell fusion, have more recently been described. GTAs are gene delivery systems that are integrated into a host chromosome and are sometimes under host regulatory control. GTAs carry small random pieces of host genome in capsids for delivery to nearby hosts. GTAs are found in both bacteria and archaea. The GTA-encoding genes do not provide an obvious benefit to the host, which donates its DNA to others, nor is the benefit to the GTA-encoding genes obvious, because the GTA does not preferentially transfer the GTA-encoding genes. The question of how these genes remain under selection for function remains enigmatic[18]. One study found that GTAs from *Rhodobacter capsulatus* were able to transfer antibiotic resistance to bacteria from different phyla; however, other studies have shown that not all bacteria, including those with the genes encoding GTAs, are able to receive gene donations via GTAs[18]. GTAs have evolved from prophages that have lost the ability to target their own DNA for packaging[18]. Most GTAs cannot package a long enough segment of DNA to transfer all the genes that are necessary to produce GTAs — that is, in contrast to phages, GTAs cannot transfer all of the genes that encode them to a new host. This is an important distinction from transduction.

Cell fusion has been observed in both Euryarchaeota (*Haloferax* spp.) and Crenarchaeota (*Sulfolobus* spp.)[19,20]. Experimentally, cell fusion has been observed on solid media where *Haloferax volcanii* forms aggregates and cells become physically joined by several small bridges of fused cell membrane[21]. Bidirectional gene transfer that is mediated through cell fusion has also been observed between different *Haloferax* species[22]. The bidirectionality of this method of gene exchange means that it is more similar to sexual reproduction in eukaryotes than it is to conjugation in prokaryotes.

Circumstances that facilitate HGT in eukaryotes. The development of the nucleus sequestered genetic material in eukaryotes made gene exchange a more complicated process, although physical association over extended periods of time can facilitate HGT. Obligate endosymbiosis as a stable form of physical association often leads to the presence of foreign genes in eukaryotic genomes, as is the case for mitochondria and plastids, which are eukaryotic organelles that evolved from bacterial endosymbionts[10], and many other endosymbionts that have donated genetic material to their host genomes[23]. In the absence of an endosymbiotic partner, a congruent phylogenetic signal from multiple foreign genes has also been used to infer the presence of obsolete endosymbionts in plants and other photosynthetic eukaryotes[25,30]. Notably, however, genes of endosymbiotic origin are either absent or not obviously enriched in several eukaryotes that harbour endosymbionts[24,26], suggesting that proximity alone is not enough to ensure successful HGT.

Feeding activities are also frequently linked to gene acquisition. The mechanism of the 'you are what you eat' gene transfer ratchet proposed by W. Ford Doolittle suggests that many protists acquire genes through phagotrophy[27]. This mechanism is consistent with the findings that phagotrophic microbial eukaryotes often harbour many foreign genes[28,29].

The recently proposed weak-link model suggests that weakly protected unicellular or early developmental stages, especially in oviparous species, might constitute potential entry points for foreign genes into multicellular eukaryotes[8]. These foreign genes could then be spread through mitosis to germline cells, and thus to offspring. This model could potentially explain the fact that genes are frequently acquired in plants and animals that have eggs associated with endosymbionts or exposed to exterior environments (for example, mosses, *Drosophila* spp. and nematodes)[23,31,32].

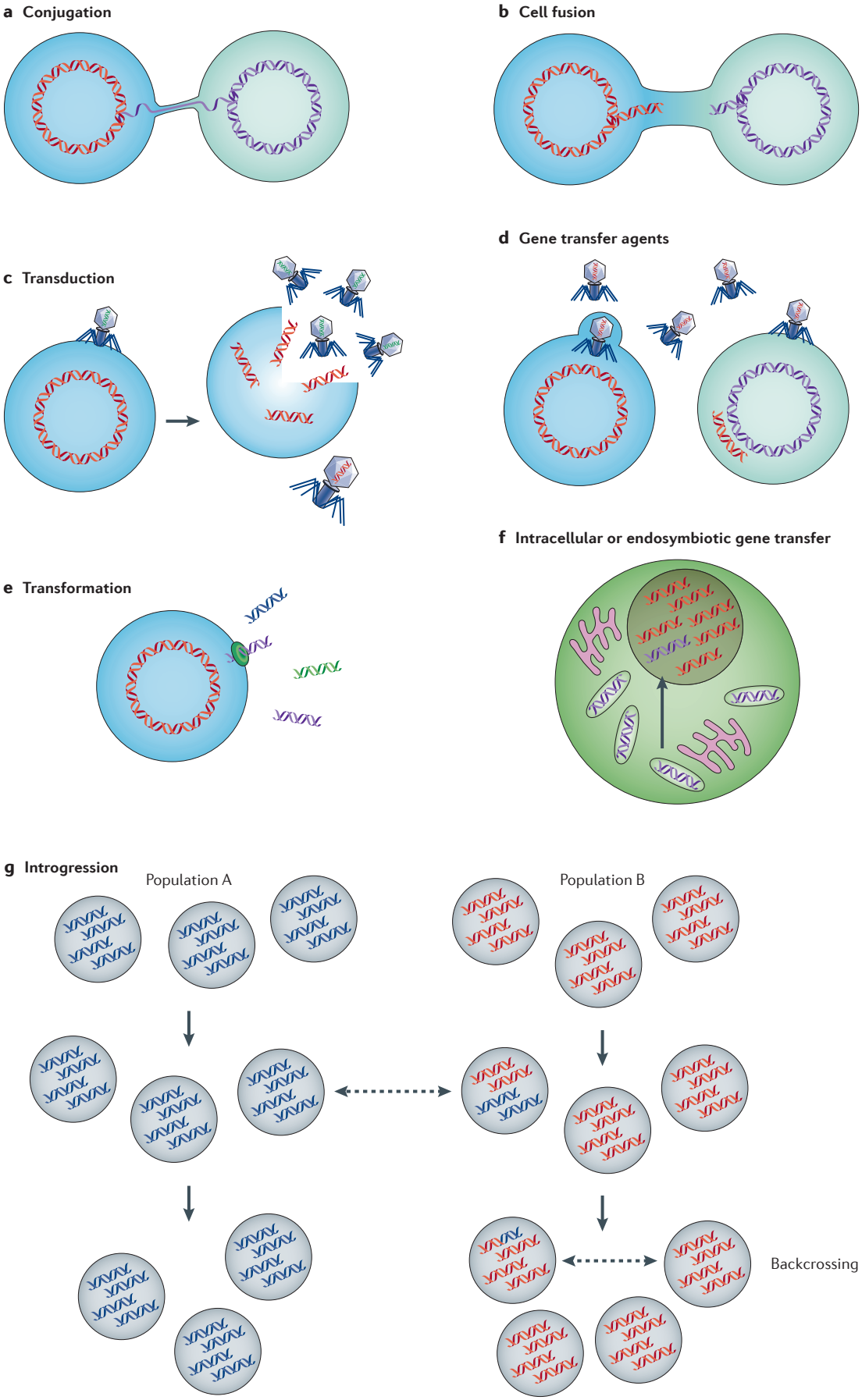
One way that genes can be exchanged between related species is through introgression — that is, gene flow due to interspecies hybridization followed by repeated backcrosses to one of the parent species. This mechanism is a major concern in transgenic crops that are grown in proximity to non-domesticated relatives[33]. Introgression of adaptive genes is not limited to plants. For example, introgression was inferred to have introduced an allele that is important in brain development from archaic to modern humans, and this transferred allele shows signs of being under positive selection in human populations[34].

Detecting HGT

Methods for detecting HGT generally rely on phylogenetic conflict, that is, conflicting branching patterns between two gene trees; usually one of these trees is considered to be an accepted species or a reference tree. Often the reference tree is assumed to represent the vertical evolution of the organisms that are being analysed; however, detecting conflict between a gene tree and the reference tree that is not due to uncertainty in phylogenetic reconstruction is sufficient to infer the transfer of either the gene or the markers used to calculate the reference tree[35]. Deviations from the branching pattern of the reference tree identify potential HGT events, and provide information about the organisms between which genes were exchanged. Species trees are often built using well-conserved housekeeping or informational genes, such as ribosomal proteins. These genes are

Microbiome
Following a definition ascribed to Joshua Lederberg this term is most often used to denote the collective genome of the indigenous microorganisms of a multicellular or unicellular host. However, the term has also been used by Lederberg and others to signify an ecological community of commensal, symbiotic and pathogenic microorganisms.

Phylogenetic conflict
Differences between the evolutionary history of a species and the evolutionary history of its genes are embodied by discrepancies in branching order between the species and the gene tree.



◀ Figure 1 | **Mechanisms of gene transfer**. Each panel represents a method of gene transfer. Conjugation (part **a**) occurs through donor–recipient cell contact, and single-stranded DNA is transferred from the donor cell to the recipient cell. Cell fusion (part **b**) differs from conjugation in that DNA is exchanged bi-directionally after cell contact and bridge formation between two cells. Gene transfer mediated by phage is known as transduction (part **c**). In the case of generalized transduction, any piece of genomic DNA may be loaded into the phage head; a general transducing phage is shown with host DNA (red). Specialized transduction occurs when an activated prophage loads a piece of genomic DNA neighbouring the prophage genome into the phage head together with the phage DNA (not shown). Gene transfer agents (GTAs) (part **d**) are phages that no longer recognize their own DNA and only carry random fragments of host DNA. Like prophage, they reside in the host cell genome. During transformation (part **e**) DNA is taken up from the surrounding environment; in the picture the DNA is depicted as entering the cell in the double stranded form, though many DNA uptake systems degrade one of the strands upon cell entry. Intracellular or endosymbiotic gene transfer (part **f**) occurs when genetic material from an endosymbiont or organelle (such as a chloroplast or mitochondrion) is incorporated into the host genome, this mainly pertains to eukaryotes. Introgression (part **g**) occurs when a hybridization event occurs between two diverging species (orange and blue populations). Backcrosses with one of the parent populations (orange) can lead to only a small piece of the divergent genome (blue) remaining in the recipient.

transferred less frequently between divergent organisms and can thus provide a good measure of vertical ancestry. Historically, the small subunit rRNA gene (SSU rRNA) has been used to determine the prokaryotic phylogeny. This practice was suggested to be problematic because several organisms have multiple divergent rRNA operons, and it was reported that homologous recombination can occur between them (see REF. 36 for a review). Multi-locus sequence analysis (MLSA) has emerged as a supplementary method for determining prokaryotic phylogeny. The aim is to minimize the phylogenetic conflict that results from the transfer of one or more of the genes by concatenating many genes. However, if the individual genes are not screened for phylogenetic conflict caused by HGT between divergent organisms, the resulting MLSA tree might not represent either a single gene tree or the organismal evolutionary history[5]. Careful screening of genes used in an MLSA data set for significant phylogenetic conflict, and using a large number of genes (such as the suite of 50 ribosomal proteins), can help to mitigate this problem. Generally, within a phylum, phylogenetic trees that are generated using MLSA are in good agreement with those made using SSU rRNA and also provide better resolution at the species level[37,38].

Quantification of bacterial and archaeal HGT is difficult because most transfers occur between closely related organisms and are difficult to distinguish owing to the genetic similarity of the host and the recipient genomes[39–41]. As mentioned above, the canonical method for detecting HGT events uses phylogenetic conflict comparing the gene history to the species history. Substantial and statistically supported conflict in the branching patterns of the gene and species trees can identify possible gene donors or the gene exchange partners if the direction of transfer cannot be interpreted. Gene duplication followed by differential gene loss is an alternative to HGT[5]; however, the more genome sequences become available, the more independent gene loss events need to be postulated and the less parsimonious the differential gene loss scenario becomes compared with an HGT explanation. Gene composition (codon usage and oligonucleotide composition) provides a tool to identify HGT candidates[42]. Composition that is different from the genome average performs especially well to identify recent transfers from distantly related donors or from phages, which have a composition that is distinct from that of the recipient[43]. Generally, the sets of identified HGTs using each of these methods (composition or phylogenetic based) are complementary rather than redundant[44].

The comparison of genomes from closely related organisms has identified large variation in gene content within a single species, especially in prokaryotic species. This variation in genome content reflects the ongoing process of gene gain and loss. Pan-genomes have been useful for studying the evolution of gene content in both prokaryotic species and genera. The pan-genome is defined as the set of all genes present in a taxon; the accessory genome contains genes that are present in only one or a few members of the taxon; and the core genome is the set of genes present in every member of the taxon. Each individual genome thus represents a sample from the pan-genome (BOX 1). An analysis of 61 *Escherichia coli* genomes revealed that only 6% of gene families were present in all genomes[45]. Pan-genomes were originally developed to explore the fluidity of prokaryotic genomes[46]; however, because HGT is more frequent between close relatives, the pan-genome also represents the set of genes that is potentially available via HGT to any member of the group. The eukaryotic pan-genome has been less extensively studied than the prokaryotic pan-genome, possibly because the impact of HGT is less well understood and the genomes are much larger. However, the pan-genome of *Emiliania huxleyi*, a globally distributed haptophyte phytoplankton species, has been studied. Although the accessory genome accounts for approximately one-third of genes present in the reference genome *E. huxleyi* CCMP1516, much of the variation in the pan-genome is related to intron tandem repeats and exon swapping, rather than HGT[47]. These data suggest that HGTs may be less frequent or that transferred genes may be less likely to persist in eukaryotes.

HGT in evolution

Mobile selfish genetic elements promote HGT. HGT enables innovations that evolved in one group of organisms to be shared across the web of life. Many HGTs provide a selective advantage to the recipient but, as described above, some transferred genes seem to be initially neutral or nearly neutral to the recipient. HGT of self-splicing selfish genetic elements such as introns and inteins provide examples of nearly neutral mobile genetic elements. Although the self-splicing activity minimizes the cost to the host organism, the additional DNA, RNA and protein synthesis associated with the selfish genetic element provide an additional burden to the host[48]. These elements persist because their success in invading new hosts compensates for the fitness cost to the host. Once

Genome streamlining
The reduction of genome size through relaxed selection and eventual loss of loci that are superfluous to the niche occupied by the organism.

Mobilome
The aggregate of mobile genetic elements in a genome, population or environment of interest.

Genome architecture imparting sequences
Strand-biased sequence motifs that are enriched towards the termini of replication; thought to direct proteins towards the termini.

established, these elements can provide material for variation, increased complexity and innovations. For example, in *Saccharomyces cerevisiae* the HO endonuclease, which evolved from an intein, functions as a mating-type switch cleaving at the MAT locus. Split inteins have become an integral part of synthesizing the DNA polymerase in marine picocyanobacteria. The group 2 introns evolved into spliceosomal introns, which now enable alternative splicing and fine-tuned regulation in most eukaryotes (see REF. 4 for a review). Thus, HGT disseminates beneficial, neutral and nearly neutral genes; subsequent selection can act on the variations that occur in the transferred genes, leading in some cases to their integration into cellular regulatory and metabolic networks.

Selfish genetic elements are commonly involved in promoting HGT and genome rearrangements, as well as facilitating the acquisition of genes that provide a selective advantage for recipients[49]. One example is the localization of antibiotic resistance genes in compound selfish elements such as plasmids, integrative conjugative elements (ICEs) and even group 2 introns[50]. These compound structures can contain a large repertoire of genes with unrelated functions. Compound selfish elements are often associated with toxin resistance genes, metabolic genes, virulence factors and a wide range of secreted factors[50]. The acquisition of a useful gene repertoire could offset the cost of maintaining and transferring a large selfish element such as a conjugal plasmid. The traits carried on compound mobile elements can be used as a gene reservoir in times of adversity[50,51]. Genome streamlining is common in prokaryotic populations, and thus the mobility of adaptive genes associated with the mobilome becomes an important evolutionary strategy. Studies of the mobilome in different populations might provide information about the selective pressures (FIG. 2) that act on these populations and that influence gene distribution via HGT.

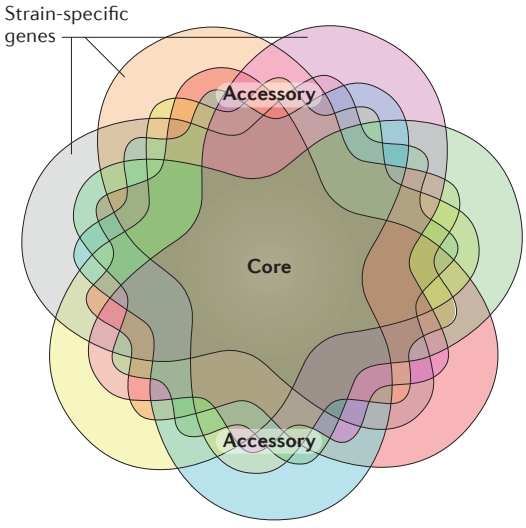
Selfish genetic elements are common in large multicellular eukaryotic genomes. Long terminal repeats often flank selfish elements and have been frequently co-opted to either increase or decrease gene expression in different tissues[52]. Syncytin genes that have a key role in trophoblast cell fusion during placenta development were repeatedly derived from retroviral envelope protein genes[52,53]. In organisms with distinct somatic and germline cells, phenotypic ingenuity often depends on the result of changes in the copy number or expression of a gene, which are often the result of selfish element dynamics in the germ line[54]. These changes can lead to divergence among or within species.

Biased gene transfer and highways of HGT. Successful HGTs frequently occur between closely related organisms[55], and the compositional similarity between the donor and the recipient genomes promotes homologous recombination that leads to homologous replacement with divergent alleles from close relatives. Additionally, the similarity between genome architecture imparting sequences in closely related organisms (same species or genera) leads to streamlined integration of the imported material[56]. In an analysis of 21 haloarchaeal genomes, over 90% of the HGTs identified through phylogenetic conflict were integrated into the recipient genome through homologous recombination[39]. The frequency of successful HGTs between pairs of Haloarchaea was shown to decrease exponentially with the phylogenetic distance (FIG. 3), probably due to the reduced efficiency of homologous recombination between genetically divergent organisms.

It was long thought that orthologous replacement through homologous recombination would be limited to the exchange of very similar gene sequences; however, the discovery of divergent isofunctional genes (known as homeoalleles) that can replace a divergent homologue in the recipient genome illustrated that homologous replacement can occur through homologous recombination in the conserved region flanking the divergent homeoalleles[40]. Divergent homeoalleles

Box 1 | **Pan-genome**

This depiction (see the figure) of the pan-genome and core genome is based on Edward's Venn cogwheel[104], and was designed by O. Zhaxybayeva, Dartmouth College, USA. The pan-genome of a group refers to the sum of all the genes that are present in members of the group. Pan-genomes comprise the core genome, which comprises the genes found in all members of a group of interest, and the accessory genome — genes that are present in only one or a few members of the group.
The concept of a pan-genome has led to the idea that steps in metabolic pathways may be distributed over several individuals within a community. The Black Queen hypothesis[105] suggests that the combination of leaky functions — genes that produce a product that is shared with others in the community — combined with a selection for small genomes, will lead to a situation in which leaky functions are encoded in the genomes of only a fraction of community members that produce this function as a common good. The pan-genomes of many taxa seem to be open (that is, of an unlimited size)[106–108], although the combination of limited population size and limited time of divergence from a common ancestor certainly limits the numbers of genes actually present in a given taxon. Estimated pan-genome sizes taking population size and divergence time into consideration can be large; for example, the *Prochlorococcus* pan-genome has been estimated to contain approximately 58,000 genes[109], whereas the individual genomes of the members of this genus encode only about 2,000 genes each.



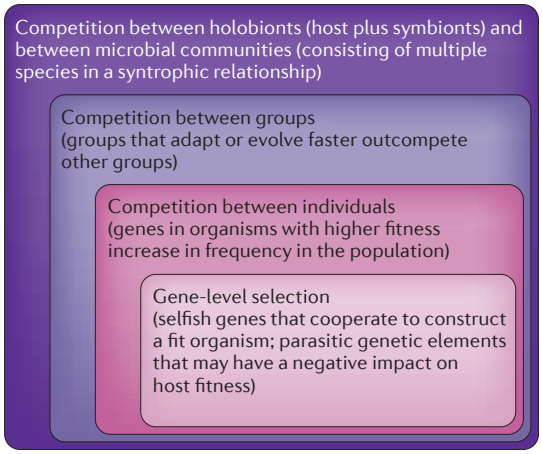


Figure 2 | **Nested levels of selection on gene content.** Each coloured box represents a different level of selection that can act on gene content.

of aminoacyl tRNA synthetases (aaRSs) provide an example of gene transfer that would go undetected by phylogenetic and compositional HGT detection methods. For many aaRSs, divergent forms evolved early in bacterial and archaeal evolution, and thus the diversity among aaRSs is easy to detect. The two or three forms with the same amino acid specificity frequently replace one another among both archaeal and bacterial species; however, because the transfers occur between related species, the gene tree of each type of aaRS remains in good agreement with the species tree[40]. Only the patchy distribution of each type reveals gene transfers and losses. Surprisingly, replacement with the divergent form was found to sometimes occur through homologous recombination in the more conserved flanking regions[40].

The frequency and bias of HGT makes it difficult to understand how adaptations might be maintained in ecological niches that are in close physical proximity[41]. At least during the initial divergence of ecotypes, genes are transferred between organisms that are adapted to different niches. It is possible that the higher frequency of within-ecotype HGT than between-ecotype HGT maintains ecotype adaptation. However, genes that adapt an organism to a particular niche are also transferred between niche boundaries[57], and such HGTs might help recipients to integrate into a new ecological niche (FIG. 4).

HGT enables key metabolic innovations. The enormous pan-genome size of many microbial species illustrates the importance of additive gene transfer, which is the process of the integration of novel genetic material into a genome. Integration into the genome can occur by non-homologous recombination or through homologous recombination involving the genes neighbouring the transferred gene (for example, see REF. 41). An additive transfer from a close relative of a gene that has an orthologue in the recipient genome leads to two similar copies being present in the recipient genome, an outcome that is similar to a gene duplication[4]. The methylaspartate cycle, for example, combines genes from several bacterial metabolic pathways that were transferred to the haloarchaeal ancestor from different bacterial donors and incorporated into a novel pathway for carbon assimilation[58]. Other examples of HGT contributing to the assembly or extension of metabolic pathways are acetoclastic methanogenesis in *Methanosarcina* spp. and the assembly of two photosystems functioning in series in oxygen-producing photosynthesis (see discussion in REF. 4 for details). In addition to frequently exchanging genes within and between genera, Haloarchaea also exchange genes with bacteria[39,59]. Haloarchaea are aerobic heterotrophs, although they evolved from methanogens — an anaerobic chemolithotrophic lineage. More than 1,000 genes were identified as imports from bacteria into Haloarchaea, including those for carbon assimilation, respiratory chain complexes, membrane transporters and cofactor biosynthesis[59]. The influx of these bacterial genes allowed the haloarchaeal ancestor to move into an aerobic environment. Similarly, the influx of bacterial genes to the ancestors of 12 other major archaeal clades is thought to have provided the key innovations to the origin of these groups[60]. Debate continues about whether the transferred genes originated from one or a few donors over a short period of time, or whether these transfers involved diverse bacterial donors[112,113]. The limited distribution of these genes within single groups of archaea indicates that 'highways' of gene sharing between archaea and bacteria have promoted archaeal diversity.

HGT and the evolution of the holobiont

Many organisms rely on a complex network of symbionts for functions ranging from defence and immunity to metabolism. The symbiotic communities that are associated with larger macro-organisms provide an initial interface with the environment, thus new properties and physiological responses often occur through HGT involving these communities. The holobiont[61] is used as a collective term for the host and its associated microbiota. For many multicellular eukaryotes, the number of genes in the microbiome[62] (genes that are present in the microbiota) dwarfs the number of genes in the nuclear genome of the host and provides an important source of genetic diversity.

The composition of human gut microbiota is affected by the diet and ecology of the human host, and by competition between members of the microbiota[62]. For example, bacteria in the gut of Japanese people can break down polysaccharides from the cell walls of seaweeds that are commonly present in the Japanese diet. The genes encoding the polysaccharide-digesting enzymes were transferred from parasites of marine algae to the gut bacteria[63,64]. This HGT has enabled Japanese people to use carbohydrates from algal cell walls as a nutrient source, whereas other populations cannot. It is tempting to interpret this as selection acting on the holobiont; however, it is more likely to reflect gut bacteria evolving to fill an available ecological niche (FIG. 2).

Ecotypes
Genetically distinct subsets of organisms within a population or species, usually genetic differences correspond to niche adaptation.

Holobiont
A multicellular or unicellular host and its collective symbionts.

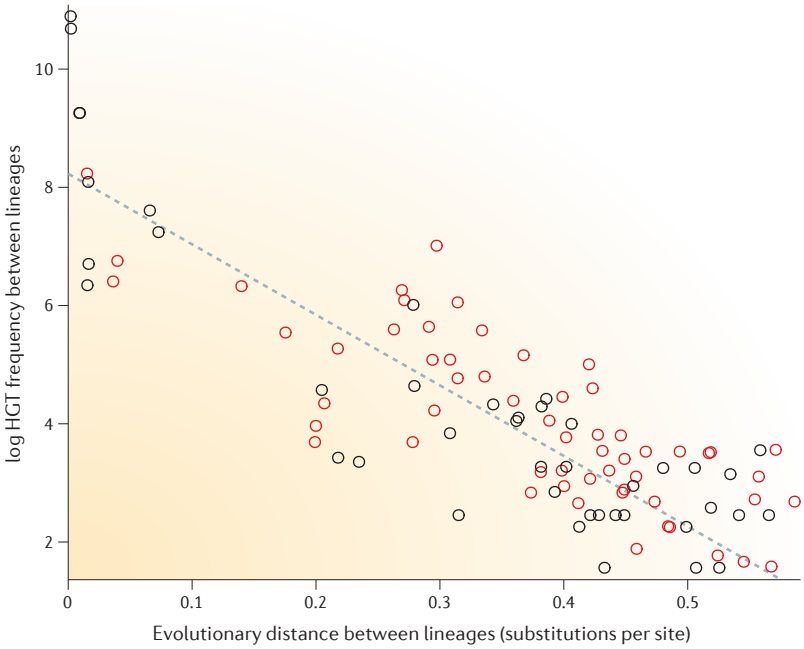


Figure 3 | **HGT is more frequent between closely related species.** The frequency of horizontal gene transfer (HGT) events in haloarchaea is plotted against evolutionary distance. Gene transfers were detected through phylogenetic conflict between the gene's phylogeny and the reference phylogeny calculated from ribosomal proteins. HGTs between terminal edges of the reference phylogeny are shown in black and those between internal edges are shown in red. Similar inverse log-linear relationships between recombination rate and divergence were also observed for bacterial genera. Reprinted from Williams, D., Gogarten, J. P. and Papke, R. T. Quantifying homologous replacement of loci between haloarchaeal species. Genome Biol. Evol. (2012). **4**, 1223–1244 by permission of Oxford University Press.

The results of recent research on the human microbiome have demonstrated the importance of the microbiota in nutrient acquisition and immune defence in humans. In an analysis that investigated recent gene transfers in the human microbiome, HGT was shown to be 25-fold more frequent between pairs of human-associated organisms than between pairs of organisms in different hosts or in aquatic or terrestrial environments[49]. Moreover, HGT between pairs of human-associated organisms isolated from the same body site are 50-fold more likely to exchange genes than pairs from other environments[49]. The surprising extent of gene transfer in human microbiota compared with other environments could indicate that environmental fluctuations that promote frequent adaptive changes are more prevalent in holobiont ecology, especially in the human holobiont. Notably, however, quantification of HGT is difficult, and sampling bias between environments (in that particular study, for example, 53% of the samples were of the human holobiont and the remaining 47% were split between aquatic, terrestrial and other host-associated environments[49]) could falsely inflate the rate of detection of HGT in well-sampled environments (humans) compared with that in environments with less available data.

HGT in eukaryotic evolution

Although still fragmented, the available data indicate that HGT is widespread in all major eukaryotic groups and has been ongoing throughout evolutionary time[7,8,65]. As stated above, the sequestration of genetic material to the nucleus requires distinct mechanisms for HGT in eukaryotes. Nevertheless, HGT is important in conferring beneficial phenotypes that may lead to the origin of major lineages. Furthermore, changes brought about by HGT may prompt the adaptive radiation of other groups through organismal interactions and genetic integration in a co-evolving web of life.

HGT in the origin of plastids and Plantae. The plant lineage is ripe with examples of HGTs that have conferred novel functions (FIG. 5). Plastids, the hallmark of photosynthetic eukaryotes, are derived from cyanobacterial endosymbionts in a eukaryotic host. With the only exception of chromatophores in amoeboid *Paulinella* spp., the well-founded belief is that all other photosynthetic eukaryotes trace their plastids to a single cyanobacterial endosymbiosis[66]. The transformation of a free-living cyanobacterium into a permanent organelle required both genetic and metabolic integration between the two partners. Several analyses identified 20–50 genes from chlamydiae, a group of obligate intracellular bacteria, in various photosynthetic eukaryotes[30,67,68]. These findings led to the suggestion that cyanobacterial and chlamydial endosymbionts coexisted in an early eukaryotic host cell, and that this tripartite relationship was responsible for the transformation of cyanobacterial endosymbionts into modern-day plastids[30,67,69,70]. Although it has been argued that these chlamydiae-related genes could have resulted from phylogenetic artefacts or could have existed in the cyanobacterial progenitor of plastids[71–73], some of these genes are only adaptive in parasitic or heterotrophic bacteria and are not found in extant cyanobacteria, suggesting that chlamydial involvement in plastid establishment is plausible[30,67,68,74]. Non-cyanobacterial prokaryotes other than chlamydiae also contributed genes for plastid genesis and functionality[69,75–77].

The establishment of cyanobacterial endosymbionts or plastids triggered the origin of Plantae: red algae, glaucophytes and green plants. Recent investigations have indicated that all three of these lineages have been affected by HGT during their evolution[69,78–80]. The glaucophyte *Cyanophora paradoxa* acquired more than 400 genes from bacteria[69]. In red algae, HGTs contributed to at least 5% of protein-coding genes in *Galdieria sulphuraria* and many others in *Porphyridium purpureum*[78,80]. Evidence of HGT has also been found in green algae[79] and land plants[81,82] (see below). For example, the moss *Physcomitrella patens* acquired genes from various sources, including fungi, bacteria, viruses and aquatic animals[32,83,84]. In most of these cases, acquired genes expanded the metabolic capabilities of recipients and had a key role in their adaptation to new environments, such as those with high salinity or acidity, extreme temperatures, or toxic substances.

HGT between plants and other eukaryotes. The origin of plastids and Plantae also spawned the emergence of other photosynthetic eukaryotes through secondary or higher-level endosymbioses. In addition, Plantae, which are rich in complex carbohydrates, generated new niches and resources for other organisms to exploit. Particularly, plant cell walls are the most abundant biomass on earth. Both the prevalence and novelty of this insoluble stored energy enhanced adaptive pressure to take advantage of novel resources free of competition. To effectively utilize plant biomass, other organisms often share genes or metabolic capabilities. For example, numerous soil bacteria reside in the rhizosphere and rely on root exudes as their primary nutrient source. An increase in exude production leads to active bacterial growth and thus more frequent plasmid transfer among rhizobacteria[85]. Choanoflagellates and rotifers, both of which live in aquatic environments, acquired numerous genes from plants and miscellaneous algae[86,87], frequently related to complex carbohydrate degradation[28]. In rumen ciliates, 46 genes related to the degradation of complex carbohydrates, such as plant biomass, were acquired by HGT, many of them from the gut bacteria of ruminant animals[88]. Beyond choanoflagellates and rumen ciliates, the ability to degrade plant biomass has been independently acquired by many other eukaryotic groups such as oomycetes, fungi and nematodes[89,90]. The widespread and diverse mechanisms for degrading complex carbohydrates in plants in so many different lineages highlight the convergent evolution through HGT for adaptation.

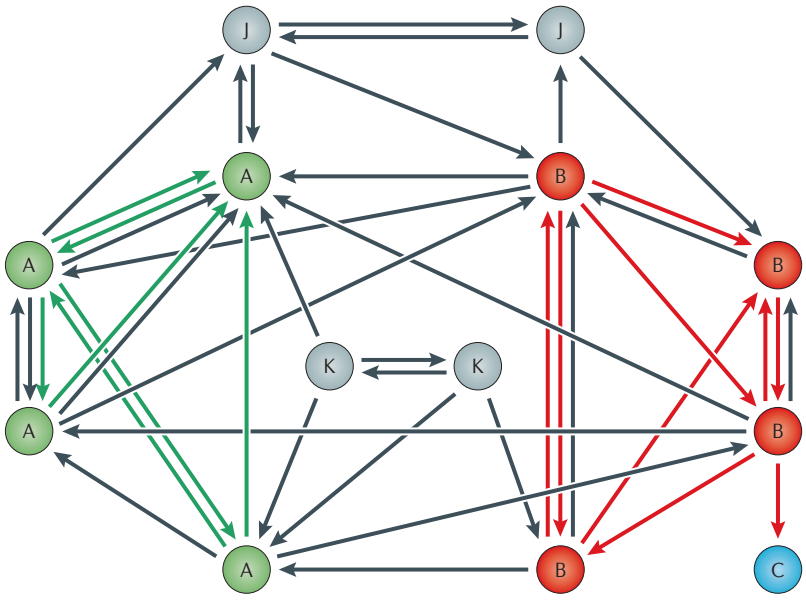


Figure 4 | **Structured exchange community.** Prokaryotic members of two distinct niches are shown as green and red circles (A and B); grey circles (K and J) are related species occupying different niches. Genes that enable the adaptation of their hosts to these niches are mostly exchanged between members of the same niche (green and red arrows), but they might also be shared with recent niche invaders (blue circle; C), accelerating the adaptation of the invader to a new habitat. Adapted with permission from REF. 57, (AAAS).

Lepidopterans are the largest group of plant-feeding insects, and their diversification coincided with the emergence of flowering plants. In an analysis of HGT in lepidopteran insects, most of the acquired genes were shown to be distributed in multiple lepidopteran groups and related to nutritional metabolism and detoxification[91]. The production of toxins by plants and the corresponding genes for detoxification in lepidopterans, and other phytophagous arthropods, exemplifies a genetic 'arms race' fuelled by HGT. Many plants can produce cyanogenic glucosides, which can be converted to highly toxic hydrogen cyanide as a defence against herbivores. Conversely, phytophagous arthropods not only sequester hydrogen cyanide as a defence against their own predators, but also counteract cyanide poisoning through detoxification genes that were originally recruited from bacteria[92].

HGT between multicellular eukaryotes. Many cases of HGT were reported between parasitic plants and their hosts[93–96]. In almost all of these cases, the direction of HGT is consistent with the direction of nutrient transfer from the host to the parasitic plant. HGT also occurs between multicellular eukaryotes with less obvious physical associations. For example, the moss *P. patens* acquired an actinoporin gene that is involved in desiccation resistance from metazoans[83]. *Alloteropsis* grasses switched to C_4 photosynthesis at least four times in the past 10 million years through the acquisition of genes from other C_4 grasses[97]. A photoreceptor gene was transferred from hornworts to ferns, allowing modern ferns to thrive in low-light conditions under the canopy[98]. Sturgeons, lampreys, which have been known to feed on sturgeons, and paddle fishes all share a transposable element, probably the result of HGT mediated by the exchange of fluids during lamprey feeding[99]. The sporadic distribution of type II antifreeze protein (AFP) genes in herring, smelt and sea raven was also mediated by HGT, allowing these fish to adapt to icy water[31].

For a long time, mitochondria were considered uniparentally inherited and subject to Muller's ratchet[100]. For many groups of organisms, this assumption seems to be correct[101]; however, plant, algal and fungal mitochondrial genomes are known to be dynamic and promiscuous, varying greatly among species in structure and gene content[102]. The transfer of mitochondrial genes between plant species can be massive and widespread. In an extreme case, *Amborella trichopoda*, a basal flowering plant, acquired at least four whole mitochondrial genomes from mosses and green algae, as well as many mitochondrial and, to a lesser degree, plastidal fragments from other flowering plants[103]. This example of HGT is not known to be associated with an adaptive benefit and is instead an important example of neutral or nearly neutral gene transfer in eukaryotes.

The mode of HGT between multicellular eukaryotes remains controversial. Are individual genes transferred, or are the transfers the consequence of between-species hybridization followed by backcrosses to one of

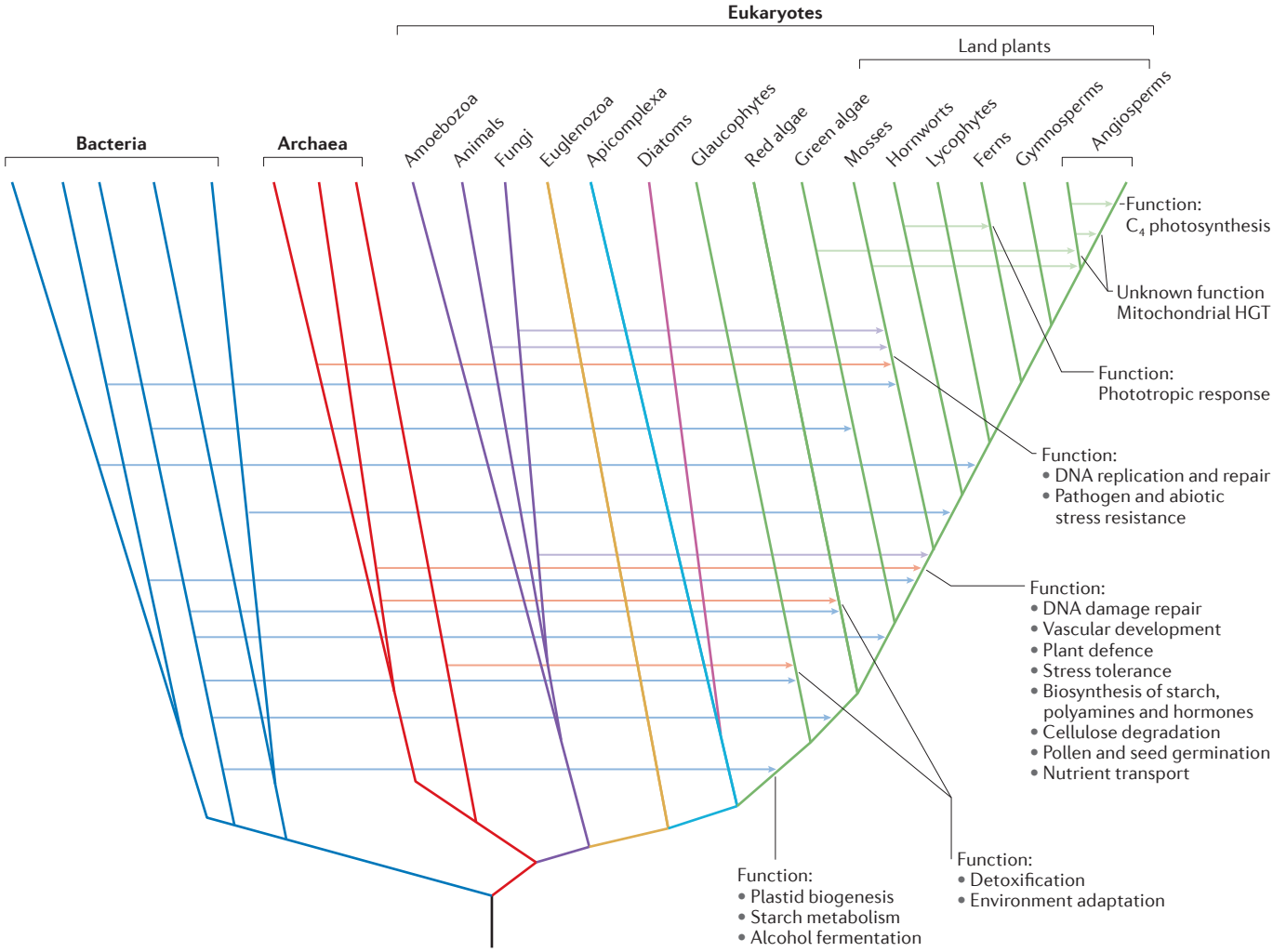


Figure 5 | **HGT to the plant lineage.** Arrows are coloured based on the origin of the gene transferred. Lines at the tips of the arrows indicate the gain of function for the plant lineage that acquired the genetic material. HGT, horizontal gene transfer. Figure modified from REF. 32, Nature Publishing Group.

the parents[7]? In many instances, such as the transfer of AFP genes from herring to smelt[31], donor and recipient diverged more than 200 million years ago, making hybridization an unlikely scenario. The conservation of introns between donor and recipient argues against independent transfers from bacterial symbionts. Sperm-mediated gene transfer between fish is one possible scenario[31]. In the moss *P. patens*, eggs and embryos that are exposed to bacteria and fungi in the environment might have facilitated gene acquisition. The large-scale acquisitions of mitochondrial genes in *Amborella trichopoda* probably occurred through mitochondrial genome fusion mediated by regenerated meristems from wounded areas.

Perspective

In this Review, we have discussed examples that illustrate how HGT shapes gene content in bacteria, archaea and unicellular eukaryotes (see Supplementary information S1 (table)). Even in multicellular eukaryotes, HGT from symbionts and between mitochondria occurs frequently and can have an important impact on gene content. Currently, we have a good understanding of the mechanisms by which prokaryotes exchange genes, including through GTAs and cell fusion in archaea; however, the mechanisms by which multicellular eukaryotes exchange genes with one another and with prokaryotes are less clear. The weak-link model, sperm-mediated gene transfer and introgression are possible gene transfer pathways, but more work is needed to explore the specific mechanisms involved. Importantly, comparisons between closely related strains will lead to a more accurate characterization of HGTs. Improvements in HGT detection based on the growing collection of sequence data will result in a more realistic estimation of HGT rates. However, accounting for false negatives and various types of transfer over different phylogenetic distances remains a challenge. Nevertheless, the surprising density of the web of life woven through genetic exchange is becoming visible.

1. Tatum, E. L. & Lederberg, J. Gene recombination in the bacterium *Escherichia coli*. *J. Bacteriol.* **53**, 673–684 (1947).
2. Went, F. W. Parallel evolution. *Taxon* **20**, 197–226 (1971).
3. Treangen, T. J. & Rocha, E. P. C. Horizontal transfer, not duplication, drives the expansion of protein families in prokaryotes. *PLoS Genet.* **7**, e1001284 (2011).
4. Swithers, K. S., Soucy, S. M. & Gogarten, J. P. The role of reticulate evolution in creating innovation and complexity. *Int. J. Evol. Biol.* **2012**, 418964 (2012).
5. Gogarten, J. P. & Townsend, J. P. Horizontal gene transfer, genome innovation and evolution. *Nat. Rev. Microbiol.* **3**, 679–687 (2005).
6. Park, C. & Zhang, J. High expression hampers horizontal gene transfer. *Genome Biol. Evol.* **4**, 523–532 (2012).
This paper examines the impact of expression level on the transferability of a gene in both environmental and laboratory populations of *E. coli*.
7. Boto, L. Horizontal gene transfer in the acquisition of novel traits by metazoans. *Proc. Biol. Sci.* **281**, 20132450 (2014).
8. Huang, J. Horizontal gene transfer in eukaryotes: the weak-link model. *Bioessays* **35**, 868–875 (2013).
This letter proposes a model for ongoing HGT in eukaryotes involving unicellular and early developmental stages to overcome the barrier of genome sequestration in eukaryotes.
9. Andersson, J. O. Gene transfer and diversification of microbial eukaryotes. *Annu. Rev. Microbiol.* **63**, 177–193 (2009).
10. 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).
11. Koonin, E. V. Darwinian evolution in the light of genomics. *Nucleic Acids Res.* **37**, 1011–1034 (2009).
12. Crisp, A., Boschetti, C., Perry, M., Tunnacliffe, A. & Micklem, G. Expression of multiple horizontally acquired genes is a hallmark of both vertebrate and invertebrate genomes. *Genome Biol.* **16**, 50 (2015).
13. Riley, D. R. *et al.* Bacteria–human somatic cell lateral gene transfer is enriched in cancer samples. *PLoS Comput. Biol.* **9**, e1003107 (2013).
14. Norman, A., Hansen, L. H. & Sørensen, S. J. Conjugative plasmids: vessels of the communal gene pool. *Phil. Trans. R. Soc. B* **364**, 2275–2289 (2009).
15. Kyndt, T. *et al.* The genome of cultivated sweet potato contains *Agrobacterium* T-DNAs with expressed genes: an example of a naturally transgenic food crop. *Proc. Natl Acad. Sci. USA* **112**, 201419685 (2015).
16. Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J.-P. Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat. Rev. Microbiol.* **12**, 181–196 (2014).
17. Chimileski, S., Dolas, K., Naor, A., Gophna, U. & Papke, R. T. Extracellular DNA metabolism in *Haloferax volcanii*. *Front. Microbiol.* **5**, 57 (2014).
18. Lang, A. S., Zhaxybayeva, O. & Beatty, J. T. Gene transfer agents: phage-like elements of genetic exchange. *Nat. Rev. Microbiol.* **10**, 472–482 (2012).
19. Naor, A. & Gophna, U. Cell fusion and hybrids in Archaea: prospects for genome shuffling and accelerated strain development for biotechnology. *Bioengineered* **4**, 126–129 (2013).
20. Schleper, C., Holz, I., Janekovic, D., Murphy, J. & Zillig, W. A multicopy plasmid of the extremely thermophilic archaeon *Sulfolobus* effects its transfer to recipients by mating. *J. Bacteriol.* **177**, 4417–4426 (1995).
21. Mevarech, M. & Werczberger, R. Genetic transfer in *Halobacterium volcanii*. *J. Bacteriol.* **162**, 461–462 (1985).
22. 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).
23. Dunning Hotopp, J. C. *et al.* Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science* **317**, 1753–1756 (2007).
This paper describes HGT between *Wolbachia* spp., an intracellular bacterial symbiont, and its multicellular eukaryotic insect hosts.
24. 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).
25. Moustafa, A. *et al.* Genomic footprints of a cryptic plastid endosymbiosis in diatoms. *Science* **324**, 1724–1726 (2009).
26. Chapman, J. A. *et al.* The dynamic genome of *Hydra*. *Nature* **464**, 592–596 (2010).
27. Doolittle, W. F. You are what you eat: a gene transfer ratchet could account for bacterial genes in eukaryotic nuclear genomes. *Trends Genet.* **14**, 307–311 (1998).
28. Yue, J., Sun, G., Hu, X. & Huang, J. The scale and evolutionary significance of horizontal gene transfer in the choanoflagellate *Monosiga brevicollis*. *BMC Genomics* **14**, 729 (2013).
29. Grant, J. R. & Katz, L. A. Phylogenomic study indicates widespread lateral gene transfer in entamoeba and suggests a past intimate relationship with parabasalids. *Genome Biol. Evol.* **6**, 2350–2360 (2014).
30. Huang, J. & Gogarten, J. P. Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? *Genome Biol.* **8**, R99 (2007).
This paper discusses a complex tripartite relationship between a eukaryotic host, a cyanobacterium and a chlamydia that may have facilitated the establishment of modern plastids.
31. Graham, L. A., Li, J., Davidson, W. S. & Davies, P. L. Smelt was the likely beneficiary of an antifreeze gene laterally transferred between fishes. *BMC Evol. Biol.* **12**, 190 (2012).
32. 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).
33. Stewart, C. N., Halfhill, M. D. & Warwick, S. I. Transgene introgression from genetically modified crops to their wild relatives. *Nat. Rev. Genet.* **4**, 806–817 (2003).
34. Evans, P. D., Mekel-Bobrov, N., Vallender, E. J., Hudson, R. R. & Lahn, B. T. Evidence that the adaptive allele of the brain size gene *microcephalin* introgressed into *Homo sapiens* from an archaic *Homo* lineage. *Proc. Natl Acad. Sci. USA* **103**, 18178–18183 (2006).
35. Williams, D. *et al.* A rooted net of life. *Biol. Direct* **6**, 45 (2011).
36. Gogarten, J. P., Doolittle, W. F. & Lawrence, J. G. Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* **19**, 2226–2238 (2002).
37. Colston, S. M. *et al.* Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using *Aeromonas* as a test case. *mBio* **5** e02136-14 (2014).
38. Delamuta, J. R. M., Ribeiro, R. A., Menna, P., Bangel, E. V. & Hungria, M. Multilocus sequence analysis (MLSA) of *Bradyrhizobium* strains: revealing high diversity of tropical diazotrophic symbiotic bacteria. *Braz. J. Microbiol.* **43**, 698–710 (2012).
39. Williams, D., Gogarten, J. P. & Papke, R. T. Quantifying homologous replacement of loci between haloarchaeal species. *Genome Biol. Evol.* **4**, 1223–1244 (2012).
40. Andam, C. P. & Gogarten, J. P. Biased gene transfer in microbial evolution. *Nat. Rev. Microbiol.* **9**, 543–555 (2011).
41. Polz, M., Alm, E. & Hanage, W. Horizontal gene transfer and the evolution of bacterial. **29**, 170–175 (2015).
This paper investigates the interplay between HGT, population structure and lineage divergence in bacteria and archaea.
42. Langille, M. G. I., Hsiao, W. W. L. & Brinkman, F. S. L. Detecting genomic islands using bioinformatics approaches. *Nat. Rev. Microbiol.* **8**, 373–382 (2010).
43. Daubin, V. & Ochman, H. Bacterial genomes as new gene homes: the genealogy of ORFans in *E. coli*. *Genome Res.* **14**, 1036–1042 (2004).
44. Ragan, M. A. On surrogate methods for detecting lateral gene transfer. *FEMS Microbiol. Lett.* **201**, 187–191 (2001).
45. Lukjancenko, O., Wassenaar, T. M. & Ussery, D. W. Comparison of 61 sequenced *Escherichia coli* genomes. *Microb. Ecol.* **60**, 708–720 (2010).
46. Tettelin, H. *et al.* Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial 'pan-genome'. *Proc. Natl Acad. Sci. USA* **102**, 13950–13955 (2005).
47. Read, B. A. *et al.* Pan genome of the phytoplankton *Emiliania* underpins its global distribution. *Nature* **499**, 209–213 (2013).
48. Barzel, A., Obolski, U., Gogarten, J. P., Kupiec, M. & Hadany, L. Home and away — the evolutionary dynamics of homing endonucleases. *BMC Evol. Biol.* **11**, 324 (2011).
49. Smillie, C. S. *et al.* Ecology drives a global network of gene exchange connecting the human microbiome. *Nature* **480**, 241–244 (2011).
This letter investigates the frequency of HGT in the human microbiome across body sites and across continents.
50. Rankin, D. J., Rocha, E. P. C. & Brown, S. P. What traits are carried on mobile genetic elements, and why? *Hered. (Edinb.).* **106**, 1–10 (2011).
This paper investigates the types of traits that are associated with compound selfish genetic elements and investigates the ecological scenarios that would select for specific types of traits.
51. Broaders, E., Gahan, C. G. M. & Marchesi, J. R. Mobile genetic elements of the human gastrointestinal tract: potential for spread of antibiotic resistance genes. *Gut Microbes* **4**, 271–280 (2013).
52. Feschotte, C. & Gilbert, C. Endogenous viruses: insights into viral evolution and impact on host biology. *Nat. Rev. Genet.* **13**, 283–296 (2012).
53. Cornelis, G. *et al.* Ancestral capture of *syncytin-Car1*, a fusogenic endogenous retroviral envelope gene involved in placentation and conserved in Carnivora. *Proc. Natl Acad. Sci. USA* **109**, E432–E441 (2012).
54. Schaack, S., Gilbert, C. & Feschotte, C. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. *Trends Ecol. Evol.* **25**, 537–546 (2010).
55. Skippington, E. & Ragan, M. A. Phylogeny rather than ecology or lifestyle biases the construction of *Escherichia coli–Shigella* genetic exchange communities. *Open Biol.* **2**, 120112 (2012).
56. Hendrickson, H. & Lawrence, J. G. Selection for chromosome architecture in bacteria. *J. Mol. Evol.* **62**, 615–629 (2006).
57. Papke, R. T. & Gogarten, J. P. Ecology. How bacterial lineages emerge. *Science* **336**, 45–46 (2012).
58. Khomyakova, M., Bükmez, Ö., Thomas, L. K., Erb, T. J. & Berg, I. A. A methylaspartate cycle in haloarchaea. *Science* **331**, 334–337 (2011).
59. Nelson-Sathi, S. *et al.* Acquisition of 1,000 eubacterial genes physiologically transformed a methanogen at the origin of haloarchaea. *Proc. Natl Acad. Sci. USA* **109**, 20537–20542 (2012).
60. Nelson-Sathi, S. *et al.* Origins of major archaeal clades correspond to gene acquisitions from bacteria. *Nature* **517**, 77–*80* (2014).
This paper suggests that acquisitions of genes from bacteria lead to the evolution of the major clades in archaea.
61. Guerrero, R., Margulis, L. & Berlanga, M. Symbiogenesis: the holobiont as a unit of evolution. *Int. Microbiol.* **16**, 133–143 (2013).
62. Ley, R. E., Peterson, D. A. & Gordon, J. I. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* **124**, 837–848 (2006).
63. Hehemann, J.-H. *et al.* Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature* **464**, 908–912 (2010).
64. Thomas, F. *et al.* Characterization of the first alginolytic operons in a marine bacterium: from their emergence in marine Flavobacteriia to their independent transfers to marine Proteobacteria and human gut *Bacteroides*. *Environ. Microbiol.* **14**, 2379–2394 (2012).
65. Hirt, R. P., Alsmark, C. & Embley, T. M. Lateral gene transfers and the origins of the eukaryote proteome: a view from microbial parasites. *Curr. Opin. Microbiol.* **23**, 155–162 (2015).
66. McFadden, G. I. Origin and evolution of plastids and photosynthesis in eukaryotes. *Cold Spring Harb. Perspect. Biol.* **6**, a016105 (2014).
67. Ball, S. G. *et al.* Metabolic effectors secreted by bacterial pathogens: essential facilitators of plastid endosymbiosis? *Plant Cell* **25**, 7–21 (2013).
68. Moustafa, A., Reyes-Prieto, A. & Bhattacharya, D. Chlamydiae has contributed at least 55 genes to Plantae with predominantly plastid functions. *PLoS ONE* **3**, e2205 (2008).
69. Price, D. C. *et al.* *Cyanophora paradoxa* genome elucidates origin of photosynthesis in algae and plants. *Science* **335**, 843–847 (2012).
70. Cenci, U. *et al.* Transition from glycogen to starch metabolism in archaeplastida. *Trends Plant Sci.* **19**, 18–28 (2014).
71. Deschamps, P. Primary endosymbiosis: have cyanobacteria and Chlamydiae ever been roommates? *Acta Soc. Bot. Pol.* **83**, 291–302 (2014).

72. Ku, C. *et al.* Endosymbiotic gene transfer from prokaryotic pangenomes: inherited chimerism in eukaryotes. *Proc. Natl Acad. Sci. USA* http://dx.doi.org/10.1073/pnas.1421385112 (2015).
73. Domman, D., Horn, M., Embley, T. M. & Williams, T. A. Plastid establishment did not require a chlamydial partner. *Nat. Commun.* **6**, 6421 (2015).
74. Ball, S. G. *et al.* Toward an understanding of the function of Chlamydiales in plastid endosymbiosis. *Biochim. Biophys. Acta* **1847**, 495–504 (2015).
75. Huang, J. & Gogarten, J. P. Concerted gene recruitment in early plant evolution. *Genome Biol.* **9**, R109 (2008).
76. Suzuki, K. & Miyagishima, S. Y. Eukaryotic and eubacterial contributions to the establishment of plastid proteome estimated by large-scale phylogenetic analyses. *Mol. Biol. Evol.* **27**, 581–590 (2010).
77. Qiu, H. *et al.* Assessing the bacterial contribution to the plastid proteome. *Trends Plant Sci.* **18**, 680–687 (2013).
78. Schonknecht, G. *et al.* Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* **339**, 1207–1210 (2013).
79. Blanc, G. *et al.* The genome of the polar eukaryotic microalga *Coccomyxa subellipsoidea* reveals traits of cold adaptation. *Genome Biol.* **13**, R39 (2012).
80. Bhattacharya, D. *et al.* Genome of the red alga *Porphyridium purpureum*. *Nat. Commun.* **4**, 1941 (2013).
81. Yue, J., Hu, X. & Huang, J. Origin of plant auxin biosynthesis. *Trends Plant Sci.* **19**, 764–770 (2014).
82. Yang, Z. *et al.* Ancient horizontal transfer of transaldolase-like protein gene and its role in plant vascular development. *New Phytol.* **206**, 807–816 (2015).
83. 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).
84. Maumus, F., Epert, A., Nogue, F. & Blanc, G. Plant genomes enclose footprints of past infections by giant virus relatives. *Nat. Commun.* **5**, 4268 (2014).
85. Molbak, L., Molin, S. & Kroer, N. Root growth and exudate production define the frequency of horizontal plasmid transfer in the rhizosphere. *FEMS Microbiol. Ecol.* **59**, 167–176 (2007).
86. Sun, G., Yang, Z., Ishwar, A. & Huang, J. Algal genes in the closest relatives of animals. *Mol. Biol. Evol.* **27**, 2879–2889 (2010).
87. Boschetti, C. *et al.* Biochemical diversification through foreign gene expression in Bdelloid Rotifers. *PLoS Genet.* **8**, e1003035 (2012).
88. 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).
89. 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).
90. Richards, T. A. *et al.* Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proc. Natl Acad. Sci. USA* **108**, 15258–15263 (2011).
91. Li, Z. W., Shen, Y. H., Xiang, Z. H. & Zhang, Z. Pathogen–origin horizontally transferred genes contribute to the evolution of Lepidopteran insects. *BMC Evol. Biol.* **11**, 356 (2011).
92. Wybouw, N. *et al.* A gene horizontally transferred from bacteria protects arthropods from host plant cyanide poisoning. *eLife* **3**, e02365 (2014).
93. Yoshida, S., Maruyama, S., Nozaki, H. & Shirasu, K. Horizontal gene transfer by the parasitic plant *Striga hermonthica*. *Science* **328**, 1128 (2010).
94. Xi, Z. *et al.* Horizontal transfer of expressed genes in a parasitic flowering plant. *BMC Genomics* **13**, 227 (2012).
95. Zhang, Y. *et al.* Evolution of a horizontally acquired legume gene, *albumin 1*, in the parasitic plant *Phelipanche aegyptiaca* and related species. *BMC Evol. Biol.* **13**, 48 (2013).
96. Zhang, D. *et al.* Root parasitic plant *Orobanche aegyptiaca* and shoot parasitic plant *Cuscuta australis* obtained Brassicaceae-specific *strictosidine synthase*-like genes by horizontal gene transfer. *BMC Plant Biol.* **14**, 19 (2014).
97. Christin, P.-A. *et al.* Adaptive evolution of C_4 photosynthesis through recurrent lateral gene transfer. *Curr. Biol.* **22**, 445–449 (2012).
98. Li, F.-W. *et al.* Horizontal transfer of an adaptive chimeric photoreceptor from bryophytes to ferns. *Proc. Natl Acad. Sci. USA* **111**, 6672–6677 (2014).
99. Zhang, H.-H., Feschotte, C., Han, M.-J. & Zhang, Z. Recurrent horizontal transfers of Chapaev transposons in diverse invertebrate and vertebrate animals. *Genome Biol. Evol.* **6**, 1375–1386 (2014).
100. Blanchard, J. L. & Lynch, M. Organellar genes. *Trends Genet.* **16**, 315–320 (2000).
101. Lynch, M. Mutation accumulation in transfer RNAs: molecular evidence for Muller's ratchet in mitochondrial genomes. *Mol. Biol. Evol.* **13**, 209–220 (1996).
102. Palmer, J. D. *et al.* Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proc. Natl Acad. Sci. USA* **97**, 6960–6966 (2000).
103. Rice, D. W. *et al.* Horizontal transfer of entire genomes via mitochondrial fusion in the angiosperm *Amborella*. *Science* **342**, 1468–1473 (2013).
This paper reports the acquisition of several mitochondrial genomes by the mitochondria in *Amborella trichopoda* a basal flowering plant.
104. Edwards, A. W. F. *Cogwheels of the Mind: The Story of Venn Diagrams* (JHU Press, 2004).
105. Morris, J. J., Lenski, R. E. & Zinser, E. R. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. *mBio* **3**, e00036-12 (2012).
This paper explains how the interplay between cheating and selection for streamlined genomes can give rise to shared genomic resources.
106. Lobkovsky, A. E., Wolf, Y. I. & Koonin, E. V. Estimation of prokaryotic supergenome size and composition from gene frequency distributions. *BMC Genomics* **15**, S14 (2014).
107. Lapierre, P. & Gogarten, J. P. Estimating the size of the bacterial pan-genome. *Trends Genet.* **25**, 107–110 (2009).
108. Puigbò, P., Lobkovsky, A. E., Kristensen, D. M., Wolf, Y. I. & Koonin, E. V. Genomes in turmoil: quantification of genome dynamics in prokaryote supergenomes. *BMC Biol.* **12**, 66 (2014).
109. Baumdicker, F., Hess, W. R. & Pfaffelhuber, P. The infinitely many genes model for the distributed genome of bacteria. *Genome Biol. Evol.* **4**, 443–456 (2012).
110. Hooper, L. V. & Gordon, J. I. Commensal host–bacterial relationships in the gut. *Science* **292**, 1115–1118 (2001).
111. Lederberg, J. & McCray, A. 'Ome sweet 'omics — a genealogical treasury of words. *Scientist* **15**, 8 (2001).
112. Becker, E, A. *et al.* Phylogenetically driven sequencing of extremely halophilic archaea reveals strategies for static and dynamic osmo-response.. *PLoS Genet.* **10**, e1004784 (2014).
113. Groussin, M *et al.* Origins of major archaeal clades do not correspond to gene acquisitions from bacteria. *BioRxiv* http://dx.doi:10.1101/019851(2015).

Acknowledgements
Work in the authors' laboratories was supported through grants from the National Science Foundation Grant (DEB 0830024), NASA Exobiology (NNX13AI03G), Binational Science Foundation (BSF 2013061), NSFC Oversea, Hong Kong, Macao collaborative grant (31328003), and the CAS/SAFEA International Partnership Program for Creative Research Teams. The authors would also like to thank K. Swithers for providing insightful comments and discussion pertaining to the body of this text.

Competing interests statement
The authors declare no competing interests.

SUPPLEMENTARY INFORMATION
See online article: S1 (table)
ALL LINKS ARE ACTIVE IN THE ONLINE PDF