



Published in final edited form as:

Trends Genet. 2011 April ; 27(4): 157–163. doi:10.1016/j.tig.2011.01.005.

Horizontal gene transfer between bacteria and animals

Julie C. Dunning Hotopp

Institute for Genome Science, Department of Microbiology & Immunology, University of Maryland School of Medicine, Baltimore, MD 21201

Abstract

Horizontal gene transfer is increasingly described between bacteria and animals. Such transfers that are vertically inherited have the potential to influence the evolution of animals. One classic example is the transfer of DNA from mitochondria and chloroplasts to the nucleus after the acquisition of these organelles by eukaryotes. Even today, many of the described instances of bacteria to animal transfer occur as part of intimate relationships like those of endosymbionts and their invertebrate hosts, particularly insects and nematodes, while numerous transfers are also found in asexual animals. Both of these observations are consistent with modern evolutionary theory, in particular the serial endosymbiotic theory and Muller's ratchet. While it is tempting to suggest that these particular lifestyles might promote horizontal gene transfer, it is difficult to ascertain given the non-random sampling of animal genome sequencing projects and the lack of a systematic analysis of animal genomes for such transfers.

Occurrence and significance of horizontal gene transfer

Horizontal gene transfer (HGT) is the transfer of DNA between diverse organisms which allows for the acquisition of novel traits that are unique from those inherited. The advent of large scale genome sequencing has greatly improved our understanding of the importance of HGT particularly among Eubacteria where most cases of genetic transfers have been described so far. For example, the phylogenetic analysis of 144 prokaryotic genomes has indicated that, although most genetic information flows vertically, genes are also frequently transferred horizontally between closely related taxa and between bacteria inhabiting the same environment [1]. HGT in Eubacteria has been implicated in the acquisition and evolution of many traits including antibiotic resistance, pathogenesis, and bioremediation [2,3].

Although less common, HGT events among eukaryotes have also been identified. In animals, cases of eukaryote-to-eukaryote HGT include the recent acquisition of P elements by *Drosophila melanogaster* from *Drosophila willistoni* [4,5], the horizontal transfer of lectin-like antifreeze proteins between fishes [6], and transfer of genes for carotenoid biosynthesis from fungi to pea aphids [7]. Recently, heritable HGT was even described in humans from the mitochondria-derived mini-circles in the eukaryote *Trypanosoma cruzi*, the causative agent of Chagas disease [8], suggesting HGT can occur in human germ cells.

While most described HGT events occur within a single domain of life and mainly involve bacteria-to-bacteria transfers, HGT between the different domains of life (Archaea,

Corresponding author: Dunning Hotopp, J. C. (jdhotopp@som.umaryland.edu).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Eubacteria and Eukaryota) has also been described. For example, the Eubacteria *Thermotoga maritima* harbors eighty-one archaeal genes clustered in fifteen 4–20 kbp islands [9]. In one of the best studied example involving Eubacteria and Eukaryota, *Agrobacterium tumefaciens* intentionally transfers 10–30 kbp of T-DNA from its Ti plasmid (200–800 kbp in size) to plants [10].

Many instances of interdomain HGT involve transfers between bacteria and eukaryotes, with bacteria-to-eukaryote HGT being more frequent than transfers from a eukaryote-to-bacteria. So far, bacteria-to-animal transfers appear to be limited to invertebrate animals and are particularly prevalent from endosymbiotic bacteria to their host and from bacteria to asexual animals. Here, I provide an overview of many recent examples of interdomain HGT between bacteria and animals (Figure 1) and discuss the role that HGT systems could play in testing evolutionary theories.

Bacteria-to-animal HGT

Bacteria-to-animal HGT can be divided into those arising from free-living bacteria and those arising from endosymbionts. Endosymbiont-to-animal HGT is likely more frequent because of the close and constant proximity of the cells from both organisms. Such transfers are likely to be even more common for endosymbionts that live in germ cells where transfers are likely to be passed on to future generations. Mitochondria and chloroplast are an example of such a scenario. These organelles arose from an alpha-proteobacteria endosymbiont and a cyanobacteria endosymbiont, respectively, and both are found in reproductive cells and are transmitted to offspring through germ cells. DNA transfers occur routinely from these organelles' genomes to the nuclear genome

Such organelle transfers have played important roles in the evolution of Eukaryotes [11,12]. One striking example of an organelle-to-eukaryote transfer is the transfer of 270 kbp (75%) of the nuclear mitochondrial insert (numt) to the *Arabidopsis thaliana* chromosome 2 centromere [13]. The chromosomal sequence is 99% identical to the mitochondrial genome, suggesting that the transfer occurred recently [13].

HGT is hypothesized to be a crucial step in the transition of endosymbionts to organelles as the “host” nucleus begins to encode for organelle proteins [14]. A HGT ratchet has been proposed to explain the accumulation of nuclear genes of organelle origin [15]. A similar ratchet has been proposed to explain the transfer of bacterial genes to the phagotrophic unicellular eukaryotes that feed on them [15]. Such ratchets may be applicable to all cases of repeated HGT where two organisms have an intimate and ongoing relationship (Figure 2).

Most cases of endosymbiont-to-eukaryote HGT described thus far involve *Wolbachia* endosymbionts. *Wolbachia* bacteria are endosymbionts that infect a wide range of arthropods and filarial nematodes, including 20–70% of insect species [16,17]. Like mitochondria, *Wolbachia* cells are intracellular, maternally inherited, and transferred through the egg cytoplasm [16,17], making them ideal donors for heritable HGT to eukaryotic genomes. I will now describe examples of each type of transfer beginning with those involving *Wolbachia* endosymbionts.

Wolbachia-to-arthropod HGT

Wolbachia-to-arthropod HGT have been described in the pea aphid (Hemiptera), mosquitoes (Diptera), a beetle (Coleoptera), fruit flies (Diptera), and parasitoid wasps (Hymenoptera). The first *Wolbachia*-host transfer described was a variant of the bean beetle *Callosobruchus chinensis* that contained *Wolbachia* genes in its chromosome [18]. Subsequently, ~30% of a *Wolbachia* genome was found on the X-chromosome, likely as the result of a single HGT

event [19]. Approximately half of the genes examined were transcribed, albeit at low levels [19]. It is not known if such transcription is biologically relevant (Box 1), but low level transcription is a trait commonly seen in bacteria-eukaryote HGT.

Box 1

Evidence for functional HGT

Finding bacterial DNA in an animal genome is not sufficient to imply that this DNA is functional in the host genome. Two relatively straightforward techniques are typically employed to assess function in an organism for genes not arising from HGT: (1) characterize the sequence and abundance of transcripts and (2) clone, express, and assay the corresponding protein. However, in the case of HGT, the relevance of this evidence can be questioned. In the absence of sufficient time to accumulate deleterious mutations, a cloned protein will maintain the same characteristics as in the bacteria where it was likely functional. But are these properties really sufficient for proper folding and activity in a eukaryotic cell?

Interpreting transcription data can be equally troublesome. Almost all of the recently described cases of bacteria-to-animal HGT describe some level of transcription. Although transcription is necessary for function; transcription does not imply function. For example, the ENCODE project found that most of the human genome is transcribed [59]. No criteria exist for defining a biologically relevant level of transcription. However, one might expect that tissue specific transcription might indicate functionality. In *A. viteae* and *O. flexuosa*, *in situ* hybridization revealed such tissue specific expression of DNA resulting from *Wolbachia* HGT. Yet, these transcripts themselves contained pseudogenes that would not be expected to encode a functional protein product [28]. In these organisms, almost half of all transcripts containing *Wolbachia* sequence contained pseudogenes making their function unclear [28].

The gold standard for determining functionality for genes resulting from HGT is a phenotype that correlates to the presence of those genes. A beautiful example of this is the transfer of carotenoid biosynthetic genes from fungi to aphids that correlates to natural red or green coloration in the aphids [7]. However, it is sometimes difficult to ascertain a phenotype. When an entire bacterial genome is transferred to an animal, where do you start?

Wolbachia transfers have now been identified in the genomes of diverse invertebrate taxa [20–23]. The HGT to *Drosophila ananassae* from its *Wolbachia* endosymbiont is the largest as the entire ~1.4 Mbp *Wolbachia* genome has been integrated into the *D. ananassae* Hawaii 2L chromosome [20]. *Wolbachia* nuclear inserts were detected in four lines of *D. ananassae* from Asia and the Pacific indicating that the insert may be widely distributed [20]. At least 28 *D. ananassae* genes of *Wolbachia* origin are transcribed [20], albeit at low levels as is seen in *C. chinensis*.

Another such transfer appears to involve the trio of parasitoid wasps, pox viruses, and obligate intracellular bacteria, including *Wolbachia* sp. and *Orientia tsutsugamushi* [23]. The wasp *Nasonia vitripennis* genome contains 13 ankyrin repeat proteins that also contain a C-terminal domain (PRANC) previously only identified in Pox viruses [23]. These proteins are found in diverse *Wolbachia* strains [23], and phylogenetic analysis of the PRANC-domain reveals that *Nasonia* wasps acquired one or more of these proteins from *Wolbachia* followed by gene duplication and divergence [23]. Most of the genes are transcribed in both males and females and through various life stages as assessed using tiling microarrays.

Eight of the eleven arthropod and nematode genomes available in 2007 that contained *Wolbachia* sequence showed evidence of having HGT between the endosymbiont genome and the host chromosomes [20]. Host-endosymbiont HGT was experimentally verified in 5 of these hosts [20]. Although the sample was small and not random with a high incidence of *Drosophila* genomes, it suggests that such nuclear inserts are common and that 70% of *Wolbachia*-infected hosts may have a nuclear insert. The incidence in filarial nematodes is likely higher, as HGT has been found in every organism examined thus far. Further examination of *Wolbachia*-to-animal HGT may provide opportunities to examine specific theories in evolution like Muller's ratchet and the serial endosymbiont theory (Box 2).

Box 2

Testing Theories of Evolution Using *Wolbachia*-Host HGT

The extensive levels of HGT between *Wolbachia* endosymbionts and their animal hosts may provide opportunities to test hypotheses and examine phenomena in evolutionary biology.

Muller's ratchet

Muller's ratchet suggests that in the absence of recombination through sexual reproduction, asexual organisms will accumulate mutations in an irreversible manner. In general, HGT is thought to be more prevalent in asexual organisms, particularly bacteria, where it can serve as a source of variation normally afforded sexual organisms through recombination. Yet, it is difficult when comparing vastly different life forms (e.g. bacteria and animals) to determine what role asexuality plays in HGT. *Wolbachia*-infected animals, in particularly those where *Wolbachia* induces asexual reproduction through parthenogenesis, can be exploited to test this hypothesis. Specifically, is HGT from the *Wolbachia* genome, as well as from the mitochondrial genome, more prevalent in exclusively asexual lineages of insects as compared to closely related lineages also infected with *Wolbachia* but that reproduce exclusively by sexual reproduction?

Role of HGT in the serial endosymbiont theory

Organelle HGT is hypothesized to be a crucial step in the transition of endosymbionts to organelles as the "host" nucleus begins to encode for organelle proteins [14]. Yet despite numerous endosymbiont relationships, this has only happened a handful of times. Is this because HGT from endosymbiont to host is now rarer than in the distant past, possibly because germ cells are better segregated? Or is it only when compartmentalization of particular tasks is necessary, as is the case for energy metabolism, that the organelle structure is maintained? Further analysis of the extent of HGT involving endosymbionts may provide insights into these possibilities and as such the serial endosymbiont theory. For instance, is HGT limited to gamete-associated endosymbionts like *Wolbachia*? Or is HGT to the host found more often with endosymbionts with reduced genomes?

Wolbachia-to-nematode HGT

Unlike insects, most filarial nematodes (Filarioidea) are thought to have an obligate mutualistic symbiosis with their *Wolbachia* strains [24–27]. Therefore, when the genome of *Brugia malayi* was sequenced, the genome of its *Wolbachia* endosymbiont was sequenced concomitantly, complicating identification of HGT. Despite stringent attempts to remove endosymbiont sequences from the worm genome, it is riddled with 249 contigs that contain *Wolbachia* sequence fragments; eight junctions of *Wolbachia*-nematode DNA were confirmed in these contigs [20].

The origin of the *Wolbachia*-filaria symbiosis is thought to have arisen about 50 million years ago in the ancestor of the subfamilies Onchocercinae and the Dirofilarinae [25–27]. Since that time there has been six independent losses of the *Wolbachia* symbiont [25–27]. Two endosymbiont-free lineages, *Acanthocheilonema viteae* (a rodent parasite) and *Onchocerca flexuosa* (a deer parasite), have evidence of numerous genes of *Wolbachia* ancestry [28] as ascertained by whole genome sequencing. This suggests that they were both once colonized with *Wolbachia* strains and that prior to the loss of the endosymbiont there was extensive bacteria-host HGT [28]. Transcription of HGT fragments was observed, including tissue specific transcription observed by *in situ* hybridization [28]. Many of the transcripts were from pseudogenes, including those examined by *in situ* hybridization [28]. As they are unable to encode the expected protein, the function of these transcripts, if any, is unclear.

Unlike in insects where *Wolbachia* is typically a facultative and often parasitic symbiont that is unnecessary for survival [16], an interesting scenario arises for co-evolving obligate mutualistic symbionts, like *Wolbachia* in filarial nematodes [24–27]. When endosymbiont genes essential for nematode survival are transferred to the nuclear chromosome through a HGT ratchet, the endosymbiont can now be lost (Figure 2). In the case of a formerly obligate endosymbiont, this would result in the presence of endosymbiont-free lineages in clades of organisms with co-evolving obligate symbionts. This is seen for filarial nematodes [25–27], although changes in the nematode biology, like a host shift that make the endosymbiont obsolete, may also explain this scenario. This type of HGT could have important medical implications, as patients infected with filarial nematodes are currently being treated with antibiotics targeting *Wolbachia* cells [29,30]. Under this scenario of HGT facilitating endosymbiont loss, antibiotic treatment could drive endosymbiont loss in these medically important filarial nematodes, although, importantly, this has not yet been observed.

Bacteria-to-animal HGT involving endosymbionts other than *Wolbachia*

Wolbachia bacteria may be the most ubiquitous of intracellular endosymbionts, but numerous endosymbiotic relationships exist in nature that may provide the opportunity for HGT. One of the best studied endosymbiotic relationships is that between the γ -proteobacteria *Buchnera aphidicola* and its host, the pea aphid *Acyrtosiphon pisum*. Aphids house *B. aphidicola* in the cytoplasm of specialized cells called bacteriocytes [31]. These obligate mutualists have co-evolved to depend on one another for survival [31]. At 416-642 kbp, *B. aphidicola* genomes are highly reduced in size lacking the ability to encode proteins that are essential in other bacteria [31]. This genome reduction is analogous to that seen in organelles. Therefore, one might anticipate significant HGT from the *Buchnera* genome to the aphid genome as has been seen for organelles. In addition to the obligate mutualistic symbiosis the aphids have with *B. aphidicola*, they also have a facultative symbiosis with a *Wolbachia* endosymbiont [31]. This provides an opportunity to compare the levels of HGT between an obligate mutualist with a severely reduced genome to a facultative symbiont with a significantly larger genome.

An exhaustive analysis of HGT in the pea aphid, *Acyrtosiphon pisum*, revealed only twelve genes or gene fragments of bacterial origin [22]. Only two of these are from *B. aphidicola*, and both are truncated with no evidence for transcription [22]. Five of the genes are most closely related to the rickettsial bacteria *Wolbachia* sp. and *Orientia tsutsugamushi* [22]. The ancestry of the remaining five genes, which resulted from a single HGT followed by duplication, is not resolved [22]. While one might anticipate that more *Buchnera* genes would be transferred, given the obligate symbiosis and reduced genome as discussed above, this does not appear to be the case. It is not clear why functional *Buchnera* genes are not found in the nuclear genome, particularly those for amino acid biosynthesis. This may

merely reflect the opportunity each endosymbiont has for HGT in host germ cells. While *Wolbachia* cells are known to be at a high density in the gametes, *B. aphidicola* cells are largely confined to bacteriocytes and are only found in germ cells when being transmitted to offspring [31]. As such *B. aphidicola* may not have enough opportunity to transfer its genes to its host genome.

Interestingly, one of the gene transfers of rickettsial origin, *ldcA*, is transcribed in the bacteriocyte and is involved in bacterial cell wall metabolism [22]. This protein is essential in *Escherichia coli*, preventing lysis in stationary phase [32]. *B. aphidicola* has the components to synthesize the bacterial cell wall, but lacks *ldcA* [32]. Presumably expression of this nuclear-encoded gene in the bacteriocyte allows aphids to control cell wall metabolism and thus growth of this organism. This raises an intriguing scenario of aphid proteins of α -proteobacteria origin controlling growth of a γ -proteobacteria via cell wall synthesis/degradation by the aphid [32].

Other bacteria-to-animal HGT events

Bacteria-animal HGT is not limited to endosymbiont interactions. *Hydra* are simple freshwater animals that can reproduce asexually by budding. The *Hydra magnipapillata* genome contains 71 candidates for HGT that show a closer relation to bacterial genes than metazoan genes, and 70% of these are supported by ESTs [33]. Many of these 71 genes are strong candidates for HGT as they only have bacterial homologs [33]. In bdelloid rotifers, which are also small freshwater asexual animals, DNA from bacteria, fungi, and plants appear to be concentrated in telomeric regions along with mobile genetic elements [34]. While some of the genes resulting from HGT are not functional, others are transcribed. Two genes for bacterial peptidoglycan synthesis are transcribed with splicing of introns not present in bacteria [34]. One of these proteins was overexpressed in *E. coli* and found to be a functional enzyme [34].

These two asexual animal genomes show evidence for extensive HGT from diverse organisms including bacteria. This may suggest a correlation between asexuality and the presence and/or persistence of DNA resulting from HGT. This is consistent with Muller's ratchet, which suggests that in the absence of recombination through sexual reproduction, asexual organisms will accumulate mutations in an irreversible manner [35-37] including DNA from HGT. The contribution of asexuality to HGT in animals can be examined using *Wolbachia* and the reproductive phenotypes it induces, which include asexuality through parthenogenesis (Box 2).

Numerous HGT events have also occurred during the evolution of plant parasitic nematodes. It is thought that such transfers allowed the free-living ancestors of these nematodes to exploit plants as a new ecological niche. Sixty genes for six protein families in the *Meloidogyne incognita* genome were found to be the result of HGT from diverse bacterial sources followed by subsequent gene duplication and expansion [38]. These include cellulases, pectate lyases, and expansin-like proteins not typically found in other animals but that are involved in degrading or modifying the plant cell wall [38]. The proteins have been biochemically characterized, are secreted into plant tissues, and are involved in parasitism [38]. Another plant parasitic nematode, *Heterodera glycines*, contains a biosynthetic pathway for vitamin B6 of bacterial origin [39]. Despite the apparent prokaryotic origin of these genes, they contain introns, have polyadenylated transcripts, and map to the nematode chromosome [39]. Phylogenetic analysis places them closest to *Protochlamydia amoebophila*, an endosymbiont of amoeba [39].

Given that there are many more bacterial cells relative to eukaryotic ones and that eukaryotic cells are frequently in proximity to bacterial cells, it is not surprising that HGT is

increasingly described between bacteria and animals. For example, the human body contains more bacterial cells than human cells (for review see [40]), and many of these bacteria are in intimate contact with human mucosal cells. In addition, many of these same bacteria have mechanisms like the Type IV secretion system to transfer DNA to the environment and to other organisms. Therefore, there are many opportunities and avenues for HGT from bacteria to animals. This type of HGT is of interest as it has the potential to provide novel functions to animals, allow adaptation to novel niches, and affect their evolution. Such transfers from the human microbiome also have the potential to induce somatic cell mutations, and hence be important to human health (Box 3).

Box 3

Bacteria-to-human HGT and the microbiome

During the original public effort to sequence the human genome, an analysis of the inheritance patterns of exons [57] refuted a previous analysis describing 223 human proteins arising from bacterial HGT [55]. Yet, in this analysis forty genes were found to be shared exclusively between bacteria and humans [57], as would genes resulting from bacteria HGT. A subsequent phylogenetic analysis of human proteins for HGT [60] overlooked comparisons with prokaryotes. Additionally, HGT of units smaller than a gene [61] would be obscured in both analyses. All of these points might suggest further analysis is warranted that does not focus solely on whole exons or genes and makes use of the many bacterial genomes sequenced subsequently.

Overall, it seems clear that inherited HGT is not ubiquitous from bacteria to humans. However, with the enormous size of the human microbiome, is there an equally enormous opportunity for HGT to somatic cells that could have profound effects in human health and disease? Such somatic cell mutations would largely be overlooked in whole genome assemblies.

The American Cancer Society reports that 15–20% of cancers worldwide and 10% of cancers in developed countries are linked to viral, bacterial, or parasite infections. Although chronic inflammation is increasingly implicated as a mechanism for cancer development (for review see [62]), proto-oncogene disruption by bacterial DNA could provide yet another mechanism. Another avenue for disease following HGT would be through the transcription and expression of bacterial DNA. *Agrobacterium tumefaciens* induces the production of a plant tumor through the expression and transcription of bacterial DNA [10,63,64]. It, therefore, seems plausible that expression and transcription of bacterial DNA could induce tumors in animals.

Although disadvantageous outcomes of HGT are easy to conceptualize, is there room for advantageous outcomes? Bacteria in the digestive tract are associated with the fermentation of non-digestible dietary residues and with the production of vitamin K (for a review see [65]). Could HGT in our intestinal cells allow gain of these properties? With the decreasing costs of human genome sequencing, we may soon be able to compare the DNA content of individual cells in frequent contact with bacteria to those from more sterile portions of body to address these highly speculative questions.

Eukaryote-to-bacteria HGT

While bacteria-to-eukaryote HGT has garnered attention recently, fewer eukaryote-to-bacteria transfers have been described. These examples include bacteria belonging to diverse phyla. For example, *Legionella pneumophila* has been found to encode more than 100 eukaryotic-like proteins [41] with phylogenetic analysis revealing that 29 are of eukaryotic ancestry [41]. The cyanobacterium *Microcystis aeruginosa* has a shell-like structure thought

to allow it to adapt to frequent salinity changes that is encoded by the eukaryotic proteins actin and the actin-binding protein profilin [42]. The ancient transfer of two tubulin genes and a kinesin light chain gene to *Prostheobacter dejongeii* has also been described [43,44]. Metagenomic analysis of Antarctic plankton identified a δ -proteobacteria with a homologue to Cys-rich eukaryotic proteins [45]. Lastly, some *Wolbachia* strains [46,47] and numerous mosquito species share phylogenetically related genes that encode salivary gland surface (SGS) proteins in mosquitoes [48,49]. These genes lack introns, are evolving under purifying selection, and are transcribed in both mosquitoes and *Wolbachia* strains [50,51] with HGT quite possibly happening in both directions.

It is expected that animal-to-bacteria HGT would be more prevalent as bacteria readily participate as both donors and recipients in HGT and they lack the barriers to HGT that animals have, like a nucleus and segregated germ cells. In an ad hoc search of Genbank, potential candidates of eukaryote-bacteria HGT can be identified that are not described in the literature and that need further robust phylogenetic investigation. However, they are by no means ubiquitous. One possible barrier to HGT from eukaryotes may be the presence of spliceosomal introns in eukaryotic genes. However, transfer of single exon genes would still be possible or transfer could be mediated through a spliced RNA intermediate in a mechanism analogous to retrotransposition. Alternatively, it has been proposed that eukaryotes may just not offer novel functionalities that bacteria want and that they cannot obtain elsewhere [52]. In that case, HGT may happen but not provide a selective advantage or otherwise become maintained in the genome facilitating its detection.

Concluding remarks and future perspectives

Nature has lots of chances for making even the most improbable events happen. It is likely that many HGT events described merely reflect a random occurrence happening at a low frequency in the population. At some other low frequencies these transfers may become transcribed, the proteins expressed, and a phenotype manifest. Those genes that have a function and can be properly regulated will be maintained in the genome while the rest will be lost over time. One challenge in future research on this topic will be to find a suitable model system to explore all of the stages of this process. Thus far the field has mainly focused on describing such transfers and attempting to ascertain a function. There is also great interest in the mechanism by which the DNA is transferred out of the bacteria, into the nucleus, incorporated into eukaryotic chromosomes, and develops a function as this has many practical implications in biotechnology, medicine, and evolution.

However, there continue to be several barriers to description of HGT in the literature, including some limitations in the genome sequencing data (Box 4). As the prevailing paradigm is to assume vertical inheritance when other evidence is lacking, HGT in eukaryotes will continue to be overlooked and under-estimated. This is necessary in order to avoid HGT being described inappropriately. However, this notion, as well as high profile erroneous reports of HGT in humans and other animals (e.g.[53–57]), has likely had a chilling effect on the field.

Box 4

Problems identifying bacteria-to-animal HGT in existing genome data

Exacerbating other barriers that limit the description of HGT in the literature, prokaryotic sequences are routinely removed from eukaryotic genome projects without description of these methods or experimental validation. For example, the *D. ananassae* genome assembly in Genbank lacks significant *Wolbachia* sequence, despite the presence of an entire *Wolbachia* genome in an autosome. Compounding this problem, the method of

removal of the prokaryotic sequences is often not described in the genome articles. *D. ananassae* is just one example, as this is ubiquitous. For example, an analysis of structural variation in humans proceeded after removing common sources of contamination including bacterial sequences [66]. Although removal of *E. coli* sequences is justified and necessary for cloned libraries, is there justification to remove all bacterial sequences? A circular argument ensues where eukaryotic genome sequencing projects remove all bacterial sequences because animals do not have HGT from bacteria, and examination of those same genomes for HGT reinforces the notion that HGT from bacteria to animals does not occur. Resources like the Sequence Read Archive and Trace Archive offer opportunities for assembly-independent analyses, but only when the raw experimental data is deposited before “contaminant” removal. In addition, while journals require submission of genome assemblies and annotation, the raw data is often overlooked, and exactly what is deposited can be unclear.

As nematodes and insects are among the most abundant animals on Earth, and both seem to have significant levels of interdomain HGT from bacteria, the view that HGT is unimportant in animal evolution needs to be re-evaluated. It is not clear what levels of HGT will be found in other animal genomes, particularly vertebrate ones. The recognition of extensive HGT from bacteria to insects and nematodes may merely reflect the interests of researchers investigating this phenomenon. For example, the *Wolbachia* research community embraced HGT from endosymbionts to hosts quite quickly, possibly due to a long standing interest in the serial endosymbiosis theory. In contrast, the increased incidence of HGT described in nematodes and insects may reflect a higher abundance of such transfers in these organisms. They are phylogenetically closely related phyla, and as such may share common features that accommodate or facilitate HGT. Alternatively, the HGT identified in arthropods and nematodes may reflect an abundance of sequence data. Next to humans and other mammals, arthropods and nematodes are among the best represented animals in terms of genome sequencing projects. Although next-generation sequencing has revolutionized microbial whole genome sequencing and eukaryote re-sequencing, it has done less to advance genomics of organisms that lack a closely related representative genome sequence. While mammals are well sequenced, other animals (e.g. birds, reptiles, amphibians, and fish) are poorly represented by genome sequencing projects. Hopefully, the promises of third generation single molecule sequencing will yield a more diverse sampling of vertebrate genomes, facilitate the differentiation of contaminants from HGT, and improve our understanding of the extent of HGT events in a wide array of diverse organisms.

Acknowledgments

We would like thank Gustavo C. Cerqueira and Michael Clark for their thoughtful discussions on trypanosomes and Muller’s ratchet, respectively. This work was funded by the National Institutes of Health through the NIH Director’s New Innovator Award Program (1-DP2-OD007372), the NSF Microbial Sequencing Program (EF-0826732), and the Bill & Melinda Gates Foundation (OPP1015708).

BIBLIOGRAPHY

1. Beiko RG, et al. Highways of Gene Sharing in Prokaryotes. *Proc Natl Acad Sci U S A*. 2005; 102:14332–14337. [PubMed: 16176988]
2. Boucher Y, et al. Lateral Gene Transfer and the Origins of Prokaryotic Groups. *Annu Rev Genet*. 2003; 37:283–328. [PubMed: 14616063]
3. de la Cruz F, Davies J. Horizontal Gene Transfer and the Origin of Species: Lessons from Bacteria. *Trends Microbiol*. 2000; 8:128–133. [PubMed: 10707066]
4. Daniels SB, et al. Evidence for Horizontal Transmission of the P Transposable Element between *Drosophila* Species. *Genetics*. 1990; 124:339–355. [PubMed: 2155157]

5. Engels WR. Invasions of P Elements. *Genetics*. 1997; 145:11–15. [PubMed: 9017385]
6. Graham LA, et al. Lateral Transfer of a Lectin-Like Antifreeze Protein Gene in Fishes. *PLoS One*. 2008; 3:e2616. [PubMed: 18612417]
7. Moran NA, Jarvik T. Lateral Transfer of Genes from Fungi Underlies Carotenoid Production in Aphids. *Science*. 2010; 328:624–627. [PubMed: 20431015]
8. Hecht MM, et al. Inheritance of DNA Transferred from American Trypanosomes to Human Hosts. *PLoS One*. 2010; 5:e9181. [PubMed: 20169193]
9. Nelson KE, et al. Evidence for Lateral Gene Transfer between Archaea and Bacteria from Genome Sequence of *Thermotoga maritima*. *Nature*. 1999; 399:323–329. [PubMed: 10360571]
10. Gelvin SB. *Agrobacterium*-Mediated Plant Transformation: The Biology Behind the "Gene-Jockeying" Tool. *Microbiol Mol Biol Rev*. 2003; 67:16–37. [PubMed: 12626681]
11. Gray MW, et al. The Origin and Early Evolution of Mitochondria. *Genome Biol*. 2001; 2:REVIEWS1018. [PubMed: 11423013]
12. Richly E, Leister D. Numts in Sequenced Eukaryotic Genomes. *Mol Biol Evol*. 2004; 21:1081–1084. [PubMed: 15014143]
13. Lin X, et al. Sequence and Analysis of Chromosome 2 of the Plant *Arabidopsis thaliana*. *Nature*. 1999; 402:761–768. [PubMed: 10617197]
14. Margulis, L. Origin of Eukaryotic Cells. Yale University Press; 1970.
15. Doolittle WF. You Are What You Eat: A Gene Transfer Ratchet Could Account for Bacterial Genes in Eukaryotic Nuclear Genomes. *Trends Genet*. 1998; 14:307–311. [PubMed: 9724962]
16. Stouthamer R, et al. *Wolbachia Pipientis*: Microbial Manipulator of Arthropod Reproduction. *Annu Rev Microbiol*. 1999; 53:71–102. [PubMed: 10547686]
17. Werren JH. Biology of *Wolbachia*. *Annu Rev Entomol*. 1997; 42:587–609. [PubMed: 15012323]
18. Kondo N, et al. Genome Fragment of *Wolbachia* Endosymbiont Transferred to X Chromosome of Host Insect. *Proc Natl Acad Sci U S A*. 2002; 99:14280–14285. [PubMed: 12386340]
19. Nikoh N, et al. *Wolbachia* Genome Integrated in an Insect Chromosome: Evolution and Fate of Laterally Transferred Endosymbiont Genes. *Genome Res*. 2008; 18:272–280. [PubMed: 18073380]
20. Dunning Hotopp JC, et al. Widespread Lateral Gene Transfer from Intracellular Bacteria to Multicellular Eukaryotes. *Science*. 2007; 317:1753–1756. [PubMed: 17761848]
21. Fenn K, et al. Phylogenetic Relationships of the *Wolbachia* of Nematodes and Arthropods. *PLoS Pathog*. 2006; 2:e94. [PubMed: 17040125]
22. Nikoh N, et al. Bacterial Genes in the Aphid Genome: Absence of Functional Gene Transfer from *Buchnera* to Its Host. *PLoS Genet*. 2010; 6:e1000827. [PubMed: 20195500]
23. Werren JH, et al. Functional and Evolutionary Insights from the Genomes of Three Parasitoid *Nasonia* Species. *Science*. 2010; 327:343–348. [PubMed: 20075255]
24. McGarry HF, et al. Population Dynamics of *Wolbachia* Bacterial Endosymbionts in *Brugia Malayi*. *Mol Biochem Parasitol*. 2004; 135:57–67. [PubMed: 15287587]
25. Taylor MJ, et al. *Wolbachia* Bacterial Endosymbionts of Filarial Nematodes. *Adv Parasitol*. 2005; 60:245–284. [PubMed: 16230105]
26. Casiraghi M, et al. Mapping the Presence of *Wolbachia pipientis* on the Phylogeny of Filarial Nematodes: Evidence for Symbiont Loss During Evolution. *Int J Parasitol*. 2004; 34:191–203. [PubMed: 15037105]
27. Fenn K, Blaxter M. Are Filarial Nematode *Wolbachia* Obligate Mutualist Symbionts? *Trends Ecol Evol*. 2004; 19:163–166. [PubMed: 16701248]
28. McNulty SN, et al. Endosymbiont DNA in Endobacteria-Free Filarial Nematodes Indicates Ancient Horizontal Genetic Transfer. *PLoS One*. 2010; 5:e11029. [PubMed: 20543958]
29. Debrah AY, et al. Doxycycline Reduces Plasma Vegf-C/Svegfr-3 and Improves Pathology in Lymphatic Filariasis. *PLoS Pathog*. 2006; 2:e92. [PubMed: 17044733]
30. Turner JD, et al. Macrofilaricidal Activity after Doxycycline Only Treatment of *Onchocerca volvulus* in an Area of *Loa loa* Co-Endemicity: A Randomized Controlled Trial. *PLoS Negl Trop Dis*. 2010; 4:e660. [PubMed: 20405054]

31. Brinza L, et al. Systemic Analysis of the Symbiotic Function of *Buchnera aphidicola*, the Primary Endosymbiont of the Pea Aphid *Acyrtosiphon pisum*. *C R Biol.* 2009; 332:1034–1049. [PubMed: 19909925]
32. Nikoh N, Nakabachi A. Aphids Acquired Symbiotic Genes Via Lateral Gene Transfer. *BMC Biol.* 2009; 7:12. [PubMed: 19284544]
33. Chapman JA, et al. The Dynamic Genome of Hydra. *Nature.* 2010; 464:592–596. [PubMed: 20228792]
34. Gladyshev EA, et al. Massive Horizontal Gene Transfer in Bdelloid Rotifers. *Science.* 2008; 320:1210–1213. [PubMed: 18511688]
35. Felsenstein J. The Evolutionary Advantage of Recombination. *Genetics.* 1974; 78:737–756. [PubMed: 4448362]
36. Muller HJ. Some Genetic Aspects of Sex. *American Naturalist.* 1932; 66:118–138.
37. Muller HJ. The Relation of Recombination to Mutational Advance. *Mutat Res.* 1964; 106:2–9. [PubMed: 14195748]
38. Danchin EGJ, et al. Multiple Lateral Gene Transfers and Duplications Have Promoted Plant Parasitism Ability in Nematodes. *Proc Natl Acad Sci U S A.* 2010; 107:17651–17656. [PubMed: 20876108]
39. Craig JP, et al. Analysis of a Horizontally Transferred Pathway Involved in Vitamin B-6 Biosynthesis from the Soybean Cyst Nematode *Heterodera glycines*. *Molecular Biology and Evolution.* 2008; 25:2085–2098. [PubMed: 18586696]
40. Robinson CJ, et al. From Structure to Function: The Ecology of Host-Associated Microbial Communities. *Microbiol Mol Biol Rev.* 2010; 74:453–476. [PubMed: 20805407]
41. Lurie-Weinberger MN, et al. The Origins of Eukaryotic-Like Proteins in *Legionella pneumophila*. *Int J Med Microbiol.* 2010; 300:470–481. [PubMed: 20537944]
42. Guljamow A, et al. Horizontal Gene Transfer of Two Cytoskeletal Elements from a Eukaryote to a Cyanobacterium. *Curr Biol.* 2007; 17:R757–759. [PubMed: 17803924]
43. Jenkins C, et al. Genes for the Cytoskeletal Protein Tubulin in the Bacterial Genus *Prostheco bacter*. *Proc Natl Acad Sci U S A.* 2002; 99:17049–17054. [PubMed: 12486237]
44. Schlieper D, et al. Structure of Bacterial Tubulin Btuba/B: Evidence for Horizontal Gene Transfer. *Proc Natl Acad Sci U S A.* 2005; 102:9170–9175. [PubMed: 15967998]
45. Moreira D, et al. Metagenomic Analysis of Mesopelagic Antarctic Plankton Reveals a Novel Deltaproteobacterial Group. *Microbiology.* 2006; 152:505–517. [PubMed: 16436438]
46. Ishmael N, et al. Extensive Genomic Diversity of Closely Related *Wolbachia* Strains. *Microbiology.* 2009; 155:2211–2222. [PubMed: 19389774]
47. Iturbe-Ormaetxe I, et al. Distribution, Expression, and Motif Variability of Ankyrin Domain Genes in *Wolbachia pipientis*. *J Bacteriol.* 2005; 187:5136–5145. [PubMed: 16030207]
48. Arca B, et al. An Updated Catalogue of Salivary Gland Transcripts in the Adult Female Mosquito, *Anopheles gambiae*. *J Exp Biol.* 2005; 208:3971–3986. [PubMed: 16215223]
49. Korochkina S, et al. A Mosquito-Specific Protein Family Includes Candidate Receptors for Malaria Sporozoite Invasion of Salivary Glands. *Cell Microbiol.* 2006; 8:163–175. [PubMed: 16367875]
50. Klasson L, et al. Horizontal Gene Transfer between *Wolbachia* and the Mosquito *Aedes aegypti*. *BMC Genomics.* 2009; 10:33. [PubMed: 19154594]
51. Woolfit M, et al. An Ancient Horizontal Gene Transfer between Mosquito and the Endosymbiotic Bacterium *Wolbachia pipientis*. *Mol Biol Evol.* 2009; 26:367–374. [PubMed: 18988686]
52. Keeling PJ, Palmer JD. Horizontal Gene Transfer in Eukaryotic Evolution. *Nat Rev Genet.* 2008; 9:605–618. [PubMed: 18591983]
53. Heritable Integration of Kdna Minicircle Sequences from *Trypanosoma cruzi* into the Avian Genome: Insights into Human Chagas Disease. *Cell.* 2005; 122:969. [PubMed: 16220614]
54. Grunau C, Boissier J. No Evidence for Lateral Gene Transfer between Salmonids and Schistosomes. *Nat Genet.* 2010; 42:918–919. [PubMed: 20980981]
55. Lander ES, et al. Initial Sequencing and Analysis of the Human Genome. *Nature.* 2001; 409:860–921. [PubMed: 11237011]

56. Melamed P, et al. Evidence for Lateral Gene Transfer from Salmonids to Two Schistosome Species. *Nat Genet.* 2004; 36:786–787. [PubMed: 15284843]
57. Salzberg SL, et al. Microbial Genes in the Human Genome: Lateral Transfer or Gene Loss? *Science.* 2001; 292:1903–1906. [PubMed: 11358996]
58. Hedges SB. The Origin and Evolution of Model Organisms. *Nat Rev Genet.* 2002; 3:838–849. [PubMed: 12415314]
59. Birney E, et al. Identification and Analysis of Functional Elements in 1% of the Human Genome by the Encode Pilot Project. *Nature.* 2007; 447:799–816. [PubMed: 17571346]
60. Huerta-Cepas J, et al. The Human Phylome. *Genome Biol.* 2007; 8:R109. [PubMed: 17567924]
61. Chan CX, et al. Are Protein Domains Modules of Lateral Genetic Transfer? *PLoS One.* 2009; 4:e4524. [PubMed: 19229333]
62. Chang AH, Parsonnet J. Role of Bacteria in Oncogenesis. *Clin Microbiol Rev.* 2010; 23:837–857. [PubMed: 20930075]
63. Gelvin SB. *Agrobacterium* and Plant Genes Involved in T-DNA Transfer and Integration. *Annu Rev Plant Physiol Plant Mol Biol.* 2000; 51:223–256. [PubMed: 15012192]
64. Tzfira T, et al. Nucleic Acid Transport in Plant-Microbe Interactions: The Molecules That Walk through the Walls. *Annu Rev Microbiol.* 2000; 54:187–219. [PubMed: 11018128]
65. Guarner F, Malagelada JR. Gut Flora in Health and Disease. *Lancet.* 2003; 361:512–519. [PubMed: 12583961]
66. Kidd JM, et al. Mapping and Sequencing of Structural Variation from Eight Human Genomes. *Nature.* 2008; 453:56–64. [PubMed: 18451855]



Figure 1. Illustration of HGT from bacteria to animals described in the main text

HGT events from bacteria to animals are illustrated by blue arrows with the width representing the abundance of transfers. Organelle transfers (chloroplast, green; mitochondria, red) are shown at the base of the appropriate lineage. Transfers from bacteria that cannot be attributed to a specific bacterial lineage are shown as arising from the middle of the Eubacteria branch. Other types of HGT are not shown, including eukaryote-to-bacteria transfers and transfers from bacteria to eukaryotes other than animals. The tree itself is used only for illustration purposes and does not accurately reflect a true phylogeny. The phylogeny of the eukaryotic and eubacterial lineages as well as issues related to resolving particular lineages is described elsewhere (e.g. [58]).

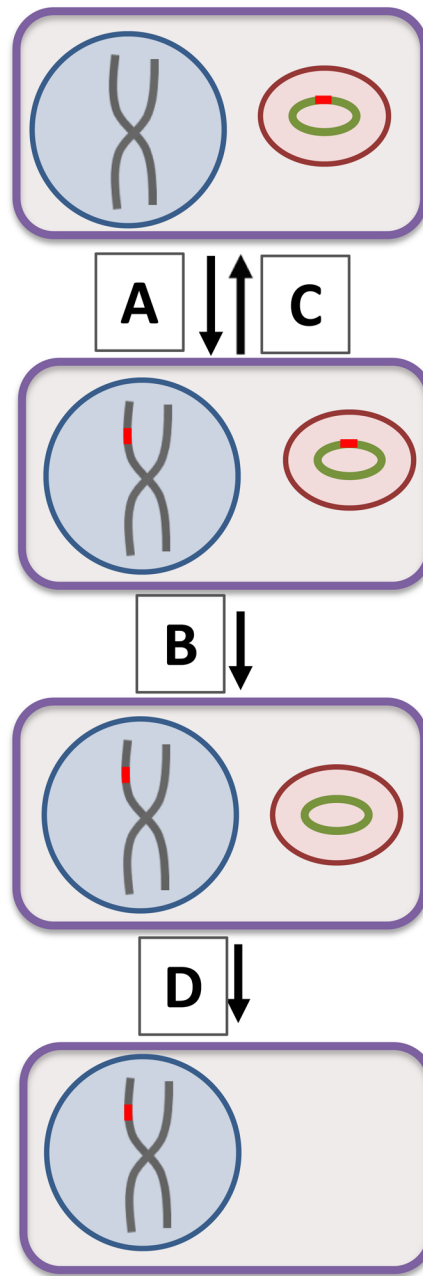


Figure 2. Endosymbiont HGT ratchet

A HGT ratchet has been proposed [15] to explain the accumulation of nuclear genes of organelle origin. This model can be extended to include HGT involving endosymbionts. At low frequency, a gene (red) is transferred from the endosymbiont genome (green circular chromosome) to the nuclear genome (gray linear chromosomes) (A). With some even lower frequency, such transfers will occur in a manner which will allow the gene to be functional in its new location. Subsequently, either the nuclear or endosymbiont version will be lost. Loss of the endosymbiont version is nearly irreversible and as such the gene becomes fixed in the nuclear genome (B). Loss of the nuclear version returns the endosymbiont to its original state and the process can repeat itself (C). The result is that over time, all of the genes that can insert into the nuclear genome will do so. When maintenance of

endosymbiont structure is not essential and the HGT now fulfills all of the needs of the host, the endosymbiont becomes obsolete and can also be lost (D).