



Published in final edited form as:

Exp Cell Res. 2017 September 15; 358(2): 421–426. doi:10.1016/j.yexcr.2017.02.009.

Lateral gene transfer between prokaryotes and eukaryotes

Karsten B. Sieber¹, Robin E. Bromley¹, and Julie C. Dunning Hotopp^{1,2,3,*}

¹Institute for Genome Science, University of Maryland School of Medicine, Baltimore, MD 21201, USA

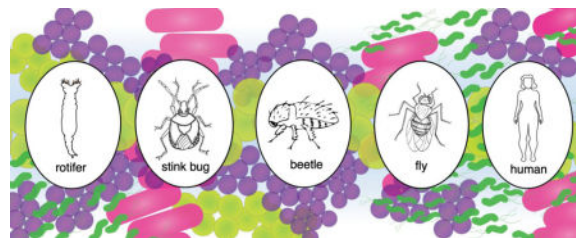
²Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, MD 21201, USA

³Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, MD 21201, USA

Abstract

Lateral gene transfer (LGT) is an all-encompassing term for the movement of DNA between diverse organisms. LGT is synonymous with horizontal gene transfer, and the terms are used interchangeably throughout the scientific literature. While LGT has been recognized within the bacteria domain of life for decades, inter-domain LGTs are being increasingly described. LGTs between bacteria and complex multicellular organisms are of interest because they challenge the long-held dogma that such transfers could only occur in closely-related, single-celled organisms. Scientists will continue to challenge our understanding of LGT as we sequence more, diverse organisms, as we sequence more endosymbiont-colonized arthropods, and as we continue to appreciate LGT events, both young and old.

Graphical abstract



Keywords

Lateral gene transfer; horizontal gene transfer; antibiotic resistance; serial endosymbiosis theory; genomics

*Corresponding author: Address: 801 W. Baltimore St., Institute for Genome Science, University of Maryland School of Medicine, Baltimore, MD 21201, USA. 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.

Lateral gene transfer as a driving evolutionary force

Sexual reproduction is considered an evolutionary advantage because offspring have increased genetic diversity. Given that bacteria reproduce asexually, bacterial offspring lack genetic diversity from the sexual reproduction of two parents. In the absence of sexual reproduction, the transfer of DNA between organisms independent of sexual reproduction via lateral gene transfer (LGT) enables bacteria to increase genetic diversity and therefore potentially increase evolutionary fitness.

Over time, novel genotypes and phenotypes arise in all organisms through a gradual process of sequential *de novo* mutations that gain prevalence through selection [1]. LGT accelerates this process through a rapid introduction of genetic diversity within a single generation, whereby a donor organism transfers a gene encoding a novel trait, or multiple traits, to a recipient organism in a single event [1]. The concept of LGT was first described by Frederick Griffith in 1928 when he demonstrated that heat-killed virulent *Streptococcus pneumoniae* were able to transfer an unknown factor to live non-virulent *S. pneumoniae*, and this unknown factor conferred virulence [2]. It was not until 1944 that Avery, MacLeod, and McCarty demonstrated that DNA was transforming factor described by Griffith [3].

The ability of LGT to act as a driving evolutionary force is epitomized by the rapid spread of antibiotic resistance genes. Between 1930 and 1945, the first three classes of antibiotics were being used therapeutically and ushered in a new era of modern medicine with the ability to treat life-threatening infections. By 1955, strains of multidrug resistance bacteria were reported [4]. It became apparent that the rate at which bacteria were obtaining resistance to these antibiotics was quicker than the expected rate of *de novo* mutations [5]. By 1960, it was shown that bacteria transferred antibiotic resistance through LGT (for review: [4, 6]). The use of antibiotics placed a strong selective pressure on bacteria to propagate the antibiotic resistance genes, and LGT enabled the bacteria to quickly respond to the selective pressure and propagate the antibiotic resistance genes throughout bacterial populations. More recently, bacteria have acquired deadly combinations of antibiotic resistance genes via LGT, such as vancomycin-resistant, methicillin-resistant *Staphylococcus aureus* [7] (Figure 1).

Originally, LGT was thought to occur primarily between closely related bacterial species through three primary mechanisms: transformation, transduction, and conjugation. Transformation describes the ability of some cells to acquire foreign DNA from the environment outside the cell, and potentially, incorporate it into the genome of the cell. Transduction occurs when a phage incorporates into the genome. Lastly, conjugation requires cell-to-cell contact for a donor cell to transfer DNA to a recipient cell. It was thought that closely related bacteria have more compatible systems for conjugation, higher potential success rate for homologous recombination, and similar codon usage [5, 8]. However, evidence has accumulated that demonstrates that distantly related bacteria can exchange DNA [9–12], demonstrating that LGT is a widespread evolutionary driving force.

Bacteria have even acquired genetic material from the human genome. A 685-bp fragment with 98–100% identity to the human L1 element was identified in 11% of *Neisseria*

gonorrhoeae strains [13]. This was specific to *N. gonorrhoeae* and not found in closely related *Neisseria meningitidis* or other commensal *Neisseria* isolates [13]. This integration is proposed to have occurred relatively recently via non-homologous end joining [13]. The integrated DNA is transcribed but a consistent difference in phenotype could not be found between strains with and without this LGT.

LGT from prokaryotes to eukaryotes

Until recently, the evolutionary impact of LGT from prokaryote donors to eukaryote recipients was less clear. With the recent development of sequencing technologies that have led to decreasing sequencing costs and of bioinformatic technologies that enable detection of LGT, the number of identified LGTs from prokaryotes to eukaryotes has increased dramatically in the past 10 years.

The most widespread instances of LGT from bacteria to the eukaryotes are the nuclear acquisitions of genes from the mitochondria and chloroplast organelles. These eukaryotic organelles originated from α -proteobacteria and Cyanobacteria, respectively [14]. Inside the cell cytoplasm, in proximity to the nucleus, these organelles have the relatively uncommon opportunity to be poised to transfer DNA to the nuclear eukaryotic genome and be inherited by future generations of cells.

Like organelles, some bacteria are intracellular, residing within cells of the eukaryotic host. These eukaryotic hosts range from single cell organisms to multicellular eukaryotic plants and animals. The bacterial endosymbiont *Wolbachia pipientis* colonizes a wide variety insects and select nematodes. Some estimates suggest that 70% of these hosts contain LGT from *Wolbachia* [15]. In the case of *Drosophila ananassae*, multiple copies of the entire 1.4 Mbp *Wolbachia* genome has been transferred to the *Drosophila* genome [16, 17]. However, the functional consequences of these *Wolbachia* LGTs, if any, remain unclear.

LGT in eukaryotes is not limited to organelles or endosymbionts. The bdelloid rotifer has extensive LGT in the telomeric regions from bacteria, fungi, and plants [18]. Specifically, ten protein-coding sequences were identified as putative LGTs. Interestingly, three of the bacterial coding sequences have spliceosomal introns [18]. A bacterial IS5-like DNA transposon has also been identified in the telomeric region of the rotifer [19]. The IS5-like transposon integration has only one copy in the haploid genome, suggesting that it was unable to further mobilize after the original integration event [19].

The coffee berry borer beetle, *Hypothenemus hampei*, has a LGT that is functional, essential, and is thought to have enabled the beetle to adapt to a new niche. The primary food source of *H. hampei* is the coffee berry, which stores carbohydrates as galactomannan [20]. The bacterial *HhMANI* gene that hydrolyzes the breakdown of galactomannan has been transferred to the beetle via LGT. This LGT is specific to *H. hampei* since close relatives do not have the *HhMANI* gene and are unable to colonize the coffee berries [20]. This class of enzyme was previously not found in any insect [20], although subsequently a putative analogous LGT was also proposed to be important in the brown marmorated stink bug [21].

Bacteria can also use LGT to create an advantageous niche and food source for their own use. *Agrobacterium tumefaciens* uses a type IV secretion system to inject bacterial proteins and its tumor inducing (Ti) plasmid into plant cells [22, 23]. Once inside the plant cells, the bacterial proteins use the plant cell machinery to transport the Ti plasmid inside the nucleus. Once inside the nucleus, through illegitimate recombination, the Ti plasmid integrates into the plant genome [22, 23]. The integrated plasmid then uses eukaryotic promoter sequences to express bacterial proteins that transform the plant cell to produce a specific carbon source for *A. tumefaciens* [22, 23]. As a result of the plant transformation, the plant develops tumor-like growths, characteristic of crown gall disease, where the bacteria grow and thrive [22, 23].

Most examples of LGT in eukaryotes involve the relatively straightforward transfer of a single gene or pathway from a single donor to a single recipient. In contrast, the *Planococcus citri* mealybug is an example of complex LGT biology [24]. Many insects in the order Hemiptera, like the mealybug, rely on endosymbionts to produce amino acids that are lacking in the plant sap on which they feed. The mealybug, *Phenacoccus avenae*, contains a *Tremblaya* endosymbiont that encodes genes for the biosynthetic pathways of eight amino acids—tryptophan, phenylalanine, histidine, arginine, isoleucine, methionine, threonine, and diaminopimelic acid [24]. In contrast, *Planococcus citri*, contains a *Tremblaya* endosymbiont with a more severely reduced genome that lacks the necessary genes to synthesize these amino acids [24]. This *Tremblaya* endosymbiont is also a host for the bacterial symbiont *Moranella endobia* (Figure 2), and it was thought that *M. endobia* may contain the missing genes and would enable synthesis of these amino acids [24]. However, it turns out that in *Pl. citri*, the biosynthetic pathways for these eight amino acids are encoded by a combination of genes in the *Tremblaya* endosymbiont, the *Moranella* endosymbiont, and at least 22 transcribed putative LGTs to the *Pl. citri* nuclear genome from three diverse bacterial taxa, α -Proteobacteria, γ -Proteobacteria, and Bacteroidetes [24]. It is not yet clear how all of the protein products of these genes in different compartments can produce functional pathways.

The serial endosymbiosis theory posits that after an early eukaryote acquired a beneficial, energy-producing, bacterial endosymbiont, the accumulation of endosymbiont genes via LGT in the nuclear genome transitioned the endosymbiont to organelle [25]. A molecular ratchet is proposed whereby all genes that can be acquired by the nuclear genome will be gradually lost by the organelle genome [26]. In both mitochondria and chloroplasts, it is thought that only mitochondria/chloroplast genes were transitioned to the nucleus. However, the *Tremblaya*/mealybug example illustrates that genes may be lost from the endosymbiont or organelle that are functionally replaced in the nucleus with functional homologues from other taxa. It raises the possibility that there are alternative paths to the formation of organelles.

LGT in the human genome

The search for LGT in the human genome has not been without controversy. In the first draft of the human genome, 223 proteins were identified with significant protein sequence similarity to bacterial proteins [27]. These proteins had no significant similarity to yeast,

worm, fly, mustard weed, or other nonvertebrate eukaryotes proteins available at the time, suggesting that they arose via LGT [27]. This finding was quickly refuted with an argument suggesting that ~180 of the 223 genes were likely not from LGT and that as more diverse eukaryotic and prokaryotic genomes were sequenced, the remaining ~40 putative LGT genes would probably be excluded as LGT candidates [28]. Instead, alternate evolutionary explanations, such as gene loss, were put forth as being more likely [28].

More than a decade later, a subsequent examination of LGT in the human genome concluded that not only were some of the previously reported LGT genes likely true LGT, but that there are an additional 128 putative LGTs [29]. One reason for the difference is the plethora of genomes from diverse organisms that the later analysis could use for its analysis. For example, the human *HASI* gene is more closely related to fungi than other metazoan genes suggesting that it may have arisen from LGT [29].

The previous studies exclusively focused on LGT into the human genome that may have an impact in an evolutionary context. These studies did not address the possibility, or potential consequences, of bacterial LGT into the somatic human genome. While somatic mutations are not important within the context of evolution, they can alter human biology. For example, human cancers typically have an accumulation of somatic mutations that alter the normal biology of cells to proliferate uncontrollably. These somatic mutations range from small single nucleotide changes [30–33] to large chromosomal rearrangements [34–36]. The human genome is also susceptible to exogenous elements causing DNA damage such as somatic integration of DNA.

The mitochondrial genome frequently integrates into the human nuclear cancer genome, with detected integrations ranging in size from 148 bp to the entire 16.5 kb mitochondrial genome [37]. These integrations were significantly enriched near the origin of replication on the heavy strand of the mitochondrial genome and were associated with other structural variations in the human genome [37]. While some of the mitochondrial integrations were identified near nuclear genes, the functional consequence of such integrations is unclear.

Viruses are also able to integrate into the human genome. The integration of human papillomavirus into the human genome is possibly the best-studied example, since the integration is a key step in promoting the development of cervical cancer [38, 39] (Figure 3). In addition, using next-generation sequencing, there is growing evidence that the integration of hepatitis B virus into the genome of hepatocellular carcinomas is frequent and carcinogenic [40].

Recent research has raised the possibility that DNA inside the cell may integrate into the human genome through a process termed “template sequence insertion” [41]. Template sequence insertion is the integration of DNA to patch repair DNA double stranded breaks. The resulting template sequence insertion lesion has hallmarks of either L1-mediated retrotransposition or nonhomologous end joining repair [42] and occurs through an RNA intermediate [41, 43].

Identification of bacterial DNA integrations into the human somatic genome

The overwhelming number of microbes in the human body provides another large source of potential DNA to integrate into the human genome, in addition to the mitochondrial genome and viral genomes described above. While human germ cells are thought to be protected from interacting with the microbiome, human somatic cells are exposed to the microbiome. Given that there are somatic integrations of viral and mitochondrial DNA into the human genome, and the large amounts of bacterial DNA in the human body, it stands to reason that bacterial DNA integrations (BDIs) may occur in the human somatic genome. BDIs into terminally differentiated cells could prove difficult to identify as only a single copy would exist, and once interrogated by sequencing, would be destroyed. In contrast, cancer cells excel at replication, with each cell replicating the mutations of the parental cell. In this way, sequencing of cancer cells may enable detection of BDIs. Bacterial DNA may integrate into safe regions of the human genome, but there is also the possibility that the BDI could cause deleterious mutations that promote carcinogenesis. Currently, large projects such as The Cancer Genome Atlas (TCGA) are using next-generation sequencing to characterize the genomic landscape of many cancers to better understand the biology driving tumorigenesis. These large publicly available sequencing projects provide a comprehensive dataset that can be used to evaluate if bacterial DNA integrates into the somatic human cancer genome.

An early release of TCGA data from the Sequence Read Archive that included sequencing data from 10 cancer types, 632 tumor samples, 220 of which had normal samples, had evidence for bacteria-human LGT in the somatic genome [44] (Figure 3). The highest number of reads supporting putative BDIs was found in acute myeloid leukemia. These BDI reads support the integration of *Acinetobacter-like* 16S and 23S rRNA gene fragments into the human mitochondrial genome [44].

The second highest number of putative BDI reads were found in stomach adenocarcinoma [44]. These BDI reads support integration of *Pseudomonas-like* 16S and 23S rRNA gene fragments into the 5'-UTR of *CEACAM5*, *CEACAM6*, *CD74*, and *TMSB10* [44]. While the BDIs are enriched in the 5'-UTR of these genes, the BDIs differ in the both the absolute and relative position of the transcriptional start site [45]. Characterization of the integrated bacterial sequence has shown that the sequences originated from stem-loop structures in the native bacterial rRNA genes [45]. As such, the BDIs may have the propensity to form complex secondary structure that have the potential to alter the human gene expression.

Moving forward

The use of public data has been key to the discovery of many LGTs, including those described above in the human genome. These discoveries are the result of secondary data analysis. LGT was thought to be a rare event, and still is by some. Therefore, many sequenced genomes were not, and are not, analyzed for LGT. Through the sharing of genome sequencing data, it is possible to perform subsequent secondary analyses to identify LGT. For example, the secondary analysis of the *Drosophila ananassae* genome identified extensive *Wolbachia* and was a seminal finding in expanding our understanding of the extent of LGT between prokaryotes and eukaryotes [46].

However, robust standards for the basic identification and verification of LGT are still needed. Over the past two decades there have been many proposed LGTs that have subsequently been disproven. Recently, a draft genome of the tardigrade was published that reported that ~1/6 of the genome can be attributed to LGT from bacteria, plants, fungi, and Archaea [47]. The data supporting the draft tardigrade genome was made publicly available, and other groups quickly published their own analyses and conclusions demonstrating that the draft tardigrade genome likely had contamination that inflated the abundance of LGT in the genome, with the latest estimates ranging from 1.9% to 4.5% LGT [48–50].

While the true amount of LGT in the tardigrade genome is still uncertain, what is certain is that the tardigrade draft genome is a success story for modern data sharing and open science. As LGT detection tools become more widely accessible and applied to more genomes, more instances of LGT will be identified, and the extent to which LGT plays a role in shaping our complex and interesting biological world will become more clear.

Acknowledgments

Our research and the preparation of this manuscript was supported by the National Science Foundation Advances in Biological Informatics (ABI-1457957) and the National Institutes of Health through the NIH Director's New Innovator Award Program (1-DP2-OD007372) and an NIH Director's Transformative Research Award (1-R01-CA206188).

References

1. Fournier GP, Andam CP, Gogarten JP. Ancient horizontal gene transfer and the last common ancestors. *BMC Evol Biol.* 2015; 15:70. [PubMed: 25897759]
2. Griffith F. The Significance of Pneumococcal Types. *The Journal of hygiene.* 1928; 27:113–159. [PubMed: 20474956]
3. Avery OT, Macleod CM, McCarty M. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types: Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med.* 1944; 79:137–158. [PubMed: 19871359]
4. Davies J. Vicious circles: looking back on resistance plasmids. *Genetics.* 1995; 139:1465–1468. [PubMed: 7789752]
5. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature.* 2000; 405:299–304. [PubMed: 10830951]
6. Davies J, Davies D. Origins and evolution of antibiotic resistance. *Microbiol Mol Biol Rev.* 2010; 74:417–433. [PubMed: 20805405]
7. Chang S, Sievert DM, Hageman JC, Boulton ML, Tenover FC, Downes FP, Shah S, Rudrik JT, Pupp GR, Brown WJ, Cardo D, Fridkin SK, Vancomycin-Resistant T. *Staphylococcus aureus* Investigative, Infection with vancomycin-resistant *Staphylococcus aureus* containing the vanA resistance gene. *N Engl J Med.* 2003; 348:1342–1347. [PubMed: 12672861]
8. Beiko RG, Harlow TJ, Ragan MA. Highways of gene sharing in prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America.* 2005; 102:14332–14337. [PubMed: 16176988]
9. Doolittle WF. Phylogenetic classification and the universal tree. *Science.* 1999; 284:2124–2129. [PubMed: 10381871]
10. Cordero OX, Hogeweg P. The impact of long-distance horizontal gene transfer on prokaryotic genome size. *Proc Natl Acad Sci U S A.* 2009; 106:21748–21753. [PubMed: 20007373]
11. Nakamura Y, Itoh T, Matsuda H, Gojobori T. Biased biological functions of horizontally transferred genes in prokaryotic genomes. *Nat Genet.* 2004; 36:760–766. [PubMed: 15208628]

12. Lerat E, Daubin V, Ochman H, Moran NA. Evolutionary origins of genomic repertoires in bacteria. *PLoS Biol.* 2005; 3:e130. [PubMed: 15799709]
13. Anderson MT, Seifert HS. Opportunity and means: horizontal gene transfer from the human host to a bacterial pathogen. *mBio.* 2011; 2:e00005–00011. [PubMed: 21325040]
14. Boucher Y, Douady CJ, Papke RT, Walsh DA, Boudreau ME, Nesbo CL, Case RJ, Doolittle WF. Lateral gene transfer and the origins of prokaryotic groups. *Annu Rev Genet.* 2003; 37:283–328. [PubMed: 14616063]
15. J.C. Dunning Hotopp. Horizontal gene transfer between bacteria and animals. *Trends Genet.* 2011; 27:157–163. [PubMed: 21334091]
16. J.C. Dunning Hotopp, Clark ME, Oliveira DC, Foster JM, Fischer P, Torres MC Munoz, Giebel JD, Kumar N, Ishmael N, Wang S, Ingram J, Nene RV, Shepard J, Tomkins J, Richards S, Spiro DJ, Ghedin E, Slatko BE, Tettelin H, Werren JH. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science.* 2007; 317:1753–1756. [PubMed: 17761848]
17. Klasson L, Kumar N, Bromley R, Sieber K, Flowers M, Ott SH, Tallon LJ, Andersson SG, J.C. Dunning Hotopp. Extensive duplication of the Wolbachia DNA in chromosome four of *Drosophila ananassae*. *BMC Genomics.* 2014; 15:1097. [PubMed: 25496002]
18. Gladyshev EA, Meselson M, Arkhipova IR. Massive horizontal gene transfer in bdelloid rotifers. *Science.* 2008; 320:1210–1213. [PubMed: 18511688]
19. Stavrinidou E, Gabrielsson R, Gomez E, Crispin X, Nilsson O, Simon DT, Berggren M. Electronic plants. *Science advances.* 2015; 1:e1501136. [PubMed: 26702448]
20. Acuna R, Padilla BE, Florez-Ramos CP, Rubio JD, Herrera JC, Benavides P, Lee SJ, Yeats TH, Egan AN, Doyle JJ, Rose JK. Adaptive horizontal transfer of a bacterial gene to an invasive insect pest of coffee. *Proc Natl Acad Sci U S A.* 2012; 109:4197–4202. [PubMed: 22371593]
21. Ioannidis P, Lu Y, Kumar N, Creasy T, Daugherty S, Chibucos MC, Orvis J, Shetty A, Ott S, Flowers M, Sengamalay N, Tallon LJ, Pick L, J.C. Dunning Hotopp. Rapid transcriptome sequencing of an invasive pest, the brown marmorated stink bug *Halyomorpha halys*. *BMC Genomics.* 2014; 15:738. [PubMed: 25168586]
22. Pitzschke A, Hirt H. New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. *EMBO J.* 2010; 29:1021–1032. [PubMed: 20150897]
23. Gelvin SB. *Agrobacterium*-Mediated Plant Transformation: the Biology behind the “Gene-Jockeying” Tool. *Microbiology and Molecular Biology Reviews.* 2003; 67:16–37. [PubMed: 12626681]
24. Husnik F, Nikoh N, Koga R, Ross L, Duncan RP, Fujie M, Tanaka M, Satoh N, Bachtrog D, Wilson AC, von Dohlen CD, Fukatsu T, McCutcheon JP. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. *Cell.* 2013; 153:1567–1578. [PubMed: 23791183]
25. Margulis, L. *Origin of Eukaryotic Cells.* Yale University Press; New Haven: 1970.
26. 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]
27. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, Funke R, Gage D, Harris K, Heaford A, Howland J, Kann L, Lehoczky J, LeVine R, McEwan P, McKernan K, Meldrim J, Mesirov JP, Miranda C, Morris W, Naylor J, Raymond C, Rosetti M, Santos R, Sheridan A, Sougnez C, Stange-Thomann N, Stojanovic N, Subramanian A, Wyman D, Rogers J, Sulston J, Ainscough R, Beck S, Bentley D, Burton J, Clee C, Carter N, Coulson A, Deadman R, Deloukas P, Dunham A, Dunham I, Durbin R, French L, Grafham D, Gregory S, Hubbard T, Humphray S, Hunt A, Jones M, Lloyd C, McMurray A, Matthews L, Mercer S, Milne S, Mullikin JC, Mungall A, Plumb R, Ross M, Shownkeen R, Sims S, Waterston RH, Wilson RK, Hillier LW, McPherson JD, Marra MA, Mardis ER, Fulton LA, Chinwalla AT, Pepin KH, Gish WR, Chissoe SL, Wendl MC, Delehaunty KD, Miner TL, Delehaunty A, Kramer JB, Cook LL, Fulton RS, Johnson DL, Minx PJ, Clifton SW, Hawkins T, Branscomb E, Predki P, Richardson P, Wenning S, Slezak T, Doggett N, Cheng JF, Olsen A, Lucas S, Elkin C, Uberbacher E, Frazier M, Gibbs RA, Muzny DM, Scherer SE, Bouck JB, Sodergren EJ, Worley KC, Rives CM, Gorrell JH, Metzker ML, Naylor SL, Kucherlapati RS, Nelson DL, Weinstock GM, Sakaki Y, Fujiyama A, Hattori M, Yada T, Toyoda A, Itoh T, Kawagoe C, Watanabe H, Totoki Y, Taylor T, Weissenbach J, Heilig R, Saurin W, Artiguenave F, Brottier P, Bruls T, Pelletier E, Robert C,

Wincker P, Smith DR, Doucette-Stamm L, Rubenfield M, Weinstock K, Lee HM, Dubois J, Rosenthal A, Platzer M, Nyakatura G, Taudien S, Rump A, Yang H, Yu J, Wang J, Huang G, Gu J, Hood L, Rowen L, Madan A, Qin S, Davis RW, Federspiel NA, Abola AP, Proctor MJ, Myers RM, Schmutz J, Dickson M, Grimwood J, Cox DR, Olson MV, Kaul R, Shimizu N, Kawasaki K, Minoshima S, Evans GA, Athanasiou M, Schultz R, Roe BA, Chen F, Pan H, Ramser J, Lehrach H, Reinhardt R, McCombie WR, Bastide M de la, Dedhia N, Blocker H, Hornischer K, Nordsiek G, Agarwala R, Aravind L, Bailey JA, Bateman A, Batzoglu S, Birney E, Bork P, Brown DG, Burge CB, Cerutti L, Chen HC, Church D, Clamp M, Copley RR, Doerks T, Eddy SR, Eichler EE, Furey TS, Galagan J, Gilbert JG, Harmon C, Hayashizaki Y, Haussler D, Hermjakob H, Hokamp K, Jang W, Johnson LS, Jones TA, Kasif S, Kasprzyk A, Kennedy S, Kent WJ, Kitts P, Koonin EV, Korf I, Kulp D, Lancet D, Lowe TM, McLysaght A, Mikkelsen T, Moran JV, Mulder N, Pollara VJ, Ponting CP, Schuler G, Schultz J, Slater G, Smit AF, Stupka E, Szustakowski J, Thierry-Mieg D, Thierry-Mieg J, Wagner L, Wallis J, Wheeler R, Williams A, Wolf YI, Wolfe KH, Yang SP, Yeh RF, Collins F, Guyer MS, Peterson J, Felsenfeld A, Wetterstrand KA, Patrinos A, Morgan MJ, de Jong P, Catanese JJ, Osoegawa K, Shizuya H, Choi S, Chen YJ. Initial sequencing and analysis of the human genome. *Nature*. 2001; 409:860–921. [PubMed: 11237011]

28. Salzberg SL, White O, Peterson J, Eisen JA. Microbial genes in the human genome: lateral transfer or gene loss? *Science*. 2001; 292:1903–1906. [PubMed: 11358996]
29. 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*. 2015; 16:50. [PubMed: 25785303]
30. Huang FW, Hodis E, Xu MJ, Kryukov GV, Chin L, Garraway LA. Highly recurrent TERT promoter mutations in human melanoma. *Science*. 2013; 339:957–959. [PubMed: 23348506]
31. Horn S, Figl A, Rachakonda PS, Fischer C, Sucker A, Gast A, Kadel S, Moll I, Nagore E, Hemminki K, Schadendorf D, Kumar R. TERT promoter mutations in familial and sporadic melanoma. *Science*. 2013; 339:959–961. [PubMed: 23348503]
32. Banerji S, Cibulskis K, Rangel-Escareno C, Brown KK, Carter SL, Frederick AM, Lawrence MS, Sivachenko AY, Sougnez C, Zou L, Cortes ML, Fernandez-Lopez JC, Peng S, Ardlie KG, Auclair D, Bautista-Pina V, Duke F, Francis J, Jung J, Maffuz-Aziz A, Onofrio RC, Parkin M, Pho NH, Quintanar-Jurado V, Ramos AH, Rebollar-Vega R, Rodriguez-Cuevas S, Romero-Cordoba SL, Schumacher SE, Stransky N, Thompson KM, Uribe-Figueroa L, Baselga J, Beroukhi R, Polyak K, Sgroi DC, Richardson AL, Jimenez-Sanchez G, Lander ES, Gabriel SB, Garraway LA, Golub TR, Melendez-Zajgla J, Toker A, Getz G, Hidalgo-Miranda A, Meyerson M. Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature*. 2012; 486:405–409. [PubMed: 22722202]
33. Newburger DE, Kashef-Haghighi D, Weng Z, Salari R, Sweeney RT, Brunner AL, Zhu SX, Guo X, Varma S, Troxell ML, West RB, Batzoglu S, Sidow A. Genome evolution during progression to breast cancer. *Genome Res*. 2013; 23:1097–1108. [PubMed: 23568837]
34. Malhotra A, Lindberg M, Faust GG, Leibowitz ML, Clark RA, Layer RM, Quinlan AR, Hall IM. Breakpoint profiling of 64 cancer genomes reveals numerous complex rearrangements spawned by homology-independent mechanisms. *Genome Res*. 2013; 23:762–776. [PubMed: 23410887]
35. Korbel JO, Urban AE, Affourtit JP, Godwin B, Grubert F, Simons JF, Kim PM, Palejev D, Carriero NJ, Du L, Taillon BE, Chen Z, Tanzer A, Saunders AC, Chi J, Yang F, Carter NP, Hurles ME, Weissman SM, Harkins TT, Gerstein MB, Egholm M, Snyder M. Paired-end mapping reveals extensive structural variation in the human genome. *Science*. 2007; 318:420–426. [PubMed: 17901297]
36. Welch JS, Westervelt P, Ding L, Larson DE, Klco JM, Kulkarni S, Wallis J, Chen K, Payton JE, Fulton RS, Veizer J, Schmidt H, Vickery TL, Heath S, Watson MA, Tomasson MH, Link DC, Graubert TA, DiPersio JF, Mardis ER, Ley TJ, Wilson RK. Use of whole-genome sequencing to diagnose a cryptic fusion oncogene. *JAMA*. 2011; 305:1577–1584. [PubMed: 21505136]
37. Ju YS, Tubio JM, Mifsud W, Fu B, Davies HR, Ramakrishna M, Li Y, Yates L, Gundem G, Tarpey PS, Behjati S, Papaemmanuil E, Martin S, Fullam A, Gerstung M, I.B.C.W. Group, I.B.C.W. Group, I.B.C.W. Group. Nangalia J, Green AR, Caldas C, Borg A, Tutt A, Lee MT, van't Veer LJ, Tan BK, Aparicio S, Span PN, Martens JW, Knappskog S, Vincent-Salomon A, Borresen-Dale AL, Eyfjord JE, Flanagan AM, Foster C, Neal DE, Cooper C, Eeles R, Lakhani SR, Desmedt C, Thomas G, Richardson AL, Purdie CA, Thompson AM, McDermott U, Yang F, Nik-Zainal S,

- Campbell PJ, Stratton MR. Frequent somatic transfer of mitochondrial DNA into the nuclear genome of human cancer cells. *Genome Res.* 2015; 25:814–824. [PubMed: 25963125]
38. Pett M, Coleman N. Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis? *J Pathol.* 2007; 212:356–367. [PubMed: 17573670]
39. Schiffman M, Castle PE, Jeronimo J, Rodriguez AC, Wacholder S. Human papillomavirus and cervical cancer. *The Lancet.* 2007; 370:890–907.
40. Sung WK, Zheng H, Li S, Chen R, Liu X, Li Y, Lee NP, Lee WH, Ariyaratne PN, Tennakoon C, Mulawadi FH, Wong KF, Liu AM, Poon RT, Fan ST, Chan KL, Gong Z, Hu Y, Lin Z, Wang G, Zhang Q, Barber TD, Chou WC, Aggarwal A, Hao K, Zhou W, Zhang C, Hardwick J, Buser C, Xu J, Kan Z, Dai H, Mao M, Reinhard C, Wang J, Luk JM. Genome-wide survey of recurrent HBV integration in hepatocellular carcinoma. *Nature Genetics.* 2012; 44:765–769. [PubMed: 22634754]
41. Onozawa M, Zhang Z, Kim YJ, Goldberg L, Varga T, Bergsagel PL, Kuehl WM, Aplan PD. Repair of DNA double-strand breaks by templated nucleotide sequence insertions derived from distant regions of the genome. *Proc Natl Acad Sci U S A.* 2014; 111:7729–7734. [PubMed: 24821809]
42. Iltumur K, Yavavli A, Apak I, Ariturk Z, Toprak N. Elevated plasma N-terminal pro-brain natriuretic peptide levels in acute ischemic stroke. *American heart journal.* 2006; 151:1115–1122. [PubMed: 16644347]
43. Karabulut A, Iltumur K, Yalcin K, Toprak N. Hepatopulmonary syndrome and right ventricular diastolic functions: an echocardiographic examination. *Echocardiography.* 2006; 23:271–278. [PubMed: 16640703]
44. Riley DR, Sieber KB, Robinson KM, White JR, Ganesan A, Nourbakhsh S, J.C. Dunning Hotopp. Bacteria-human somatic cell lateral gene transfer is enriched in cancer samples. *PLoS Comput Biol.* 2013; 9:e1003107. [PubMed: 23840181]
45. Sieber KB, Gajer P, J.C. Dunning Hotopp. Modeling the integration of bacterial rRNA fragments into the human cancer genome. *BMC Bioinformatics.* 2016; 17:134. [PubMed: 27001685]
46. J.C. Dunning Hotopp, Clark ME, Oliveira DC, Foster JM, Fischer P, Torres MC, Giebel JD, Kumar N, Ishmael N, Wang S, Ingram J, Nene RV, Shepard J, Tomkins J, Richards S, Spiro DJ, Ghedin E, Slatko BE, Tettelin H, Werren JH. Widespread lateral gene transfer from intracellular bacteria to multicellular eukaryotes. *Science.* 2007; 317:1753–1756. [PubMed: 17761848]
47. Boothby TC, Tenlen JR, Smith FW, Wang JR, Patanella KA, E. Osborne Nishimura, Tintori SC, Li Q, Jones CD, Yandell M, Messina DN, Glasscock J, Goldstein B. Evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proc Natl Acad Sci U S A.* 2015; 112:15976–15981. [PubMed: 26598659]
48. Koutsovoulos G, Kumar S, Laetsch DR, Stevens L, Daub J, Conlon C, Maroon H, Thomas F, Aboobaker AA, Blaxter M. No evidence for extensive horizontal gene transfer in the genome of the tardigrade *Hypsibius dujardini*. *Proc Natl Acad Sci U S A.* 2016; 113:5053–5058. [PubMed: 27035985]
49. Arakawa K. No evidence for extensive horizontal gene transfer from the draft genome of a tardigrade. *Proc Natl Acad Sci U S A.* 2016; 113:E3057. [PubMed: 27173901]
50. Bemm F, Weiss CL, Schultz J, Forster F. Genome of a tardigrade: Horizontal gene transfer or bacterial contamination? *Proc Natl Acad Sci U S A.* 2016; 113:E3054–3056. [PubMed: 27173902]

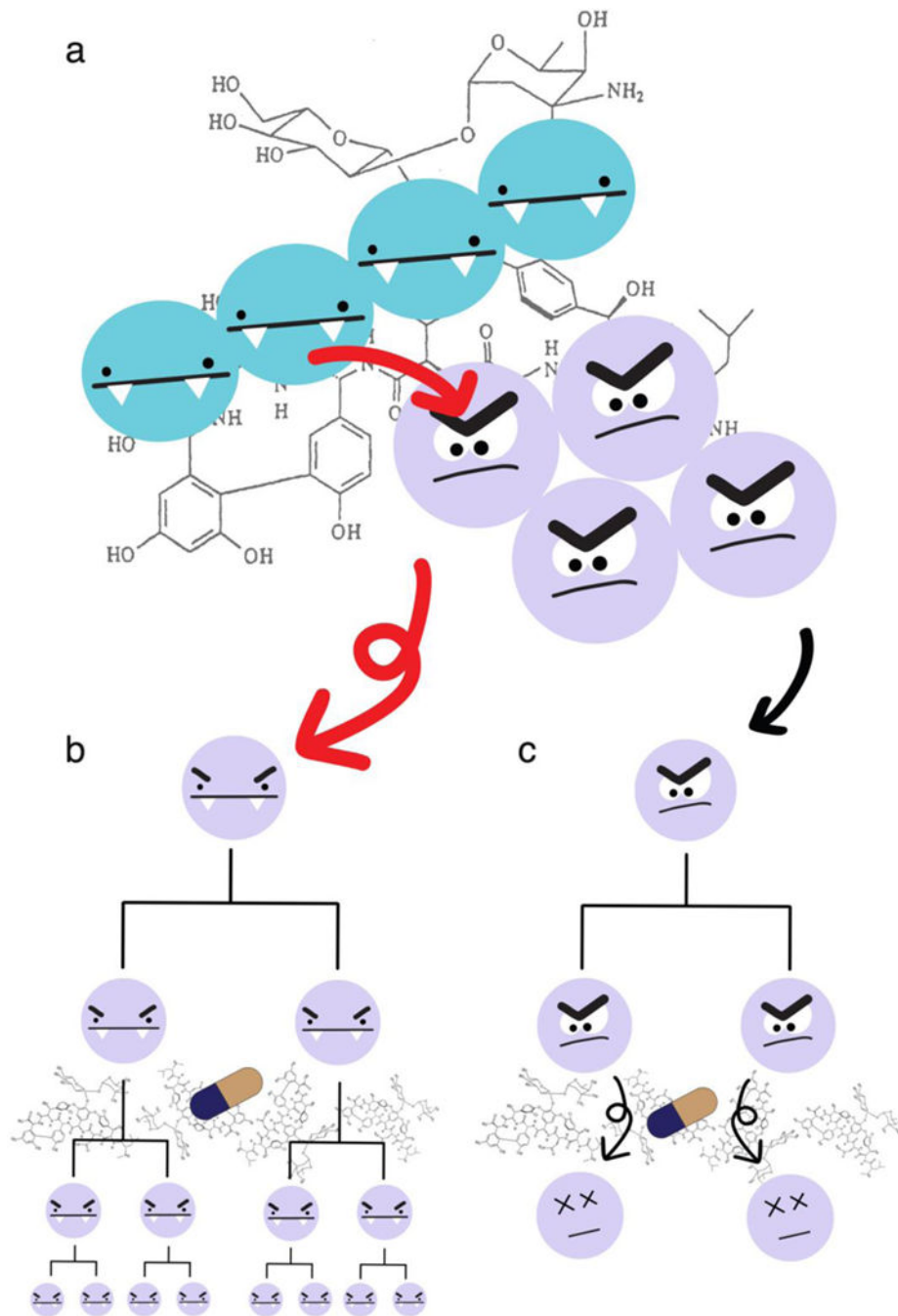


Figure 1. Representation of the spread of vancomycin resistance genes via inter-species LGT
Enterococcus faecalis (blue) and *Staphylococcus aureus* (lavender) can sometimes be found together, particularly in clinical settings. In these settings, they can be exposed to strong selective agents, like the antibiotic vancomycin. (a) In 2002, bacteria were discovered in two clinical settings where the genes for vancomycin resistance were transferred via LGT from vancomycin-resistant *E. faecalis* to methicillin-resistant *S. aureus*. Vancomycin resistance is represented by beady eyes and fangs, methicillin resistance is represented by furloughed eyebrows, and the LGT of vancomycin resistance genes from VRE to MRSA is represented

by red arrows. (b) LGT from VRE to MRSA results in a vancomycin-resistant and methicillin-resistant *Staphylococcus aureus* (VRSA) which vertically passes resistance onto daughter cells which continue to propagate in the presence of vancomycin. (c) The MRSA that did not receive vancomycin resistance via LGT remain susceptible to vancomycin and die in the continued presence of vancomycin. The evolution of vancomycin resistance by way of genetic mutation is not addressed in this figure.

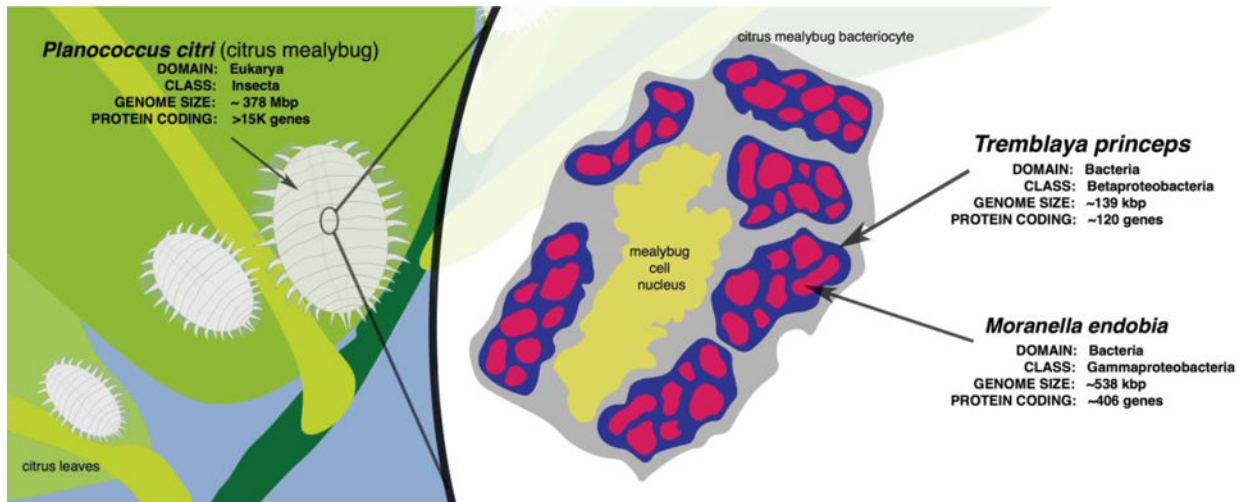


Figure 2. Mealybug endosymbiosis

The citrus mealybug *P. citri* has two endosymbionts, *T. princeps* and *M. endobia*. *M. endobia* is an endosymbiont of *T. princeps*, and *T. princeps* is in turn an endosymbiont of the mealybug. These *T. princeps* in turn reside in the mealybug bacteriocyte, a specialized cell where the insect houses endosymbionts.

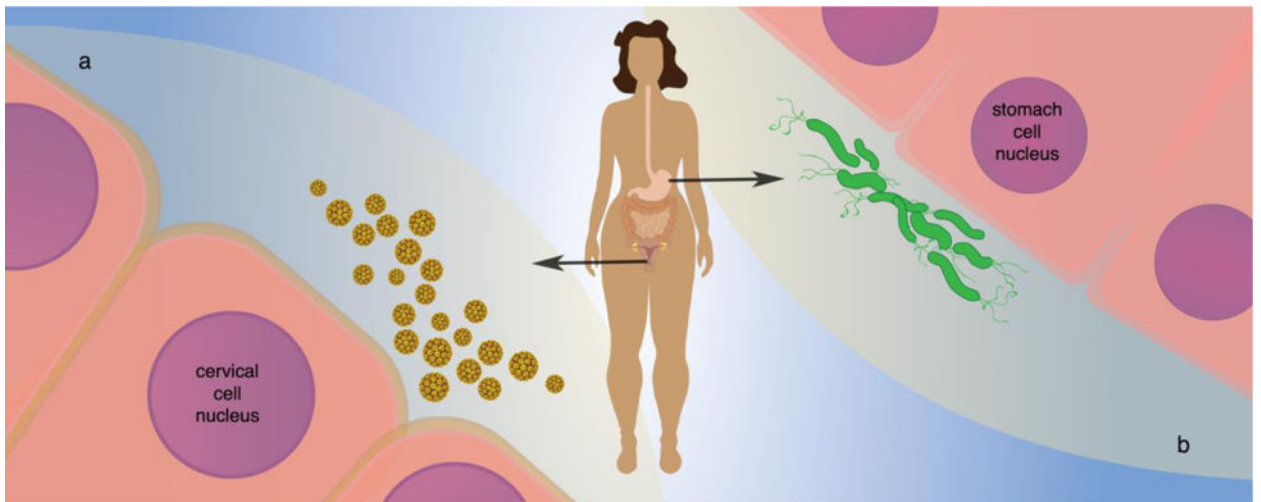


Figure 3. Lateral gene transmission to human somatic cells

(a) Integration of viral DNA into the host cell nuclear chromosome via LGT is a well-documented part of the human papilloma virus cancer cycle. (b) Analysis of sequences from the Cancer Genome Atlas supports bacterial DNA integration into the nuclear chromosomes of stomach cells.