Chapter 8.2

Horizontal Gene Transfer in Microbial Evolution

Johann Peter Gogarten and R. Thane Papke

lds orcid.org/0000-0001-6459-6518 orcid.org/0000-0003-1371-2341

CONTENTS

8.2.1 Introduction

Horizontal Gene Transfer (HGT) can be described as the transfer of genetic information between cells that are not in an ancestor-descendant relationship. In bacteria this sharing of genetic information was first discovered in the 1940s (Tatum and Lederberg 1947). The comparison of molecular phylogenies to those of ribosomal RNA revealed that most gene families are not in complete agreement with the ribosomal tree of life (Hilario and Gogarten 1993; Woese et al. 2000). The best explanation for these conflicts is HGT. While gene transfers between divergent lineages allow to correlate evolutionary events in different parts of the tree of life (Gogarten 1995), frequent and especially biased gene transfer also poses a problem for phylogenetic reconstruction (Gogarten et al. 2002). Advances in genome sequencing have revealed HGT as the most important driving force in microbial evolution see; however, gene transfer also occurs between eukaryotes, and from bacteria and archaea to eukaryotes, often in the context of a close symbiotic association (see Soucy et al. 2015 for a recent review). In bacteria and archaea,

This is the authors' version. The final version is published in the
Handbook of Astrobiology, edited by Vera Kolb, CRC Press, 2019

transfer of small DNA fragments that integrate into the recipient genome through homologous recombination, prevent genome wide selective sweeps that would purge within population diversity; the transfer of genes and operons can provide new capabilities to the recipient and allow them to move into new ecological niches. HGT also is an innovative force in microbial evolution (see Swithers et al. 2012 for a recent review): it is the main process for gene family expansion in bacteria (Treangen and Rocha 2011), and it played a role in extending existing and assembling new metabolic pathways (Boucher et al. 2003; Fournier and Gogarten 2008; Khomyakova et al. 2011).

8.2.2. Definitions of pan- and core genomes of taxonomic genomes (strict and extended)

Only a decade ago most biologists, including microbiologists, expected that genomes from members of the same species had about the same gene content. Species definitions in plants, animals, algae and fungi are linked to type specimen; and this rule has been extended to microbial taxonomy: type strains are used in the description of bacterial and archaeal species, they are deposited in culture collections, and they and their genomes serve as important taxonomic and phylogenomic reference points (Kyrpides et al. 2014). However, in contrast to eukaryotes, the genome of an individual type specimen captures only a tiny fraction of the gene content within a population. The pan-genome of a taxon or group refers to the sum of all genes present in each individual (Tettelin et al. 2005; Lapierre and Gogarten 2009). Pan-genomes comprise the core genome, i.e., the genes that are found in all members, and the accessory genome, *i.e.*, genes that are present in only one or a few members of the group. Welch *et al.* provided the first illustration that genome content in bacteria changes rapidly (Welch et al. 2002). Comparing three *Escherichia coli* strains they found the shared core to be less than 40% of the gene families present in all three genomes. More recently the size of this core was further reduced to less than 10% of gene families present in 61 *Escherichia coli* genomes (Lukjancenko et al. 2010). When the number of gene families per genome encountered in a taxon is plotted as function of the number of genomes sampled, the resulting curve approaches a line that continues to rise. Pan-genomes that are characterized by these non-saturating rarefaction curves are considered open (Lapierre and Gogarten 2009; Lobkovsky et al. 2014; Puigbò et al. 2014). Obviously, pan-genomes cannot really have an infinite size and rarefaction curves will saturate eventually. Baumdicker et al. provided a realistic estimation of the size of the pan-genome: Taking population size and time since divergence into account they estimated that the *Prochlorococcus* pan-genome contains about 58,000 genes (Baumdicker et al. 2012) - recently this estimate was increased to 84,872 genes (Biller et al. 2015), whereas the individual *Prochlorococcus* genomes encode only about 2000 genes each.

The pan-genome concept was originally developed to describe the fluidity of prokaryotic genomes (Tettelin et al. 2005). Because HGT is more frequent between close relatives (Ravin 1963; Dykhuizen and Green 1991; Vulic et al. 1997; Gogarten et al. 2002; Andam and Gogarten 2011), the pan-genome also represents a set of genes that is more readily available via HGT to any member of the group. The pan-genome may then be thought of as a shared genetic resource of a population.

Large "open" pan-genomes are found for most bacterial and archaeal species. The most detailed study of population genomics was performed in *Prochlorococcus* (Kashtan et al. 2014; Biller et al. 2015), which showed that even in the presence of gene transfer within and between subpopulations, divergent sub-populations exist for long periods of time. These marine nanocyanobacteria provide about 50% of CO_2 fixation on this planet, and constitute the largest know bacterial population (Lynch and Conery 2003; Baumdicker et al. 2012; Kashtan et al. 2014), the overall population is differentiated in high light and low light ecotypes, but within each of these ecotypes many stable subpopulations exist (Kashtan et al. 2014). Gene flow between *Prochlorococcus* subpopulations, and between low light adapted *Prochlorococcus* and *Synecchococcus* occurs frequently, transferring even DNA that encodes genes for ribosomal proteins and rRNA (Zhaxybayeva et al. 2009; Shapiro et al. 2012). For example the ITS region in the ribosomal rRNA operon, while usually a good predictor and classifier for subpopulations, failed in two out of 96 instances compared to a strong coherent signal from whole genome sequences (Kashtan et al. 2014).

8.2.3. Description of processes that lead to genes with limited distribution

8.2.3.1 The black queen hypotheses

The black queen hypothesis proposed by Morris *et al.* (Morris et al. 2012) is built on the premise of "leaky" common good functions, *i.e.*, gene encoded processes that lead to products that benefit a population or community, not only the producer. This hypothesis suggests that leaky functions combined with selection for small genomes may lead to a situation in which these leaky functions are encoded in only a fraction of the genomes comprising the community. The name "Black Queen" originates from the card game "hearts" where two winning strategies exist, one is to acquire all hearts, and the most valuable queen of spades, the other strategy is to divest as many heart and the queen of spades as possible. Applied to genomes, these strategies correspond to (A) maintaining a keystone genome that includes all necessary genes, or (B) to lose as many genes encoding leaky functions as possible.

A) The cell can retain all genes encoding leaky functions (in the game of hearts this strategy is known as "shooting the moon"). The cost is a large genome, and consequently a lower growth rate and a decreasing frequency in the population. The advantage is that following a population bottleneck all genes encoding leaky functions are available in the genome. If they exist, these members of a community may be thought of as analogous to keystone species.

(B) The cell loses some or all of its leaky functions and increases its growth rate (in hearts, this represents the usual strategy of taking as few point cards as possible). The cell that no longer produces the leaky function relies on a common good provided by other cells. If a bottleneck occurs, a cell following this strategy is unlikely to survive on its own, provided the leaky product is not provided by other species. A possible outcome of all cells in a population following strategy B is that all members of a population cheat on some leaky functions.

Division of labor rather than cheating may be a more appropriate description, especially in those cases where a keystone genome is no-longer present in the population and all individuals in the population are dependent on some common good produced by others (strong Black Queen

Hypothesis (Morris et al. 2012; Fullmer et al. 2015). Wide spread cheating can lead to the tragedy of the commons; Oliveira et al. found that stable cooperation through reciprocal exchange was unlikely to emerge in evolution under the models and parameters tested(Oliveira et al. 2014); however, experimental work by Morris et al. has shown that producers and consumers of the leaky product, in their case H_2O_2 detoxification, can enable the stable co-existence of two very similar organisms that use the same resources (Morris et al. 2014). Additionally, most bacterial and archaeal cells live in biofilms or small aggregates (Stemmann and Boss 2012; Kolter and Greenberg 2006) and therefore are more likely to be close to cells with whom they share recent ancestry and more likely to have the same genotype with respect to leaky functions. These neighborhood relations are expected to increase frequency dependent selection on genes encoding these functions. For example, Drescher et al showed that *Vibrio cholerae* can avoid the public goods dilemma by strengthening relationships between cells of the same genotype through creation of a thick biofilm, leading to larger benefits of the producers, when the overall concentration of the public good decreases (Drescher et al. 2014).

Not all cheating leads to mutual collaboration. In addition to frequency dependent selection other mechanisms exist to slow the emergence of cheaters. In the production of common goods, for which lower levels of synthesis are detrimental for the whole population and that are under the control of quorum sensing, the emergence of cheaters that ignore the quorum sensing signal was found to be counter selected through quorum sensing signals also controlling the synthesis of private goods (Dandekar et al. 2012), in this case adenosine catabolism.

8.2.3.2 The red queen hypothesis - genes altering the interactions with viruses and other parasites

Van Valen proposed the red queen hypothesis was proposed to describe the necessity for ongoing evolution in the arms race between biological species (Van Valen 1973). Bacteria and archaea are under severe virus predation (Thurber 2009); in addition, members of the same population compete for limited resources. Therefore, selection pressure to evade predation and outcompete niche rivals causes a constant genetic arms race in cellular and viral populations, hence the analogy to the Red Queen from Lewis Carroll's *Through the Looking-Glass* (Carroll 1871; Carroll and Gardner 1993), who states that "*it takes all the running you can do, to keep in the same place*" (Van Valen 1973). The analysis of phage metagenomes and rank abundance curves indicated that phage predation follows the *kill the winner* strategy (Hoffmann et al. 2007), where successful strains in a population are targeted more frequently. The surprising stability of species composition despite phage predation suggests that cycling between different susceptible target cells is more frequent within a population than between populations from different species (Rodriguez-Brito et al. 2010). Consequently, within a population, host genes that are utilized by phage and virus to enter the cell are expected to turn over quickly, creating within population diversity (Chaturongakul and Ounjai 2014), as will genes that are costly otherwise, but provide anti-viral resistance (Hille and Charpentier 2016).

Chemicals and peptides that act as antibiotics are often produced by separate members of the same species and population resulting in a fitness advantage of resistant bacteria. Conflict is

more commonly observed between isolates from different locations, and cooperation dominates between conspecifics (Cordero et al. 2012); however, this does not mean that conflict does not arise in populations: The observation by Cordero et al. of conflict between conspecifics being largely absent may be due to selection for antibiotic resistance having already successfully occurred within populations (*i.e.*, the non-resistant members of the population have died out).

8.2.3.3 Selfish genetic elements and selectively nearly neutral genes randomly acquired

In most lineages, genomes constantly acquire and lose genes. Some genes provide a selective advantage to the recipient and will become fixed in the population. However, many of the transferred genes do not find permanent homes in recipient genomes (Lawrence and Ochman 1997; Mira et al. 2001). Among these genes are parasites (prophages) and selfish genetic elements. Most, but certainly not all (Lobkovsky et al. 2013), of the transferred genes are selectively neutral or nearly neutral to the recipient (Gogarten and Townsend 2005; Baumdicker et al. 2010; Haegeman and Weitz 2012). These genes do not persist in the genomes they "visit", and their loss also has little impact on fitness (Bolotin and Hershberg 2015; Bolotin and Hershberg 2016).

8.2.3.4. Niche adapting genes and weakly selected functions.

In the case when a population is present in different neighboring ecological niches, genes that adapt their carrier to a particular niche, such as virulence genes and genes that encode metabolic pathways, may be present in a subpopulation only. In addition to spatial heterogeneity, many genes are only used temporarily under some circumstances; e.g., a phosphatase that releases phosphate from extracellular macromolecules is only under purifying selection when inorganic phosphate is a limiting nutrient. Lawrence and Roth (Lawrence and Roth 1996) in developing the selfish operon theory describe these genes as encoding a weakly selected function.

8.2.3.5. Collaboration and units of selection.

For many decades, especially since the formulation of the Gaia hypothesis by Lovelock and Margulis (Lovelock 1972; Lovelock and Margulis 2011), some considered microbial communities, and in the extreme all bacteria and archaea existing on Earth, as a single superorganism. This idea was spearheaded by Sorin Sonea (Sonea 1988a; Sonea 1988b) and is based on the observation that bacteria and archaea can share genes between cells that are not in a parent - descendent relationship. Related is the question on units and levels of selection (gene, individual organism, population, community)(Soucy et al. 2015). Dawkins introduced a gene centered view of evolution (Dawkins 1976), in which all genes are selfish although most express their selfishness in collaboration with other genes to build an organisms with increased fitness. Evolution can be studied and described by selection acting on individuals resulting in changing gene frequencies within populations (Graur 2016). In the case of selfish genetic elements (non-cooperating genes, molecular parasites), at least initially, the benefit or cost to the individual are small, especially in case of self-splicing elements, and the element can spread in the population as a parasite, limited only by its transmissibility. Levels of selection often are intertwined, and

selection at the gene level (e.g., spread of metal or antibiotic resistance genes, or Ti plasmids that allow utilization of a particular food source) often also provides a strong benefit at the group level (e.g., avoiding severe bottlenecks, thereby maintaining within population diversity), although in most instances it is doubtful that group level selection was the driving force to share these genes (Olendzenski and Gogarten 2009; Naor et al. 2016). Life is a densely-woven fabric: viruses, plasmids, and selfish genetic elements contribute to constructing the fibers, composition, and pattern of life's fabric.

8.2.4. Mechanisms for gene transfer

Genes can be transferred horizontally by a variety of mechanisms. Traditionally, conjugation (DNA is transferred between cells through a specialized machinery), transduction (DNA is transferred through phages), and transformation (DNA is taken up from the environment into a cell) are distinguished. Related to these are Gene Transfer Agents (GTAs), and cell fusion, and in eukarytotes introgression and endosymbiotic (aka, intracellular) gene transfer. Some of these categories are not clearly separated, and in many instances rather different processes and machineries are lumped together into a single category. In case of multicellular eukaryotes the fact that genes were transferred has been established in many instances, (e.g. (e.g., Stewart et al. 2003; Graham et al. 2012). Mechanisms for these transfers have been proposed (e.g., the weak-link model (e.g., the weak-link model, Huang 2013); however, the details for these transfers remain to be established.

8.2.4.1 Conjugation

Conjugation was first described by (Lederberg and Tatum 1946). Cells make contact through a pilus, and a newly synthesized strand of DNA is transferred through a pilus into the recipient cell. The dedicated conjugation machinery often is plasmid encoded, and often only the plasmid is transferred into the recipient. However, if the plasmid is integrated into the chromosome, or if the machinery is encoded on the chromosome (Derbyshire and Gray 2014), chromosomal DNA is transferred.

A single bacterial cell can possess multiple conjugation machineries, e.g., *Agrobacterium tumefaciens* possesses a machinery that it uses to transfer T-DNA into plant cells (Joos et al. 1983) and another system for conjugation with other bacteria (Alt-Mörbe et al. 1996) that is under control of quorum sensing and following the successful transformation of a plant catalyzes the transfer of the Ti-plasmid between different *Agrobacterium* strains (White and Winans 2007).

8.2.4.2 Cell Fusion

Many archaeal cells often form intricate networks with connections between individual cells (Rosenshine et al. 1989; Stetter 2013). During mating in *Haloferax volcanii* these connections lead to exchange of cytoplasm between the mating cells. Fused heterodiploid cells can be recovered, in which recombination between the two parental chromosomes has been detected (Naor et al. 2012).

8.2.4.3 Transduction

Phage particles can pack host DNA in addition or instead of the phage DNA. In case of generalized transduction a random piece of Host DNA is packaged into the phage particle an delivered to a recipient cell; in specialized transduction a prophage imprecisely excises itself from the host genome from the host on delivers neighboring genes to the host. In addition, many phage include genes in their genome that do not play a function in phage replication. These genes have been termed morons, and in case of lysogenic phage these genes can have a large impact on the host bacteria (Hendrix et al. 2000; Cumby et al. 2012). For example, the toxins in botulism, diphtheria, and cholera all are encoded as morons in prophage genomes (Brussow et al. 2004), reflecting the interrelation between selection on the phage, the bacterial symbiont, and the eukaryotic host.

8.2.4.4 Gene transfer agents

Gene transfer agents (GTAs) can be described as prophage that have lost the ability to recognize and preferentially package their own DNA. They were described in five different groups of bacteria and archaea. The best studied GTA is the one in *Rhodobacter capsulatus*. See (Lang et al. 2017) for a recent review. The frequency of occurrence in some groups of Alphaproteobacteria argues that the GTA are more than defective prophage; however, the mechanisms of GTA maintenance remains under debate (Lang et al. 2012; Omer et al. 2017). The GTA head is too small to package all of the GTA encoding genes, and the release of GTAs is accompanied by the lysis of the bacterium. The latter thus constitutes an ultimate fitness cost to the bacterium activating its GTA, the former makes it impossible that a complete GTA gene cluster can be transferred to a new host.

8.2.4.5 Transformation

Transformation is the uptake of exogenous DNA from the environment. This was first demonstrated by Fredrick Griffith in *Streptococcus pneumoniae* (Griffith 1928), the transforming principle was later shown to be DNA (Avery et al. 1944). Transformation has been described for both archaea and bacteria (Chimileski et al. 2014; Johnston et al. 2014). Some bacteria are naturally competent (Johnston et al. 2014), others can be transformed after special treatment. Naturally competent bacteria often have an uptake bias in favor of DNA similar to their own, preferentially taking up DNA that contains motifs that bind to the uptake machinery (Mell and Redfield 2014). In other cases, DNA uptake is biased by being under the control of quorum sensing and occurring (by definition) only when there are many nearby conspecific donors that often were the victims of fratricide (Claverys and Håvarstein 2007; Borgeaud et al. 2015).

8.2.4.6 Introgression

In eukaryotes, hybridization between species, followed by repeated backcrosses to one of the parent species allows for gene flow across species boundaries. This process is of concern in case of transgenic crops that grow in the neighborhood of their non-domesticated relatives (Stewart et al. 2003). Adaptive introgression also has occurred in human evolution through hybridization between archaic human lineages and modern humans (Racimo et al. 2016).

8.2.4.7 Endosymbiotic or intracellular gene transfer

Gene transfer in eukaryotes frequently is between symbionts and their host. In case of the endosymbionts that evolved into mitochondria and plastids, only a few genes remain in the organellar genome, most genes were been transferred to the nucleus, where they are transcribed, translated in the host cell's cytoplasma, and the encoded proteins are transported into the organelle. This transfer from endosymbiont to host has been described as intracellular or endosymbiotic gene transfer (Adams et al. 1999; Timmis et al. 2004). Many other transfers likely have happened from symbiont to host (see Soucy et al. 2015 for discussion); for example, many chlamydial genes found in archaeplastida (plants and algae with primary plastids) were suggested to have contributed to integrating plastid and host (Huang and Gogarten 2007; Tyra et al. 2007), and in case of secondary plastids that evolved from endosymbiotic eukaryotic algae gene transfer from many other sources occurred (Archibald et al. 2003); however, in many other cases, once the endosymbiont is lost from a lineage, only the transfer from a bacterium to a eukaryote remains detectable.

8.2.5. Conclusion

Gene transfer has turned the evolution of genomes into a network. The Tree of Cells is embedded in this network. Gene transferred is biased and can create patterns of apparent relationships between genomes that are undistinguishable from patterns due to shared ancestry. However, because a strong bias is towards close relatives, gene transfer tends to reinforce the patterns due to shared ancestry. Gene transfer allows for the spread of selfish genetic elements and the acquisition of neutral and slightly deleterious genes, but is also allows for the long-term persistence of weakly selected functions turning the population pan-genome into shared genetic resource.

References

Adams, K. L., K. Song, P. G. Roessler, J. M. Nugent, J. L. Doyle, J. J. Doyle, and J. D. Palmer. 1999. Intracellular gene transfer in action: dual transcription and multiple silencings of nuclear and mitochondrial cox2 genes in legumes. Proceedings of the National Academy of Sciences of the United States of America 96: 13863–8.

Alt-Mörbe, J., J. L. Stryker, C. Fuqua, P. L. Li, S. K. Farrand, and S. C. Winans. 1996. The conjugal transfer system of Agrobacterium tumefaciens octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. Journal of Bacteriology 178: 4248–57.

Andam, C. P., and J. P. Gogarten. 2011. Biased gene transfer in microbial evolution. Nature Reviews. Microbiology 9: 543–55. doi:10.1038/nrmicro2593.

Archibald, J. M., M. B. Rogers, M. Toop, K.-I. Ishida, and P. J. Keeling. 2003. Lateral gene transfer and the evolution of plastid-targeted proteins in the secondary plastid-containing alga Bigelowiella natans. Proceedings of the National Academy of Sciences of the United States of America 100: 7678–83. doi:10.1073/pnas.1230951100.

Avery, O. T., C. M. Macleod, and M. McCarty. 1944. 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. The Journal of

Experimental Medicine 79: 137-158
Baumdicker, F., W. R. Hess, and P. Pfaffelhuber. 2010. The diversity of a distributed genome in bacterial populations. The Annals of Applied Probability 20: 1567–1606.
Baumdicker, F., W. R. Hess, and P. Pfaffelhuber. 2012. The infinitely many genes model for the distributed genome of bacteria. Genome Biology and Evolution 4: 443–56. doi:10.1093/gbe/evs016.
Biller, S. J., P. M. Berube, D. Lindell, and S. W. Chisholm. 2015. Prochlorococcus: the structure and function of collective diversity. Nature Reviews. Microbiology 13: 13–27. doi:10.1038/nrmicro3378.
Bolotin, E., and R. Hershberg. 2015. Gene Loss Dominates As a Source of Genetic Variation within Clonal Pathogenic Bacterial Species. Genome Biology and Evolution 7: 2173–2187. doi:10.1093/gbe/evv135.
Bolotin, E., and R. Hershberg. 2016. Bacterial intra-species gene loss occurs in a largely clocklike manner mostly within a pool of less conserved and constrained genes. Scientific Reports 6: 35168. doi:10.1038/srep35168.
Borgeaud, S., L. C. Metzger, T. Scrignari, and M. Blokesch. 2015. The type VI secretion system of Vibrio cholerae fosters horizontal gene transfer. Science 347: 63–7. doi:10.1126/science.1260064.
Boucher, Y., C. J. Douady, R. T. Papke, D. A. Walsh, M. E. Boudreau, C. L. Nesbo, R. J. Case, and W. F. Doolittle. 2003. Lateral gene transfer and the origins of prokaryotic groups. Annual Review of Genetics 37: 283–328.
Brussow, H., C. Canchaya, and W.-D. Hardt. 2004. Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. Microbiology and Molecular Biology Reviews 68: 560–602. doi:10.1128/MMBR.68.3.560-602.2004.
Carroll, L. 1871. Through the Looking-Glass. first edit. London, United Kingdom: Mamillan.
Carroll, L., and M. Gardner. 1993. The Annotated Alice: Alice's Adventures in Wonderland & Through the Looking Glass. New York City, NY: Random House Value Publishing.
Chaturongakul, S., and P. Ounjai. 2014. Phage-host interplay: examples from tailed phages and Gram-negative bacterial pathogens. Frontiers in Microbiology 5: 442. doi:10.3389/fmicb.2014.00442.
Chimileski, S., K. Dolas, A. Naor, U. Gophna, and R. T. Papke. 2014. Extracellular DNA metabolism in Haloferax volcanii. Frontiers in Microbiology 5: 57. doi:10.3389/fmicb.2014.00057.
Claverys, J.-P., and L. S. Håvarstein. 2007. Cannibalism and fratricide: mechanisms and raisons d'être. Nature Reviews. Microbiology 5: 219–29. doi:10.1038/nrmicro1613.
Cordero, O. X., H. Wildschutte, B. Kirkup, S. Proehl, L. Ngo, F. Hussain, F. Le Roux, T. Mincer, et al. 2012. Ecological populations of bacteria act as socially cohesive units of antibiotic production and resistance. Science 337: 1228–31. doi:10.1126/science.1219385.
Cumby, N., A. R. Davidson, and K. L. Maxwell. 2012. The moron comes of age. Bacteriophage 2: e23146. doi:10.4161/bact.23146.
Dandekar, A. A., S. Chugani, and E. P. Greenberg. 2012. Bacterial quorum sensing and metabolic incentives to cooperate. Science 338: 264–6. doi:10.1126/science.1227289.
Dawkins, R. 1976. The Selfish Gene. Oxford, Great Britain: Oxford University Press.

Derbyshire, K. M., and T. A. Gray. 2014. Distributive Conjugal Transfer: New Insights into Horizontal Gene Transfer and Genetic Exchange in Mycobacteria. Microbiology Spectrum 2. doi:10.1128/microbiolspec.MGM2-0022-2013.

Drescher, K., C. D. Nadell, H. A. Stone, N. S. Wingreen, and B. L. Bassler. 2014. Solutions to the public goods dilemma in bacterial biofilms. Current Biology 24: 50–5. doi:10.1016/j.cub.2013.10.030.

Dykhuizen, D. E., and L. Green. 1991. Recombination in Escherichia coli and the definition of biological species. Journal of Bacteriology 173: 7257–7268.

Fournier, G. P., and J. P. Gogarten. 2008. Evolution of acetoclastic methanogenesis in Methanosarcina via horizontal gene transfer from cellulolytic Clostridia. Journal of Bacteriology 190: 1124–7. doi:10.1128/JB.01382-07.

Fullmer, M. S., S. M. Soucy, and J. P. Gogarten. 2015. The pan-genome as a shared genomic resource: mutual cheating, cooperation and the black queen hypothesis. Frontiers in Microbiology 6: 728. doi:10.3389/fmicb.2015.00728.

Gogarten, J. P. 1995. The early evolution of cellular life. Trends in Ecology and Evolution 10: 147–151.

Gogarten, J. P., and J. P. Townsend. 2005. Horizontal gene transfer, genome innovation and evolution. Nature Reviews. Microbiol 3: 679–87.

Gogarten, J. P., W. F. Doolittle, and J. G. Lawrence. 2002. Prokaryotic evolution in light of gene transfer. Molecular Biology and Evolution 19: 2226–38.

Graham, L. A., J. Li, W. S. Davidson, and P. L. Davies. 2012. Smelt was the likely beneficiary of an antifreeze gene laterally transferred between fishes. BMC Evolutionary Biology 12: 190. doi:10.1186/1471-2148-12-190.

Graur, D. 2016. Molecular and Genome Evolution. Sinauer Associates, Inc., Sunderland, MA, USA.

Griffith, F. 1928. The Significance of Pneumococcal Types. The Journal of Hygiene 27: 113–59.

Haegeman, B., and J. S. Weitz. 2012. A neutral theory of genome evolution and the frequency distribution of genes. BMC Genomics 13: 196. doi:10.1186/1471-2164-13-196.

Hendrix, R. W., J. G. Lawrence, G. F. Hatfull, and S. Casjens. 2000. The origins and ongoing evolution of viruses. Trends in Microbiology 8: 504–8.

Hilario, E., and J. P. Gogarten. 1993. Horizontal transfer of ATPase genes--the tree of life becomes a net of life. Biosystems 31: 111–9.

Hille, F., and E. Charpentier. 2016. CRISPR-Cas: biology, mechanisms and relevance. Philosophical transactions of the Royal Society of London. Series B, Biological Sciences 371. doi:10.1098/rstb.2015.0496.

Hoffmann, K. H., B. Rodriguez-Brito, M. Breitbart, D. Bangor, F. Angly, B. Felts, J. Nulton, F. Rohwer, et al. 2007. Power law rank-abundance models for marine phage communities. FEMS Microbiology Letters 273: 224–8. doi:10.1111/j.1574-6968.2007.00790.x.

Huang, J. 2013. Horizontal gene transfer in eukaryotes: the weak-link model. BioEssays 35: 868–75. doi:10.1002/bies.201300007.

Huang, J., and J. P. Gogarten. 2007. Did an ancient chlamydial endosymbiosis facilitate the establishment of primary plastids? Genome Biology 8: R99. doi:10.1186/gb-2007-8-6-r99.

Johnston, C., B. Martin, G. Fichant, P. Polard, and J.-P. Claverys. 2014. Bacterial

transformation: distribution, shared mechanisms and divergent control. Nature reviews. Microbiology 12: 181–96. doi:10.1038/nrmicro3199.

Joos, H., B. Timmerman, M. V Montagu, and J. Schell. 1983. Genetic analysis of transfer and stabilization of Agrobacterium DNA in plant cells. The EMBO Journal 2: 2151–60.

Kashtan, N., S. E. Roggensack, S. Rodrigue, J. W. Thompson, S. J. Biller, A. Coe, H. Ding, P. Marttinen, et al. 2014. Single-cell genomics reveals hundreds of coexisting subpopulations in wild Prochlorococcus. Science 344: 416–20. doi:10.1126/science.1248575.

Khomyakova, M., Ö. Bükmez, L. K. Thomas, T. J. Erb, and I. A. Berg. 2011. A methylaspartate cycle in haloarchaea. Science 331: 334–7. doi:10.1126/science.1196544.

Kolter, R., and E. P. Greenberg. 2006. Microbial sciences: the superficial life of microbes. Nature 441: 300–302. doi:10.1038/441300a.

Kyrpides, N. C., P. Hugenholtz, J. A. Eisen, T. Woyke, M. Göker, C. T. Parker, R. Amann, B. J. Beck, et al. 2014. Genomic encyclopedia of bacteria and archaea: sequencing a myriad of type strains. PLoS Biology 12: e1001920. doi:10.1371/journal.pbio.1001920.

Lang, A. S., O. Zhaxybayeva, and J. T. Beatty. 2012. Gene transfer agents: phage-like elements of genetic exchange. Nature reviews. Microbiology 10: 472–82. doi:10.1038/nrmicro2802.

Lang, A. S., A. B. Westbye, and J. T. Beatty. 2017. The Distribution, Evolution, and Roles of Gene Transfer Agents in Prokaryotic Genetic Exchange. Annual Review of Virology 4: 87–104. doi:10.1146/annurev-virology-101416-041624.

Lapierre, P., and J. P. Gogarten. 2009. Estimating the size of the bacterial pan-genome. Trends in Genetics 25: 107–10. doi: 10.1016/j.tig.2008.12.004.

Lawrence, J. G., and H. Ochman. 1997. Amelioration of bacterial genomes: rates of change and exchange. Journal of Molecular Evolution 44: 383–97.

Lawrence, J. G., and J. R. Roth. 1996. Selfish operons: horizontal transfer may drive the evolution of gene clusters. Genetics 143: 1843–1860.

Lederberg, J., and E. L. Tatum. 1946. Gene recombination in Escherichia coli. Nature 158: 558.

Lobkovsky, A. E., Y. I. Wolf, and E. V Koonin. 2013. Gene frequency distributions reject a neutral model of genome evolution. Genome Biology and Evolution 5: 233–42. doi:10.1093/gbe/evt002.

Lobkovsky, A. E., Y. I. Wolf, and E. V Koonin. 2014. Estimation of prokaryotic supergenome size and composition from gene frequency distributions. BMC Genomics 15: S14. doi:10.1186/1471-2164-15-S6-S14.

Lovelock, J. E. 1972. Gaia as seen through the atmosphere. Atmospheric Environment (1967) 6: 579–580. doi:10.1016/0004-6981(72)90076-5.

Lovelock, J. E., and L. Margulis. 2011. Atmospheric homeostasis by and for the biosphere: the gaia hypothesis. Tellus A 26. doi:10.3402/tellusa.v26i1-2.9731.

Lukjancenko, O., T. M. Wassenaar, and D. W. Ussery. 2010. Comparison of 61 sequenced Escherichia coli genomes. Microbial Ecology 60: 708–20. doi:10.1007/s00248-010-9717-3.

Lynch, M., and J. S. Conery. 2003. The origins of genome complexity. Science 302: 1401–4.

Mell, J. C., and R. J. Redfield. 2014. Natural Competence and the Evolution of DNA Uptake Specificity. Journal of Bacteriology 196: 1471–1483. doi:10.1128/JB.01293-13.

Mira, A., H. Ochman, and N. A. Moran. 2001. Deletional bias and the evolution of bacterial genomes. Trends in Genetics 17: 589–96.

Morris, J. J., R. E. Lenski, and E. R. Zinser. 2012. The Black Queen Hypothesis: evolution of dependencies through adaptive gene loss. mBio 3: e00036-12-. doi:10.1128/mBio.00036-12.

Morris, J. J., S. E. Papoulis, and R. E. Lenski. 2014. Coexistence of evolving bacteria stabilized by a shared black queen function. Evolution 68: 2960–71. doi:10.1111/evo.12485.

Naor, A., P. Lapierre, M. Mevarech, R. T. Papke, and U. Gophna. 2012. Low species barriers in halophilic archaea and the formation of recombinant hybrids. Current Biology : CB 22: 1444–8. doi:10.1016/j.cub.2012.05.056.

Naor, A., N. Altman-Price, S. M. Soucy, A. G. Green, Y. Mitiagin, I. Turgeman-Grott, N. Davidovich, J. P. Gogarten, et al. 2016. Impact of a homing intein on recombination frequency and organismal fitness. Proceedings of the National Academy of Sciences of the United States of America 113: E4654-61. doi:10.1073/pnas.1606416113.

Olendzenski, L., and J. P. Gogarten. 2009. Evolution of genes and organisms: the tree/web of life in light of horizontal gene transfer. Annals of the New York Academy of Sciences 1178: 137–145.

Oliveira, N. M., R. Niehus, and K. R. Foster. 2014. Evolutionary limits to cooperation in microbial communities. Proceedings of the National Academy of Sciences of the United States of America 111: 17941–17946. doi:10.1073/pnas.1412673111.

Omer, S., T. J. Harlow, and J. P. Gogarten. 2017. Does Sequence Conservation Provide Evidence for Biological Function? Trends in Microbiology 25. doi:10.1016/j.tim.2016.09.010.

Puigbò, P., A. E. Lobkovsky, D. M. Kristensen, Y. I. Wolf, and E. V Koonin. 2014. Genomes in turmoil: Quantification of genome dynamics in prokaryote supergenomes. BMC Biology 12: 66. doi:10.1186/s12915-014-0066-4.

Racimo, F., D. Marnetto, and E. Huerta-Sánchez. 2016. Signatures of Archaic Adaptive Introgression in Present-Day Human Populations. Molecular Biology and Evolution 34: msw216. doi:10.1093/molbev/msw216.

Ravin, A. W. 1963. Experimental approaches to the study of bacterial phylogeny. The American Naturalist 97: 307–318. doi:10.2307/2458469.

Rodriguez-Brito, B., L. Li, L. Wegley, M. Furlan, F. Angly, M. Breitbart, J. Buchanan, C. Desnues, et al. 2010. Viral and microbial community dynamics in four aquatic environments. The ISME Journal 4: 739–751. doi:10.1038/ismej.2010.1.

Rosenshine, I., R. Tchelet, and M. Mevarech. 1989. The mechanism of DNA transfer in the mating system of an archaebacterium. Science 245: 1387–9.

Shapiro, B. J., J. Friedman, O. X. Cordero, S. P. Preheim, S. C. Timberlake, G. Szabó, M. F. Polz, and E. J. Alm. 2012. Population genomics of early events in the ecological differentiation of bacteria. Science 336: 48–51. doi:10.1126/science.1218198.

Sonea, S. 1988a. A bacterial way of life. Nature 331: 216.

Sonea, S. 1988b. The global organism: A new view of bacteria. The Sciences 28: 38–45.

Soucy, S. M., J. Huang, and J. P. Gogarten. 2015. Horizontal gene transfer: Building the web of life. Nature Reviews. Genetics 16. doi:10.1038/nrg3962.

Stemmann, L., and E. Boss. 2012. Plankton and Particle Size and Packaging: From Determining Optical Properties to Driving the Biological Pump. Annual Review of Marine Science 4: 263–290. doi:10.1146/annurev-marine-120710-100853.

Stetter, K. O. 2013. A brief history of the discovery of hyperthermophilic life. Biochemical Society Transactions 41: 416–420. doi:10.1042/BST20120284.

Stewart, C. N., M. D. Halfhill, and S. I. Warwick. 2003. Transgene introgression from genetically modified crops to their wild relatives. Nature Reviews. Genetics 4: 806–17. doi:10.1038/nrg1179.

Swithers, K. S., S. M. Soucy, and J. P. Gogarten. 2012. The role of reticulate evolution in creating innovation and complexity. International Journal of Evolutionary Biology 2012: 418964. doi:10.1155/2012/418964.

Tatum, E. L., and J. Lederberg. 1947. Gene Recombination in the Bacterium Escherichia coli. Journal of Bacteriology 53: 673–84.

Tettelin, H., V. Masignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L. Ward, S. V. Angiuoli, J. Crabtree, et al. 2005. Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: implications for the microbial "pan-genome." Proceedings of the National Academy of Sciences of the United States of America 102: 13950–5.

Thurber, R. V. 2009. Current insights into phage biodiversity and biogeography. Current Opinion in Microbiology 12: 582–7. doi:10.1016/j.mib.2009.08.008.

Timmis, J. N., M. A. Ayliffe, C. Y. Huang, and W. Martin. 2004. Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. Nature Reviews. Genetics 5: 123–135. doi:10.1038/nrg1271.

Treangen, T. J., and E. P. C. Rocha. 2011. Horizontal Transfer, Not Duplication, Drives the Expansion of Protein Families in Prokaryotes. Plos Genetics 7. doi:e1001284 10.1371/journal.pgen.1001284.

Tyra, H. M., M. Linka, A. P. Weber, and D. Bhattacharya. 2007. Host origin of plastid solute transporters in the first photosynthetic eukaryotes. Genome Biology 8: R212.

Van Valen, L. 1973. A new evolutionary law. Evolutionary Theory 1: 1–30.

Vulic, M., F. Dionisio, F. Taddei, and M. Radman. 1997. Molecular keys to speciation: DNA polymorphism and the control of genetic exchange in enterobacteria. Proceedings of the National Academy of Sciences of the United States of America 94: 9763–9767. doi:10.1073/pnas.94.18.9763.

Welch, R. A., V. Burland, G. Plunkett, P. Redford, P. Roesch, D. Rasko, E. L. Buckles, S. R. Liou, et al. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America 99: 17020–17024.

White, C. E., and S. C. Winans. 2007. Cell-cell communication in the plant pathogen Agrobacterium tumefaciens. Philosophical Transactions of the Royal Society B: Biological Sciences 362: 1135–1148. doi:10.1098/rstb.2007.2040.

Woese, C. R., G. J. Olsen, M. Ibba, and D. Soll. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. Microbiology and Molecular Biology Reviews 64: 202–36.

Zhaxybayeva, O., W. F. Doolittle, R. T. Papke, and J. P. Gogarten. 2009. Intertwined evolutionary histories of marine Synechococcus and Prochlorococcus marinus. Genome Biology and Evolution 2009: 325–339. doi:10.1093/gbe/evp032.