

## Chapter 8.1

### The progenote, LUCA, and the root of the cellular tree of life

Johann Peter Gogarten

Id [orcid.org/0000-0001-6459-6518](https://orcid.org/0000-0001-6459-6518)

#### CONTENTS

##### 8.1.1 Introduction

##### 8.1.1.1 The Tree Metaphor to Describe Evolution

##### 8.1.1.2 The Tree in Light of Gene Transfer

##### 8.1.1.3 Most Recent Common Ancestors, Roots, and Stem Groups

##### 8.1.2 Placing the Root in the Tree of Cellular Life

##### 8.1.3 The Progenote Concept

##### 8.1.4 Inferring Properties of Last Universal Common Ancestor

##### References

### 8.1.1 Introduction

#### *8.1.1.1. The tree metaphor to describe evolution*

Trees have a long history to depict genealogies of individuals (family trees) and interactions between living systems, including the biblical tree of life, the world tree in Mesoamerican cultures and Yggdrasil in Norse mythology (Gogarten et al. 2008; Anonymous 2017). Mark Ragan in a well-illustrated article (Ragan 2009) describes trees and network diagrams that were used before Darwin to depict relationships between plants and animals. The first explicit description of evolution as a tree goes back to Jean-Baptiste Lamarck, although Peter Simon Pallas's description of the gradation between organisms came close to depicting evolution. Lamarck's tree-like diagram overlaid multiple chains-of-progress, each for a different lineage into a single diagram (Lamarck 1809). Charles Darwin in this *Origin of Species* (Darwin 1859) provides a poetic description of the Tree-of-Life as an image of evolution, and the tree diagrams in his notebooks (Darwin 1836) illustrate the use of trees to depict shared ancestry. In his notebook Darwin expressed concern about the tree metaphor, because in a botanical tree the whole tree is alive, including the root, whereas in the tree-of-life image only the top layer is represented by living organisms (Darwin 1836; Olendzenski and Gogarten 2009), and his concern has gained new prominence with the realization that molecular based phylogenies usually only access data from extant or recently extinct species, and that most organisms that ever existed on this planet belong to lineages that are now extinct (Zhaxybayeva and Gogarten 2004; Fournier et al. 2009; Fournier et al. 2015; Weigel 2017).

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

### ***8.1.1.2. The tree in light of gene transfer***

The exchange of genetic information between independent lines of descent and the fusion of independent lineages are deviations from a tree-like structure. The evolution of genomes undoubtedly is a highly-reticulated network. While gene transfer is indeed rampant in most bacterial and archaeal lineages, and even appears to be frequent in eukaryotes (Soucy et al. 2015), most genes are transferred between closely related organisms (Andam et al. 2011). The transfer rate to organisms belonging to different orders and classes is several orders of magnitude lower than that of within genus transfers (Williams et al. 2012). Over short periods of time most genes are passed on vertically (from mother to daughter cell). This majority signal over short periods of time can define the organismal lineage, even in the presence of rampant gene transfer. Garry Olsen used the metaphor of a rope to describe this. In a rope (the lineage) each fiber (the gene traveling through time) extends only over a very short distance. Nevertheless, the rope is a continuous reality. By analogy, the same can be said for the organismal lineage: Even if not a single gene travels through a lineage from beginning to end, the lineage can still be defined through the majority of genes being passed on over short time intervals (Zhaxybayeva and Gogarten 2004).

Instances of lineage fusion that occurred in case of the endosymbionts that evolved into mitochondria and plastids violate the tree paradigm, if one considers the host and the symbiont as equal contributors. However, these reticulations in the tree of life are rare (Martin and Herrmann 1998), in many or even all instances the host cell's genome dominates the gene content persisting in the symbiosis. In most instances the phylogenetic signals resulting from shared ancestry and from biased gene transfer reinforce one another (Andam et al. 2010; Pace et al. 2012). Therefore, while the history of genomes is highly reticulated, hope persists that careful examination of the phylogenetic signal retained in molecular phylogenies may result in an improved inference of life's early history (Williams et al. 2011). The phylogenetic information provided through transferred genes that persist in the recipient lineage might improve the reconstruction of early cellular evolution (Huang and Gogarten 2006; Szollosi et al. 2012) and allows to correlate evolutionary events in different parts of the tree of life (Gogarten 1995). However, emphasis has to be placed on "careful", because the uncritical application of computational screening tools may lead to unwarranted conclusions (Gogarten and Deamer 2016).

### ***8.1.1.3. Most recent common ancestors, roots, and stem groups.***

With the introduction of ribosomal RNA as a marker molecule it became possible to place unicellular anucleate organisms onto a tree-like phylogeny (Woese and Fox 1977a). One immediate realization was that prokaryotes fall into two groups, now known as Bacteria and Archaea, previously labelled as Eubacteria and Archaeobacteria. Willi Hennig developed a natural taxonomy in which proper groups are defined by a common ancestor that is only ancestor to members of this group (Hennig 1966). Such a group is known as monophyletic. A group whose members all trace back to the same ancestor, but this ancestor also gave rise to organisms that are not part of the group is known as paraphyletic, and usually not considered as

monophyletic [However, see (Ashlock 1971) who defines monophyletic to include both para- and holophyletic groups.] In the now traditional tree of life, the Archaea are the sistergroup to the eukaryotic nucleocytoplasm, and the ancestor of the bacteria and archaea is also the ancestor to the eukaryotes. Therefore, in a cladistics classification system, the prokaryotes are a paraphyletic group, and the name prokaryotes describes a grade (developmental stage) and not a proper group in a natural taxonomic system. However, naming groups based on properties has a long history, often makes intuitive sense, and many of these group labels are in common use (e.g., reptiles, whose most recent common ancestor also is ancestor to birds and mammals). Before the availability of molecular data, grades often were the only available classification scheme. For further discussion see the exchange of opinions between Ernst Mayr and Carl Woese (Mayr 1998; Woese 1998).

As an aside, a similar debate concerns the archaea: if the eukaryotic nucleocytoplasm emerged from within the archaeal domain [(Zaremba-Niedzwiedzka et al. 2017); see (Da Cunha et al. 2017; Levasseur et al. 2017)for conflicting opinions], then the archaea will either need to be considered paraphyletic, and/or eukaryotes and archaea need to be united into a single taxonomic group.

In considering a tree, or part of a tree, the most recent common ancestor (MRCA) of a group is the organism placed at the deepest (earliest) split between lineages leading to members of the group. In the field of evolutionary biology, this MRCA is often described as the root of the group. In case of the tree of life that includes all cellular organisms, the root of the tree of life refers to that organism that existed at the deepest split, and this organism is also known as the Last Universal Common Ancestor, the MRCA of all cellular life, or as the organismal cenancestor (Fitch and Upper 1987). This root of a phylogenetic tree is very different from a botanical tree, where the root is located at the base of the stem.

The stem group of a taxon is defined as the group of those extinct organisms that branch off the lineage leading to the MRCA of that taxon. In case of the organismal LUCA, molecular evidence suggests that deeper branching lineages existed and contributed some genes through horizontal gene transfer to extant lineages (Fournier et al. 2009; Fournier et al. 2015).

### **8.1.2 Placing the Root in the Tree of Cellular Life.**

The tree of cells is embedded into network of gene trees. To find the root of a phylogenetic tree, different approaches are employed. The most common and widely accepted approach is to use an outgroup. In case of the tree of life, no organism can function as an outgroup, because in case of the tree of life all organisms by definition are part of the ingroup. However, ancient gene duplications that occurred before LUCA can provide an outgroup to molecular phylogenies (Gogarten et al. 1989; Iwabe et al. 1989; Brown and Doolittle 1995; Gribaldo and Cammarano 1998). Margaret Dayhoff first had suggested the approach to use an ancient gene duplication to provide an outgroup (Dayhoff 1972). For most ancient duplicated genes, this approach places the root between bacteria on one side of the root, and the archaea and the eukaryotic nucleocytoplasm on the other side. Using the unrooted tree of the translation machinery as reference, the root is

placed on the branch leading from the central trifurcation to the bacterial domain; therefore, this placement of the root is also known as placing the root on the bacterial branch. The placement of LUCA on the branch connecting bacteria and archaea was also inferred using ancestral sequence reconstruction for ribosomal proteins and identifying the place in the unrooted tree of ribosomal proteins that has an amino acid composition that likely reflects the signal from the assembly of the genetic code, i.e., proteins that evolved before the genetic code evolved to include all of today's genetically encoded amino acids should not include the later added amino acids in conserved positions (Fournier and Gogarten 2010). A few scientist place LUCA in different places in the tree of life, either within the bacteria (e.g., (Cavalier-Smith 2002; Skophammer et al. 2007)), or within the archaea (Di Giulio 2007; Kim and Caetano-Anollés 2011)). Often these rootings are based on singular characters, e.g., split genes (Di Giulio 2007) or the outer membrane of the Gram negative bacteria (Cavalier-Smith 2002), or novel approaches to reconstruct evolution (Cejchan 2004; Skophammer et al. 2007; Swithers et al. 2011). While the placement of the root remains under debate, most scientist, including former critics of the rooting on the bacterial branch, agree that the root of the cellular tree of life is located between Bacteria and Archaea, and with the eukaryotic nucleocytoplasm grouping on the archaeal side of the tree (Puigbò et al. 2013; Forterre 2015).

### 8.1.3 The Progenote Concept

Woese and Fox (1977b) defined the progenote as a hypothetical stage in evolution that existed before a strict coupling between geno- and phenotype had emerged, i.e., before the typical features characteristic of today's prokaryotes had emerged: *"Eucaryotes did arise from procaryotes, but only in the sense that the procaryotic is an organizational, not a phylogenetic distinction. In analogous fashion procaryotes arose from simpler entities. The latter are properly called progenotes, because they are still in the process of evolving the relationship between genotype and phenotype"*. The progenote concept is related to the Darwinian threshold (Woese 2002) or Darwinian transition (Goldenfeld et al. 2017). After living system passed this threshold, natural selection acting on cells drove the optimization of transcription and translation of genetically encoded systems. Below this threshold life is assumed to have been characterized by a high rate of gene exchange. A high gene transfer rate in communal entities also allows for smaller genomes (Lawrence 1999) and may help avoid an error catastrophe (Eigen 1971; Biebricher and Eigen 2005). Otto Kandler described the same concept under the term pre-cell populations (Kandler 1994).

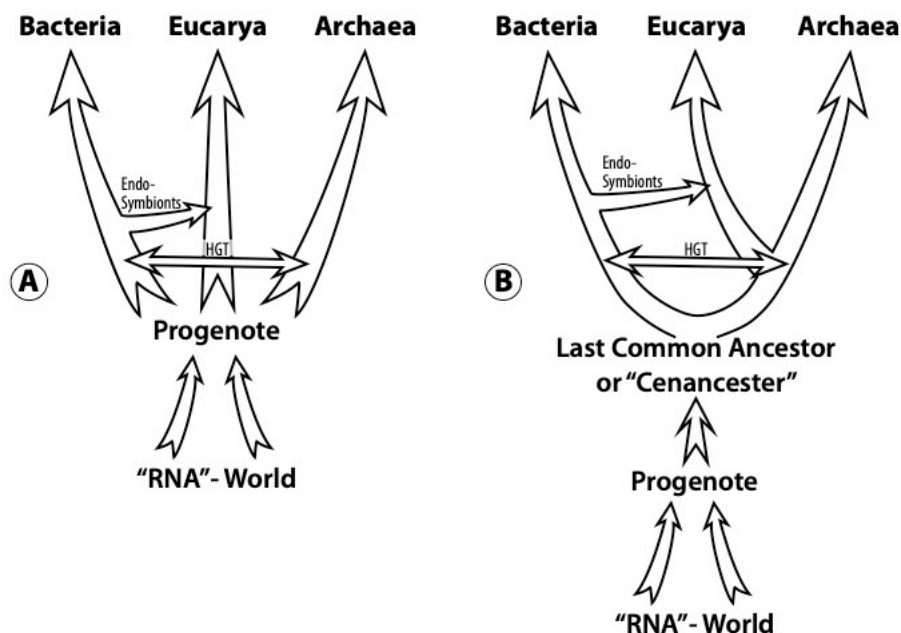


Figure 8.1.1. Two versions of the tree of life indicating the possible relations between the Last Common Ancestor or cenancestor of all extant cellular life, and the progenote-phase of early evolution. The progenote describes a period in the evolution of cells before a strict coupling between geno- and phenotype had been established. During this phase cells or pre-cells frequently exchanged genes with one another (see text for discussion). The sketch in panel A reflects the view popularized through the work of Carl Woese and Otto Kandler (Woese and Fox, 1977b; Kandler, 1994). In this version of the tree of life, the three domains of life evolved independently from the progenote. The sketch in panel B recognizes that LUCA was a complex organism with an established coupling between geno- and phenotype (Gogarten and Taiz, 1992; Delaye et al., 2005; Kim and Caetano-Anollés, 2011; Goldman et al., 2013), and depicts the progenote as an earlier phase preceding LUCA.

#### 8.1.4 Inferring LUCA's Properties

Confusion regarding the term progenote resulted from Woese and Fox assuming that the most recent ancestor of the three domains was a progenote (Woese and Fox 1977). In their and Kandler's vision (Kandler 1994) all three cellular domains independently trace their ancestry back to the progenote phase of life (see Figure 8.1.1 panel A).

In contrast, the many properties shared by all cells, and the fact that all ancient duplicated genes, including the aminoacyl tRNA synthetases themselves (Wolf et al. 1999; Woese et al. 2000;

Fournier et al. 2011), make use of the same twenty genetically encoded amino acids, and the observation that among the ancient duplicated genes are two subunits of the ATP synthase that translates transmembrane gradients for ions (sodium or hydrogen ions) into chemical energy stored in the ATP molecule) (Gogarten et al. 1989) suggest that LUCA was a cellular entity not too dissimilar from a modern prokaryote that possessed ribosomes for mRNA directed protein biosynthesis, aminoacyl tRNA synthetase proteins to charge tRNAs, and membranes that were used for chemiosmotic coupling (Gogarten and Taiz 1992) (see Figure 8.1.1 panel B).

As discussed in section 8.1.1.2., gene transfer has not erased all information that can be used to trace the evolutionary history of cell, and in many instances gene transfer has generated signals that allow to define groups and correlate evolutionary events in different parts of the tree. However, when trying to infer LUCA's properties from molecular phylogenies, wrong conclusions can be reached, when gene transfer events are mistaken for lines of vertical descent. In case of proteins with nearly universal distribution, such as amino acyl tRNA synthetase or the F/V/A-ATPases careful analysis and comparison to other conserved proteins allows identification of branches in a molecular phylogeny that correspond to gene transfer event (Ibba et al. 1999; Wolf et al. 1999; Zhaxybayeva et al. 2005; Lapierre et al. 2006). However, this task is more challenging, when genes are considered that do not have near universal distribution. The automated detection of genes that were present in LUCA faces two difficulties: either too many gene families are removed from consideration, because they show signs of horizontal gene transfer, or a gene transfer event that took place between archaea and bacteria is mistaken for the branch containing LUCA (Gogarten and Deamer 2016). Weiss et al. (Weiss et al. 2016) used a computational pipeline to identify genes present in LUCA. Many of the genes they identified as being present in LUCA mistook a gene transfer event between the domains for the location of LUCA. This is even true for genes that according to more careful analyses were present in LUCA, such as the phenylalanine amino acyl tRNA synthetase (pheRS) subunits. The LUCA branch identified in Weiss et al. (Weiss et al. 2016) for these gene families corresponds to a known transfer of the archaeal version of this enzyme to Spirochetes (Andam and Gogarten 2011), whereas most of the bacterial pheRS genes were not included in the analysis. At present, reliable inference of genes to be present in LUCA is limited to genes with nearly universal distribution. Analysis of less conserved genes encoding functions in metabolism faces hurdles, such as the decay of phylogenetic information and difficulties to identify gene transfer events. However, progress in gene tree - species tree reconciliation (Boussau et al. 2013; Sjostrand et al. 2014; Bansal et al. 2015; Wen and Nakhleh 2017) may make further progress possible.

Another source for inferring properties of early life is the analysis of genes that diverged before LUCA. At present these analyses are restricted to ATP synthase and aminoacyl tRNA synthetase subunits (Fournier and Gogarten 2007), and information gleaned from these data is restricted to the assembly of the genetic code, and the way tRNAs were charged during the early expansion of the genetic code (Fournier et al. 2011). While sequence data for most gene families may not be sufficiently conserved for this type of analyses, consideration of conserved protein folds might provide alternative avenues to characterize pre LUCA evolution (see (Koonin et al. 2006)).

## References

- Andam, C. P., and J. P. Gogarten. 2011. Biased gene transfer and its implications for the concept of lineage. *Biology direct* 6. BioMed Central Ltd: 47. doi:10.1186/1745-6150-6-47.
- Andam, C. P., D. Williams, and J. P. Gogarten. 2010. Natural taxonomy in light of horizontal gene transfer. *Biology & Philosophy* 25: 589–602. doi:10.1007/s10539-010-9212-8.
- Andam, C. P., G. P. Fournier, and J. P. Gogarten. 2011. Multilevel populations and the evolution of antibiotic resistance through horizontal gene transfer. *FEMS Microbiology Reviews* 35: 756–767. doi:10.1111/j.1574-6976.2011.00274.x.
- Anonymous. 2017. Tree of life. [https://en.wikipedia.org/wiki/Tree\\_of\\_life](https://en.wikipedia.org/wiki/Tree_of_life). Accessed July 27.
- Ashlock, P. D. 1971. Monophyly and associated terms. *Systematic Zoology* 20: 63–69.
- Bansal, M. S., Y.-C. Wu, E. J. Alm, and M. Kellis. 2015. Improved gene tree error correction in the presence of horizontal gene transfer. *Bioinformatics* 31: 1211–1218. doi:10.1093/bioinformatics/btu806.
- Biebricher, C. K., and M. Eigen. 2005. The error threshold. *Virus Res* 107: 117–27.
- Boussau, B., G. J. Szollosi, L. Duret, M. Gouy, E. Tannier, and V. Daubin. 2013. Genome-scale coestimation of species and gene trees. *Genome Research* 23: 323–330. doi:10.1101/gr.141978.112.
- Brown, J. R., and W. F. Doolittle. 1995. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proceedings of the National Academy of Sciences* 92: 2441–5.
- Cavalier-Smith, T. 2002. The neomuran origin of archaeobacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int J Syst Evol Microbiol* 52: 7–76.
- Cejchan, P. A. 2004. LUCA, or just a conserved Archaeon?: comments on Xue et al. *Gene* 333: 47–50.
- Da Cunha, V., M. Gaia, D. Gadelle, A. Nasir, and P. Forterre. 2017. Lokiarchaea are close relatives of Euryarchaeota, not bridging the gap between prokaryotes and eukaryotes. Edited by Antonis Rokas. *PLoS genetics* 13: e1006810. doi:10.1371/journal.pgen.1006810.
- Darwin, C. 1836. *Charles Darwin's notebooks, 1836-1844, transcription published in 1987*. Edited by Paul H. Barrett ... [et al.]. Ithaca, N.Y.: Cornell University Press.
- Darwin, C. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. London: John Murray, Albemarle Street.
- Dayhoff, M. O. 1972. *Atlas of protein sequence and structure, Volume 5*. Washington, DC: National Biomedical Research Foundation.
- Eigen, M. 1971. Selforganization of matter and the evolution of biological macromolecules. *Die Naturwissenschaften* 58: 465–523.
- Fitch, W. M., and K. Upper. 1987. The phylogeny of tRNA sequences provides evidence for ambiguity reduction in the origin of the genetic code. *Cold Spring Harb Symp Quant Biol* 52: 759–67.
- Forterre, P. 2015. The universal tree of life: an update. *Frontiers in microbiology* 6. Frontiers Media SA: 717. doi:10.3389/fmicb.2015.00717.

- Fournier, G. P., and J. P. Gogarten. 2007. Signature of a primitive genetic code in ancient protein lineages. *J Mol Evol* 65: 425–36. doi:10.1007/s00239-007-9024-x.
- Fournier, G. P., and J. P. Gogarten. 2010. Rooting the Ribosomal Tree of Life. *Molecular Biology and Evolution* 27: 1792–1801. doi:10.1093/molbev/msq057.
- Fournier, G. P., J. Huang, and J. P. Gogarten. 2009. Horizontal gene transfer from extinct and extant lineages: biological innovation and the coral of life. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364: 2229–39. doi:10.1098/rstb.2009.0033.
- Fournier, G. P., C. P. Andam, E. J. Alm, and J. P. Gogarten. 2011. Molecular Evolution of Aminoacyl tRNA Synthetase Proteins in the Early History of Life. *Origins of Life and Evolution of Biospheres* 41. Springer Netherlands: 621–632. doi:10.1007/s11084-011-9261-2.
- Fournier, G. P., C. P. Andam, and J. P. Gogarten. 2015. Ancient horizontal gene transfer and the last common ancestors. *BMC Evolutionary Biology* 15: 70. doi:10.1186/s12862-015-0350-0.
- Di Giulio, M. 2007. The tree of life might be rooted in the branch leading to Nanoarchaeota. *Gene* 401: 108–113.
- Gogarten, J. P. 1995. The early evolution of cellular life. *Trends in Ecology and Evolution* 10: 147–151.
- Gogarten, J. P., and D. Deamer. 2016. Is LUCA a thermophilic progenote? *Nature Microbiology* 1. doi:10.1038/nmicrobiol.2016.229.
- Gogarten, J. P., and L. Taiz. 1992. Evolution of proton pumping ATPases: Rooting the tree of life. *Photosynthesis Research* 33: 137–146. doi:10.1007/BF00039176.
- Gogarten, J. P., H. Kibak, P. Dittrich, L. Taiz, E. J. Bowman, B. J. Bowman, M. F. Manolson, R. J. Poole, et al. 1989. Evolution of the vacuolar H<sup>+</sup>-ATPase: implications for the origin of eukaryotes. *Proceedings of the National Academy of Sciences of the United States of America* 86: 6661–5.
- Gogarten, J. P., G. P. Fournier, and O. Zhaxybayeva. 2008. Gene Transfer and the Reconstruction of Life's Early History from Genomic Data. *Space Science Reviews* 135: 115–131.
- Goldenfeld, N., T. Biancalani, and F. Jafarpour. 2017. Universal biology and the statistical mechanics of early life. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 375: 20160341. doi:10.1098/rsta.2016.0341.
- Gribaldo, S., and P. Cammarano. 1998. The root of the universal tree of life inferred from anciently duplicated genes encoding components of the protein-targeting machinery. *Journal of molecular evolution* 47: 508–16.
- Hennig, W. 1966. *Phylogenetic systematics*. Edited by R. Zangerl. Urbana: University of Illinois Press.
- Huang, J., and J. P. Gogarten. 2006. Ancient horizontal gene transfer can benefit phylogenetic reconstruction. *Trends Genet* 22: 361–6.



- Ibba, M., A. W. Curnow, J. Bono, P. A. Rosa, C. R. Woese, and D. Soll. 1999. Archaeal aminoacyl-tRNA synthesis: unique determinants of a universal genetic code? *Biol Bull* 196: 335-6; discussion 336-7.
- Iwabe, N., K.-I. Kuma, M. Hasegawa, S. Osawa, T. M. Source, M. Hasegawat, S. Osawat, and T. Miyata. 1989. Evolutionary Relationship of Archaeobacteria, Eubacteria, and Eukaryotes Inferred from Phylogenetic Trees of Duplicated Genes. *Proceedings of the National Academy of Sciences of the United States of America Evolution* 86: 9355–9359. doi:10.1073/pnas.86.23.9355.
- Kandler, O. 1994. The early diversification of life. In *Early Life on Earth*, ed. S. Bengtson, 152–509. New York, NY, USA: Columbia Univ. Press.
- Kim, K. M., and G. Caetano-Anollés. 2011. The proteomic complexity and rise of the primordial ancestor of diversified life. *BMC evolutionary biology* 11: 140. doi:10.1186/1471-2148-11-140.
- Koonin, E. V., T. G. Senkevich, and V. V. Dolja. 2006. The ancient Virus World and evolution of cells. *Biology direct* 1: 29.
- Lamarck, J.-B. 1809. *Philosophie zoologique, Part II, p 463*. Paris: Dentu, et l’Auteur.
- Lapierre, P., R. Shial, and J. P. Gogarten. 2006. Distribution of F- and A/V-type ATPases in *Thermus scotoductus* and other closely related species. *Syst Appl Microbiol* 29: 15–23.
- Lawrence, J. G. 1999. Gene transfer and minimal genome size. In *Size Limits of Very Small Microorganisms*, ed. A. KNOLL, M. J. OSBORN, J. BAROSS, H. C. BERG, N. R. PACE, and M. SOGIN, 32–38. Washington, D.C: National Research Council.
- Levasseur, A., V. Merhej, E. Baptiste, V. Sharma, P. Pontarotti, and D. Raoult. 2017. The Rhizome of Lokiarchaeota Illustrates the Mosaicity of Archaeal Genomes. *Genome Biology and Evolution* 9: 2635–2639. doi:10.1093/gbe/evx208.
- Martin, W., and R. G. Herrmann. 1998. Gene transfer from organelles to the nucleus: how much, what happens, and Why? *Plant physiology* 118. American Society of Plant Biologists: 9–17.
- Mayr, E. 1998. Two empires or three? *Proceedings of the National Academy of Sciences of the United States of America* 95: 9720–3.
- Olendzenski, L., and J. P. Gogarten. 2009. Evolution of genes and organisms: the tree/web of life in light of horizontal gene transfer. *Ann N Y Acad Sci* 1178: 137–145. doi:NYAS4998 [pii]10.1111/j.1749-6632.2009.04998.x.
- Pace, N. R., J. Sapp, and N. Goldenfeld. 2012. Phylogeny and beyond: Scientific, historical, and conceptual significance of the first tree of life. *Proceedings of the National Academy of Sciences* 109: 1011–1018. doi:10.1073/pnas.1109716109.
- Puigbò, P., Y. I. Wolf, and E. V Koonin. 2013. Seeing the Tree of Life behind the phylogenetic forest. *BMC biology* 11. BioMed Central: 46. doi:10.1186/1741-7007-11-46.
- Ragan, M. A. 2009. Trees and networks before and after Darwin. *Biology direct* 4. BioMed Central: 43; discussion 43. doi:10.1186/1745-6150-4-43.
- Sjostrand, J., A. Tofigh, V. Daubin, L. Arvestad, B. Sennblad, and J. Lagergren. 2014. A Bayesian Method for Analyzing Lateral Gene Transfer. *Systematic Biology* 63: 409–420. doi:10.1093/sysbio/syu007.

- Skophammer, R. G., J. A. Servin, C. W. Herbold, and J. A. Lake. 2007. Evidence for a gram-positive, eubacterial root of the tree of life. *Mol Biol Evol* 24: 1761–8. doi:msm096 [pii] 10.1093/molbev/msm096.
- Soucy, S. M., J. Huang, and J. P. Gogarten. 2015. Horizontal gene transfer: building the web of life. *Nature Reviews in Genetics* 16: 472–482.
- Swithers, K. S., G. P. Fournier, A. G. Green, J. P. Gogarten, and P. Lapierre. 2011. Reassessment of the lineage fusion hypothesis for the origin of double membrane bacteria. *PLoS ONE* 6. doi:10.1371/journal.pone.0023774.
- Szollosi, G. J., B. Boussau, S. S. Abby, E. Tannier, and V. Daubin. 2012. Phylogenetic modeling of lateral gene transfer reconstructs the pattern and relative timing of speciations. *Proceedings of the National Academy of Sciences* 109: 17513–17518. doi:10.1073/pnas.1202997109.
- Weigel, C. 2017. Small Things Considered: To Not Get Lost in the Trees of Life (ToL).
- Weiss, M. C., F. L. Sousa, N. Mrnjavac, S. Neukirchen, M. Roettger, S. Nelson-Sathi, and W. F. Martin. 2016. The physiology and habitat of the last universal common ancestor. *Nature Microbiology* 1. Nature Publishing Group: 16116. doi:10.1038/nmicrobiol.2016.116.
- Wen, D., and L. Nakhleh. 2017. Coestimating Reticulate Phylogenies and Gene Trees from Multilocus Sequence Data. *Systematic Biology*. doi:10.1093/sysbio/syx085.
- Williams, D., G. P. P. Fournier, P. Lapierre, K. S. S. Swithers, A. G. G. Green, C. P. P. Andam, and J. P. P. Gogarten. 2011. A rooted net of life. *Biology direct* 6. BioMed Central: 45. doi:10.1186/1745-6150-6-45.
- Williams, D., J. P. Gogarten, and R. T. Papke. 2012. Quantifying homologous replacement of loci between haloarchaeal species. *Genome biology and evolution* 4: 1223–1244. doi:10.1093/gbe/evs098.
- Woese, C. R. 1998. Default taxonomy: Ernst Mayr's view of the microbial world. *Proc Natl Acad Sci U S A* 95: 11043–6.
- Woese, C. R. 2002. On the evolution of cells. *Proc Natl Acad Sci U S A* 99: 8742–7.
- Woese, C. R., and G. E. Fox. 1977a. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Sciences - USA* 74: 5088–90.
- Woese, C. R., and G. E. Fox. 1977b. The concept of cellular evolution. *J Mol Evol* 10: 1–6.
- Woese, C. R., G. J. Olsen, M. Ibba, and D. Soll. 2000. Aminoacyl-tRNA synthetases, the genetic code, and the evolutionary process. *Microbiol Mol Biol Rev* 64: 202–36.
- Wolf, Y. I., L. Aravind, N. V. Grishin, and E. V. Koonin. 1999. Evolution of aminoacyl-tRNA synthetases--analysis of unique domain architectures and phylogenetic trees reveals a complex history of horizontal gene transfer events. *Genome Res* 9: 689–710.
- Zaremba-Niedzwiedzka, K., E. F. Caceres, J. H. Saw, D. Bäckström, L. Juzokaite, E. Vancaester, K. W. Seitz, K. Anantharaman, et al. 2017. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541. Nature Research: 353–358. doi:10.1038/nature21031.
- Zhaxybayeva, O., and J. P. Gogarten. 2004. Cladogenesis, coalescence and the evolution of the three domains of life. *Trends in Genetics* 20: 182–187.

Zhaxybayeva, O., P. Lapierre, and J. P. Gogarten. 2005. Ancient gene duplications and the root(s) of the tree of life. *Protoplasma* 227: 53–64.