Mobile Genetic Elements 1:4, 304–308; November/December 2011; © 2011 Landes Bioscience

Goods-thinking vs. tree-thinking

Finding a place for mobile genetic elements

James O. McInerney,* Carla A. Cummins and Leanne S. Haggerty

Bioinformatics and Molecular Evolution Unit; Department of Biology; National University of Ireland Maynooth, Co.; Kildare, Ireland

While it has become increasingly clear that the Tree of Life hypothesis has limitations in its ability to describe the evolution of all evolving entities on the planet, there has been a marked reluctance to move away from the tree-based language. Ironically, while modifying the idea of the Tree of Life to the extent that it is only very distantly related to its original descriptions, there has been a very careful attempt to retain the language of tree-thinking. The recent movement away from a tree-thinking language toward a goods-thinking language and perspective is a significant improvement. In this commentary, we describe how goods-thinking can provide better descriptions of evolution, can integrate evolution with environment more closely and can offer an equal place for Mobile Genetic Elements and chromosomal elements in discussions of evolutionary history.

Keywords: phylogeny, horizontal gene transfer, public goods hypothesis, phage, virus, plasmid, mobile genetic element

Submitted: 11/10/11

Revised: 12/21/11

Accepted: 12/22/11

http://dx.doi.org/10.4161/mge.19153

*Correspondence to: James O. McInerney; Email: james.o.mcinerney@nuim.ie

Background

In the analysis of the overall evolution of life on the planet the Tree of Life (ToL) hypothesis has dominated for almost two centuries,[1] though networks, ladders and other kinds of structures have been employed both before and since the ToL hypothesis became established.[2] Naturally, because the ToL hypothesis predated the discovery of mobile genetic elements (MGEs) or even the discovery of genes, the initial formulation of the ToL hypothesis specifically dealt with, and was synthesized using the observed phenotypes of cellular life.

The discoveries of conjugation,[3] transduction,[4] transformation,[5] plasmids,[6] bacteriophage,[7] gene transfer agents[8,9] and nanotubes[10] have presented the ToL hypothesis with its greatest challenges because these processes and associated mobile genetic elements have the potential to disrupt the vertical inheritance pattern that is expected from the ToL hypothesis—they facilitate horizontal gene transfer (HGT). The past decade has seen a significant amount of debate concerning whether or not HGT is important,[11,12] irrelevant,[13,14] or inbetween.[15,16] One particularly interesting fact that has emerged from the sequencing of genomes came from the analysis of 10 million protein-coding genes and gene tags in sequenced eubacterial, archaebacterial and eukaryotic genomes as well as metagenomes. It was observed from this analysis that genes encoding transposases are the most abundant kinds of genes in nature.[17] These genes are responsible for facilitating the horizontal transfer of genetic material and testify to the importance, or at the very least, the success of such processes.

The upshot of our genome-level analyses is that HGT can easily be shown to be almost ubiquitous, frequent in some kinds of genes, less frequent in others, performed between cellular life forms and mobile genetic elements.[18,19] In fact, it now seems that one of the major restraints on HGT has nothing to do with phylogeny, rather it is the degree of a protein in its protein-protein interaction network.[20] In other words, proteins have a strong tendency to be involved in HGT, with this process being mitigated or moderated simply by the degree to which a protein interacts with other proteins. If it is

positioned centrally in the network (as judged by its degree), then it is less likely to be involved in successful HGT. Therefore, the ToL hypothesis has now been well and truly tested and shown to be an inadequate model for all life on the planet. The largest "statistical trend" has been suggested to be approximately 1.5% of the data,[16] though Leigh and coworkers,[21,22] have shown that even within the set of trees displaying this trend, there is significant incongruence. This, surely, is not what is envisioned by the ToL hypothesis.

One of the most obvious shortcomings of the ToL hypothesis is that it does not deal with all the evolving entities on the planet. Mobile elements have normally been frozen out of discussions of the grand schemes of evolution of life. They simply did not feature in the Tree of Life hypothesis and with notable exceptions (e.g., refs. 19, 23 and 24) they have not been included in Tree of Life diagrams or in discussions of the evolutionary relationships between cellular organisms and MGEs. This might seem to be permissible if MGEs played a very small role in the evolutionary history of life on the planet, but when the data are examined, we can see that MGEs have played an enormous role. Phage, which are important agents of HGT, for instance, are the most abundant life forms on earth, with ~10^{30} tailed phage particles on the planet and are responsible for 10^{25} infections per second.[25] We can see that cells (and in particular, prokaryotic cells) are hugely influenced by MGEs and we also see MGEs themselves are greatly influenced by cells.[18,26] Therefore, there surely must be a better means of thinking about the evolving entities on the planet than simply focusing on the ToL hypothesis and only considering small portions of the genomes of a fraction of the evolving entities. The evolutionary history of life on the planet is full of vertical and horizontal connections between all kinds of evolving entities. In **Figure 1** we demonstrate an ever more frequently seen kind of network diagram[27,28] that is displaying a very common motif in evolutionary biology. This diagram depicts the connections between genes that are found in enteric bacteria and some mobile genetic elements, in this case, plasmids.

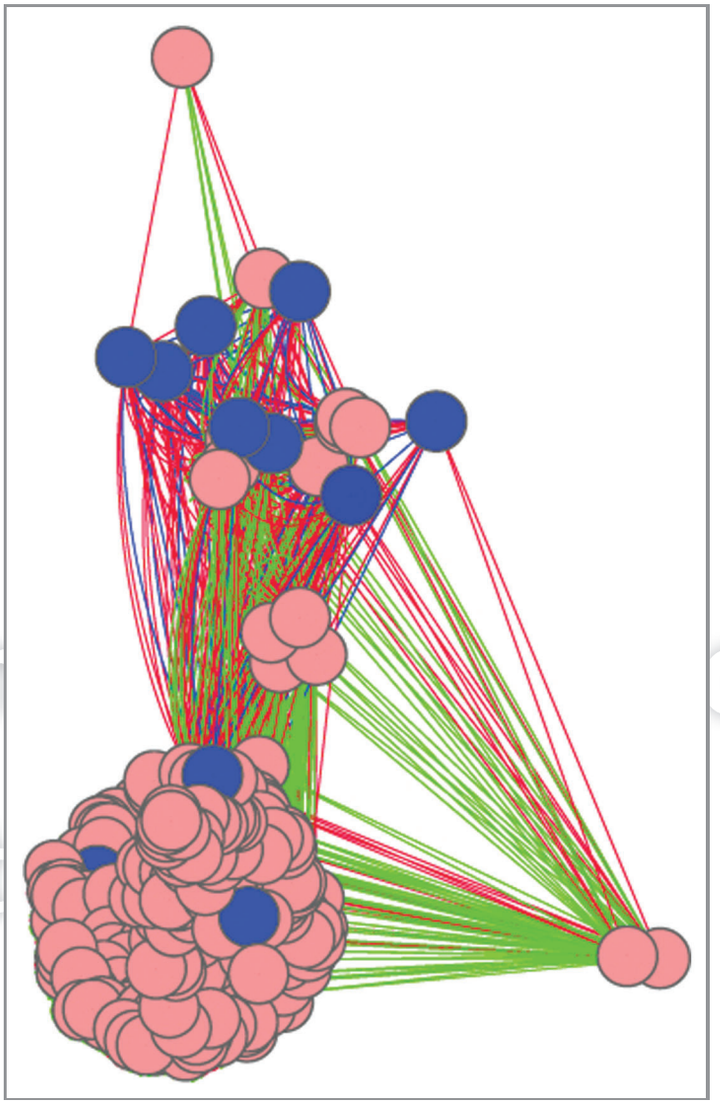

Figure 1. This is a network where the nodes represent genes and the edges represent links between genes where the sequence similarity is greater than 95% at the nucleotide level. The blue nodes are genes that are found on plasmids, while the brown nodes are genes that are found in cellular chromosomes. The blue edges link MGE genes with other MGE genes, the green edges link chromosomal genes with chromosomal genes and the red genes link MGE genes with chromosomal genes.

Every node in the network is a gene and every edge is a statement that the two connected nodes manifest greater than 95% sequence similarity. The length of the edge is determined by the actual sequence similarity. The interesting thing is that we can see all possible kinds of connections—chromosomal genes connected to chromosomal genes, MGE genes to MGE genes and chromosomal genes connected to MGE genes. This figure is clearly not depicting constantly diverging cellular organisms; it is a network showing the sharing of genes between all kinds of evolving entities.

Moving Away from Tree-Thinking

A useful way to construct hypotheses is to start at uncontroversial starting points—axioms that we can all agree on. With a goal to describe life in its most fundamental way, we might consider that the first axiom might be that all evolving entities are included in any hypothesis that dealt with the most fundamental

description of evolution. Next if we refer to the hundreds of thousands of experiments and observations that have concerned the inheritance of genetic information, we have made the observation that genes can be acquired by both vertical and horizontal transmission. So, this is the second axiom—the theory must encompass horizontal and vertical acquisition of genetic material.

Starting with these two axioms, we can easily see that ToL models have handicaps that are difficult to overcome. Many ToL models do not even accommodate these two most basic of axioms. In its place, we have proposed that that the best way to describe evolution is to consider evolving entities such as nucleotides or genes or operons or even genomes as genetic goods in the same sense as goods are often viewed in the discipline of economics.[29] We have called this the "public goods" hypothesis, for reasons we shall elaborate on later.[30] This is a very broad and fundamental view of evolutionary history. It encompasses all the patterns and processes we see in nature and at no time does it try to ignore patterns or evolving entities. The public goods hypothesis gets away from the vocabulary of tree-thinking and its associated history and does not require ad hoc amendments or qualifications in order to incorporate the observed data. Current incarnations of the ToL require footnotes relating to horizontal gene transfer, hand-waving in relation to mobile genetic elements and dismissal or avoidance of fusions of cellular and/or genetic elements.

In the field of economics, goods may be exchanged, moved, modified and amalgamated into larger goods, transported or even destroyed. Furthermore, goods can be classified according to the properties of excludability or rivalry (also known as subtractability). An excludable good is one where it is relatively easy to prevent others from accessing that good and a rivalrous good is one whose use by one individual effectively prevents its use by another. Naturally, the objects in evolutionary biology have their own properties and how we view them is likely to differ from how economists view goods, nonetheless, goods-thinking provides a useful perspective on evolutionary biology.

For the most part, it is difficult to see that genetic goods might be excludable. The nucleotides that are used by all evolving entities are the same and the genetic code always consists of triplets of codons, genes have promoters, start codons and stop codons and recombination (breaking and joining of nucleic acids) is cosmopolitan. The machinery of DNA replication and of translation of protein-coding genes is pretty universal—the exceptions for translation are the alternative genetic codes. However, for the most part, genetic material is not excludable and this is irrespective of whether the genetic material is found in a cell, on a plasmid or in a virus. **Figure 1** shows a network of gene sharing. There are three kinds of evolving entities on this network, represented by the nodes and the edges represent identifiably homologous regions that are shared. As can be seen, we have sharing between cellular life forms, between mobile genetic elements and between cellular and mobile elements. We know of no formulation of the ToL hypothesis that encompasses this kind of situation, however a goods-thinking approach is more than adequate to encompass the observed data and in this particular case, the genetic goods do not appear to be excludable.

It remains to be seen whether effective excludability is possible for some genetic goods and this line of thinking automatically suggests a program of research.

We then ask whether genes might be rivalrous or subtractable. This property refers to whether a good is available to be used by one individual simultaneous with its use by another individual. How we view rivalry might depend on whether we wish to view a gene as all copies of its orthologs or whether we wish to focus on a single copy of the gene itself. If we view a gene as all copies of the gene, then that gene is non-rivalrous, whereas the latter might indicate that it is rivalrous. Clearly, because the mechanism of replicating genes is found in all cellular life forms, then any kind of gene can be potentially replicated and therefore, it is difficult to make any kind of gene rivalrous.

As a consequence of the difficulty in making genes excludable or rivalrous, we view genes to be public goods for the most part. They are free to be inherited vertically from parent to offspring and they are free to be acquired horizontally. This does not mean that there are no constraints on gene movement, but this does not affect the definition of genes as public goods.

Just because we might view genes as public goods in a fundamental way does not mean that evolving entities have not privatised them in some ways. We have mentioned the alternative genetic codes, but we might also consider toxin-antitoxin genes[31] as being somewhat dependent on one another and therefore, they can effectively exclude other genes or genomes from having one without the other. Likewise, with plasmid incompatibility systems,[32] we find that plasmids are rivalrous for the cells in which they replicate, meaning that a particular cell might become a club for only one kind of plasmid with a particular incompatibility system. It is outside the scope of this manuscript to detail other situations where genetic goods might be privatized or brought into clubs or coalitions, but it is likely that such situations exist.

In a brief aside, we might consider the public availability of genes and proteins to be something that is a general feature of both genetic and proteic material. It would be wrong to consider proteins that are largely contained within a bacterial cell to be private goods for that bacterium. A suite of proteins capable of breaking down phenylacetate,[33] for instance, might all be contained within the cell that produced the proteins, however, removal of phenylacetate from the environment might benefit other organisms directly. Therefore, we might consider that even though these phenylacetate-degrading proteins appear to be private goods, the consequence of their existence is a public good and because the genes can be acquired by other organisms through appropriate vectors or transformation, the genes can also be considered public goods.

How does this Affect Mobile Genetic Elements?

In 2004, the discovery of a 1.2 megabase virus[24] led to speculation on whether this might represent a fourth 'domain' of life

(*sensu* Woese). A technical comment in response to this publication showed a phylogeny with one mimivirus gene occupying a phylogenetic position within the eukaryotes and not as a separate group outside the three domains.[34] The response to this comment was that "[…] the tradition is to deny viruses the status of bona fide living organisms and to a priori doubt the capacity of phylogenetic analyses to investigate their deepest origin."[35] However, since then it has been shown that mimivirus genes occupy many different phylogenetic positions and the positioning of mimivirus as a fourth domain of life is probably not sensible.[36] This exchange encapsulates the two sides of the argument—on one side it is held that cellular organisms are real life, whereas viruses are not and on the other side, viruses have all the traits necessary to be included in any discussions about life on the planet. Standard definitions of viruses usually cite their size or an inability to replicate autonomously as being the feature that separates them from cellular organisms. But with giant viruses like mimivirus, megavirus[37] and mamavirus[38] that are larger than many bacteria and with many intracellular bacteria being unable to replicate or live autonomously, then the distinctions are no longer so clear, yet the ToL treats them as entirely separate. Both kinds of evolving entity (cellular and viral) can contain genes for transcription, translation, replication, metabolisms of all kinds and so forth.

An analysis of *Escherichia coli*, has identified almost 16,000 different gene families that are to be found in the various genomes of this species,[39] with only approximately 6% of these gene families being found in all sequenced genomes of *E. coli*. However, just like Mimivirus, many of these genes have phylogenetic affinities that are incompatible with one another and many of the genes have only been found in one strain of *E. coli*. In many respects, mimivirus and *E. coli* have similar features. The only way in which they differ is that *E. coli* catalyzes its own replication.

Therefore, it seems unusual from this perspective to completely exclude viruses from hypotheses governing life on the planet. Goods-thinking makes no distinction between viruses and cellular organisms. In both cases, we view genes as goods that can be acquired vertically or horizontally and this includes all kinds of mobile genetic elements.

The Integration of Ecology with Evolution

The public goods hypothesis has the particular feature that ecology is allowed to play a much greater role in evolution than is generally acknowledged under a tree model. Under a tree model, we might expect an evolving entity to adapt to a particular environment in a gradual piecemeal fashion and even it might be expected that different parts of the phylogenetic tree would be adapted to different environments. What we see instead are high levels of diversity of organisms with mosaic genomes. In the public goods view of evolution, the environment would play a very big role and effectively a suite of genes that are adapted to a particular environment would operate in that environment, irrespective of the phylogenetic assignment of the cell that was replicating, transcribing and translating those genes. In other words, the environment, combined with population processes would select the collection of goods that existed in that environment, irrespective of which organisms were replication, transcribing and translating those goods and irrespective of whether the goods were on the chromosomes of cellular organisms, on plasmids or on viruses.

While the ToL hypothesis did not imply any particularly strong role for the environment, apart from the selection of fitter genetic variants, the public goods hypothesis implies that the environment is intimately involved in the rapid evolution of mosaic genotypes in order to procure the phenotypes that are best adapted.

Relativism is Unhelpful

In ascribing a more important role to MGEs, we would be at pains to point out that the evolutionary history of life on the planet has seen contributions from both vertical and horizontal gene transfer. Often it has been said that if we count the actual number of vertical vs. horizontal transfers of genes and we include the numbers of genes involved (usually thousands of genes are inherited via cell division at every cell division and only a small number are acquired via horizontal acquisition and only infrequently), then vertical transmission of genes has occurred much more often than horizontal transfer. However, this line of thinking can be very unhelpful if we wish to discover long-term consequences of genome change. Counting numbers of genes and prioritizing vertical gene transfer has led to an unwillingness to see how it is often the horizontal transfer of genes that has led to rapid response to environmental change in a lineage. A very important example has been the spread of plasmid-borne antibiotic resistance[40] in the past 50 y as a consequence of large-scale production of antibiotics. In the absence of MGEs, many lineages of pathogens would surely nowadays be almost extinct. However, acquisition of these goods has led to the proliferation of the organisms with these goods.

Conclusion

Darwin's paradigm has broadened and now includes descent with modification due to error-prone polymerases, genomic modification by horizontal gene transfer, natural selection on new variants and in the case of some cellular entities, such as animals, speciation as envisioned by people like Ernst Mayr.[41]

Acknowledgments

This work was funded in part by Science Foundation Ireland Research Frontiers Programme grants 09/RFP/2510 and 07/RFP/EEEOBF654.

Glossary

Goods: Commodities merchandise or wares. The word "goods" is quite different to the word good and does not necessarily connote 'beneficial', instead it connotes objects that can be obtained, exchanged or incorporated into other goods.

Private goods: goods that are excludable (relatively easy to prevent others from obtaining the good) and rivalrous (when the good is being used by one individual, it is not available to another individual).

A cup of coffee might be considered a private good.

Club goods: goods that are excludable, but not rivalrous.

Common goods: goods that are not excludable, but are rivalrous.

Public goods: goods that are neither excludable nor rivalrous.

Gene-sharing network: A network diagram where every node is either a gene or a genome and every edge is a statement of homology linking the nodes.

References

1. McInerney JO, Cotton JA, Pisani D. The prokaryotic tree of life: past, present... and future? Trends Ecol Evol 2008; 23:276-81; PMID:18367290; http://dx.doi.org/10.1016/j.tree.2008.01.008
2. Ragan MA. Trees and networks before and after Darwin. [discussion]. Biol Direct 2009; 4:43, discussion 43; PMID:19917100; http://dx.doi.org/10.1186/1745-6150-4-43
3. Lederberg J, Tatum EL. Gene recombination in Escherichia coli. Nature 1946; 158:558; PMID:21001945; http://dx.doi.org/10.1038/158558a0
4. Zinder ND, Lederberg J. Genetic exchange in Salmonella. J Bacteriol 1952; 64:679-99; PMID:12999698
5. Griffith F. The Significance of Pneumococcal Types. J Hyg (Lond) 1928; 27:113-59; PMID:20474956; http://dx.doi.org/10.1017/S0022172400031879
6. Hayes W. [Observations on a transmissible agent determining sexual differentiation in Bacterium coli]. J Gen Microbiol 1953; 8:72-88; PMID:13035034
7. Duckworth DH. "Who discovered bacteriophage?". Bacteriol Rev 1976; 40:793-802; PMID:795414
8. Lang AS, Beatty JT. Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of Rhodobacter capsulatus. Proc Natl Acad Sci U S A 2000; 97:859-64; PMID:10639170; http://dx.doi.org/10.1073/pnas.97.2.859
9. McDaniel LD, Young E, Delaney J, Ruhnau F, Ritchie KB, Paul JH. High frequency of horizontal gene transfer in the oceans. Science 2010; 330:50; PMID:20929803; http://dx.doi.org/10.1126/science.1192243
10. Dubey GP, Ben-Yehuda S. Intercellular nanotubes mediate bacterial communication. Cell 2011; 144:590-600; PMID:21335240; http://dx.doi.org/10.1016/j.cell.2011.01.015
11. Beiko RG, Harlow TJ, Ragan MA. Highways of gene sharing in prokaryotes. Proc Natl Acad Sci U S A 2005; 102:14332-7; PMID:16176988; http://dx.doi.org/10.1073/pnas.0504068102
12. Doolittle WF. Phylogenetic classification and the universal tree. Science 1999; 284:2124-9; PMID:10381871; http://dx.doi.org/10.1126/science.284.5423.2124
13. Daubin V, Moran NA, Ochman H. Phylogenetics and the cohesion of bacterial genomes. Science 2003; 301:829-32; PMID:12907801; http://dx.doi.org/10.1126/science.1086568
14. Kurland CG. Something for everyone. Horizontal gene transfer in evolution. EMBO Rep 2000; 1:92-5; PMID:11265763; http://dx.doi.org/10.1093/embo-reports/kvd042
15. Creevey CJ, Fitzpatrick DA, Philip GK, Kinsella RJ, O'Connell MJ, Pentony MM, et al. Does a tree-like phylogeny only exist at the tips in the prokaryotes? Proc Biol Sci 2004; 271:2551-8; PMID:15615680; http://dx.doi.org/10.1098/rspb.2004.2864
16. Puigbò P, Wolf YI, Koonin EV. The tree and net components of prokaryote evolution. Genome Biol Evol 2010; 2:745-56; PMID:20889655; http://dx.doi.org/10.1093/gbe/evq062
17. Aziz RK, Breitbart M, Edwards RA. Transposases are the most abundant, most ubiquitous genes in nature. Nucleic Acids Res 2010; 38:4207-17; PMID:20215432; http://dx.doi.org/10.1093/nar/gkq140
18. Brilli M, Mengoni A, Fondi M, Bazzicalupo M, Liò P, Fani R. Analysis of plasmid genes by phylogenetic profiling and visualization of homology relationships using Blast2Network. BMC Bioinformatics 2008; 9:551; PMID:19099604; http://dx.doi.org/10.1186/1471-2105-9-551
19. Halary S, Leigh JW, Cheaib B, Lopez P, Bapteste E. Network analyses structure genetic diversity in independent genetic worlds. Proc Natl Acad Sci U S A 2010; 107:127-32; PMID:20007769; http://dx.doi.org/10.1073/pnas.0908978107
20. Cohen O, Gophna U, Pupko T. The complexity hypothesis revisited: connectivity rather than function constitutes a barrier to horizontal gene transfer. Mol Biol Evol 2011; 28:1481-9; PMID:21149642; http://dx.doi.org/10.1093/molbev/msq333
21. Leigh JW, Lapointe FJ, Lopez P, Bapteste E. Evaluating phylogenetic congruence in the post-genomic era. Genome Biol Evol 2011; 3:571-87; PMID:21712432; http://dx.doi.org/10.1093/gbe/evr050
22. Leigh JW, Schliep K, Lopez P, Bapteste E. Let them fall where they may: congruence analysis in massive phylogenetically messy data sets. Mol Biol Evol 2011; 28:2773-85; PMID:21527387; http://dx.doi.org/10.1093/molbev/msr110
23. Brüssow H. The not so universal tree of life or the place of viruses in the living world. Philos Trans R Soc Lond B Biol Sci 2009; 364:2263-74; PMID:19571246; http://dx.doi.org/10.1098/rstb.2009.0036
24. Raoult D, Audic S, Robert C, Abergel C, Renesto P, Ogata H, et al. The 1.2-megabase genome sequence of Mimivirus. Science 2004; 306:1344-50; PMID:15486256; http://dx.doi.org/10.1126/science.1101485
25. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol 2005; 3:722-32; PMID:16138100; http://dx.doi.org/10.1038/nrmicro1235
26. Poullain V, Gandon S, Brockhurst MA, Buckling A, Hochberg ME. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. Evolution 2008; 62:1-11; PMID:18005153
27. Bittner L, Halary S, Payri C, Cruaud C, de Reviers B, Lopez P, et al. Some considerations for analyzing biodiversity using integrative metagenomics and gene networks. Biol Direct 2010; 5:47; PMID:20673351; http://dx.doi.org/10.1186/1745-6150-5-47
28. Beauregard-Racine J, Bicep C, Schliep K, Lopez P, Lapointe FJ, Bapteste E. Of woods and webs: possible alternatives to the tree of life for studying genomic fluidity in E. coli. [discussion]. Biol Direct 2011; 6:39-, discussion 39; PMID:21774799; http://dx.doi.org/10.1186/1745-6150-6-39
29. Samuelson PA. The Pure Theory of Public Expenditure. Rev Econ Stat 1954; 36:387-9; http://dx.doi.org/10.2307/1925895
30. McInerney JO, Pisani D, Bapteste E, O'Connell MJ. The public goods hypothesis for the evolution of life on Earth. Biol Direct 2011; 6:41; PMID:21861918; http://dx.doi.org/10.1186/1745-6150-6-41
31. Gerdes K, Christensen SK, Løbner-Olesen A. Prokaryotic toxin-antitoxin stress response loci. Nat Rev Microbiol 2005; 3:371-82; PMID:15864262; http://dx.doi.org/10.1038/nrmicro1147
32. Novick RP. Plasmid incompatibility. Microbiol Rev 1987; 51:381-95; PMID:3325793
33. Martin FJ, McInerney JO. Recurring cluster and operon assembly for Phenylacetate degradation genes. BMC Evol Biol 2009; 9:36; PMID:19208251; http://dx.doi.org/10.1186/1471-2148-9-36
34. Moreira D, López-García P. Comment on "The 1.2-megabase genome sequence of Mimivirus". [author reply]. Science 2005; 308:1114-, author reply 1114; PMID:15905382; http://dx.doi.org/10.1126/science.1110820
35. Moreira D, López-García P, Raoult D, Claverie JM. Comment on "The 1.2-megabase genome sequence of Mimivirus". Science 2005; 308:1114-, author reply 1114; PMID:15905382; http://dx.doi.org/10.1126/science.1110820
36. Moreira D, Brochier-Armanet C. Giant viruses, giant chimeras: the multiple evolutionary histories of Mimivirus genes. BMC Evol Biol 2008; 8:12; PMID:18205905; http://dx.doi.org/10.1186/1471-2148-8-12
37. Arslan D, Legendre M, Seltzer V, Abergel C, Claverie JM. Distant Mimivirus relative with a larger genome highlights the fundamental features of Megaviridae. Proc Natl Acad Sci U S A 2011; 108:17486-91; PMID:21987820; http://dx.doi.org/10.1073/pnas.1110889108
38. Colson P, Yutin N, Shabalina SA, Robert C, Fournous G, La Scola B, et al. Viruses with more than 1,000 genes: Mamavirus, a new Acanthamoeba polyphaga mimivirus strain, and reannotation of Mimivirus genes. Genome Biol Evol 2011; 3:737-42; PMID:21705471; http://dx.doi.org/10.1093/gbe/evr048
39. Lukjancenko O, Wassenaar TM, Ussery DW. Comparison of 61 sequenced Escherichia coli genomes. Microb Ecol 2010; 60:708-20; PMID:20623278; http://dx.doi.org/10.1007/s00248-010-9717-3
40. Cohen SN, Chang AC, Hsu L. Nonchromosomal antibiotic resistance in bacteria: genetic transformation of Escherichia coli by R-factor DNA. Proc Natl Acad Sci U S A 1972; 69:2110-4; PMID:4559594; http://dx.doi.org/10.1073/pnas.69.8.2110
41. Mayr E. Animal Species and Evolution. Cambridge, Mass.: Belknap Press, 1963.