The Evolution of Life Compared to Computers:
Coincidences, Similarities or Logic? A Simple Way to Understand the Phylogenetic Incongruence

Karla Suemy Clemente Yotoko*

Laboratory of Bioinformatics and Evolution, Federal University of Viçosa, Av. Peter Henry Rolfs, Campus Universitário, 36570-000 Viçosa, MG, Brazil
*Corresponding author: karla.yotoko@gmail.com

Victor Satler Pylro

Laboratory of Environmental Biotechnology and Biodiversity, Microbiology Department, Federal University of Viçosa, Av. Peter Henry Rolfs, Campus Universitário, 36570-000 Viçosa, MG, Brazil

Copyright© 2013 K. S. C. Yotoko and V. S. Pylro. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Overview: “Recent studies involving whole genome comparisons (phylogenomics) have shown that living creatures and computers function much more similarly than one might imagine based on a superficial analysis. These results suggest that the separate evolution of the two systems had similar processes in which exchange of information was decisive in the speed with which the diversity and complexity of both systems increased. From this perspective we can easily discuss phylogenetic incongruence in a concise and simple way.”

Computers and living creatures show some clear functional similarities. For instance, both systems capture external “stimuli” (the environment in the case of organisms, and the user’s desire in the case of computers) and perform “tasks” that provide different “answers” to these stimuli. It is widely known that bacteria only express a particular gene when a specific nutrient is present in the medium as well as a computer only run a particular program once the user accesses it and
enters data to be evaluated. Given these similarities, is it possible to draw a parallel and argue that these systems have also evolved in similar ways? As the evolution of computers is relatively recent and well documented, we just need to unravel the evolutionary history of living beings to try to answer this question. But have evolutionists already solved this issue? Unfortunately it has not yet been done.

Ever since Darwin postulated that organisms evolved from a common ancestor [2], evolutionists try to unravel the evolutionary history and build the tree of life similarly historians, i.e., gathering evidence and documents to help them reconstruct the story of life on Earth.

The search for the evolutionary history of life on earth (phylogenetic reconstruction) often involves the analysis of fossil organisms as well as molecules and morphology of existing organisms [5]. Fossils provide evidence of how and when ancestral forms lived. Morphological characters provide clues of evolutionary relationships, giving that two similar organisms should share a recent common ancestor. Regarding molecules, generally proteins and nucleic acids, it is well known that direct sequence comparison is a powerful tool for determining whether two species are closely or distantly related. Given the availability of such documentation, why we still do not have a complete and reliable phylogeny yet? Because each document tells a different story and, what is worse, the results can also vary if different methodologies are applied to the same data set [12].

Evolutionary inference based on DNA sequences was initially applied by comparing homologous versions of a single gene in different organisms. Currently, the accepted universal tree of life is still based on the small ribosomal RNA subunit (SSU rRNA) [3], which is widely distributed and highly conserved [14]. Nevertheless, although rare, this gene may undergo HGT (Horizontal Gene Transfer) or recombination [10,11,15], sometimes providing insufficient information to resolve phylogenies [8]. Nevertheless, some researchers believe that there is a set of genes that have never or rarely been transferred, and expect to reconstruct, for example, bacterial phylogeny based on these genes. These genes comprise the set of genes responsible for processing and storing the information of the cells being essential for their survival and reproduction.

Recently, several research groups, supported by public and private institutions, are engaged in sequence genomes for, among other things, trying to reconstruct the tree of life. These studies have been accelerated by the advent of new technology high yield, especially those called "Next Generation Sequencing" (NGS).

The work began! The scientific community, including biologists and computer scientists, have joined efforts to try to deal with these large databases. The results were published [1], but no conclusive answers emerged [13]. The conflicts continued: the reliability of the analysis is high, but still not sufficient, and the results indicated by different methods of analysis rarely coincide. The frustration at the inability to recover the tree of life has not diminished. Instead, it has only increased. We believed that using whole genomes would be the key to unlocking the whole story [9], however, we realized that this is not true for most
cases, and worse, we don’t know if such key exist... From all this frustration, however, a pattern emerged: not a few, but many of the genes contained in a particular bacterial species seem to come from other species, which is evidenced by the fact that different genes tell different stories, as if they had not evolved together from a common ancestor. It is true that biologists, in particular microbiologists, have long known that bacteria can transfer some genes (sometimes large sets of genes) to other species by HGT events, but we did not know that this phenomenon was so frequent.

Coincidently or not, computers also work because of the existence of an operating system responsible for managing the processing and storage of information, which is rarely modified and without which you cannot run other programs. On the other hand, there are different software, which can vary among computers carrying the same operating system, depending on the needs, desires and sometimes vulnerabilities (such as lack of a good antivirus) of each user. Thus, the operating system could be compared to genes that are not (or rarely) exchanged by the bacteria while the software would be equivalent to genes that are transferred from one species to another. So the same species of bacteria is now able to survive in different environments, according to the set of "software" that it can "run". Importantly, the genes exchanged between living beings are not limited to those that are needed by the organisms that received them. Some genes simply do not have any effect on the life of a bacterium, while others may cause problems (obviously its carriers are properly disposed by natural selection), and even so they are transferred, exactly as happens with some computers programs exchanged via Internet. In short, the study of genomes explained one more similarity between organisms and computers operation: the existence of fixed sets (operating system or genes that were not or rarely transferred) and transient information (software or genes that are exchanged frequently), also called the Pan-genome, which is composed of a “core genome” and “dispensable genome” [4, 8].

The results obtained so far allow the inference of an evolutionary scenario that began about 3.5 billion years ago, over which bacteria have been swapping genes that eventually, helped them adapt to a wide range of environments and modes of life.

This entire “network” of genetic information exchange certainly contributed to the formation of the millions of bacterial species we can find (or not) today, so that there is life adapted to virtually anywhere on the planet. Assuming that all organisms evolved from a single common ancestor, we can imagine that all organisms present nowadays are descendents of chimeras formed by genes from different organisms, which were mixed and reorganized several times throughout evolution. It is possible that without this, 3.5 billion years would not be enough time to generate the current planetary organic complexity. As for computers, it is possible that without the constant exchange of information between companies (works in collaboration, fellowship and, of course, industrial espionage), less than a century would not be enough time to make the technology available in cheap pocket versions, with high processing power.
Finally, several types of molecular data are available on the World Wide Web for anyone who is willing to analyse it (e.g. http://ncbi.nlm.nih.gov/, http://earthmicrobiome.org, and in a near future http://brmicrobiome.org). For an evolutionist, the mere existence of a database containing information about all the genetic material of several hundred organisms from different environments leaves the comfortable feeling that we will finally be able to resolve the evolutionary history of species on this planet, enabling us to draw the tree of life and show this history in details. The great hope lies in the massive amounts of data that can be analysed simultaneously, which in theory can produce results with high statistical reliability.

References


The evolution of life compared to computers


Received: April 12, 2013