New 'Tree of Life' established for one of the largest groups of bacteria

May 17 2010

A new "tree of life" has been constructed by researchers at the Virginia Bioinformatics Institute (VBI) at Virginia Tech for the gamma-proteobacteria, a large group of medically and scientifically important bacteria that includes Escherichia coli, Salmonella typhimurium, and other disease-causing organisms.* By building powerful phylogenetic trees, scientists are able to quickly identify similarities and differences between the make-up of many different organisms, crucial information in the search for treatments to fight anything from the bugs that cause food poisoning to the pathogens that cause life-threatening diseases such as cholera and the plague.

A "tree of life," or phylogenetic tree, is a way to visualize the evolutionary relationships among different biological species that have descended from a common ancestor. The gamma-proteobacteria tree developed by VBI researchers was reconstructed using powerful computers from as many as 30 million data points of bacterial sequence information.

Kelly Williams, Research Investigator at VBI, remarked: "Ribosomal RNA is one of the central components of the ribosome, the protein manufacturing machinery of all living cells. In the past, researchers have often depended on looking at a single ribosomal RNA gene to construct evolutionary relationships for their tree-building efforts. The method we use to make our tree of life uses hundreds of different genes and integrates much more information than can be gleaned from the traditional single gene approach. We firmly believe that the multi-gene

or phylogenomics approach should become the standard for tree-building when several genome sequences are available, which is now the case for most bacterial groups."

The researchers selected 108 available genomes from the more than 200 complete and partial sequences available for the gamma-proteobacteria, placing the emphasis on the diversity of the bacterial species and quality of the original sequence data. Allan Dickerman, Assistant Professor at VBI, remarked: "The consensus tree that we have put together for the gamma-proteobacteria is a powerful tool that can be used to predict shared biology and analyze, for example, the novel ways that bacteria have adapted to their living environments. Phylogenomics provides for very accurate reconstructions of inheritance from common ancestors."

The researchers looked at a very large class of bacteria that lack a well-resolved phylogenetic tree. By placing emphasis on searches for single-copy genes, the scientists were able to radically improve the resolution of the evolutionary tree. Said Williams, "Some parts of our tree were still not fully resolved, but we believe that future work will improve our method further to handle these deficiencies."

Bruno Sobral, Director of the CyberInfrastructure Section at VBI, commented: "The work described in this paper was inspired and funded by the needs of our PATRIC 2.0 project. The effort is part of the on-going work of PATRIC 2.0 team members to build a comprehensive, state-of-the-art bioinformatics resource for bacteria that serves the biomedical research community. Because of the exponentially growing number of bacterial genomes that PATRIC needs to handle, we are now in a phase where whole-genome phylogenetic analysis is both possible and necessary. PATRIC is integrating the very latest phylogenomic information and tools, such as those in this paper and a preceding publication that developed a phylogenetic tree for the alpha-proteobacteria**, into our system." He added: "This work is a great

example of how PATRIC implements and deploys an infrastructure that will allow any person to develop these results in the future by going to the PATRIC site."

In October 2009, The National Institute of Allergy and Infectious Diseases (NIAID), part of the National Institutes of Health (NIH), awarded a 5-year, $27,670,448 contract to Dr. Sobral's CyberInfrastructure Group of VBI to support the biomedical research community's work on infectious diseases. The funding is being used to integrate vital information on pathogens, provide key resources and tools to scientists, and help researchers to analyze genomic, proteomic and other data arising from infectious disease research.

  More information: For more information on PATRIC 2.0 please consult patricbrc.vbi.vt.edu/portal/portal/patric/Home

* Williams KP, Gillespie JJ, Sobral BW, Nordberg EK, Snyder EE, Shallom JM, Dickerman AW (2010) Phylogeny of Gammaproteobacteria. Journal of Bacteriology 192(9): 2305-2314. [PMID: 20207755]

** Williams KP, Sobral BW, Dickerman AW (2007) A robust species tree for the alphaproteobacteria. Journal of Bacteriology 189: 4578-4586. [PMID: 17483224]

Provided by Virginia Tech