GENOME WATCH

Reshaping the tree of life

Eva Heinz and Daryl Domman

This month's Genome Watch highlights how metagenomics is continuing to reveal the diversity of microorganisms in the environment and how it is challenging and expanding our understanding of how life evolved on Earth.

Understanding the evolutionary relationships between organisms has been a long-standing challenge in the field of evolutionary biology. In 1977, Carl Woese and colleagues revolutionized the field by proposing a new third domain of life known as the Archaea. This proposal led to a reorganization of the tree of life into three separate domains — the Eukarya, Bacteria and Archaea. Similarly, the discovery of large DNA viruses (also known as 'giant viruses'), which have genome sizes that are larger than some bacterial genomes (>1 Mb)¹, prompted a discussion on a possible fourth domain of life. Two recent metagenomics studies provide new insights into how the tree of life might be shaped, either by proposing a new origin of the Eukarya branch in the tree or by considering whether the tree should include giant viruses in a fourth domain.

In the three-domain tree of life that was proposed by Woese, the Eukarya is an

independent lineage from the Archaea and Bacteria; however, an alternative hypothesis suggests that eukaryotes originated from within the archaeal domain. The first evidence for an archaeal last common ancestor of eukaryotes was reported in a study that compared ribosome structures from the three domains of life², which was further supported by phylogenetic studies that used advanced statistical models. Using the power of metagenomics coupled with diverse environmental sampling, a recent study by Zaremba-Niedzwiedzka *et al.*³ was able to reconstruct the genomes of novel archaea, which comprise two new phyla called Odinarchaeota and Heimdallarchaeota. These phyla are related to the recently discovered Thorarchaeota and Lokiarchaeota⁴, which altogether comprise the Asgard superphylum. Remarkably, phylogenetic and gene content analyses demonstrated that the Asgard superphylum is the closest related lineage to eukaryotes. These findings provide strong support for the hypothesis that the Eukarya arose from within the Archaea. If correct, the new tree of life would have Bacteria and Archaea as the two main branches.

Another challenge in evolutionary biology is the placement of viruses in the tree of life, with some studies suggesting an ancient cellular ancestry for viruses¹. When the first giant virus (Mimivirus)

genome was sequenced in 2004, many attributes that are found in cellular organisms, such as genes that encode components of the protein translation machinery and DNA repair proteins, were found to be encoded in the Mimivirus genome¹. On the basis of these analyses, a fourth domain of life was hypothesized, which places viruses emerging from the branch that connects the Archaea and Eukarya. However, this hypothesis is controversial, as other studies support

their evolution from other viruses through the acquisition of genes through horizontal gene transfer. A recent metagenomics survey discovered a novel giant virus, called Klosneuvirus, in a wastewater treatment plant in Austria⁵. The authors then expanded their dataset by screening nearly 7,000 previously published metagenomics datasets to search for similar giant viruses. The authors also discovered three additional giant viruses that were closely related to Klosneuvirus, which now form a new proposed family of giant viruses. The group then found that the genomes of these new giant viruses arose from much smaller viral genomes that had acquired genes from their eukaryotic hosts, such as protists, throughout evolutionary history, rather than originating from a shared cellular ancestor.

Together, these studies highlight that metagenomic sampling of diverse environments can reveal the hidden diversity of microorganisms, and increase our understanding of the deep evolutionary relationships between the major domains of life.

Eva Heinz and Daryl Domman are at the Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK.
e-mail: microbes@sanger.ac.uk

doi:10.1038/nrmicro.2017.51
Published online 12 May 2017

1. Raoult, D. *et al.* The 1.2-megabase genome sequence of Mimivirus. *Science* **306**, 1344–1350 (2004).
2. Lake, J. A. *et al.* Eocytes: a new ribosome structure indicates a kingdom with a close relationship to eukaryotes. *Proc. Natl Acad. Sci. USA* **81**, 3786–3790 (1984).
3. Zaremba-Niedzwiedzka, K. *et al.* Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017).
4. Spang, A. *et al.* Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
5. Schulz, F. *et al.* Giant viruses with an expanded complement of translation system components. *Science* **356**, 82–85 (2017).

Competing interests statement

The authors declare no competing interests.