PUBLISHED: 25 JULY 2016 | ARTICLE NUMBER: 16139 | DOI: 10.1038/NMICROBIOL.2016.139

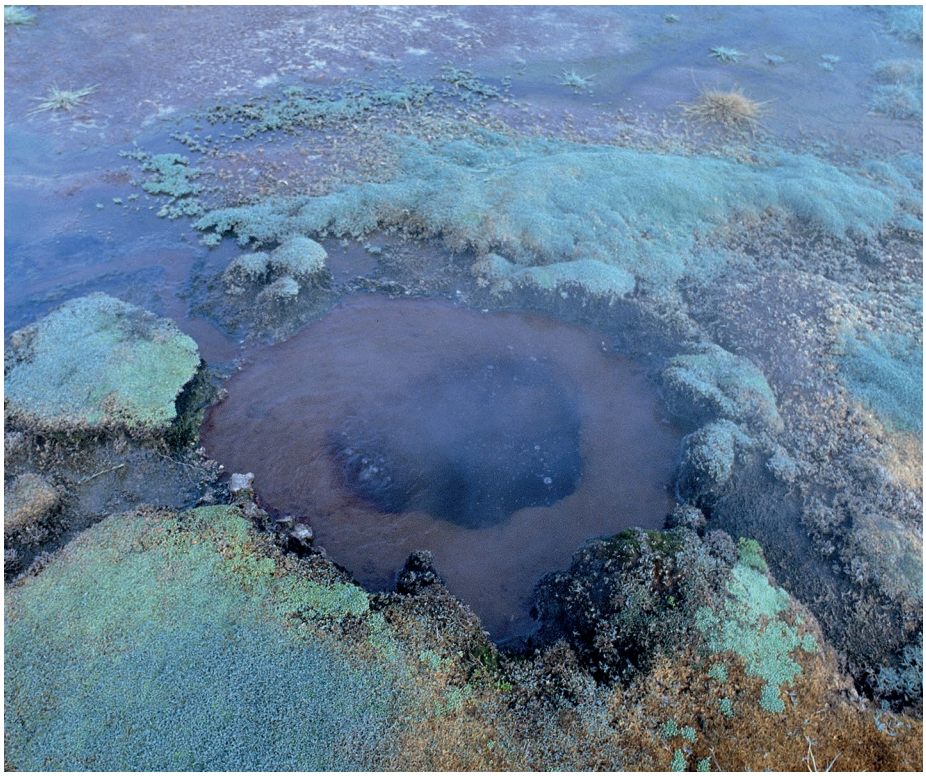
EVOLUTION

A four billion year old metabolism

Inspection of more than 286,000 gene families has shed light on the most recent common ancestors of all life. The last universal common ancestor was likely to have been a thermophilic, anaerobic, N_2-fixing organism that used the Wood–Ljungdahl pathway to fix CO_2, using H_2 as an electron donor.

James O. McInerney

The history of life on the planet is mostly the history of single-celled prokaryotes. Until recently, it was accepted that life was divided into three 'domains' — forms of life of equal standing in a taxonomic sense[1]. However, we now know that the deepest division of life on the planet separates Archaea (Archaebacteria) from Bacteria (Eubacteria), while Eukaryotes are a more recent, secondary group[2–4]. The organism from which Bacteria and Archaea emerged is termed the last universal common ancestor (LUCA). Although it is certain that LUCA had full DNA replication, transcription and translation systems[5], other facets of its lifestyle have so far remained difficult to reconstruct. Given that genomes are usually well adapted to their environment, knowing LUCA's genome could also lend insight into the environmental conditions in which LUCA lived. In this issue of *Nature Microbiology*, Weiss *et al.*[6] have significantly advanced our understanding of what LUCA did for a living. What has emerged is a picture of heat-loving organisms, living in a world without oxygen, dependent on hydrogen as an electron donor, using CO_2 as electron acceptor and capable of nitrogen fixation.



PHOTOALTO

A major complicating factor for understanding evolutionary history is horizontal gene transfer (HGT). A gene can be transferred from one organism to another and, with a few caveats[7], this gene can often function quite well in the host organism[8]. If HGT didn't exist and we saw that a gene family was widely distributed among a set of taxa, then we might infer its origin to be ancient. However, because of HGT, this assumption is invalid and a gene family is just as likely to be common because it is frequently horizontally transferred.

Therefore, the strategy employed by Weiss *et al.*[6] for pinpointing LUCA's genes involved identifying genes whose evolutionary history was mostly vertical. Weiss *et al.* constructed phylogenetic trees from every gene family that stood a fighting chance of having been in LUCA — the gene family had to be in at least two species of Bacteria and two species of Archaea[6]. Of the initial six million genes grouped into 286,514 gene families, only 11,093 families had a broad enough distribution in modern organisms to have the potential to have been in LUCA. However, when phylogenetic trees were constructed from these families, only 355 phylogenies recovered the two prokaryote groups as being monophyletic. That is to say that only 0.1% of the data had a chance of speaking to this research question. It's important to note that LUCA was probably a fully functioning, though perhaps somewhat primitive cell, so it is highly likely that LUCA harboured many more genes than the 355 families identified by Weiss *et al.* Interestingly though, this set of 355 families was not a random subset of the total data. Rather, it turns out that they are a very specific set of genes that imply a very specific lifestyle that does not point to a chemoheterotrophic lifestyle.

The presence of the thermophile-specific enzyme reverse gyrase in the dataset of LUCA genes implies that LUCA was a thermophile. A rotator–stator ATP synthase subunit suggests LUCA was able to harness ion gradients for energy. The only energy pathway enzymes present were those of the Wood–Ljungdahl pathway, which is a pathway that uses H_2 as an electron donor and CO_2 as electron acceptor. The H_2 must have come from geological sources, since it could not have been made through fermentation. Analysis of the phylogenetic trees constructed from the 355 protein families places Clostridia and methanogens as the earliest-diverging organisms — both of which are anaerobic, H_2-dependent and use the Wood–Ljungdahl pathway. The implication of this work is that LUCA was very much dependent on abiotic sources of H_2 to provide it with energy.

So, what was happening on the planet at the time of LUCA? It was very different to today. The amount of oxygen available for biological cells was negligible and all life was anaerobic. The late heavy bombardment (LHB) of Earth by comets and asteroids approximately 4–3.8 billion years ago[9] probably resulted in Earth being periodically heated to the point that the oceans were vaporized and probably led to bottlenecks in the diversity of life at the time, meaning that only hyperthermophiles survived[10]. The presence of a reverse gyrase, a thermophile-specific gene[11], in LUCA lends weight to this scenario, leading Weiss and co-workers infer that life began at hydrothermal vents (pictured). The data is certainly compatible with this idea, but it's important to remember that LUCA was not the first form of life. Thus, the data is equally compatible with LUCA being a life-form that could make it through the LHB.

When we look at the inferred metabolism of LUCA, we are looking at the dominant and most successful kind of metabolism on the planet before the Bacteria and Archaea diverged. This new study provides us with a very intriguing insight into life 4 billion years ago. However, what Weiss *et al.* cannot elucidate is whether the 355 gene families identified were in the same cell at the same time, nor what other genes were present. To answer these questions, future work should aim to further disentangle HGT from vertical evolution in widely distributed gene families. ❐

James O. McInerney is in the Faculty of Biology, Medicine and Health, Michael Smith building, The University of Manchester, Manchester M13 9PL, UK.
e-mail: james.mcinerney@manchester.ac.uk

References

1. Woese, C. R., Kandler, O. & Wheelis, M. L. *Proc. Natl Acad. Sci. USA* **87,** 4576–4579 (1990).
2. Williams, T. A. *et al. Nature* **504,** 231–236 (2013).
3. Cox, C. J. *et al. Proc. Natl Acad. Sci. USA* **105,** 20356–20361 (2008).
4. McInerney, J. O., O'Connell, M. J. & Pisani, D. *Nature Rev. Microbiol.* **12,** 449–455 (2014).
5. Doolittle, W. F. & Brown, J. R. *Proc. Natl Acad. Sci. USA* **91,** 6721–6728 (1994).
6. Weiss, M. C. *et al. Nature Microbiol.* **1,** 16116 (2016).
7. McInerney, J. O. & Pisani, D. *Science* **318,** 1390–1391 (2007).
8. McInerney, J. O. *et al. Biol. Direct* **6,** 41 (2011).
9. Cohen, B. A., Swindle, T. D. & Kring, D. A. *Science* **290,** 1754–1756 (2000).
10. Nisbet, E. G. & Sleep, N. H. *Nature* **409,** 1083–1091 (2001).
11. Kikuchi, A. & Asai, K. *Nature* **309,** 677–681 (1984).