



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/114978/>

Version: Accepted Version

Article:

McInerney, JO and O'Connell, MJ (2017) Microbiology: Mind the gaps in cellular evolution. *Nature*, 541 (7637). pp. 297-299. ISSN: 0028-0836

<https://doi.org/10.1038/nature21113>

(c) 2017, Macmillan Publishers Limited, part of Springer Nature. All rights reserved. This is an author produced version of a paper published in *Nature*. Uploaded in accordance with the publisher's self-archiving policy.

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Mind the gaps in cellular evolution

JAMES O. MCINERNEY & MARY J. O'CONNELL

Eukaryotic cells, with complex features such as membrane-bound nuclei, evolved from prokaryotic cells that lack these components. A newly identified prokaryotic group reveals intermediate steps in eukaryotic-cell evolution. See Article p.XXX

Eukaryotic cells contain membrane-bound organelles such as a nucleus, and complex cellular components, including protein-based transport systems that move molecules around the cell. About 1.8 billion years ago¹, eukaryotic cells arose from cells that lack these features, known as prokaryotic cells. However, gaps in the cellular family tree mean that limited evidence exists that might illuminate how some eukaryote-specific features arose during the earliest intermediate steps in the evolution of eukaryotic cells. In a paper online in *Nature*, Zaremba-Niedzwiedzka *et al.*² identify a superphylum branch of the prokaryotic family tree that contains some genes previously thought to be eukaryote-specific.

On the basis of their gene content and aspects of their cellular physiology, prokaryotes are classified into two domains: Bacteria and Archaea. Eukaryotes are thought to be derived from a merger that occurred when an archaeal cell engulfed a bacterial cell related to modern alphaproteobacteria³. It has been proposed⁴ that, within these early eukaryotic cells, the internalized proteobacterium eventually evolved to form the membrane-bound organelles known as mitochondria that provide energy for the cell. However, early events in eukaryotic evolution have remained poorly understood. Few species have been identified whose genome content could provide insight into steps in the transition between prokaryotic and eukaryotic cells.

Until now, the archaea that have been identified as the most closely related ancestors of eukaryotic cells come from the group known as Lokiarchaeota⁵; these were identified by genome sequencing of organisms found in deep-sea sediments. Lokiarchaeota contain features that were previously thought to be eukaryote-specific, including several genes distantly related to those involved in eukaryotic protein transport. However, to understand the transition from prokaryotic to eukaryotic cellular life, a more complete picture is needed of the genes present in the archaeal cells that gave rise to eukaryotes.

DNA encoding 16S ribosomal RNA sequences is often used to determine the genetic relationships (taxonomy) between species. Twenty-five years ago, a study⁶ of the DNA encoding 16S RNA sequences obtained from seawater organisms revealed the presence of archaeal species related only distantly to the known archaea cultured in

laboratories. This work hinted that cultured prokaryotes capture just a small fraction of global prokaryotic diversity. Zaremba-Niedzwiedzka *et al.* also investigated samples of underwater organisms, and their results reveal that major groups of life can still be discovered.

The authors obtained samples of aquatic sediments from seven locations worldwide. They extracted short fragments of DNA, representing a mixture of the species present, and sequenced more than 644 billion nucleotides. The short fragments of sequence were assembled into longer pieces, and sequences containing at least six genes from an evolutionarily conserved ribosomal-protein gene cluster were identified. These were analysed to determine the taxonomic relationships of the sampled genomes. The researchers identified the sequences that were most similar to the previously sequenced Lokiarchaeota⁵ and Thorarchaeota⁷ (another archaeal species related to eukaryotes).

Zaremba-Niedzwiedzka *et al.* used a statistical method⁸ to classify the ribosomal sequences on the basis of similarities in the patterns of nucleotides used. From this analysis, they identified a superphylum of Archaea containing four major lineages of genetic material: Lokiarchaeota, Thorarchaeota and the newly found groupings of Odinararchaeota and Heimdallarchaeota. The authors named this superphylum Asgard. The exact family-tree relationships between the groupings within the Asgard superphylum are difficult to ascertain, because of uncertainty in the phylogenetic trees that were constructed. But the analyses support the existence of the superphylum as a whole.

The position of the Asgard group within the family tree of cellular life suggests that its members form the closest known archaeal sister lineage to eukaryotes (Fig. 1). This finding is consistent³ with the view that the three-domains hypothesis, which suggests that eukaryotes arose independently of bacterial or archaeal diversification, has fallen. The work by Zaremba-Niedzwiedzka *et al.* adds to the body of evidence that eukaryotes arose through a merger of cells from within Archaea and Bacteria.

The authors found several eukaryotic-like genes that function in protein transport, signalling and protein degradation and that were previously reported to be present in Lokiarchaeota⁵. Their work also reveals expanded repertoires of these genes in the newly discovered archaeal groups. The Asgard members are certainly not eukaryotes. However, they harbour types of genes that, until now, were thought only to have originated early in eukaryotic evolution. This research therefore pushes back the origins of parts of some cellular machinery.

Of particular interest in the Asgard group are the genes that relate to the cytoskeleton: a framework of structural proteins that gives a cell its shape and helps to organize its internal structure. Tubulin protein can form filaments that are key components of the cytoskeleton and are essential for eukaryotic cell division. The presence of tubulin in Archaea has previously been reported⁹. However, Zaremba-Niedzwiedzka *et al.* identified tubulin in Odinararchaeota that has sequence similarity closer to that of eukaryotic tubulin. They also found sequences encoding other

cytoskeletal components — including lokiactins and proteins that contain gelsolin and profilin domains — as well as sequences relating to the protein complex ARP2/3, which regulates the cytoskeleton. The sequences identified do not represent the blueprint for a complete eukaryotic cytoskeletal complex; but the authors indicate that these cellular machines, previously thought to be specific to eukaryotes, might instead have archaeal origins.

Asgard sequences also contain components of DNA-processing machinery that were previously thought to be exclusively eukaryotic. These include versions of the ϵ subunit of DNA polymerase enzymes. In eukaryotes, this polymerase subunit contains three domains. In the Asgard group, however, it contains two and is missing the evolutionarily conserved domain of unknown function.

Another surprising discovery was that versions of genes encoding eukaryotic membrane-trafficking proteins are found in the Asgard superphylum. In eukaryotes, such proteins are involved in moving proteins around the cell, between membrane-bound organelles known as the endoplasmic reticulum and the Golgi complex. But the small size of prokaryotic cells suggests that such intracellular trafficking mechanisms might not be necessary. Given that, and the absence in such cells of an endoplasmic reticulum and Golgi complex, it is unclear how these genes might function in Archaea. Of course, shared sequence similarity does not guarantee evolutionarily conserved function. The common cellular ancestor of Asgard and eukaryotic cells existed 1.8 billion years ago³, so these genes have independently evolved over a total of 3.6 billion years, a time span that could have allowed many functional differences to arise in the encoded proteins.

The distribution of genes formerly thought to be eukaryote-specific is patchy in the Asgard superphylum, and no individual archaeal group seems to have a full set of such genes. So, although eukaryotes usually need all these genes together, the members of the Asgard group do not. Members of the Asgard superphylum have not been observed through the microscope yet, nor has their laboratory culture been reported. Therefore, many key details about their ecology, evolution and cell biology remain to be described. Further investigation of this superphylum might shed more light on the early origins of eukaryotic cells.

James O. McInerney is in the Division of Evolution and Genomic Sciences, School of Biological Sciences, and Manchester Academic Health Science Centre, University of Manchester, Manchester M13 9PL, UK. **Mary J. O'Connell** is in the Computational and Molecular Evolutionary Biology Group, School of Biology, Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK.

e-mail: james.mcinerney@manchester.ac.uk

1. McInerney, J. O., O'Connell, M. J. & Pisani, D. *Nature Rev. Microbiol.* **12**, 449–455 (2014).

2. Zaremba-Niedzwiedzka K. *et al. Nature* <http://dx.doi.org/10.1038/nature21031> (2017).

3. McNerney, J., Pisani, D. & O'Connell, M. J. *Phil. Trans. R. Soc. B* **370**, 20140323 (2015).
4. Martin, W. F., Neukirchen S., Zimorski, V., Gould, S. B. & Sousa, F. L. *BioEssays* **38**, 850–856 (2016).
5. Spang, A. *et al. Nature* **521**, 173–179 (2015).
6. Fuhrman, J. A., McCallum, K. & Davis, A. A. *Nature* **356**, 148–149 (1992).
7. Seitz, K. W., Lazar, C. S., Hinrichs, K. U., Teske, A. P. & Baker, B. J. *ISME J.* **10**, 1696–1705 (2016).
8. Brady, A. & Salzberg, S. L. *Nature Methods* **6**, 673–676 (2009).
9. Yutin, N. & Koonin, E. V. *Biol. Direct* <http://doi.org/10.1186/1745-6150-7-10> (2012).

Figure 1: A family tree of cellular life forms. Eukaryotic cells have complex cellular features such as membrane-bound nuclei and evolved from prokaryotic cells lacking these features. Eukaryotic cells are thought to have arisen when an archaeal prokaryotic cell merged with a type of bacterial prokaryotic cell related to modern-day alpha-proteobacteria³. However, many aspects of the early evolution of eukaryotic cells remain unknown, such as how eukaryotic-specific features might have arisen. Zaremba-Niedzwiedzka *et al.*² analysed genomic sequences from deep-sea organisms which they used to identify a superphylum of Archaea that they named the Asgard group. Their genetic analysis placed this group near to the eukaryotic cell lineage. Members of this superphylum contain versions of genes previously thought to have been eukaryote-specific.