

Lokiarchaeon is hydrogen dependent

Filipa L. Sousa¹, Sinje Neukirchen¹, John F. Allen², Nick Lane², William F. Martin^{1*}

¹ Institute of Molecular Evolution, Heinrich-Heine University, 40225 Düsseldorf, Germany

² Research Department of Genetics, Evolution and Environment, Darwin Building, University College London, Gower Street, London U.K., WC1E 6BT

*Author for correspondence: bill@hhu.de

Abstract:

The nature of host that acquired the mitochondrion at eukaryote origin (the host) is an important microbial evolutionary issue. Modern phylogenetics indicate that the host was an archaeon. The metagenome sequence of *Candidatus* Lokiarchaeon identified it as the closest relative of the host yet known. Here we report comparative genomic evidence indicating that Lokiarchaeon is hydrogen-dependent, as one theory for eukaryote origin — the hydrogen hypothesis — predicted for the host lineage.

Main text

The origin of eukaryotes is a critical evolutionary transition¹. Recent metagenome data indicate that *Candidatus* Lokiarchaeon (Loki) is, by the measure of ribosomal protein phylogeny, the closest known free-living archaeal relative to the host² that acquired the mitochondrion. Loki's genome sequence lists several protein families that link archaea with eukaryotes, including ESCRT proteins, ubiquitin domain

proteins and eukaryotic-type Ras GTPases, prompting discussion of the possibility that the mitochondrion might have been acquired by an archaeon that possessed primitive phagocytotic capabilities^{2,3}. Here we show that Loki's genome harbours a complete tetrahydromethanopterin (H₄MPT) dependent Wood-Ljungdahl (WL) pathway and enzymes revealing it to be an H₂-dependent, strictly anaerobic, and very probably autotrophic archaeon, attributes of the host that are specifically predicted by the hydrogen hypothesis for the origin of eukaryotes⁴. In archaea studied to date, ESCRT proteins are involved in cell division^{3,5}, not phagocytosis, while the presence of an H₄MPT dependent WL pathway indicates an H₂-dependent autotrophic lifestyle. Loki's genome data, though incomplete², shed light on the physiology of archaeal lineages thought to be closely related to the host and implicate a role for hydrogen dependence⁴ and anaerobic syntrophy⁴, rather than phagotrophy^{2,3}, in acquisition of the mitochondrion at eukaryote origin.

In search of evidence for Lokiarchaeon's physiology in its metagenome, we first asked whether it harbours the hallmark pathway of H₂-dependent anaerobic autotrophs: the WL (or acetyl-CoA) pathway⁶. The acetyl-CoA pathway exists in two forms, a tetrahydrofolate (H₄F) dependent bacterial version and a tetrahydromethanopterin (H₄MPT) dependent archaeal version⁶. Loki possesses a complete methyl synthesis branch of the H₄MPT-dependent acetyl-CoA pathway (Fig. 1). It also possesses an almost complete pathway for the biosynthesis of H₄MPT, the typical archaeal one-carbon carrier and the cofactor that underpins the archaeal acetyl-CoA pathway⁶ (Fig. 1). Two subunits of bifunctional CODH-ACS in the acetyl-CoA pathway are missing, but the Lokiarchaeon genome is estimated to be only about 90% complete².

The H₄MPT-dependent acetyl-CoA pathway is characteristic for H₂-dependent archaeal anaerobic autotrophs^{6,7}. The Loki genome lacks almost all genes that are specific to methanogenesis and anaerobic methane oxidation; of the ion-pumping MtrA-H complex only MtrH (a methyl-H₄MPT dependent methyltransferase subunit)⁸ is present, and methyl-CoM reductases essential for methanogenesis and anaerobic methane oxidation⁹ are missing. Autotrophs that fix CO₂ via the acetyl-CoA pathway use pyruvate synthase (pyruvate:ferredoxin oxidoreductase; PFO) to generate pyruvate⁶, Loki possesses an archaeal PFO (Fig. 1).

Further evidence for Lokiarchaeon's anaerobic lifestyle comes from the presence of several enzymes that perform flavin based electron bifurcation (supplemental data Table 2), a newly characterized mechanism of energetic coupling widespread among anaerobes⁷. Loki's electron bifurcating enzymes include NADH-dependent ferredoxin: NADP oxidoreductase⁷ (NfnAB) fused to a coenzyme F₄₂₀ hydrogenase subunit B (FrhB) domain having high similarity with the sulfide dehydrogenase (SudHI/II) from *Pyrococcus furiosus*, the electron transfer flavoprotein Bcd/EtfBC⁷, and a complete H₂-dependent electron bifurcating MvhADG-HdrABC hydrogenase system including the Hyp NiFe maturation system⁷ required for nickel insertion, cyanide ligand synthesis from carbamoyl phosphate, and cyanide transfer to the active site of the MvhA subunit. The MvhADG-HdrABC hydrogenase, which is so far specific for anaerobes⁷, provides additional evidence for Loki's H₂-dependence.

We detected no genes for membrane complexes involved in chemiosmotic energy coupling, including those typical for anaerobic autotrophs: Rnf/Nqr complexes⁷, the Mrp/hydrogenase family of complexes that includes complex I, Ech⁷ and Mbh were missing, as were components of mitochondrial-type respiratory chains as well as the HdrE/NarI/FdnI cytochrome-containing enzymes typical of anaerobic respiratory chains. An incomplete A₀A₁ ATP-synthase (subunits *a*, A(β), B(α), and D(γ)) was present, but a eukaryotic type vacuole-acidifying (protease-activating) vacuolar ATPase, required in free-living eukaryotes for acidifying food vacuoles to activate digestive enzymes during phagotrophy¹⁰, was not. Neither quinone interacting proteins nor heme-containing enzymes were detected, although siroheme proteins are present in the current Loki genome, suggesting that it might be capable of sulfur reduction¹¹.

The enzymes of Loki's H₄MPT-dependent Wood-Ljungdahl pathway were most similar to to methanogen homologues (supplemental data Table 2), providing important evidence for the existence of methanogen-like metabolism outside the euryarchaeotes². Metagenomic data have uncovered the existence of further methanogen-like metabolisms outside the euryarchaeotes⁹, but, as with Loki, the energy metabolism of the organisms behind these new metagenomic lineages is

unknown. Loki's central intermediary metabolism (Supplementary Information) resembles that of *Candidatus Methanoplasma termitum*¹².

Our characterization of Loki as an anaerobic, H₂-dependent autotroph, as predicted by the hydrogen hypothesis, does not support the view that the host for the origin of mitochondria was a phagotroph². This distinction needs to be seen in the light of three decades of dispute about the identity of the host cell that acquired mitochondria^{13,14}. The archezoa hypothesis argued that the host cell was a phagotroph that engulfed and digested food particles such as bacteria^{13,14}. It predicted that at least some modern archezoa were evolutionary intermediates, survivors from early eukaryotic evolution before the acquisition of mitochondria. This hypothesis was disproved with the demonstration that all the supposed archezoa in fact arose by reductive evolution from more complex eukaryotic ancestors, with their mitochondria specializing to become hydrogenosomes or mitosomes¹⁴. Phylogenetic studies, including the Loki study², have since established that the host cell was an archaeon, and therefore a *bona fide* prokaryote¹. Yet the idea that the host cell somehow had to be a phagocyte to have acquired mitochondria has risen again with the Lokiarchaeon genome report².

How an archaeal host cell, presumably with a cell wall, acquired mitochondria, is unknown; but examples of walled prokaryotes that have acquired bacterial endosymbionts do exist¹⁵, so it is plainly possible. There is no need to invoke a phagotrophic host cell, and no evidence that one ever existed. On the contrary, bioenergetic considerations indicate that the host cell could hardly have been phagotrophic¹⁵. Phagocytosis requires a fully-fledged eukaryotic cytoskeleton, food-vacuole formation, and complex intracellular membrane vesicle trafficking¹⁰, none of which is known among prokaryotes¹⁰. The machinery of phagocytosis has a high energetic cost, not only in terms of cytoskeletal motility, cell movement and dynamic vesicle trafficking, but even more importantly in its requirement for high expression of many hundreds if not thousands of genes³. The energetic costs for the *de novo* evolution of complex traits such as phagocytosis in terms of exploration of protein sequence space and gene expression must be substantially greater than the costs of maintaining the system once it has evolved. Yet bacteria and archaea are strictly limited in their size and energy availability per gene by their bioenergetic architecture, in particular their use of the plasma membrane for chemiosmotic coupling¹⁵.

In contrast, mitochondria internalize respiration, and always retain tiny specialized bioenergetic genomes and ribosomes¹⁶. This uniquely eukaryotic bioenergetic architecture enables the local control of chemiosmotic coupling, increasing energy availability per gene by orders of magnitude relative to prokaryotes¹⁵. Though a recent theoretical study of ATP cost accounting concluded that mitochondria might not have been essential to eukaryote origin after all¹⁷, the modeled cell had an unlimited (unconstrained) ATP supply, thereby missing the role of mitochondria at eukaryote origin altogether¹⁸, while the cost accounting furthermore failed to include ribosomes¹⁷, the energetically most expensive component of eukaryotic cells¹⁸. Only the enormous energy boost provided by mitochondria could power the *de novo* evolution of the machinery of phagocytosis, explaining why such an apparently advantageous mode of living is completely absent in prokaryotes¹⁵.

The data presented (Fig. 1) here indicate that Loki is H₂-dependent, strictly anaerobic and probably autotrophic. Its H₂-dependence needs to be included in discussions about the metabolic context of mitochondrial origin^{2,3}. For thermodynamic reasons¹⁹, H₂-dependence would preclude a heterotrophic carbon metabolism for Loki. Because phagocytosis requires heterotrophy it is therefore unlikely that Loki possesses a level of cellular complexity² that would include phagocytotic abilities.

By the measure of ribosomal protein phylogeny, Loki is indeed the closest known relative of the host identified so far². The available genomic evidence has it that Loki is H₂-dependent, bearing out central predictions of the hydrogen hypothesis⁴ with regard to the nature of the host lineage. The further predictions are clear: under a microscope, Loki will not be a large archaeon with a eukaryote-like feeding habit^{2,3}, it will, instead, be a strictly anaerobic, hydrogen-dependent⁴ archaeon.

Methods

Sequences for 394 proteins involved in carbon and energy metabolism across diverse prokaryotic groups were used to query Loki's genome using BLAST²⁰ using as threshold 25% amino acid identity and an E-value of 10⁻¹⁰ (query sequences and

references are listed in Supplementary table 1). Lokiarchaeon genes showing matches at the specified threshold were blasted against RefSeq (version 72), the first 10 matches were parsed against NCBI taxonomy. All results are summarized in Supplementary Table 1.

2210 words

Author contributions

F.L.S and W.F.M designed the research. F.L.S. and S.N. performed the analysis. J.F.A. and N. L. assessed and commented on the results and conclusion. All authors discussed the results. F.L.S., J.F.A., N.L. and W.F.M wrote the paper.

Competing financial interests

The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to W.F.M. (bill@hhu.de).

References

1. Williams, T. A., Foster, P. G., Cox, C. J. & Embley, T. M. An archaeal origin of eukaryotes supports only two primary domains of life. *Nature* **504**, 231–236 (2013).
2. Spang, A. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
3. Koonin, E. V. Origin of eukaryotes from within archaea, archaeal eukaryome and bursts of gene gain: eukaryogenesis just made easier? *Phil. Trans. R. Soc. B* 370: 20140333 (2015).
4. Martin, W. & Müller, M. The hydrogen hypothesis for the first eukaryote. *Nature* **392**, 37–41 (1998).
5. Samson, R. Y. & Bell, S. D. Ancient ESCRTs and the evolution of binary fission. *Trends Microbiol.* **17**, 507–513 (2009).
6. Fuchs, G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631–658 (2011).

7. Buckel, W. & Thauer, R. K. Energy conservation via electron bifurcating ferredoxin reduction and proton/Na⁺ translocating ferredoxin oxidation. *Biochim. Biophys. Acta* **1827**, 94–113 (2013).
8. Hippler, B. & Thauer, R. K. The energy conserving methyltetrahydromethanopterin: coenzyme M methyltransferase complex from methanogenic archaea: function of the subunit MtrH. *FEBS Lett.* **449**, 165–168 (1999).
9. Evans, P. N. et al. Methane metabolism in the archaeal phylum Bathyarchaeota revealed by genome-centric metagenomics. *Science* **350**, 434–438 (2015).
10. Flannagan, R. S., Jaumouillé, V. & Grinstein, S. The cell biology of phagocytosis. *Annu. Rev. Pathol.* **7**, 61–98 (2012).
11. Rabus, R., Venceslau, S. S., Wöhlbrand, L., Wall, J. D. & Pereira, I. A. C. A post-genomic view of the ecophysiology, catabolism and biotechnological relevance of sulphate-reducing prokaryotes. *Adv. Microb. Phys.* **66**, 55-321 (2015).
12. Lang, K., Schuldes, J., Klingl, A., Poehlein, A., Daniel, R. & Bruner A. New mode of energy metabolism in the seventh order of methanogens as revealed by comparative genome analysis of “*Candidatus Methanoplasma termitum*”. *Appl. Environ. Microbiol.* **81**, 1338–1352 (2015).
13. Embley, T.M. & Martin, W. Eukaryotic evolution, changes and challenges. *Nature* **440**, 623–630 (2006).
14. Archibald, J.M. Endosymbiosis and eukaryotic cell evolution. *Curr. Biol.* **25**, R911–R21 (2015).
15. Lane, N. & Martin, W. The energetics of genome complexity. *Nature* **467**, 929–934 (2010).
16. Allen, J.F. Why chloroplasts and mitochondria retain their own genomes and genetic systems: colocation for redox regulation of gene expression. *Proc. Natl. Acad. Sci. USA.* **112**, 10231–10238 (2015).
17. Lynch, M. & Marinov, G.K. The bioenergetic costs of a gene. *Proc. Natl. Acad. Sci. USA* **112**, 15690–15695 (2015).
18. Lane, N. & Martin, W.F. Mitochondria, complexity and evolutionary deficit spending. *Proc. Natl. Acad. Sci. USA* (2016) (doi/10.1073/pnas.1522213113) *in press*.

19. Schönheit, P., Buckel, W. & Martin, W.F. On the origin of heterotrophy. *Trends Microbiol.* **24**, 12–24 (2016).
20. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

Figure Legend

Figure 1. Typical genes of H₂-dependent, anaerobic autotrophs, were identified by using query sequences for well-characterized proteins specific to the pathways indicated in the figure in sequence comparisons, and scoring whether the proteins in the Lokiarchaeon genome annotation² were more similar to archaeal or bacterial homologues (Supplementary Table 1). Filled ovals indicate that the gene is present in the Lokiarchaeon genome at an E-value threshold of 10⁻¹⁰ and at least 25% local amino acid identity; empty ovals indicate that the gene was missing at that threshold or that the Lokiarchaeon sequence was more similar to homologues from the domain indicated. Abbreviations: fdhAB, formate dehydrogenase; fhs, 10-formyl-H₄F-synthetase; folD, 5,10-methenyl-H₄F cyclohydrolase/dehydrogenase; metF/MTHFR, 5,10-methylene-H₄F reductase; acsE, corrinoid iron–sulfur protein methyltransferase; cdhδ/acsD, corrinoid iron–sulfur protein CoFeS; cdh(α/β)/acsB, cdhγ/acsC carbon monoxide dehydrogenase/acetyl-CoA synthase; fmd/fwdABCDEF formyl-methanofuran dehydrogenase; ftr formyl transferase; mch, 5,10-methenyl-H₄-MPT cyclohydrolase; mtd, 5,10-methylene-H₄-MPT dehydrogenase; mer, 5,10-methylene-H₄-MPT reductase; cdhγ/acsC methyl-H₄-MPT:corrinoid iron–sulfur protein methyltransferase; corrinoid iron–sulfur protein cdhγδ/acsD; CO dehydrogenase/acetyl-CoA synthase cdhα, cdhβ □ cdhε; porABCD, nifJ, pyruvate:ferredoxin oxidoreductase; folE, 7,8-dihydroneopterin triphosphate synthase; PTPSIII, pyruvoyltetrahydropterin synthase; folB, 7,8-dihydroneopterin aldolase; folK, diphosphokinase; folE2, GTP cyclohydrolase IB; mptA, Fe(II)-dependent-GTP cyclohydrolase IB; PTPSVI, pyruvoyltetrahydropterin synthase; mptB, Fe(II)-dependent-cyclic phosphodiesterase; mptD, 7,8-dihydroneopterin aldolase; mptE 7,8-dihydro-6-hydroxymethylpterin diphosphokinase; RFAP synt, β-

D-ribofuranosylaminobenzene-5-phosphate synthase; mptH, dihydropteroate synthase; folP, dihydropteroate synthase; folC, dihydrofolate synthase; folA, folM, dihydrofolate reductase; HdrABC and HdrED, heterodisulfide reductase; mvhADG, methyl viologen hydrogenase complex; frhABG, coenzyme F₄₂₀ hydrogenase; hydAB [FeFe] hydrogenase; hmd, [Fe] H₂-dependent methylene-tetrahydromethanopterin dehydrogenase. * Sequences from this family can include ketoisovalerate oxidoreductase (vorABCD) and 2-oxoacid:ferredoxin oxidoreductase (korABCD).^a - the detected mvhD gene may be fused with one HdrA domain.