RESEARCH ARTICLE

A *Thermococci*-to-*Clostridia* Pathway for the Evolution of the Bacteria Domain

Tze-Fei Wong[1*], Chung-Kwon Chan[2], and Hong Xue[1]

[1]Division of Life Science, Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
[2]Institute of Space and Earth Information Science, Chinese University of Hong Kong, Hong Kong, China

*bcjtw@ust.hk

ABSTRACT

With the identification of an archaeal Last Universal Common Ancestor (LUCA) related to the archaeon *Methanopyrus*, the origin of Bacteria became a choice between autonomic development versus descent from Archaea. The similarity bitscores between paralogous valyl-tRNA synthetase (VARS) and isoleucyl-tRNA synthetase (IARS) suggest that the five oldest bacteria were *Mahella australiensis, Thermincola potens, Halobacteroides halobius, Desulfosporosinus orientis* and *Caldicellulosiruptor lactoaceticus*, which were all *Clostridia* species and hydrogen producers. A search for archaea that could be a candidate Progenitor of Bacteria endowed with an Emden-Myerhof-Parnas type glycolytic pathway and a clostridial-like dark-fermentation mechanism for generating hydrogen pointed to such a role for the anaerobic chemoorganotroph *Thermococci*, which were known to engage in rapid evolutionary changes at high-biodiversity sites abundant in sugars and other small molecular substrates, and constitute together with *Clostridia* the two most powerful microbial generators of hydrogen. Moreover, two-domain maximum-likelihood and maximum-parsimony phylogenetic trees for VARS showed that *Thermococci* and *Clostridia* formed sister clades on both trees, and close similarity was evident between their VARS sequences, which were consistent with kinship between them. On this basis, it was proposed that the Bacteria domain emerged from a thermocococal or thermococcal-like Progenitor of Bacteria possibly at some high-biodiversity site, through the formation of a *Thermincola*-proximal Last Bacterial Common Ancestor (LBCA).

Introduction

The postulate of a three-domain structure of life has brought precious unity to biological evolution, even though it placed the root of life, or LUCA, within the Bacteria domain due to its reliance on a reported paralogous-protein tree built without adequate archaeal representation[1]. This placement has become unsupportable with the identification of a LUCA closely related to the methanogenic archaeon *Methanopyrus kandleri* (Mka) on the basis of alloacceptor tRNA analysis of substantial numbers of genomes from all three living domains[2], which has been confirmed by the primitivity of the wobble rules in Mka for translation using the GNN and UNN anticodons to decode tetra-codon family boxes, and encoding of neighboring tRNAs in sequence space to read the well separated UCN and AGY codons for Ser in Mka[3], the oldest age of Mka at 2.8 Gya among the archaeal lineages[4], and the high-yield thermal synthesis of the intermediates and end products of the acetyl-CoA pathway in hydrothermal vents[5]. More recently, application of the paralog-divergence method for estimating the relative ages of organisms[6] to the VARS and IARS sequences of more than 5,000 species of archaea, bacteria and eukaryotes showed that Mka displayed the highest VARS-IARS similarity bitscore, and the formation of an Ancestral Bacterial Cluster centered at clostridial *Mahella* containing the oldest bacteria[7]. These findings on LUCA established both the importance of identifying the roots of the three domains so that the momentous undertaking of constructing the Tree of Life (TOL) can proceed based on an accurate foundation, and the ability of the paralog-divergence method using VARS-IARS similarity bitscores as the single criterion for estimating the ages of species in phylogenetic analysis.

In the search for the root of Eukarya, the one hundred year old postulate of endosymbiosis-based eukaryotic development initiated prolonged debate over whether the mitochondria-early scenario was essential, and if so how cellular space could be found for its accommodation in an archaeal or bacterial host, which became resolved only when it was realized that the small archaeon *Aciduliprofundum boone* (Abo; 1.4 Mb genome) had an eukaryogenesis approach that sidestepped the space issue. By exploiting its unsurpassed capability to assemble tranches of genes from a wide spectrum of archaea and bacteria, this archaeon grew in size until it could acquire and accommodate a mitochondria-generating alphaproteobacterial endosymbiont (6.5 Mb genome) to become the Progenitor of Eukarya. It gave rise to a Microsporidial Last Eukaryotic Common Ancestor (LECA) and eventually mammals where the complex proteome of the brain still includes a range of microbial-originated proteins including Abo proteins[8].

Hitherto, searches for LBCA have given rise to diverse bacterial groups as candidates including *Aquifex* and *Thermotoga*[9]; *Plantomycetales*[10]; Gram-positive bacteria[11]; a bacterium near Fusobacteriota[12]; and a bacterium with clostridial-like metabolism[13]. These diverse candidates led to the notions that the search for a single TOL was a quixotic pursuit dependent on the unproven assumption of the natural order being an inclusive hierarchy[14]; the conclusion that TOL has not been falsified by horizontal gene transfers

(HGTs)[15]; and the proposal of segmenting TOL into sister clades of prokaryotes and eukaryotes[16] or multiple regional trees[17]. In view of such dissension, the objective of the present study was to expand our earlier study on the use of VARS-IARS bitscores to estimate the relative ages of species which had revealed the presence of *Clostridia* among the ancient bacteria[7], so that useful knowledge can be gained regarding the oldest bacteria, and therefore possibly the nature of the Progenitor of the Bacteria domain. This method readily pointed to the cross-domain kinship between primitive *Clostridia* and *Thermococci*. Because the *Clostridia* have evolved so extensively over time, this potential kinship would be much more obscure if the recently evolved clostridial pathogens were compared instead with the *Thermococci*.

Methods

VARS sequences and VARS-IARS similarity bitscores were analyzed as described[7], and multiple sequences were aligned using the clustalw command from CLUSTAL v2.1. Phylogenetic analyses of maximum likelihood (ML) was inferred using iqtree2 v.2.2.0 under LG+C10+F+G model[18] chosen with 5,000 ultrafast bootstep replicates. Figtree was used to view phylogenetic tree graphically from the result generated by iqtree2. Phylogenetic analyses of maximum parsimony (MP) was performed by using PAUP ver. 4.0b10 through heuristic search with random addition of sequences and branch-swapping with tree-bisection-reconnection (TBR)[19].

Results

BIOLOGICAL PROPERTIES OF LBCA

Based on the VARS-IARS similarity bitscores ("bit") of 3,621 bacterial species, the bacteria with the five highest bitscores were all anaerobic *Clostridia* and hydrogen-producers[7], viz. *Mahella australiensis*[20] (bit378); *Thermincola potens*, which was isolated from a thermophilic cathode catalyzing sustainable hydrogen production[21] (bit377) and whose close relative *T. carboxydiphila* could form carbon dioxide and hydrogen[22]; *Halobacteroides halobius*[23] (bit375); *Desulfosporosinus orientis*[24] (bit372); and *Caldicelllulosiruptor lactoaceticus*[25] (bit368). Their much lower VARS-IARS bitscores than the top bitscore of 473 for archaea, and much higher VARS-IARS bitscores than the top bitscore of 240 for eukaryotes indicated that LBCA was either an archaeon or an autonomically developed microbe.

CANDIDATE PROGENITOR OF BACTERIA

On account of the limited availability of sugar substrates in the pre-bacterial age, not many anaerobic archaea were endowed with a functional glycolytic pathway. Notable exceptions were *Thermoproteus tenax*, *Desulfurococcus amylolyticus*, *Archaeoglobus fulgidus* and *Thermococci* exemplified by *Pyrococcus furiosus*[27]. Among them, *T. tenax* oxidized glucose by sulfur respiration, and glucose degradation was accompanied by sulfate reduction in *A. fulgidus*, neither of which would furnish a likely template for the EMP pathway of bacteria.

T. tenax and *P. furiosus* could degrade 100% of added glucose to pyruvate[26,27]; however, *T. tenax* was a weak hydrogen producer. On the other hand, the ten intermediates in the 'modified' mo-EMP pathway of *P. furiosus* were identical to the ten intermediates utilized by the 'classical' cl-EMP pathway of *Caldicellulosiruptor saccharolyticus* and

Caldanaerobacter tengcongengsis, and each of these three species contained side branches on their EMP sequences that led to ferredoxin- or NADH2-mediated formation of hydrogen by hydrogenase[28]. These three species, one thermococcal and two clostridial, ranked among the most powerful microbial hydrogen-producers known[29].

Although the usage of ADP instead of ATP as substrates by some archaeal mo-EMP kinases was a distinctive departure from the ATP-utilizing bacterial cl-EMP kinases, the retention of 20% residual activity by the phosphofructokinase (PFK) of *T. kodakarensis* upon switching from ADP substrate to ATP suggests that such switches were not too difficult to implement[30].

HIGH-BIODIVERSITY ENVIRONMENT
Previously, a survey of sixteen *Thermococci* species showed that they widely employed peptides as energy substrates, whereas only the species that originated from hydrothermal vents, viz. *T. aggregans*, *T. guamasensis*, *T. fumicolans*, *T. hydrothermalis*, *T. profundus* and *P. furiosis* metabolized maltose[31,32]. These findings suggest that high-biodiversity sites such as the hydrothermal vents accelerated the evolution of the *Thermococci* toward novel directions including the acquisition of bacterial-like sugar metabolism. In this regard, analysis of the vents at Guaymas Basin revealed that they produced and released an abundance of low molecular-weight carbohydrates, peptides and hydrocarbons when magmatic sills intruded into the organic-rich sediments[33,34], resulting thereby in a high level of biodiversity[35]. In a single study, five novel *Thermococcus* species were identified from these vents based on SSU rRNA,

elongation factors EF-1 alpha and EF-2, which testified to a rapid rate of thermococcal evolution[36]. Based on 37 single-copy genes including the VARS and IARS genes, a series of metagenome-assembled genomes (MAGs) of *Thermococcus* strains also have been isolated from these vents[37].

TWO-DOMAIN VARS TREES
To assess whether the phenotypic *Thermococci-Clostridia* resemblances were the consequence of vertical gene transmission between them or extraneous factors like convergent evolution and HGT, two-domain VARS phylogenetic trees were built using the maximum-likelihood (ML) and maximum-parsimony (MP) methods covering the five oldest *Clostridia* and the slightly younger *Moorella thermoacetica*[38] (bit361) and *Caldanaerobacter tengcongensis*[39] (another strong hydrogen producer; bit351), as well as a series of *Thermococci* species, a couple of MAG *Thermococcus* strains, and a wide spectrum of other archaea with high VARS-IARS bitscores. Because any species with a high bitscore would contain VARS and IARS of relatively ancient origin, whenever its VARS or IARS came to be replaced by an extraneous VARS or IARS through HGT, its VARS-IARS bitscore would likely be reduced. Therefore, the high-bitscore species tended not to contain HGT-derived VARS. Figures 1 and 2 showed that *Thermococci* and *Clostridia* formed sister clades on both the ML and MP trees for VARS, in accord with phylogenetic kinship between them. In addition, striking conservation of VARS sequences between these two groups of microbes was illustrated by the decapeptide GRKMSKSLGN in Figure 3: the seven primitive clostridial sequences

were totally conserved among themselves, differed from seven out of eight thermococcal sequences at only a single residue, and identical to the eighth thermococcal sequence of *Paelococccus, pacificus*.
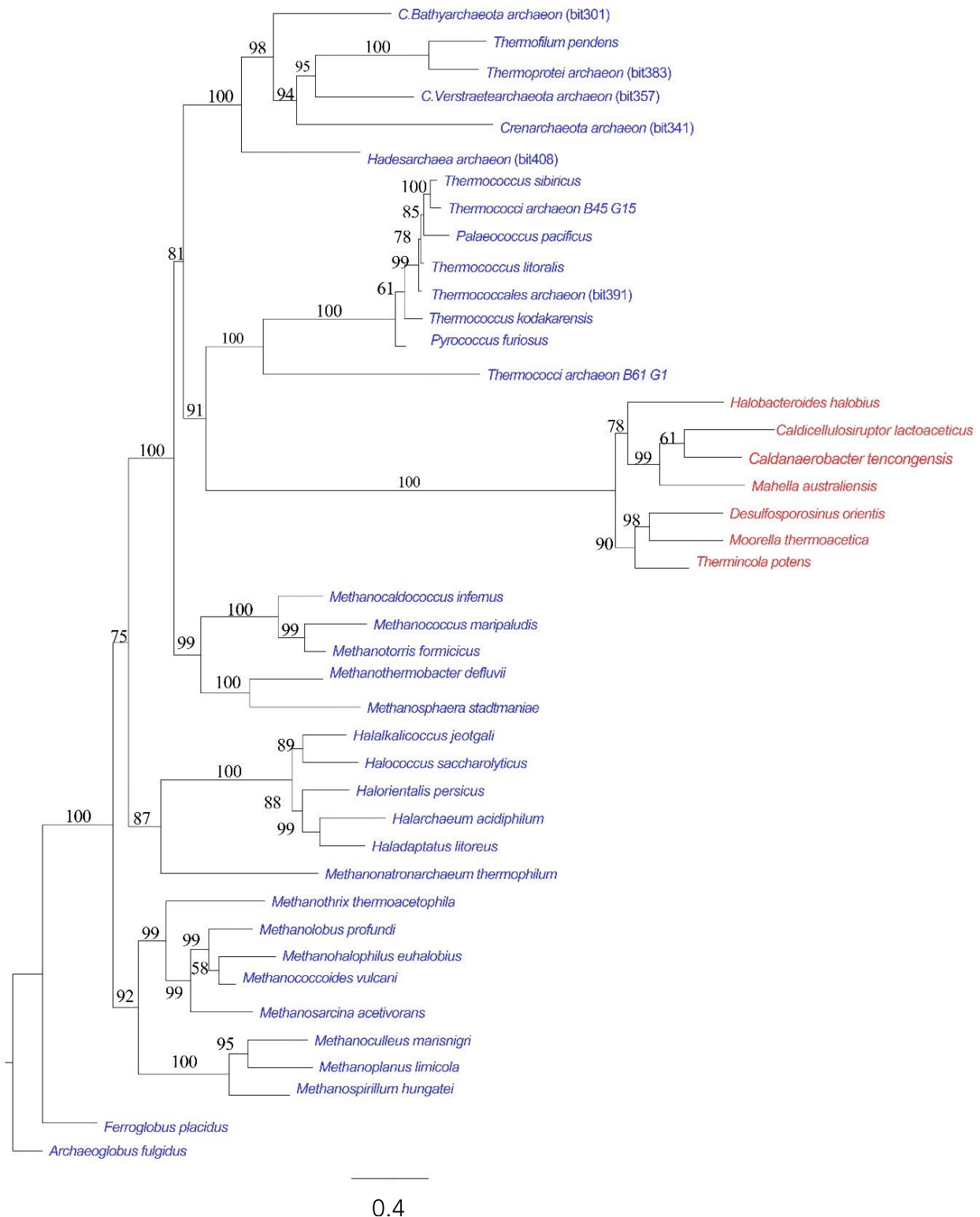

Figure 1. Two-domain VARS phylogenetic tree built using the maximum likelihood method. MAG genomes of *Thermococcus* isolated from Guaymas Basin[37] were designated by Bxx Gyy.
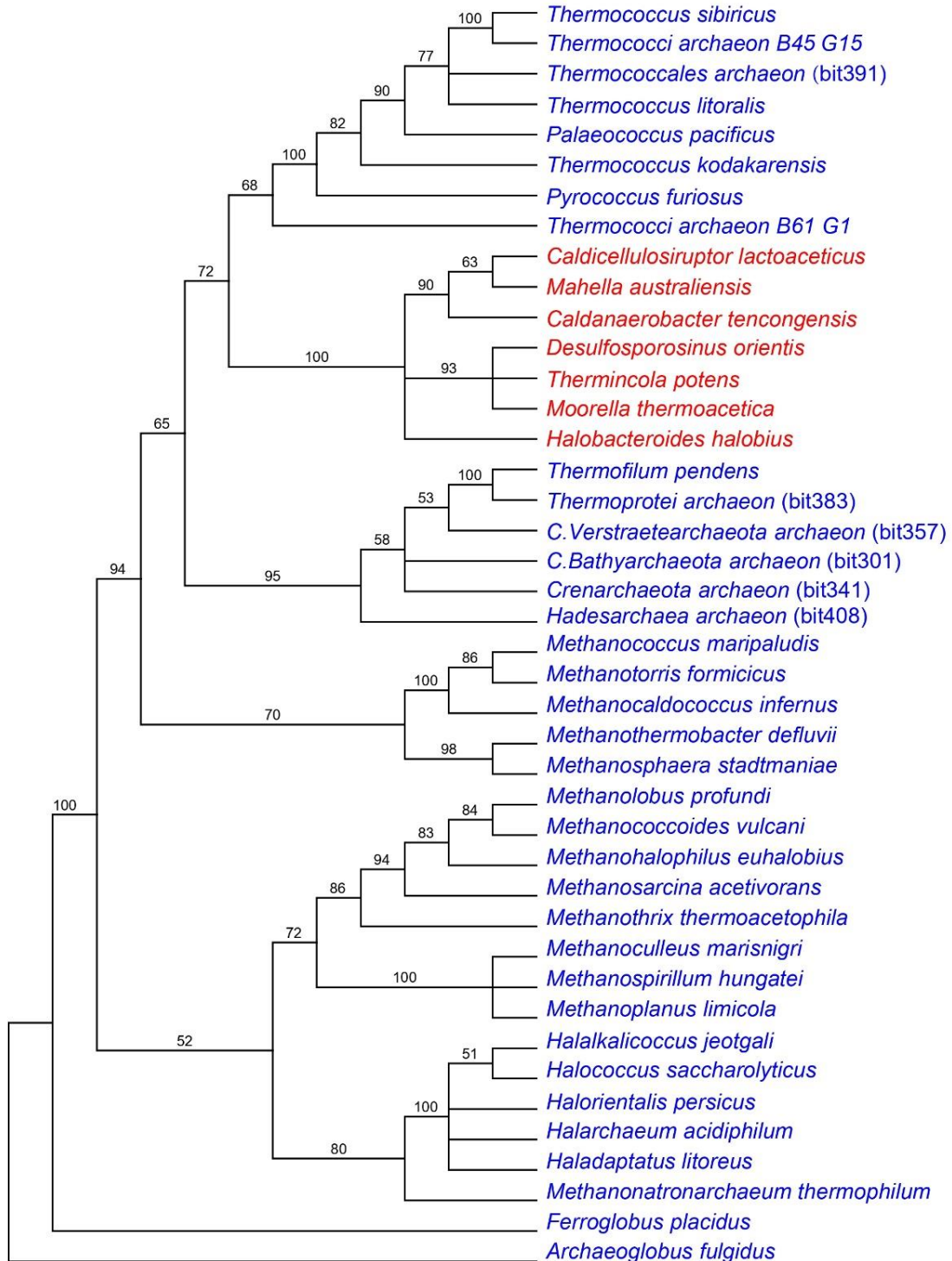
Bootstrap consensus tree

Figure 2. Two-domain VARS phylogenetic tree built using the maximum parsimony method.

Species	Sequence	
Thermococcus sibiricus	GRKMSKSYGN	
Thermococci archaeon B45 G15	GRKMSKSYGN	
Thermococcales archaeon (bit391)	GRKMSKSYGN	
Thermococcus litoralis	GRKMSKSYGN	
Palaeococcus pacificus	GRKMSKSLGN	*
Thermococcus kodakarensis	GRKMSKSYGN	
Pyrococcus furiosus	GRKMSKSYGN	
Thermococci archaeon B61 G1	GRKMSKSYGN	
Caldicellulosiruptor lactoaceticus	GRKMSKSLGN	*
Mahella australiensis	GRKMSKSLGN	*
Caldanaerobacter tencongensis	GRKMSKSLGN	*
Desulfosporosinus orientis	GRKMSKSLGN	*
Thermincola potens	GRKMSKSLGN	*
Moorella thermoacetica	GRKMSKSLGN	*
Halobacteroides halobius	GRKMSKSLGN	*
Thermofilum pendens	GRMMHKSYGN	
Thermoprotei archaeon (bit383)	GRMMHKSYGN	
C.Verstraetearchaeota archaeon (bit357)	GRMMHKSYGN	
C.Bathyarchaeota archaeon (bit301)	GRKMSKSLGN	*
Crenarchaeota archaeon (bit341)	GRMMHKSYGN	
Hadesarchaea archaeon (bit408)	GRKMSKSLGN	*
Methanococcus maripaludis	GFKMSKSRGN	
Methanotorris formicicus	GHKMSKSRGN	
Methanocaldococcus infernus	GYKMSKSRGN	
Methanothermobacter defluvii	GHKMSKSRGN	
Methanosphaera stadtmaniae	GHKMSKSLGN	
Methanolobus profundi	GHKMSKSLGN	
Methanococcoides vulcani	GHKMSKSLGN	
Methanohalophilus euhalobius	GHKMSKSRGN	
Methanosarcina acetivorans	GHKMSKSLGN	
Methanothrix thermoacetophila	GRKMSKSLNN	
Methanoculleus marisnigri	GFKMSKSRSN	
Methanospirillum hungatei	GFKMSKSRGN	
Methanoplanus limicola	GFKMSKSRGN	
Halalkalicoccus jeotgali	GNKMSKSRGN	
Halococcus saccharolyticus	GNKMSKSKDN	
Halorientalis persicus	GNKMSKSRGN	
Halarchaeum acidiphilum	GNKMSKSRGN	
Haladaptatus litoreus	GNKMSKSRGN	
Methanonatronarchaeum thermophilum	GNKMSKSLGN	
Ferroglobus placidus	GRKMSKSLGN	*
Archaeoglobus fulgidus	GRKMSKSLGN	*

Figure 3. Sequence conservation in VARS. The aligned sequences of a VARS decapeptide illustrated the close relationship between *Clostridia* and *Thermococci*. The first residue of the decapeptide corresponds to residue 582 on the VARS of *Crenarchaeota archaeon*, which possesses the longest N-terminal polypeptide among the species analyzed. The C-termini of all the sequences identical to that of *T. potens* are marked by asterisks.

Within the thermococcal clade on the ML tree, the VARS sequences of the MAG genome of *T. archaeon B61 G1* and *Pyrococcus furiosus* were lowest branching, suggesting that the clostridial branch diverged from the archaea in the vicinity of these two archaeons to form the Progenitor of Bacteria. Within the clostridial clade on the same tree,

Thermincola potens represented the lowest branch, which was consistent with the location of LBCA near this species. Since *T. potens* and *M. australiensis* were the two bacterial species with the highest VARS-IARS bitscores, there was evident consonance between the ML tree branchings and the VARS-IARS bitscores in their support of a *Thermincola*-proximal LBCA.

Discussion

The sister clades formed by *Thermococci* and *Clostridia* on the two-domain ML and MP trees of VARS (Figures 1 and 2), and the ten identical intermediates of the mo-EMP and cl-EMP pathways, were in accord with an evolutionary precursor-product relationship between these two microbial groups. On the other hand, that only four of the ten cl-EMP enzymes, viz. triosephosphate isomerase, phosphoglycerate mutase, enolase and pyruvate kinase had orthologues among the mo-EMP enzymes opened up the possibility of convergent evolution or HGT as an alternative explanation. While convergent evolution occurred from time to time, it would be rare to encounter as many as six such events within the same pathway. Although HGTs occurred most frequently between closely related species, they might play a significant role in facilitating the adaptive radiation of novel genetic traits[40]. However, contribution from HGTs to the conversion of the thermococcal mo-EMP pathway to the clostridial cl-EMP pathway would be limited because bacteria-derived genes were nowhere to be found in the pre-LBCA eon for transfer to LBCA via HGTs. As a result, *the Thermococci* had to rely on the inventions of novel genes, or non-homologous gene replacements by borrowing genes from

diverse metabolic areas in either their own genomes or the genomes of other archaea via HGTs.

The archaeal mo-EMP enzymes, especially in the hyperthermophilic hydrogen producers like *P. furiosus,* were saddled with disadvantages that needed to be addressed, including their low working efficiency and deficient glycolytic yield of ATP[28], and their conflicting requirements for relatively rigid protein structures to combat thermal denaturation[41], yet also for elevated growth temperature for the maintenance of a high rate of hydrogen production[29]. Moreover, a preference for protein rigidity might be a factor in their abstention from usage of allosteric regulations of glucokinase, phosphofructokinase and pyruvate kinase in for example *P. furiosus* in contrast to the widespread usage of allosteric control in bacteria. Without allosteric control, it would be difficult to balance rapidly between the glycolytic and gluconeogenic fluxes through the EMP pathway. Altogether, the drastic changes from the mo-EMP pathway to the cl-EMP pathway, leaving only four of the ten mo-EMP enzymes largely unrevised constituted a timely innovation instituted by the *Thermoccoci*. Its success was made possible by both external and internal factors. Externally, large high-biodiversity sites exemplified by the Guaymas Basin, generating plentiful carbohydrates and other nutrients, represented such a rare incentive for diversification by these chemoorganotrophic microbes that they mounted a genomic response that was unprecedented in scale, especially with respect to carbohydrate utilization. Internally, the *Thermococci* were genomically adventurous, as demonstrated

by the greater similarity between the hydrogen-production pathways of thermococcal *P. furiosus* and clostridial *C. saccharolyticus* and *C. tencongensis* compared to that between the thermococcal *P. furiosus* and *T. kodakarenis*[29]. With such a distinct predisposition for variability, they were particularly suited to be the pioneers who first broke away from the Archaea to establish a novel domain of life.

In conclusion, the formation of the cl-EMP pathway of *Clostridia* was activated by necessity and such was also likely the case with their other bacterial characteristics, all adding up to the novel features of a third biological domain. The cl-EMP pathway has withstood the test of eons of time to this day as the core of bacterial and eukaryotic central metabolisms, supporting the running of marathon and 100-meter dash, as well as transcontinental migration on wings. No convergent evolution or horizontal gene transfer, and only elaborate mutation and selection, could have yielded such superb performance. Therefore, based on the sister-clade relationship between *Thermococci* and *Clostridia* on the two-domain VARS trees, the close similarity between their VARS sequences, and the resemblances in their EMP-pathways and hydrogen-formation mechanisms, it can be proposed that the emergence of the Bacteria domain was mediated by a *Thermococci*-to-*Clostridia* evolutionary pathway.

Competing interests Statement:

The authors declare no conflict of interest.

Acknowledgements Statement:

Author contributions:

Conceptualization T.F.W.; data analysis C.K.C. and H.X.; writing, T.F.W., C.K.C. and H.X. All authors have read and agreed to the published version of the manuscript.

References:

1. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990; 87(12):4576-4579. doi:10.1073/pnas.87.12.4576.

2. Xue H, Tong KL, Marck C, Grosjean H, Wong JTF. Transfer RNA paralogs: evidence for genetic code-amino acid biosynthesis coevolution and an archaeal root of life. Gene. 2003; 310:59-66. doi:10.1016/s0378- 1119 (03)00552-3

3. Wong JTF, Ng SK, Mat WK, Hu T, Xue H. Coevolution theory of the genetic code at age forty: pathway to translation and synthetic life. Life. 2016;6(1):12. doi:10.3390/life6010012.

4. Blank CE. Low rates of lateral gene transfer among metabolic genes define the evolving biogeochemical niches of archaea through deep time. Archaea 2012; 2012: 23. doi: 10.1155/2012/843539.

5. Weiss MC, Preiner M, Xavier JC, Zimorski V, Martin WF. The last universal common ancestor between ancient Earth chemistry and the onset of genetics. PLoS Genet. 2018;14(8) :e1007518.doi:10.1371/journal.pgen.1007518.

6. Schwartz RM, Dayhoff MO. Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. Science. 1978; 199(4327):395-403. doi:10.1126/science.202030

7. Long X, Xue H, Wong JT. Descent of bacteria and eukarya from an archaeal root of life. Evol Bioinform. 2020;16:117693432 0908267. doi:10.1177/1176934320908267

8. Wong JT, Chan CK, Long X, Xue H. An ectosymbiosis-based mechanism of eukaryogenesis. Med Res Arch 2023; 11(1).

Doi.org/10.18103/mro.v11i1.3365.

9. Bocchetta M, Gribaldo S, Sanangelantoni A, Cammarano P. Phylogenetic depth of the bacterial genera Aquifex and Thermotoga inferred from analysis of ribosomal protein, elongation factor, and RNA polymerase subunit sequences. J Mol Evol. 2000; 50(4):366-380. doi:10.1007/s002399910040

10. Brochier C, Philippe H. Phylogeny: a non-hyperthermophlic ancestor for bacteria. Nature 2002; 447 (6886):244.doi:10.1038/ 417244a.

11. Ciccarelli FD, Doerks T, von Mering C, Creevey CJ, Snel B, Bork P. Toward automatic reconstruction of a highly resolved tree of life. Science. 2006; 311(5765): 1283-1287. Brochierdoi:10.1126/science.1123061

12. Coleman GA, Davín AA, Mahendrarajah TA, et al. A rooted phylogeny resolves early bacterial evolution. Science. 2021;372(6542) :eabe0511. doi:10.1126/science.abe0511

13. Xavier JC, Gerhards RE, Wimmer JLE, Brueckner J, Tria FDK, Martin WF. The metabolic network of the last bacterial common ancestor. Commun Biol. 2021; 4(1):413. doi:10.1038/s42003-021-01918-4

14. Doolittle WF, Bapteste E. Pattern pluralism and the Tree of Life hypothesis. Proc Natl Acad Sci U S A. 2007;104(7):2043-2049. doi:10.1073/pnas.0610699104

15. Blais C, Archibald JM. The past, present and future of the tree of life. Curr Biol. 2021; 31(7):R314-R321. doi:10.1016/j.cub.2021.02.052

16. Harish A, Kurland CG. Empirical genome evolution models root the tree of life. Biochimie. 2017;138:137-155. doi:10.1016/j. biochi.2017.04.014

17. Estrada A, Suárez-Díaz E, Becerra A.

Reconstructing the Last Common Ancestor: epistemological and empirical challenges. Acta Biotheor 2022; 70(2):15. doi:10.1007/s10441-022-09439-1

18. Minh BQ, Lanfear R, Ly-Trong N, Trifinopoulos J, Schrempf D et al, IQ-TREE version 2.2.0 Tutorials and manual phylogenomic software by maximum likelihood. 2022; 37(5). doi: 10.1093/molbev/ msaa015

19. Swoford, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10 win32. Sunderland. MA: Sinauer.

20. Salinas BM, Fardeau ML, Thomas P, Cayol JL, Patel BKC, Ollivier B. Mahella australiensis gen. nov., sp. nov., a moderately thermophilic anaerobic bacterium isolated from an Australian oil well. Int J Syst Evol Microbiol. 2004; 54(Pt 6):2169-2173. doi:10.1099/ijs.0.02926-00.

21. Fu, Q., Kobayashi, H., Kuramochi, Y., Xu, J., Wakayama, T. et al, 2013. Bioelectrochemical analyses of a thermophilic biocathode catalyzing sustainable hydrogen production. Int. J. Hydrogen Energy 38(35), 15638-45.

22. Sokolova TG, Kostrikina NA, Chernyh NA, Kolganova TV, Tourova TP, Bonch-Osmolovskaya EA. Thermincola carboxydiphila gen. nov., sp. nov., a novel anaerobic, carboxydotrophic, hydrogenogenic bacterium from a hot spring of the Lake Baikal area. Int J Syst Evol Microbiol. 2005; 55(Pt 5):2069-2073. doi:10.1099/ijs.0.63299-0.

23. Oren A, Weisburg WG, Kessel M, Woese CR. Halobacteroides halobius gen. nov., sp. nov., a moderately halophilic anaerobic bacterium from the bottom sediments of the Dead Sea. Syst Appl Microbiol 1984;5(1): 58-70. doi:10.1016/S0723-2020(84)80051-X.

24. Agostino V, Lenic A, Bardl B, et al. Electrophysiology of the facultative autotrophic bacterium *Desulfosporosinus orientis.* Front Bioeng Biotechnol. 2020;8:457. doi:10.3389/fbioe.2020.00457.

25. Mladenovska Z, Mathrani IM, Ahring BK. Isolation and characterization of Caldicellulosiruptor lactoaceticus sp. nov., an extremely thermophilic, cellulolytic, anaerobic bacterium. Arch Microbiol 1995;163:223-230. doi:10.1007/BF00305357

26. van Niel EW. Biological processes for hydrogen production. *Adv Biochem Eng Biotechnol.* 2016; 156:155-193. doi:10.1007/10_2016_11

27. Bräsen C, Esser D, Rauch B, Siebers B. Carbohydrate metabolism in Archaea: current insights into unusual enzymes and pathways and their regulation. Microbiol Mol Biol Rev. 2014; 78(1):89-175. doi:10.1128/MMBR.00041-13

28. Verhees CH, Kengen SW, Tuininga JE, et al. The unique features of glycolytic pathways in Archaea. *Biochem J.* 2003;375(Pt 2):231-246. doi:10.1042/BJ20021472

29. Verhaart MR, Bielen AA, van der Oost J, Stams AJ, Kengen SW. Hydrogen production by hyperthermophilic and extremely thermophilic bacteria and archaea: mechanisms for reductant disposal. Environ Technol 2010; 31(8-9):993-1003. doi:10.1080/09593331003710244

30. Shakir NA, Bibi T, Aslam M, Rashid N. Biochemical characterization of a highly active ADP-dependent phosphofructokinase from Thermococcus kodakarensis. J Biosci Bioeng. 2020; 129(1):6-15. doi:10.1016/j.jbiosc.2019.06.014

31. Kobayashi, T. Genus I. Thermococcus. (2001) 'The archaea and deeply branching and phototrophic bacteria, eds', in Bergey's

Manual of systematic bacteriology. 2nd edn, pp. 342–346.

32. Fiala G, Stetter KO. Pyrococcus furiosus sp. nov. represents a novel genus of marine heterotrophic archaebacteria growing optimally at 100℃. Arch Microbiol 1986;145:56-61. doi: 10.1007/BF00413027

33. Einsele G, Gieskes JM, Curray J, et al. Intrusion of basaltic sills into highly porous sediments, and resulting hydrothermal activity. Nature 1980;283:441-445. doi:10.1038/283441a0

34. Simoneit BRT, Lonsdale PF. Hydrothermal petroleum in mineralized mounds at the seabed of Guaymas Basin. Nature 1982;5(1-2):29-40. doi:10.1038/295198a0

35. McKay L, Klokman VW, Mendlovitz HP, et al. Thermal and geochemical influences on microbial biogeography in the hydrothermal sediments of Guaymas Basin, Gulf of California. Environ Microbiol Rep 2016;8(1):150-161. doi:10.1111/1758-2229.12365

36. Liu L, Wang F, Xu J, Xiao X. Molecular diversity of Thermococcales isolated from Guaymas Basin hydrothermal vents. Acta Oceanol Sin 2013;32:75-81. doi:10.1007/s13131-013-0323-3

37. Dombrowski N, Teske AP, Baker BJ. Expansive microbial metabolic versatility and biodiversity in dynamic Guaymas Basin hydrothermal sediments. Nat Commun. 2018; 9(1):4999. doi:10.1038/s41467-018-07418-0

38. Collins MD, Lawson PA, Willems A, et al. The phylogeny of the genus Clostridium: proposal of five new genera and eleven new species combinations. *Int J Syst Bacteriol.* 1994; 44(4):812-826. doi:10.1099/00207713-44-4-812.

39. Bao Q, Tian Y, Li W, Xu Z, Yang H. A complete sequence of the T. tencongensis genome. Genome Res 2002; 12(5):689-700. Doi:10.1101/gr.219302

40. Soucy SM, Huang J, Gogarten JP. Horizontal gene transfer: building the web of life. Nat Rev Genet. 2015;16(8):472-482. doi: 10.1038/nrg3962.

41. Reed CJ, Lewis H, Trejo E, Winston V, Evilia C. Protein adaptations in archaeal extremophiles. Archea.2013; 2013:373275. doi :10.1155/2013/373275