Characterization of Reconstructed Ancestral Proteins Suggests a Change in Temperature of the Ancient Biosphere

Satoshi Akanuma

Faculty of Human Sciences, Waseda University, 2-579-15 Mikajima, Tokorozawa, Saitama 359-1192, Japan; akanuma@waseda.jp; Tel.: +81-4-2947-6727; Fax: +81-4-2947-6811

Abstract: Understanding the evolution of ancestral life, and especially the ability of some organisms to flourish in the variable environments experienced in Earth's early biosphere, requires knowledge of the characteristics and the environment of these ancestral organisms. Information about early life and environmental conditions has been obtained from fossil records and geological surveys. Recent advances in phylogenetic analysis, and an increasing number of protein sequences available in public databases, have made it possible to infer ancestral protein sequences possessed by ancient organisms. However, the in silico studies that assess the ancestral base content of ribosomal RNAs, the frequency of each amino acid in ancestral proteins, and estimate the environmental temperatures of ancient organisms, show conflicting results. The characterization of ancestral proteins reconstructed in vitro suggests that ancient organisms had very thermally stable proteins, and therefore were thermophilic or hyperthermophilic. Experimental data supports the idea that only thermophilic ancestors survived the catastrophic increase in temperature of the biosphere that was likely associated with meteorite impacts during the early history of Earth. In addition, by expanding the timescale and including more ancestral proteins for reconstruction, it appears as though the Earth's surface temperature gradually decreased over time, from Archean to present.

Keywords: ancestral sequence reconstruction; ancient biosphere; last universal common ancestor; phylogenetic analysis; Precambrian; thermophilicity

1. Introduction

There is still limited understanding of ancestral life on Earth, and the environment in which it evolved. Information about early life and the biosphere has often been obtained from fossil records and geological surveys [1,2]. In 1993, Schopf discovered fossilized stromatolite-like structures in the Apex chert from 3.5 gigayears ago (Gya) [1]. Recently, Nutman et al. reported evidence for ancient life obtained from a newly exposed outcrop of 3.7 Gya metacarbonate rocks in the Isua supracrustal belt [3]. Dodd et al. also reported putative microfossils of microorganisms that are possibly 4.3 Gya old in ferruginous sedimentary rocks from the Nuvvuagittuq belt in Quebec, Canada [4]. Evidence for the existence of methanogens and microbial sulfate reduction at 3.5 Gya [5,6] and the emergence of life at 3.8–4.1 Gya have also been reported [7,8].

A growing amount of genomic data available in public databases provide the necessary resource for the study of molecular phylogeny. By comparing a large number of homologous gene or protein sequences, we can now infer the sequences of genes and proteins that were possessed by ancestral organisms [9–12]. In addition, we can also synthesize the inferred nucleotide and amino acid sequences [9–11,13]. Since the physical properties of extant proteins are well adapted to their hosts' environment, the same must have been true for primitive proteins that existed earlier than 3.5 Gya. Therefore, the nature of ancestral organisms and their environments can be inferred by the

reconstruction and characterization of their proteins [10,12,14,15]. In this review, I will discuss the environmental temperatures experienced by ancient organisms that existed during the Precambrian era, as inferred from amino acid sequences of ancient proteins reconstructed by comparing modern homologous sequences.

2. Early Studies on the Environmental Temperatures of Ancestral Life

Although there has been a long-running debate about the environment of early life, no consensus has yet been obtained (Table 1). In particular, there has been intense debate about the environmental temperature of the last universal common ancestor. Although the last universal common ancestor is sometimes called Commonote [16] or *Commonote commonote* [17], I hereafter refer to it as LUCA, the most commonly used term. LUCA is not the oldest life, but rather the most recent common ancestor of all modern life. In the phylogenetic tree built by Furukawa et al. [18] (Figure 1), the left end corresponds to the origin of life. LUCA (indicated as 'a') is an intermediate ancestor from the origin of life to modern organisms.

Table 1. Theoretical and in silico studies for predicting the environmental temperatures of early organisms.

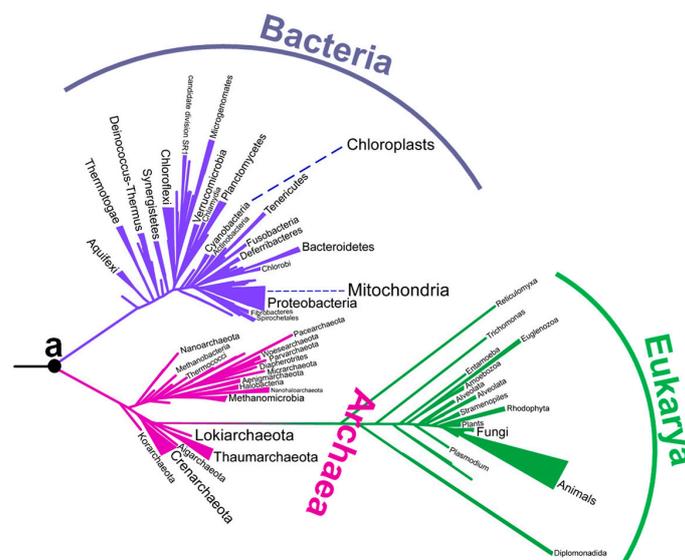| Target	Method	Conclusion	Refs.
Common ancestors of Archaea and Bacteria	rRNA tree	Hyperthermophilic	[19,20]
Bacterial common ancestors	rRNA tree	Mesophilic or thermophilic	[21]
LUCA	G + C content in rRNA	Mesophilic	[22]
LUCA	Reanalysis of the data used in [22]	Thermophilic or hyperthermophilic	[23]
LUCA	Evolution of reverse gyrase	Mesophilic or thermophilic	[24]
LUCA	A gene for reverse gyrase found in a gene set of LUCA	Hyperthermophilic	[25]
LUCA	G + C contents in rRNA and amino acid composition inferred using a non-homogeneous model	Psychrophilic or mesophilic	[26]
LUCA	Amino acid composition inferred using a non-homogeneous model	Mesophilic	[27]

Figure 1. A phylogenetic tree constructed from aminoacyl-tRNA synthetase sequences [18]. The position of Commonote (LUCA) is indicated with 'a'.

According to a frequently referenced phylogenetic tree based on small subunit ribosomal RNA sequences, branches for hyperthermophilic archaea and bacteria are concentrated near the root of the tree [28–30]. Therefore, the archaeal and bacterial common ancestors were both thought to be hyperthermophilic organisms [19,20]. In addition, Occam's razor suggests that their common ancestor, that is LUCA, was also hyperthermophilic and lived in a hot environment.

However, the theory of hyperthermophilic ancestry has often been criticized. Miller and Lazcano [31] argued that it was not likely that the earliest life was hyperthermophilic, because bio-related materials such as ATP are thermally instable. Indeed, it has been empirically demonstrated that ribose, a backbone of RNA, and its analogs quickly decompose at high temperatures [32]. Doolittle [33] pointed out that it is quite difficult to properly represent the early history of life on a tree. Therefore, an accurate tree cannot be obtained and any implications derived from the tree are hard to prove. Indeed, on a tree representing bacterial phylogeny built by Brochier and Philippe [21], the shortest and deepest branches were not for hyperthermophilic bacteria such as *Thermotogales* and *Aquificales*, but rather for mesophilic species. Therefore, they asserted that the hyperthermophilic bacteria emerged as a result of a secondary adaptation to high temperature.

Galtier et al. [22] thought that inferring the guanine plus cytosine (G + C) content of an ancestral rRNA would provide a powerful method to predict the optimum temperature of the ancestral organism. The G + C content of the stem region of prokaryotic ribosomal RNA (rRNA) and the optimum environmental temperature of the host organism are well correlated; an extant prokaryote with greater G + C content in rRNA often shows a higher optimum environmental temperature. The calculated G + C content of LUCA was not similar to the values found for organisms living at high temperatures. Accordingly, Galtier et al. proposed that LUCA was likely a mesophile. However, Di Giulio reanalyzed the same genome data set using a different computational algorithm, and obtained the contradicting conclusion that LUCA was thermophilic or hyperthermophilic [23,34,35].

Reverse gyrase, an ATP-dependent type I DNA topoisomerase, is possessed by all known hyperthermophilic species, and is therefore thought to be an essential protein for adaptation to very high temperatures [36,37]. Accordingly, the emergence of reverse gyrase might be crucial for the origin of hyperthermophilic organisms. Reverse gyrase consists of topoisomerase and helicase domains, which are evolutionarily independent of each other [38]. It is reasonably assumed that the topoisomerase and helicase families evolved independently in mesophilic or thermophilic organisms prior to the emergence of reverse gyrase [39]. Later, the domains fused to each other, and then were recruited by hyperthermophilic organisms [24]. This argument suggests that hyperthermophiles are descendants of mesophilic or thermophilic organisms. However, it cannot be ruled out that reverse gyrase had evolved prior to the emergence of LUCA. Indeed, a very recent study indicated that a gene for reverse gyrase was included in a 355-gene set that might have been possessed by LUCA [25]. Therefore, the discussion about the origin and evolution of reverse gyrase is compatible with the following scenario: at the time when the universal ancestor lived, a variety of organisms existed in a wide range of temperature environments, and when the surface temperature of Earth drastically increased due to meteorite impacts and various other reasons, only a hyperthermophilic ancestor survived [40].

3. Experimental Procedure to Reconstruct an Ancestral Protein Sequence

Since the end of the 20th century, the history of life and proteins has been traced by comparing the amino acid sequences of homologous proteins [15,17,41–48]. Ancestral protein sequences have also been inferred computationally [9,12,49,50]. Figure 2 shows a flowchart for reconstructing an ancestral protein. The first step of the reconstruction is to retrieve homologous protein sequences of the target protein from public databases. The homologous sequences are then aligned to generate a multiple sequence alignment. Our group often uses MAFFT [51] to align a set of homologous sequences, but manually corrects the positions of insertions and gaps if necessary. Next, the alignment and the homologous sequences are used to build a phylogenetic tree. An ancestral amino acid sequence is then computed using the tree topology,

the homologous sequences contained in the tree, and either a homogeneous or a non-homogeneous amino acid substitution model. The homogeneous model uses an approximation of constant global amino acid compositions in proteins throughout evolution [52]. In contrast, the non-homogeneous evolution model relaxes this constraint and allows for different global amino acid compositions at different times and for different lineages of the tree [22,53]. The positions of gaps/inserts are given in the ancestral sequence. We often use the program GASP [54] for this purpose. The gene encoding the inferred amino acid sequence is synthesized in vitro, and then expressed in a host organism such as *Escherichia coli*. Finally, the recombinant ancestral protein is purified and characterized experimentally. Theories and procedures of the reconstruction technique are described in greater detail in excellent reviews by Thornton [9], Gaucher et al. [10], and Merkl and Sterner [11].
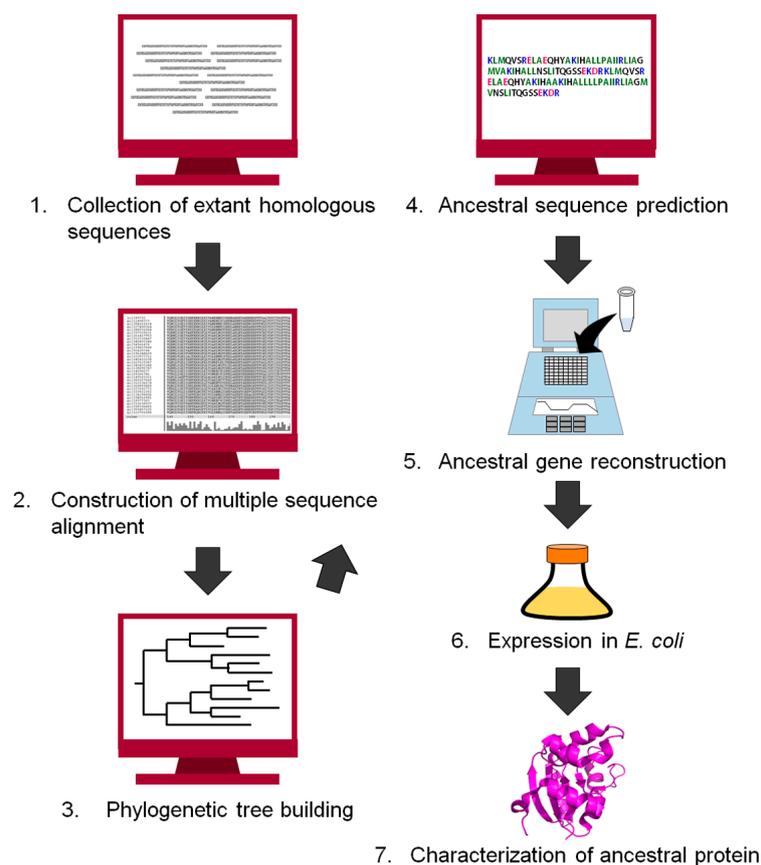

Figure 2. Flowchart of the procedure to infer an ancestral sequence and to reconstruct the ancestral protein in vitro.

4. Experimentally Testing If Ancestral Organisms Were Thermophiles

Yamagishi and coworkers developed an experimental way to test the thermophilicity of LUCA. They first inferred an ancestral amino acid sequence of a protein that might be possessed by LUCA. Then, one or a few amino acid(s) found in the inferred ancestral sequence were introduced into a protein from a modern thermophilic organism as amino acid substitution(s); then, the thermal stability of the mutant proteins was assessed. Using this method, they constructed mutants for 3-isopropylmalate dehydrogenase (IPMDH) from the hyperthermophile *Sulfolobus tokodaii* [55], isocitrate dehydrogenase from the extremely thermophile *Caldococcus noboribetus* [56], and glycyl-tRNA synthetase [57] and IPMDH [58] from the extremely thermophile *Thermus thermophilus*. From these experiments, they observed that the mutant proteins showed a trend toward greater thermal stability than the wild-type proteins. They asserted that their observations were evidence that LUCA possessed very thermostable

proteins, thus supporting the hyperthermophilicity of LUCA. Similar methods also improved the thermal stability of mesophilic proteins [59–61]. However, they compared a relatively small number of homologous amino acid sequences to infer the ancestral sequences, and it is not likely that a tree based on a small number of homologous sequences would accurately reflect the phylogeny of all modern life. Therefore, the inferred sequences might not represent the true ancestral sequences. Accordingly, the observed trend for increased thermal stability of mutant proteins may be an artifact of the sequence inference method [62].

Genes encoding entire ancestral amino acid sequences were synthesized to study the evolution of protein function. In 2005, ancestral forms of yeast alcohol dehydrogenase were reconstructed to determine the original function of the protein [42]. Thornton and coworkers intensively investigated changes in ligand specificities of steroid receptors [63–66]. The same group also used an ancestral sequence reconstruction technique to study the evolutionary process of increased complexity of eukaryotic V-ATPase proton pumps [67]. The experimental reconstruction method was also used to address the history of recruiting 20 genetically coded amino acids [46]. Thus, the reconstruction of an entire ancestral amino acid sequence in vitro is a commonly used technique to understand the histories of proteins and their host organisms.

Gaucher et al. [41] applied this technique to investigate the environmental temperature experienced by early organisms. They reconstructed two ancestral elongation factor-Tu proteins corresponding to the last bacterial common ancestor, and then characterized the optimum temperatures of the GTP hydrolysis activity. The ancestral proteins demonstrated optimum activity at temperatures similar to that of a modern thermophilic elongation factor-Tu, supporting the theory that the last bacterial common ancestor was likely a thermophile, rather than a hyperthermophile or mesophile [41]. Butzin et al. [44] also conducted similar ancestral sequence reconstruction experiments, and reported that the environmental temperatures of the most recent common ancestor of Thermotogales, an order of hyperthermophilic bacteria, were higher than those of its descendants, which are all hyperthermophilic.

As mentioned above, ancestral sequences of some proteins have been computationally predicted using a phylogenetic tree, and homologous amino acid sequences contained in the tree. However, protein sequences have evolved at different rates, and many mutations have accumulated during evolution. Therefore, for many proteins, it is very difficult to follow homology far back in time. Nucleoside diphosphate kinase (NDK) is distributed among Bacteria, Archaea and Eukarya, and most extant organisms possess the gene. Therefore, it is a reasonable assumption that ancient organisms also had an NDK gene. In addition, NDK sequences are well conserved among species, and a multiple alignment of extant NDK sequences contains few insertions/gaps that often interfere with the process of predicting reliable ancestral sequences. Therefore, one can suppose that an ancestral NDK sequence can be predicted with a high degree of confidence. However, a predicted ancestral sequence is also affected by the topology of the tree used to infer the sequence, and it is not possible to predict a definitively true tree topology. Indeed, while a number of phylogenetic and phylogenomic studies propose three domains in a universal tree of life [30,68–71], other studies instead support a two-domain hypothesis [72–75]. The tree illustrated in Figure 1 supports the two-domain hypothesis [18]. Therefore, three independent phylogenetic trees were built that differed in topology. Then, ancestral sequences of NDK were inferred using each tree, and experimentally reconstructed. The reconstructed proteins showed extreme thermal stability and high optimum temperature for catalytic activity, thus supporting the thermophilic ancestry of life [15]. The result is robust because similar characteristics were predicted for the ancestral proteins, even when using three topologically different phylogenetic trees.

Two other concerns are the reliability of the reconstructed ancestral amino acids, and the observed high thermal stability. Indeed, the reliability of some ancestral residue reconstruction has not been high enough, and therefore may not represent true ancestral residues (although most ancestral residues are strongly supported). We found that the predicted thermal stability of ancestral NDKs are valid, even if some residues in the reconstructed sequences do not represent the true ancestral residues [15].

Eick et al. [76] also reported that the observed characteristics of reconstructed ancestral proteins are robust to the uncertainty found in inferred sequences.

Sterner and coworkers reconstructed primordial enzyme complexes thought to be possessed by extinct species [47]. They resurrected the imidazole glycerol phosphate synthase (ImGPS) complex possessed by LUCA, and the tryptophan synthase (TS) complex possessed by the last bacterial common ancestor, and found that the two subunits (cyclase and glutaminase subunits) of the ancestral ImGPS, and the two subunits (α- and β- subunits) of the ancestral TS, were all thermostable. Moreover, it was observed that the ancestral cyclase and an extant glutaminase formed a complex structure and channeled ammonia from glutaminase to cyclase. The two ancestral subunits of TS also formed an $\alpha\beta\beta\alpha$ complex similar to the TS complexes from extant species. The two ancestral subunits mutually activated each other, and indole was channeled from the α subunit to the β subunit, which suggested that the sophisticated enzyme complexes responsible for substrate channeling and allosteric regulation had already been established when LUCA or the last bacterial common ancestor lived. The same research group also applied the sequence reconstruction technique to investigate the evolution of a TIM barrel protein fold [50] and identify an interface hotspot in a metabolic enzyme complex [77].

5. Ancestral Sequence Reconstruction Using a Non-Homogeneous Model

Some computational and empirical studies that support a theory of thermophilic ancestry used homogeneous amino acid substitution models to infer ancestral sequences [15,23,34,35,78]. The homogeneous substitution models assume that the global amino acid composition does not change among lineages and along phylogenetic trees. However, this does not accurately reflect evolution, because all sequences have different amino acid compositions. Therefore, homogeneous models are likely too simplistic to reliably infer ancestral sequences. In contrast, non-homogeneous amino acid substitution models relax this constraint by allowing different lineages to have different equilibrium compositions [22,53]. Some computational studies have focused on the environmental temperatures experienced by ancient life using a non-homogeneous evolution model [53]. Boussau et al. [26] suggested that the use of a non-homogeneous model was quite important in order to infer ancestral sequences more accurately. They reconstructed ancestral sequences of rRNAs and proteins in silico, and estimated the ancestral G + C contents of rRNA and the relative frequency of particular amino acid types. Based on their calculations, the last common archaeal and bacterial ancestors were thermophiles, but LUCA was a mesophilic or psychrophilic species [26]. Groussin et al. [27] also computed the relative frequency of each amino acid in proteins using the non-homogeneous models. According to their calculations, the last common archaeal and bacterial ancestors were likely to be thermophilic, but LUCA was likely to be mesophilic. However, it is possible that the early evolution of the amino acid repertoire [46,79–85] affected the frequency of each amino acid in primordial proteins. Accordingly, the thermal stability of a primordial protein—and therefore the environmental temperature of its host organism estimated from an analysis of amino acid contents—are inferential unless the stability is tested experimentally.

Using a non-homogeneous substitution model, we reanalyzed the NDK sequences that were previously used to reconstruct the ancestral NDK sequences based on a homogeneous model, and inferred additional ancestral NDK sequences [17]. The newly reconstructed ancestral NDKs also showed extremely high thermal stability, further supporting our conclusion that LUCA had a very thermally stable protein, even when ancestral sequences were inferred using a non-homogeneous substitution model. We also found that the denaturation temperature of NDK is well correlated with its host's optimum environmental temperature [15,48]. Therefore, the thermal stability of NDK works as a molecular thermometer. Using a calibration curve based on the correlation between the denaturation temperature of NDK and the optimum environmental temperature of the host, we estimated that the environmental temperature of LUCA was 97 ± 3 °C (Figure 3).
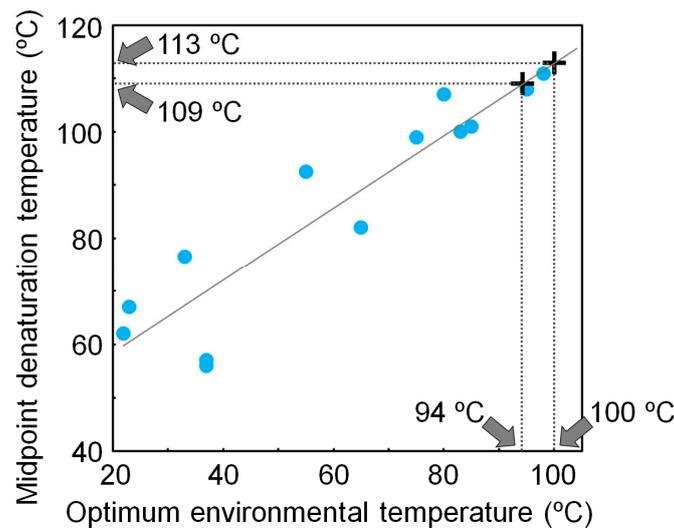
Figure 3. Relationship between the midpoint denaturation temperature of microbial nucleoside diphosphate kinases (NDKs) and their hosts' optimum environmental temperatures. The optimum environmental temperature of LUCA is estimated from the calibration curve and the denaturation temperatures of the reconstructed NDKs that might be possessed by LUCA.

6. Estimating Long-Term Change in Biosphere Temperature

Other ancestral sequence reconstruction studies expanded the timescale of targets to be reconstructed. Gaucher et al. [43] comprehensively analyzed the internal nodes of the bacterial phylogeny, and estimated the environmental temperatures of ancestral bacteria that existed 0.5–3.5 Gya. Their results suggest that the bacterial ancestor was thermophilic, and adapted later to progressively lower-temperature environments over Precambrian time. This trend is similar to a gradual cooling of the ancient ocean, as suggested by the oxygen isotope compositions of marine cherts [86,87]. A similar experiment was done by Butzin et al. [44] with proteins possessed by hyperthermophilic bacteria as the targets. They reported that the environmental temperatures of the most recent existing common ancestor of Thermotogales were higher than those of its descendants, which are all hyperthermophiles.

Groussin and Gouy [88] targeted the entire domain of Archaea. They analyzed the G + C contents of rRNAs, and the frequency of each amino acid in a set of proteins possessed by archaeal ancestors at internal nodes of the archaeal phylogeny. These data were then used to estimate environmental temperatures. The results indicated that the last common archaeal ancestor was a hyperthermophile, and extant mesophilic archaea have adapted to lower environmental temperatures over evolution.

Hart et al. [45] reported a slightly different result. The reconstructed ancestral ribonuclease H1 (RNH) suggests an unfolding temperature that was higher than that of an extant mesophilic RNH, but lower than that of an extant thermophilic RNH. They argued that the high thermal stability observed for the extant thermophilic RNH is a result of gradual adaptation to higher temperatures over time.

The high thermal stabilities of these reconstructed proteins are compatible with the high environmental temperatures of Archaean life. However, these ancestral organisms may have lived in local high-temperature environments; therefore, the results may not reflect the ambient surface temperature of ancient Earth. To overcome such a problem, Garcia et al. [48] restricted their targets to ancestral NDKs reconstructed by comparing proteins from phototrophic species that required light for growth. They reconstructed ancestral NDKs that might represent the common ancestors of cyanobacteria (oxygenic photosynthetic prokaryotes), nostocaleans (later-evolved cyanobacteria), Viridiplantae (green algae and land plants) and Embryophyta (land plants only). The ancestors of cyanobacteria, nostocaleans, Viridiplantae and Embryophyta are predicted to have existed approximately 2.7–3.1 Gya, 2.1–2.3 Gya, 0.70–0.85 Gya, and 0.44–0.46 Gya, respectively. They experimentally determined the denaturation temperatures of the reconstructed NDKs and estimated the environmental temperatures

as approximately 64–82 °C (ancestor of cyanobacteria), 36–56 °C (ancestor of nostocaleans), 24–62 °C (ancestor of Viridiplantae) and 20–58 °C (ancestor of Embryophyta). The results are quite consistent to those inferred from isotope-based data [86,87], and suggest a general cooling of Earth's surface environment over time from Archaean (~55–85 °C) to present (~15 °C) (Figure 4).
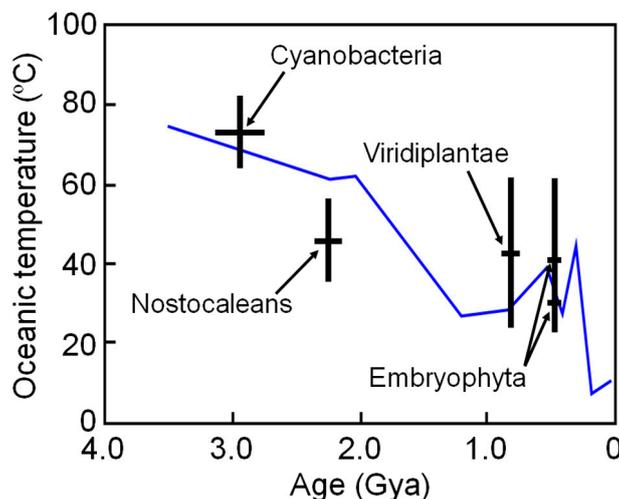

Figure 4. Environmental temperature ranges inferred from the midpoint denaturation temperatures of the reconstructed NDK as a function of age, with uncertainty when respective ancestral organisms first appeared. Paleotemperatures inferred from isotope-based evidence in marine cherts are also shown for comparison (blue line) [86].

7. Limitations of Ancestral Inference to Estimate Temperatures of Ancient Biosphere

The characterization of the physical properties of ancestral proteins reconstructed by phylogenetic approaches has provided independent evidence of the environmental conditions of ancient Earth, and complemented the data obtained from geological and paleontological studies. Techniques to infer ancestral sequences have greatly improved in the last decade, but ancestral sequences still cannot be reconstructed with absolute certainty, and there are likely a number of inaccurate ancestral residues present in inferred sequences. An inaccurate reconstructed sequence would result in an overestimate of its thermodynamic stability, especially if the sequence was inferred by maximum parsimony or maximum likelihood [62]. Their argument was based on the observation that these two algorithms tend to adopt the amino acid that occurs most frequently among modern protein sequences as the ancestral residue. The most frequent (consensus) amino acids at a site among homologous proteins have a greater contribution to thermal stability of a protein than less frequent amino acids [89–93]. Ancestral inference has an inherent, undeniable tendency to converge into the consensus amino acids at many positions, and it is therefore difficult to discriminate if the observed high stability is due to the antiquity of the residue or the consensus effect; as such, the environmental temperatures inferred from the thermal stability of reconstructed ancestral proteins could be overestimated.

High temperature may not have been the only environmental parameter requiring a high stability of ancient proteins. Tawfik et al. suggested that high oxidative pressure and radiation levels, the absence of cellular osmolytes (that are prevalent in thermophiles) and/or chaperones, or low fidelity of the transcription–translation machinery might be involved in the high stability of ancient proteins [94,95].

The position of the root of a tree used to infer an ancestral sequence would also affect the results. Placement of LUCA on a tree used to reconstruct ancestral sequences would require inclusion of two or more paralogous proteins that diverged prior to the appearance of LUCA [96]. In 1989, such composite trees were reported from two independent groups using elongation factor and H^+-ATPase, respectively [97,98]. In both studies, the position of LUCA was located within the branch connecting the archaeal and bacterial common ancestors. The universal ancestor was placed at the same position

on other composite trees based on isocitrate dehydrogenase and 3-isopropylmalte dehydrogenase [55], and aminoacyl-tRNA synthase [18,99,100]. However, it has also been suggested that LUCA was located within the Bacteria domain [101–106]. Therefore, the conclusion that LUCA was thermophilic or hyperthermophilic is dependent on the hypothesis that archaea and bacteria were derived from LUCA. In addition, the incorporation of information on gene duplication, lateral gene transfer, and gene loss would improve the accuracy of inferred ancestral sequences, and therefore assumptions about the physical properties of the reconstructed protein [107]. Future work should be done to account for these evolutionary events.

8. Conclusions

The last two decades have seen technical improvements for inferring the ancestral nucleotide and amino acid sequences of ancestral genes and proteins. Now, these improvements provide an alternative and independent method to investigate the early evolution of life and Earth's biosphere. However, computer-based studies to predict the environmental temperature of LUCA have provided conflicting results (Table 1). It is not evident that the correlation between the composition of genes or proteins, and the environmental temperature of the host found for present-day organisms, is applicable to ancient organisms. Therefore, it may not be appropriate to infer the environmental temperature of ancient life from a predicted base or amino acid composition.

Reconstruction of ancestral proteins suggests that LUCA possessed proteins with a thermal stability similar to, or even greater than, that of extant hyperthermophilic proteins (Table 2). The results were robust to modern protein datasets, the topology of the phylogenetic tree, and amino acid substitution models that were used to infer the ancestral sequences. The uncertainty of ancestral residues associated with the ancestral sequence reconstruction did not significantly affect the predicted thermal stability of the ancestral proteins. Therefore, LUCA is likely to have been a thermophile or hyperthermophile that thrived at a very high temperature, and its mesophilic descendants were adapted to lower temperatures as Earth's surface environment cooled over time.

Table 2. In vitro experimental studies for predicting the environmental temperatures of early organisms.

Target	Method	Conclusion	Refs.
LUCA	Introduction of a few amino acids into the sequence of a modern thermophilic protein	Hyperthermophilic	[55–58]
Bacterial common ancestors	Reconstruction of ancestral elongation factors	Thermophilic	[41,43]
Common ancestor of Thermotogales	Reconstruction of ancestral *Myo*-inositol-3-phospate synthases	Hyperthermophilic	[44]
LUCA	Reconstruction of ancestral NDKs using a homogeneous substitution model	Thermophilic or hyperthermophilic	[15]
LUCA	Reconstruction of ancestral NDKs using a non-homogeneous substitution model	Hyperthermophilic	[17]

Even if LUCA was a thermophile or a hyperthermophile, it does not mean that the first life on Earth was born in a high temperature environment. Rather, our conclusion is compatible with the idea that most organisms existing at the time of LUCA became extinct, and that only LUCA that was adapted to high temperatures survived when the early Earth's temperature drastically increased due to meteorite impacts [40].

Conflicts of Interest: The author declares no conflict of interest.

References

1. Schopf, J.W. Microfossils of the early Archean apex chert: New evidence of the antiquity of life. *Science* **1993**, *260*, 640–646. [CrossRef] [PubMed]

2. Wacey, D.; McLoughlin, N.; Green, O.; Parnell, J.; Stoakes, C.A.; Brasier, M.D. The ~3.4 billion-year-old strelley pool sandstone: A new window into early life on Earth. *Int. J. Astrobiol.* **2006**, *5*, 333. [CrossRef]

3. Nutman, A.P.; Bennett, V.C.; Friend, C.R.; Van Kranendonk, M.J.; Chivas, A.R. Rapid emergence of life shown by discovery of 3700-million-year-old microbial structures. *Nature* **2016**, *537*, 535–538. [CrossRef] [PubMed]

4. Dodd, M.S.; Papineau, D.; Grenne, T.; Slack, J.F.; Rittner, M.; Pirajno, F.; O'Neil, J.; Little, C.T. Evidence for early life in Earth's oldest hydrothermal vent precipitates. *Nature* **2017**, *543*, 60–64. [CrossRef] [PubMed]

5. Shen, Y.; Buick, R.; Canfield, D.E. Isotopic evidence for microbial sulphate reduction in the early Archaean era. *Nature* **2001**, *410*, 77–81. [CrossRef] [PubMed]

6. Ueno, Y.; Yamada, K.; Yoshida, N.; Maruyama, S.; Isozaki, Y. Evidence from fluid inclusions for microbial methanogenesis in the early Archaean era. *Nature* **2006**, *440*, 516–519. [CrossRef] [PubMed]

7. Rosing, M.T. ^{13}C-depleted carbon microparticles in >3700-Ma sea-floor sedimentary rocks from west Greenland. *Science* **1999**, *283*, 674–676. [CrossRef] [PubMed]

8. Bell, E.A.; Boehnke, P.; Harrison, T.M.; Mao, W.L. Potentially biogenic carbon preserved in a 4.1 billion-year-old zircon. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 14518–14521. [CrossRef] [PubMed]

9. Thornton, J.W. Resurrecting ancient genes: Experimental analysis of extinct molecules. *Nat. Rev. Genet.* **2004**, *5*, 366–375. [CrossRef] [PubMed]

10. Gaucher, E.A.; Kratzer, J.T.; Randall, R.N. Deep phylogeny—How a tree can help characterize early life on Earth. *Cold. Spring Harb. Perspect. Biol.* **2010**, *2*, a002238. [CrossRef] [PubMed]

11. Merkl, R.; Sterner, R. Ancestral protein reconstruction: Techniques and applications. *Biol. Chem.* **2016**, *397*, 1–21. [CrossRef] [PubMed]

12. Perez-Jimenez, R.; Ingles-Prieto, A.; Zhao, Z.M.; Sanchez-Romero, I.; Alegre-Cebollada, J.; Kosuri, P.; Garcia-Manyes, S.; Kappock, T.J.; Tanokura, M.; Holmgren, A.; et al. Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. *Nat. Struct. Mol. Biol.* **2011**, *18*, 592–596. [CrossRef] [PubMed]

13. Akanuma, S.; Yamagishi, A. A strategy for designing thermostable enzymes by reconstructing ancestral sequences possessed by ancient life. In *Biotechnology of Extremophiles: Advances and Challenges*; Rampelotto, P.H., Ed.; Springer: Berlin, Germany, 2016; pp. 581–596. [CrossRef]

14. Boussau, B.; Gouy, M. What genomes have to say about the evolution of the earth. *Gondwana Res.* **2012**, *21*, 483–494. [CrossRef]

15. Akanuma, S.; Nakajima, Y.; Yokobori, S.; Kimura, M.; Nemoto, N.; Mase, T.; Miyazono, K.; Tanokura, M.; Yamagishi, A. Experimental evidence for the thermophilicity of ancestral life. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 11067–11072. [CrossRef] [PubMed]

16. Yamagishi, A.; Kon, T.; Takahashi, G.; Oshima, T. From the common ancestor of all living organisms to protoeukaryotic cell. In *Thermophiles: The Keys to Molecular Evolution and the Origin of Life?* Wiegel, J., Adams, M.W.W., Eds.; Taylor & Francis: Abingdon, UK, 1998; pp. 287–295.

17. Akanuma, S.; Yokobori, S.; Nakajima, Y.; Bessho, M.; Yamagishi, A. Robustness of predictions of extremely thermally stable proteins in ancient organisms. *Evolution* **2015**, *69*, 2954–2962. [CrossRef] [PubMed]

18. Furukawa, R.; Nakagawa, M.; Kuroyanagi, T.; Yokobori, S.I.; Yamagishi, A. Quest for ancestors of eukaryal cells based on phylogenetic analyses of aminoacyl-tRNA synthetases. *J. Mol. Evol.* **2017**, *84*, 51–66. [CrossRef] [PubMed]

19. Pace, N.R. Origin of life—Facing up to the physical setting. *Cell* **1991**, *65*, 531–533. [CrossRef]

20. Stetter, K.O. Hyperthermophilic procaryotes. *FEMS Microbiol. Rev.* **1996**, *18*, 149–158. [CrossRef]

21. Brochier, C.; Philippe, H. Phylogeny: A non-hyperthermophilic ancestor for bacteria. *Nature* **2002**, *417*, 244. [CrossRef] [PubMed]

22. Galtier, N.; Tourasse, N.; Gouy, M. A nonhyperthermophilic common ancestor to extant life forms. *Science* **1999**, *283*, 220–221. [CrossRef] [PubMed]

23. Di Giulio, M. The universal ancestor lived in a thermophilic or hyperthermophilic environment. *J. Theor. Biol.* **2000**, *203*, 203–213. [CrossRef] [PubMed]

24. Forterre, P. A hot topic: The origin of hyperthermophiles. *Cell* **1996**, *85*, 789–792. [CrossRef]

25. Weiss, M.C.; Sousa, F.L.; Mrnjavac, N.; Neukirchen, S.; Roettger, M.; Nelson-Sathi, S.; Martin, W.F. The physiology and habitat of the last universal common ancestor. *Nat. Microbiol.* **2016**, *1*, 16116. [CrossRef] [PubMed]

26. Boussau, B.; Blanquart, S.; Necsulea, A.; Lartillot, N.; Gouy, M. Parallel adaptations to high temperatures in the Archaean eon. *Nature* **2008**, *456*, 942–945. [CrossRef] [PubMed]

27. Groussin, M.; Boussau, B.; Charles, S.; Blanquart, S.; Gouy, M. The molecular signal for the adaptation to cold temperature during early life on Earth. *Biol. Lett.* **2013**, *9*, 20130608. [CrossRef] [PubMed]

28. Woese, C.R. Bacterial evolution. *Microbiol. Rev.* **1987**, *51*, 221–271. [PubMed]

29. Achenbach-Richter, L.; Gupta, R.; Zillig, W.; Woese, C.R. Rooting the archaebacterial tree: The pivotal role of *Thermococcus celer* in archaebacterial evolution. *Syst. Appl. Microbiol.* **1988**, *10*, 231–240. [CrossRef]

30. Woese, C.R.; Kandler, O.; Wheelis, M.L. Towards a natural system of organisms: Proposal for the domains archaea, bacteria, and eucarya. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 4576–4579. [CrossRef] [PubMed]

31. Miller, S.L.; Lazcano, A. The origin of life—Did it occur at high temperatures? *J. Mol. Evol.* **1995**, *41*, 689–692. [CrossRef] [PubMed]

32. Larralde, R.; Robertson, M.P.; Miller, S.L. Rates of decomposition of ribose and other sugars: Implications for chemical evolution. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8158–8160. [CrossRef] [PubMed]

33. Doolittle, W.F. Phylogenetic classification and the universal tree. *Science* **1999**, *284*, 2124–2129. [CrossRef] [PubMed]

34. Di Giulio, M. The universal ancestor and the ancestor of bacteria were hyperthermophiles. *J. Mol. Evol.* **2003**, *57*, 721–730. [CrossRef] [PubMed]

35. Di Giulio, M. The universal ancestor was a thermophile or a hyperthermophile: Tests and further evidence. *J. Theor. Biol.* **2003**, *221*, 425–436. [CrossRef] [PubMed]

36. Forterre, P. A hot story from comparative genomics: Reverse gyrase is the only hyperthermophile-specific protein. *Trends. Genet.* **2002**, *18*, 236–237. [CrossRef]

37. Atomi, H.; Matsumi, R.; Imanaka, T. Reverse gyrase is not a prerequisite for hyperthermophilic life. *J. Bacteriol.* **2004**, *186*, 4829–4833. [CrossRef] [PubMed]

38. Declais, A.C.; Marsault, J.; Confalonieri, F.; de La Tour, C.B.; Duguet, M. Reverse gyrase, the two domains intimately cooperate to promote positive supercoiling. *J. Biol. Chem.* **2000**, *275*, 19498–19504. [CrossRef] [PubMed]

39. Heine, M.; Chandra, S.B. The linkage between reverse gyrase and hyperthermophiles: A review of their invariable association. *J. Microbiol.* **2009**, *47*, 229–234. [CrossRef] [PubMed]

40. Nisbet, E.G.; Sleep, N.H. The habitat and nature of early life. *Nature* **2001**, *409*, 1083–1091. [CrossRef] [PubMed]

41. Gaucher, E.A.; Thomson, J.M.; Burgan, M.F.; Benner, S.A. Inferring the palaeoenvironment of ancient bacteria on the basis of resurrected proteins. *Nature* **2003**, *425*, 285–288. [CrossRef] [PubMed]

42. Thomson, J.M.; Gaucher, E.A.; Burgan, M.F.; De Kee, D.W.; Li, T.; Aris, J.P.; Benner, S.A. Resurrecting ancestral alcohol dehydrogenases from yeast. *Nat. Genet.* **2005**, *37*, 630–635. [CrossRef] [PubMed]

43. Gaucher, E.A.; Govindarajan, S.; Ganesh, O.K. Palaeotemperature trend for Precambrian life inferred from resurrected proteins. *Nature* **2008**, *451*, 704–707. [CrossRef] [PubMed]

44. Butzin, N.C.; Lapierre, P.; Green, A.G.; Swithers, K.S.; Gogarten, J.P.; Noll, K.M. Reconstructed ancestral Myo-inositol-3-phosphate synthases indicate that ancestors of the Thermococcales and Thermotoga species were more thermophilic than their descendants. *PLoS ONE* **2013**, *8*, e84300. [CrossRef] [PubMed]

45. Hart, K.M.; Harms, M.J.; Schmidt, B.H.; Elya, C.; Thornton, J.W.; Marqusee, S. Thermodynamic system drift in protein evolution. *PLoS Biol.* **2014**, *12*, e1001994. [CrossRef] [PubMed]

46. Fournier, G.P.; Alm, E.J. Ancestral reconstruction of a pre-LUCA aminoacyl-tRNA synthetase ancestor supports the late addition of Trp to the genetic code. *J. Mol. Evol.* **2015**, *80*, 171–185. [CrossRef] [PubMed]

47. Busch, F.; Rajendran, C.; Heyn, K.; Schlee, S.; Merkl, R.; Sterner, R. Ancestral tryptophan synthase reveals functional sophistication of primordial enzyme complexes. *Cell Chem. Biol.* **2016**, *23*, 709–715. [CrossRef] [PubMed]

48. Garcia, A.K.; Schopf, J.W.; Yokobori, S.I.; Akanuma, S.; Yamagishi, A. Reconstructed ancestral enzymes suggest long-term cooling of Earth's photic zone since the Archean. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 4619–4624. [CrossRef] [PubMed]

49. Messier, W.; Stewart, C.B. Episodic adaptive evolution of primate lysozymes. *Nature* **1997**, *385*, 151–154. [CrossRef] [PubMed]

50. Richter, M.; Bosnali, M.; Carstensen, L.; Seitz, T.; Durchschlag, H.; Blanquart, S.; Merkl, R.; Sterner, R. Computational and experimental evidence for the evolution of a $(\beta\alpha)_8$-barrel protein from an ancestral quarter-barrel stabilised by disulfide bonds. *J. Mol. Biol.* **2010**, *398*, 763–773. [CrossRef] [PubMed]

51. Katoh, K.; Standley, D.M. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]

52. Yang, Z. Paml: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **1997**, *13*, 555–556. [CrossRef] [PubMed]

53. Blanquart, S.; Lartillot, N. A site- and time-heterogeneous model of amino acid replacement. *Mol. Biol. Evol.* **2008**, *25*, 842–858. [CrossRef] [PubMed]

54. Edwards, R.J.; Shields, D.C. GASP: Gapped ancestral sequence prediction for proteins. *BMC Bioinform.* **2004**, *5*, 123. [CrossRef] [PubMed]

55. Miyazaki, J.; Nakaya, S.; Suzuki, T.; Tamakoshi, M.; Oshima, T.; Yamagishi, A. Ancestral residues stabilizing 3-isopropylmalate dehydrogenase of an extreme thermophile: Experimental evidence supporting the thermophilic common ancestor hypothesis. *J. Biochem.* **2001**, *129*, 777–782. [CrossRef] [PubMed]

56. Iwabata, H.; Watanabe, K.; Ohkuri, T.; Yokobori, S.; Yamagishi, A. Thermostability of ancestral mutants of *Caldococcus noboribetus* isocitrate dehydrogenase. *FEMS Microbiol. Lett.* **2005**, *243*, 393–398. [CrossRef] [PubMed]

57. Shimizu, H.; Yokobori, S.; Ohkuri, T.; Yokogawa, T.; Nishikawa, K.; Yamagishi, A. Extremely thermophilic translation system in the common ancestor commonote: Ancestral mutants of glycyl-tRNA synthetase from the extreme thermophile *Thermus thermophilus*. *J. Mol. Biol.* **2007**, *369*, 1060–1069. [CrossRef] [PubMed]

58. Watanabe, K.; Ohkuri, T.; Yokobori, S.; Yamagishi, A. Designing thermostable proteins: Ancestral mutants of 3-isopropylmalate dehydrogenase designed by using a phylogenetic tree. *J. Mol. Biol.* **2006**, *355*, 664–674. [CrossRef] [PubMed]

59. Yamashiro, K.; Yokobori, S.; Koikeda, S.; Yamagishi, A. Improvement of *Bacillus circulans* beta-amylase activity attained using the ancestral mutation method. *Protein Eng. Des. Sel.* **2010**, *23*, 519–528. [CrossRef] [PubMed]

60. Semba, Y.; Ishida, M.; Yokobori, S.; Yamagishi, A. Ancestral amino acid substitution improves the thermal stability of recombinant lignin-peroxidase from white-rot fungi, *Phanerochaete chrysosporium* strain UAMH 3641. *Protein Eng. Des. Sel.* **2015**, *28*, 221–230. [CrossRef] [PubMed]

61. Fukuda, Y.; Abe, A.; Tamura, T.; Kishimoto, T.; Sogabe, A.; Akanuma, S.; Yokobori, S.; Yamagishi, A.; Imada, K.; Inagaki, K. Epistasis effects of multiple ancestral-consensus amino acid substitutions on the thermal stability of glycerol kinase from *Cellulomonas* sp. NT3060. *J. Biosci. Bioeng.* **2016**, *121*, 497–502. [CrossRef] [PubMed]

62. Williams, P.D.; Pollock, D.D.; Blackburne, B.P.; Goldstein, R.A. Assessing the accuracy of ancestral protein reconstruction methods. *PLoS. Comput. Biol.* **2006**, *2*, e69. [CrossRef] [PubMed]

63. Bridgham, J.T.; Carroll, S.M.; Thornton, J.W. Evolution of hormone-receptor complexity by molecular exploitation. *Science* **2006**, *312*, 97–101. [CrossRef] [PubMed]

64. Ortlund, E.A.; Bridgham, J.T.; Redinbo, M.R.; Thornton, J.W. Crystal structure of an ancient protein: Evolution by conformational epistasis. *Science* **2007**, *317*, 1544–1548. [CrossRef] [PubMed]

65. Bridgham, J.T.; Ortlund, E.A.; Thornton, J.W. An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature* **2009**, *461*, 515–519. [CrossRef] [PubMed]

66. Carroll, S.M.; Ortlund, E.A.; Thornton, J.W. Mechanisms for the evolution of a derived function in the ancestral glucocorticoid receptor. *PLoS Genet.* **2011**, *7*, e1002117. [CrossRef] [PubMed]

67. Finnigan, G.C.; Hanson-Smith, V.; Stevens, T.H.; Thornton, J.W. Evolution of increased complexity in a molecular machine. *Nature* **2012**, *481*, 360–364. [CrossRef] [PubMed]

68. Harris, J.K.; Kelley, S.T.; Spiegelman, G.B.; Pace, N.R. The genetic core of the universal ancestor. *Genome Res.* **2003**, *13*, 407–412. [CrossRef] [PubMed]

69. Ciccarelli, F.D.; Doerks, T.; von Mering, C.; Creevey, C.J.; Snel, B.; Bork, P. Toward automatic reconstruction of a highly resolved tree of life. *Science* **2006**, *311*, 1283–1287. [CrossRef] [PubMed]

70. Yutin, N.; Makarova, K.S.; Mekhedov, S.L.; Wolf, Y.I.; Koonin, E.V. The deep archaeal roots of eukaryotes. *Mol. Biol. Evol.* **2008**, *25*, 1619–1630. [CrossRef] [PubMed]

71. Rinke, C.; Schwientek, P.; Sczyrba, A.; Ivanova, N.N.; Anderson, I.J.; Cheng, J.F.; Darling, A.; Malfatti, S.; Swan, B.K.; Gies, E.A.; et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **2013**, *499*, 431–437. [CrossRef] [PubMed]

72. Rivera, M.C.; Lake, J.A. Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* **1992**, *257*, 74–76. [CrossRef] [PubMed]

73. Cox, C.J.; Foster, P.G.; Hirt, R.P.; Harris, S.R.; Embley, T.M. The archaebacterial origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 20356–20361. [CrossRef] [PubMed]

74. 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* **2013**, *504*, 231–236. [CrossRef] [PubMed]

75. Raymann, K.; Brochier-Armanet, C.; Gribaldo, S. The two-domain tree of life is linked to a new root for the archaea. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 6670–6675. [CrossRef] [PubMed]

76. Eick, G.N.; Bridgham, J.T.; Anderson, D.P.; Harms, M.J.; Thornton, J.W. Robustness of reconstructed ancestral protein functions to statistical uncertainty. *Mol. Biol. Evol.* **2017**, *34*, 247–261. [CrossRef] [PubMed]

77. Holinski, A.; Heyn, K.; Merkl, R.; Sterner, R. Combining ancestral sequence reconstruction with protein design to identify an interface hotspot in a key metabolic enzyme complex. *Proteins* **2017**, *85*, 312–321. [CrossRef] [PubMed]

78. Brooks, D.J.; Fresco, J.R.; Singh, M. A novel method for estimating ancestral amino acid composition and its application to proteins of the last universal ancestor. *Bioinformatics* **2004**, *20*, 2251–2257. [CrossRef] [PubMed]

79. Wong, J.T. A co-evolution theory of the genetic code. *Proc. Natl. Acad. Sci. USA* **1975**, *72*, 1909–1912. [CrossRef] [PubMed]

80. Baumann, U.; Oro, J. Three stages in the evolution of the genetic code. *BioSystems* **1993**, *29*, 133–141. [CrossRef]

81. Trifonov, E.N. Consensus temporal order of amino acids and evolution of the triplet code. *Gene* **2000**, *261*, 139–151. [CrossRef]

82. Akanuma, S.; Kigawa, T.; Yokoyama, S. Combinatorial mutagenesis to restrict amino acid usage in an enzyme to a reduced set. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13549–13553. [CrossRef] [PubMed]

83. Brooks, D.J.; Fresco, J.R.; Lesk, A.M.; Singh, M. Evolution of amino acid frequencies in proteins over deep time: Inferred order of introduction of amino acids into the genetic code. *Mol. Biol. Evol.* **2002**, *19*, 1645–1655. [CrossRef] [PubMed]

84. Jordan, I.K.; Kondrashov, F.A.; Adzhubei, I.A.; Wolf, Y.I.; Koonin, E.V.; Kondrashov, A.S.; Sunyaev, S. A universal trend of amino acid gain and loss in protein evolution. *Nature* **2005**, *433*, 633–638. [CrossRef] [PubMed]

85. Longo, L.M.; Lee, J.; Blaber, M. Simplified protein design biased for prebiotic amino acids yields a foldable, halophilic protein. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 2135–2139. [CrossRef] [PubMed]

86. Knauth, L.P.; Lowe, D.R. Oxygen isotope geochemistry of cherts from the Onverwacht Group (3.4 billion years), Transvaal, South Africa, with implications for secular variations in the isotopic composition of cherts. *Earth Planet. Sci. Lett.* **1978**, *41*, 209–222. [CrossRef]

87. Robert, F.; Chaussidon, M. A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. *Nature* **2006**, *443*, 969–972. [CrossRef] [PubMed]

88. Groussin, M.; Gouy, M. Adaptation to environmental temperature is a major determinant of molecular evolutionary rates in archaea. *Mol. Biol. Evol.* **2011**, *28*, 2661–2674. [CrossRef] [PubMed]

89. Steipe, B.; Schiller, B.; Pluckthun, A.; Steinbacher, S. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* **1994**, *240*, 188–192. [CrossRef] [PubMed]

90. Rath, A.; Davidson, A.R. The design of a hyperstable mutant of the Abp1p SH3 domain by sequence alignment analysis. *Protein Sci.* **2000**, *9*, 2457–2469. [CrossRef] [PubMed]

91. Lehmann, M.; Kostrewa, D.; Wyss, M.; Brugger, R.; D'Arcy, A.; Pasamontes, L.; van Loon, A.P. From DNA sequence to improved functionality: Using protein sequence comparisons to rapidly design a thermostable consensus phytase. *Protein Eng.* **2000**, *13*, 49–57. [CrossRef] [PubMed]

92. Steipe, B. Consensus-based engineering of protein stability: From intrabodies to thermostable enzymes. *Methods Enzymol.* **2004**, *388*, 176–186. [CrossRef] [PubMed]

93. Loening, A.M.; Fenn, T.D.; Wu, A.M.; Gambhir, S.S. Consensus guided mutagenesis of Renilla luciferase yields enhanced stability and light output. *Protein Eng. Des. Sel.* **2006**, *19*, 391–400. [CrossRef] [PubMed]

94. Goldsmith, M.; Tawfik, D.S. Potential role of phenotypic mutations in the evolution of protein expression and stability. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6197–6202. [CrossRef] [PubMed]

95. Trudeau, D.L.; Kaltenbach, M.; Tawfik, D.S. On the potential origins of the high stability of reconstructed ancestral proteins. *Mol. Biol. Evol.* **2016**, *33*, 2633–2641. [CrossRef] [PubMed]

96. Akanuma, S.; Yokobori, S.; Yamagishi, A. Comparative genomics of thermophilic bacteria and archaea. In *Thermophilic Microbes in Environmental and Industrial Biotechnology*; Satyanarayana, T., Litterchild, J., Kawarabayasi, Y., Eds.; Springer: Berlin, Germany, 2013; pp. 331–349.

97. Iwabe, N.; Kuma, K.; Hasegawa, M.; Osawa, S.; Miyata, T. Evolutionary relationship of archaebacteria, eubacteria, and eukaryotes inferred from phylogenetic trees of duplicated genes. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9355–9359. [CrossRef] [PubMed]

98. Gogarten, J.P.; Kibak, H.; Dittrich, P.; Taiz, L.; Bowman, E.J.; Bowman, B.J.; Manolson, M.F.; Poole, R.J.; Date, T.; Oshima, T.; et al. Evolution of the vacuolar H^+-ATPase: Implications for the origin of eukaryotes. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 6661–6665. [PubMed]

99. Brown, J.R.; Doolittle, W.F. Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 2441–2445. [CrossRef] [PubMed]

100. Fournier, G.P.; Andam, C.P.; Alm, E.J.; Gogarten, J.P. Molecular evolution of aminoacyl tRNA synthetase proteins in the early history of life. *Orig. Life Evol. Biosph.* **2011**, *41*, 621–632. [CrossRef] [PubMed]

101. Cavalier-Smith, T. The neomuran origin of archaebacteria, the negibacterial root of the universal tree and bacterial megaclassification. *Int. J. Syst. Evol. Microbiol.* **2002**, *52*, 7–76. [CrossRef] [PubMed]

102. Cavalier-Smith, T. Cell evolution and Earth history: Stasis and revolution. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2006**, *361*, 969–1006. [CrossRef] [PubMed]

103. Cavalier-Smith, T. Rooting the tree of life by transition analyses. *Biol. Direct* **2006**, *1*, 19. [CrossRef] [PubMed]

104. Lake, J.A.; Servin, J.A.; Herbold, C.W.; Skophammer, R.G. Evidence for a new root of the tree of life. *Syst. Biol.* **2008**, *57*, 835–843. [CrossRef] [PubMed]

105. Lake, J.A.; Skophammer, R.G.; Herbold, C.W.; Servin, J.A. Genome beginnings: Rooting the tree of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2009**, *364*, 2177–2185. [CrossRef] [PubMed]

106. Cavalier-Smith, T. Deep phylogeny, ancestral groups and the four ages of life. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2010**, *365*, 111–132. [CrossRef] [PubMed]

107. Groussin, M.; Hobbs, J.K.; Szollosi, G.J.; Gribaldo, S.; Arcus, V.L.; Gouy, M. Toward more accurate ancestral protein genotype-phenotype reconstructions with the use of species tree-aware gene trees. *Mol. Biol. Evol.* **2015**, *32*, 13–22. [CrossRef] [PubMed]