The hyperthermophilic origin of life revisited

D.W. Schwartzman*1 and C.H. Lineweaver†

*Department of Biology, Howard University, Washington DC 20059, U.S.A., and †Department of Astrophysics and Optics, University of New South Wales, Sydney 2052, Australia

Abstract

We revisit the case for the hyperthermophilic scenario for the origin of life and the last common ancestor. Evidence includes studies of phylogenetic trees, rRNA, G and C content, hyperthermophilic proteins, correlations between maximal temperature tolerances and genetic distances, saline stabilization of DNA/RNA, and the inferred climatic temperatures of the early Earth. Although some doubts remain, the case for hot biogenesis and the last common ancestor has gotten stronger.

A hyperthermophilic LCA (last common ancestor) of life?

A hyperthermophilic LCA of life has long been thought to be implied by the rRNA phylogenetic tree (Figure 1) [1,2]. While the deep-rootedness of hyperthermophilic Archaea still appears to be robust [3,4], some researchers argue that even deeply rooted Bacteria acquired their hyperthermophily by horizontal gene transfer from Archaea [5–7], a view under challenge [8]. All known hyperthermophiles apparently have reverse gyrase (rg) in a fused gene [9], with the exception of the newly discovered Nanoarchaeota [10]. The latter discovery suggests that fusion of separate genes to reverse gyrase occurred under hyperthermophilic, not mesophilic, conditions as previously argued [11,12]. Alternatively, a different protective mechanism was present in a hyperthermophilic LCA and early organisms [13,14].

An objection to a hyperthermophilic LCA based on G + C content of rRNA of extant organisms [15] has been challenged. In addition to G + C content, Di Giulio used a thermophily index based on the propensity of amino acids to enter more frequently into (hyper)thermophile proteins, concluding that the late stage of genetic code structuring took place in a (hyper)thermophilic organism [16,17]. These studies have used the correlation of the optimal growth temperature with the above measures to infer the temperature of the LCA, but the maximum temperature for growth ($T_{\text{max}}$) is plausibly closer to the temperature of emergence, giving a stronger inference of a hyperthermophilic LCA. Similarly, a hyperthermophilic LCA is supported by the robust correlation of $T_{\text{max}}$ with the rRNA phylogenetic distance from the LCA for (hyper)thermophilic prokaryotes (Figure 2) [18], with an inferred temperature exceeding 120°C, consistent with the newly claimed record $T_{\text{max}}$ of 121°C for life (an archaeon close to Pyrodictium) [19].

Other recent support for a hyperthermophilic LCA comes from expanded sequence and secondary structure data of rRNA [20], the tRNA sequence tree [21] and experimental tests using modified enzyme residues [22,23].

Hyperthermophilic biogenesis

A hyperthermophilic LCA does not of course require an origin of life at similar high temperatures. In particular, biogenesis could have occurred at mesophilic temperatures to be followed by a major impact event, sterilizing the surface and leaving a subsurface hyperthermophilic LCA survivor or primitive prokaryotes close to the LCA [24], although a recent re-interpretation of the lunar impact record argues that such sterilizing events may well not have occurred [25]. In any case, the possibility that biogenesis did indeed occur at hyperthermophilic temperatures should be re-examined in light of recent research. The case for a high-temperature origin of life was made by Wachtershauser [26], who argued that biogenesis and microbial evolution proceeded deterministically from hyperthermophiles to mesophiles. We have likewise argued that biogenesis and the emergence of the three domains of life was thermally inevitable, given the abiotic initial conditions of our planet [27,28].

The chemical evolution research programme continues to generate support for determinism, for example, in arguments for the likely universality of intermediary autotrophic metabolism [29] and the genetic code [30]. If future phylogenetic trees based on greater sampling and better-understood genomes confirm the apparent near-universal absence of hyperthermophiles with mesophilic ancestors (Figure 1), the asymmetric evolution of hyperthermophiles to mesophiles would be confirmed. This asymmetry is supported by the absence of any living eukaryote with a $T_{\text{max}}$ exceeding a few degrees above 60°C, in spite of at least 2 billion years of opportunity to adapt to hyperthermophily.

A plausible scenario for the origin of life requiring the presence of a hydrothermal regime and thus a proximate hyperthermophilic environment is outlined by [31], with the locus being within FeS precipitates produced by reaction of contrasting pH solutions reacting at the sea floor. They postulate a temperature for this origin at about 50°C,
Figure 1 | Phylogenetic tree based on rRNA sequences
The scale bar corresponds to 0.1 changes per nucleotide. The LCA is located on the bacterial line between the node to *Aquifex* and the Y intersection of the three domains. Maximal growth temperatures have been used to colour-code the branches. A black-ending line indicates that *T*<sub>max</sub> data are lacking for this organism. This plot shows a strong asymmetry: all mesophiles have hyperthermophilic ancestors while no hyperthermophiles or thermophiles have mesophilic ancestors with the possible exception of *Archaeoglobus*, a candidate for acquisition of hyperthermophily by horizontal gene transfer. Reprinted with permission from Pace, N.R. (1997) Science, vol. 276, pp. 734–740. © (1997) AAAS.

argued from the assumed constraint of RNA instability at higher temperatures [32]. However, saline solutions, readily available in this environment, appear to stabilize DNA/RNA at hyperthermophilic temperatures [33,34]. Therefore, the possibility of a hyperthermophilic RNA/DNA world [35] should be revisited. Other objections to a hot origin of life including the instability of amino acids and ribose at hyperthermophilic temperatures [36,37] have been addressed by supporters of a hydrothermal scenario, citing evidence for the thermodynamic favourability [38] and actual synthesis of amino acids under hydrothermal conditions [39,40].

Thus we argue that the chemical evolution to the LCA of Bacteria and Archaea could have occurred at hyperthermophilic conditions in Martin and Russell’s scenario [31] as Russell et al. argued previously [41]. Hence, the divergent
The maximum growth temperature of (hyper)thermophiles versus their phylogenetic distance (branch length) from the LCA using rRNA phylogenetic tree (Figure 1)


chemical evolution leading to Archaeal and Bacterial free-living cells as proposed by [31] can be accommodated in a temperature gradient between the hydrothermal source and ambient climatic temperature, with the upper temperature limits of Archaea (greater than or equal to about 120°C) and Bacteria (at 95°C if not higher) corresponding to the first possible self-organization of each cell type with cell membranes and walls. In this scenario, biogenesis occurred with the emergence of protocells with minimal metabolism and replication at the "edge of stability" to use the phrase in [14]. The differences between Archaeal and Bacterial membrane lipid biosynthesis and cell-wall biochemistry is central to Martin and Russell’s hypothesis. Their thermal stability might be clues to the temperature regimes prior to the emergence of free-living cells.

The deep nesting of primitive metabolisms inferred from extant hyperthermophiles [42,43] can be readily explained by this scenario. The emergence of Archaea has been linked to the metabolism of methanogenesis [44,45], a plausible natural exploitation of abundant CO$_2$ and hydrogen by early life. The other primitive metabolisms found in deeply rooted Archaea can likewise be explained as the exploitation of available raw materials (sulphur, trace metals such as Fe, Ni, Cr, W and As; see [46]) readily available on early Earth, particularly in the hydrothermal regime environment [41], first by chemoautotrophs and soon followed by heterotrophs, in the chemoautotrophic scenario for biogenesis. Plausibly the origin of these metabolisms occurred before 3.7 Gyr, the age of metasedimentary rocks from Greenland containing the earliest known evidence for life, $^{13}$C-depleted carbon [47].

While the climatic temperature for the Hadean is not well constrained other than an inference of liquid water at 4.4 Gyr [48], temperatures of $80^\circ$C at 3.8 Gyr [49] and $70^\circ$C ($\pm 15^\circ$C) at 3.5 Gyr [50] have been derived from the oxygen isotopic record of marine cherts. Thus while biogenesis plausibly occurred above late Hadean/early Archaean climatic temperatures, a climatic temperature constraint on microbial evolution prevailed through the Precambrian [27]. The robust correlation of $T_{max}$ with
the rRNA phylogenetic distance from the LCA (Figures 1 and 2) is consistent with the maximum growth temperature of thermophiles being close to the climatic temperature of each organism at emergence.

References
47 Rosing, M.T. (1999) Science 283, 674-676

Received 19 September 2003