



推断祖先年龄和代谢特征——远古生命进化研究的新趋势

滕文凯, 张传伦*

南方科技大学海洋科学与工程系, 深圳海洋地球古菌组学重点实验室, 深圳 518055

* 通讯作者, E-mail: zhangcl@sustech.edu.cn

收稿日期: 2024-12-10; 收修改稿日期: 2025-03-27; 接受日期: 2025-03-28; 网络版发表日期: 2025-04-28

国家自然科学基金项目(92351301、32393974)和深圳市科学技术创新委员会稳定支持重点项目(20200925173954005)资助

现存细胞生命共享很多核心的代谢特征, 例如利用ATP作为共同的能量货币、使用高度一致的分子机制处理遗传信息流(如生物学中心法则所展示), 表明它们应该演化自一个共同的祖先(Theobald, 2010; Hug等, 2016). 因此, 了解远古生命尤其是细胞生命的最近共同祖先(last universal common ancestor, LUCA)的年龄和代谢特征是研究地球生命起源及其进化历史的关键(Crapitto等, 2022). 最近, Moody等(2024)采用了最新的数据集和演化模型较为精确地推断出LUCA大约生活在42亿年前(即4.2Ga), 其基因组超过了2.5Mb, 可编码大约2600种蛋白质, 与现代原核生物的基因组大小相当. 该研究创新性提出LUCA是42亿年前就已经形成的地球早期生态系统的一部分, 它很可能在厌氧条件下通过代谢H₂和CO₂合成乙酸, 并为其他生物提供生态位(Moody等, 2024).

LUCA的年龄和代谢特征是生命科学和地球科学领域共同关注的瓶颈问题. 首先, 目前普遍采用松弛分子钟模型对系统发育树上的祖先节点进行定年, 该方法非常依赖化石和地球化学数据的准确性(Ho和Duchêne, 2014). 然而, 前寒武纪的化石和地球化学信息十分有限且充满模糊性, 会导致代表LUCA的系统发育树根节点的年龄推断存在巨大误差(Moody等, 2022).

根据月球形成时大碰撞的时间和最早的生物化石记录, 生命之树的根节点通常被赋予一个较大的时间范围(4.5Ga~3.4Ga)约束, 而LUCA的年龄也被估计介于4.5Ga和3.4Ga之间(Sugitani等, 2015; Betts等, 2018). 为了应对这些挑战, Moody等(2024)利用了LUCA前旁系同源基因来提高分子钟的可靠性. 就像现代生物的基因组中存在着旁系同源物, 一些基因(例如ATPase)可能在LUCA出现之前便产生了基因复制并以旁系同源物的形式存在于LUCA的基因组(Gogarten和Taiz, 1992). 使用这些LUCA出现前复制的基因可以将LUCA由系统发育树上的一个根节点变成两个后代节点, 并且相同的化石年代校准可以平行地应用在两棵子树的镜像节点上(如图1a所示), 从而实现对时间推断的交叉支撑(cross-bracing), 这大大降低了由序列进化速率差异而导致的定年的不可靠性(Gribaldo和Cammarano, 1998; Moody等, 2024). Moody等(2024)共采用了5对LUCA前复制的旁系同源基因, 使用它们的串联序列比对以及单个基因序列比对进行了多次分子钟的时间校对, 给出了目前最精确的LUCA年龄推断(95%置信区间: 4.33~4.18Ga).

Moody等(2024)也提供了关于LUCA代谢性质的新观点. 早期研究主要基于基因家族的系统发育分布

中文引用格式: 滕文凯, 张传伦. 2025. 推断祖先年龄和代谢特征——远古生命进化研究的新趋势. 中国科学: 地球科学, 55(6): 2094–2097, doi: [10.1360/SSTe-2024-0340](https://doi.org/10.1360/SSTe-2024-0340)

英文引用格式: Teng W, Zhang C. 2025. Inferring ancestral age and metabolic traits—Promises and challenges in the study of the evolution of ancient life. Science China Earth Sciences, 68(6): 2044–2047, <https://doi.org/10.1007/s11430-024-1556-6>

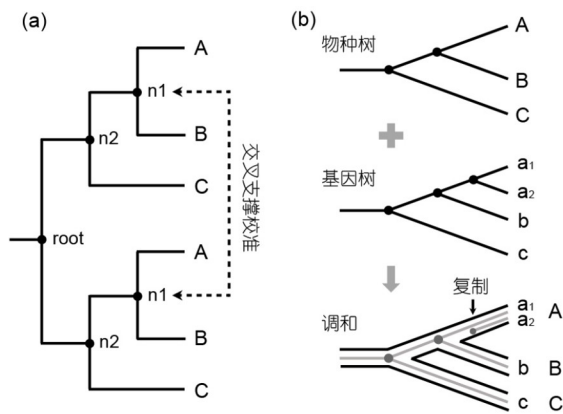


图1 分子钟的交叉支撑校准以及基因树和物种树调和示意图

(a) 以物种A、B和C的共同祖先(common ancestor, CA)为例, 使用CA出现之前复制的旁系同源基因构建系统发育树, 会产生两个通过根节点相连的均包含物种A、B和C的子树, 这时候CA不再是系统发育树的根节点(root), 而是两个后代节点(n2)。原本应用在n1节点上的时间约束(根据化石年代)可以平行应用在两棵子树上的镜像节点, 实现分子钟的交叉支撑校准(Moody等, 2024)。(b) 物种A包含一对旁系同源基因a1和a2, 物种B和C分别包含对应的同源基因b和c。将物种树和基因树进行调和, 推断出物种A所在的分支发生过一次基因复制事件(Szöllösi等, 2013)

(也叫系统发育图谱)来重建LUCA基因组(Kunin和Ouzounis, 2003)。如果只考虑垂直遗传, 科学家们会根据现存生物基因组中普遍存在的基因家族, 推断出一个过于简洁的“LUCA基因组”。这一“LUCA基因组”仅包含核糖体蛋白以及DNA复制、转录和翻译相关的蛋白, 缺乏必要的代谢基因来判断LUCA的代谢特征(Harris等, 2003)。Moody等(2024)使用了最近发展的基因树-物种树调和算法合并似然估计(amalgamated likelihood estimation, ALE)来推断每个KEGG直系同源基因家族(KEGG Orthology, KO)的进化历史。具体来说, ALE通过比较每一个基因家族的基因树和物种树(即使用共有单拷贝标记基因的串联比对构建的系统发育树), 估计该基因家族进化历史中的基因复制、基因丢失和水平基因转移事件, 以及该基因家族出现在物种树上特定节点的概率(图1b)。通过分析700个基因组(包含350个古菌和350个细菌), Moody等(2024)推断出399个很可能存在于LUCA的KO(存在概率 ≥ 0.75 , 且在古菌和细菌中均出现)。如果降低阈值(存在概率 ≥ 0.5), 可能存在于LUCA的KO高达1124个, Moody等(2024)据此进一步推断出LUCA基因组的大小约为

2.75Mb(95%置信区间: 2.49~2.99Mb)。与以往研究推断出的较为简单的LUCA基因组(Crapitto等, 2022)相比较, Moody等(2024)给我们展示了一个更加完整并且与现代原核生物类似的LUCA: 它很可能栖息于海洋表面或深海, 在无氧条件下通过Wood-Ljungdahl途径产乙酸, 通过糖异生/糖酵解获得生长能量, 甚至拥有早期的Cas酶免疫系统。鉴于LUCA的复杂性和代谢特点, 无论LUCA是自养、异样还是两者兼有, 它都不大可能是孤立存在的, 也即LUCA应该生活于一个已存在地球早期生态系统中。

综上所述, Moody等(2024)对LUCA年龄和代谢特征的研究方法与结果对探索地球生命演化历史具有重要的参考价值。目前, 基因树-物种树调和工具已经帮助推断多个谱系祖先的代谢性质, 包括全体细菌、CCTB(closely clustered thermophilic bacteria)分枝以及古菌中的Nitrososphaeria、Bathyarchaeia、Heimdal-larchaeia(真核生物潜在祖先)和Marine group II(MGII)等(Coleman等, 2021; Qi等, 2021; Eme等, 2023; Leng等, 2023; Fan等, 2024; Luo等, 2024)。结合祖先年龄的推断, 这些研究正致力于构建一个生命与地球协同演化的体系(Hou等, 2023)。不过, 当前的研究仍存在一些挑战, 这包括数目有限的LUCA前旁系同源基因以及他们未知的遗传稳定性(Zhaxybayeva等, 2005)。换句话说, 这些基因在LUCA后是否均为稳定的垂直遗传(直系同源), 仍需要进一步的探索。此外, 具有较为完善的功能描述的基因家族(被注释到KO)只占现代细菌和古菌代表基因组中所有编码基因的一部分(平均约55%), 并且这个比例因谱系而异(图2)。这导致我们对现代微生物代谢能力及其多样性的认知仍然有很大局限性, 并影响我们对祖先基因组推断的准确性(Koppel等, 2018; Crapitto等, 2022)。例如, 尽管Moody等(2024)推断出LUCA的基因组能编码超过2600个蛋白, 但只有1124个功能已知(存在概率 ≥ 0.5), 而用于LUCA代谢重构的只有其中的399个(存在概率 ≥ 0.75)。

对微生物基因组“暗物质”的挖掘以及更多谱系(尤其是那些发生了重要进化事件的谱系)的原核生物祖先基因组重建或许将有助于逐步追溯并重构一个更加真实的LUCA。除了Moody等(2024)使用的交叉支撑方法, 最近一些研究还构建了包含细菌、真核生物线粒体和质体的系统发育树, 进而利用真核生物丰富的化石资料校准原核生物进化节点, 以增强分子钟的精

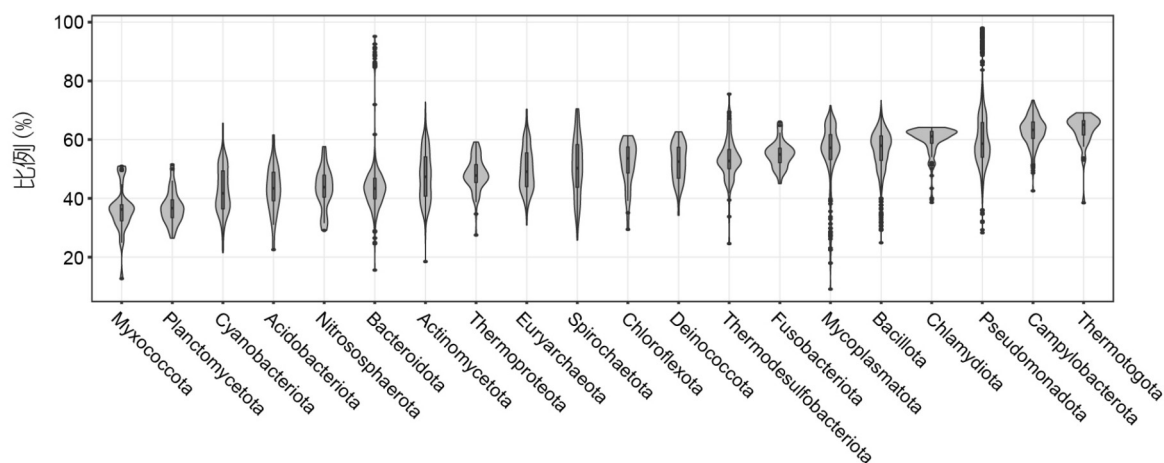


图 2 细菌和古菌不同门基因组中被分配KO的基因占所有编码基因的比例
根据2024年最新的KEGG数据库统计(<https://www.genome.jp/kegg/>)

确度(Liao等, 2024). 令人鼓舞的是, 最近一些研究利用分子钟探索了氨氧化、厌氧甲烷氧化和富羧基氧化等代谢特征的进化与重要地质事件之间的耦合关系(Ren等, 2019; Yang等, 2021; Wang等, 2022; Shang, 2023), 展示了微生物进化的地质学意义. 总之, 在目前大数据和人工智能突飞猛进发展的情况下, 不同领域科学家们的协作有望产生更新且更准确的标准化方法, 进而帮助更好地探索原核生物祖先的年龄和代谢特征以及重大地质事件中生命与地球化学过程协同演化的机制.

致谢 布里斯托大学Tom A WILLIAMS教授和Edmund RR MOODY博士、南方科技大学范陆教授、上海交通大学王寅焯教授、德克萨斯大学南海涛博士对本文提出了宝贵的修改建议, 在此一并致谢.

参考文献

Betts H C, Puttick M N, Clark J W, Williams T A, Donoghue P C J, Pisani D. 2018. Integrated genomic and fossil evidence illuminates life's early evolution and eukaryote origin. *Nat Ecol Evol*, 2: 1556–1562

Coleman G A, Davin A A, Mahendrarajah T A, Szánthó L L, Spang A, Hugenholtz P, Szöllösi G J, Williams T A. 2021. A rooted phylogeny resolves early bacterial evolution. *Science*, 372: eabe0511

Crapitto A J, Campbell A, Harris A J, Goldman A D. 2022. A consensus view of the proteome of the last universal common

ancestor. *Ecol Evol*, 12: e8930

Eme L, Tamarit D, Caceres E F, Stairs C W, De Anda V, Schön M E, Seitz K W, Dombrowski N, Lewis W H, Homa F, Saw J H, Lombard J, Nunoura T, Li W J, Hua Z S, Chen L X, Banfield J F, John E S, Reysenbach A L, Stott M B, Schramm A, Kjeldsen K U, Teske A P, Baker B J, Ettema T J G. 2023. Inference and reconstruction of the heimdallarchaeal ancestry of eukaryotes. *Nature*, 618: 992–999

Fan L, Xu B, Chen S, Liu Y, Li F, Xie W, Prabhu A, Zou D, Wan R, Li H, Liu H, Liu Y, Kao S J, Chen J, Zhu Y, Rinke C, Li M, Zhu M, Zhang C, Dupont C. 2024. Gene inversion led to the emergence of brackish archaeal heterotrophs in the aftermath of the Cryogenian Snowball Earth. *Proc Natl Acad Sci U S A Nexus*, 3: pgae057

Gogarten J P, Taiz L. 1992. Evolution of proton pumping ATPases: Rooting the tree of life. *Photosynth Res*, 33: 137–146

Gribaldo S, Cammarano P. 1998. The root of the universal tree of life inferred from anciently duplicated genes encoding components of the protein-targeting machinery. *J Mol Evol*, 47: 508–516

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

Ho S Y W, Duchêne S. 2014. Molecular-clock methods for estimating evolutionary rates and timescales. *Mol Ecol*, 23: 5947–5965

Hou J, Wang Y, Zhu P, Yang N, Liang L, Yu T, Niu M, Konhauser K, Woodcroft B J, Wang F. 2023. Taxonomic and carbon metabolic diversification of Bathyarchaeia during its coevolution history with early Earth surface environment. *Sci Adv*, 9: eadf5069

Hug L A, Baker B J, Anantharaman K, Brown C T, Probst A J, Castelle C J, Butterfield C N, Hemsdorf A W, Amano Y, Ise K, Suzuki Y, Dudek N, Relman D A, Finstad K M, Amundson R, Thomas B C,

- Banfield J F. 2016. A new view of the tree of life. *Nat Microbiol*, 1: 16048
- Koppel N, Bisanz J E, Pandelia M E, Turnbaugh P J, Balskus E P. 2018. Discovery and characterization of a prevalent human gut bacterial enzyme sufficient for the inactivation of a family of plant toxins. *eLife*, 7: e33953
- Kunin V, Ouzounis C A. 2003. GeneTRACE—reconstruction of gene content of ancestral species. *Bioinformatics*, 19: 1412–1416
- Leng H, Wang Y, Zhao W, Sievert S M, Xiao X. 2023. Identification of a deep-branching thermophilic clade sheds light on early bacterial evolution. *Nat Commun*, 14: 4354
- Liao T, Wang S, Zhang H, Stueken E E, Luo H, Battistuzzi F U. 2024. Dating ammonia-oxidizing bacteria with abundant eukaryotic fossils. *Mol Biol Evol*, 41: msae096
- Luo Z H, Li Q, Xie Y G, Lv A P, Qi Y L, Li M M, Qu Y N, Liu Z T, Li Y X, Rao Y Z, Jiao J Y, Liu L, Narsing Rao M P, Hedlund B P, Evans P N, Fang Y, Shu W S, Huang L N, Li W J, Hua Z S. 2024. Temperature, pH, and oxygen availability contributed to the functional differentiation of ancient *Nitrososphaeria*. *ISME J*, 18: wrad031
- Moody E R R, Álvarez-Carretero S, Mahendrarajah T A, Clark J W, Betts H C, Dombrowski N, Szánthó L L, Boyle R A, Daines S, Chen X, Lane N, Yang Z, Shields G A, Szöllősi G J, Spang A, Pisani D, Williams T A, Lenton T M, Donoghue P C J. 2024. The nature of the last universal common ancestor and its impact on the early Earth system. *Nat Ecol Evol*, 8: 1654–1666
- Moody E R, Mahendrarajah T A, Dombrowski N, Clark J W, Petitjean C, Offre P, Szöllősi G J, Spang A, Williams T A. 2022. An estimate of the deepest branches of the tree of life from ancient vertically evolving genes. *eLife*, 11: e66695
- Qi Y L, Evans P N, Li Y X, Rao Y Z, Qu Y N, Tan S, Jiao J Y, Chen Y T, Hedlund B P, Shu W S, Hua Z S, Li W J, Lindemann S R. 2021. Comparative Genomics Reveals Thermal Adaptation and a High Metabolic Diversity in “*Candidatus* Bathyarchaea”. *mSystems*, 6: 10.1128/msystems.00252-21
- Ren M, Feng X, Huang Y, Wang H, Hu Z, Clingenpeel S, Swan B K, Fonseca M M, Posada D, Stepanauskas R, Hollibaugh J T, Foster P G, Woyke T, Luo H. 2019. Phylogenomics suggests oxygen availability as a driving force in Thaumarchaeota evolution. *ISME J*, 13: 2150–2161
- Shang H. 2023. Dichotomous effects of oxidative metabolisms: A theoretical perspective on the dolomite problem. *Glob Planet Change*, 222: 104041
- Sugitani K, Mimura K, Takeuchi M, Lepot K, Ito S, Javaux E J. 2015. Early evolution of large micro-organisms with cytological complexity revealed by microanalyses of 3.4 Ga organic-walled micro-fossils. *Geobiology*, 13: 507–521
- Szöllősi G J, Rosikiewicz W, Boussau B, Tannier E, Daubin V. 2013. Efficient exploration of the space of reconciled gene trees. *Systatic Biol*, 62: 901–912
- Theobald D L. 2010. A formal test of the theory of universal common ancestry. *Nature*, 465: 219–222
- Wang Y, Xie R, Hou J, Lv Z, Li L, Hu Y, Huang H, Wang F. 2022. The late Archaeal to early Proterozoic origin and evolution of anaerobic methane-oxidizing archaea. *mLife*, 1: 96–100
- Yang Y, Zhang C, Lenton T M, Yan X, Zhu M, Zhou M, Tao J, Phelps T J, Cao Z, Ursula Battistuzzi F. 2021. The evolution pathway of ammonia-oxidizing archaea shaped by major geological events. *Mol Biol Evol*, 38: 3637–3648
- Zhaxybayeva O, Lapierre P, Gogarten J P. 2005. Ancient gene duplications and the root(s) of the tree of life. *Protoplasma*, 227: 53–64

(责任编辑: 石拓)