

DOI: 10.1360/yc-006-1445

系统发育基因组学——重建生命之树的一条迷人途径

于黎^{1,2}, 张亚平²

(1. 云南大学生物资源保护与利用重点实验室, 昆明 650091; 2. 中国科学院昆明动物研究所细胞与分子进化重点实验室, 昆明 650023)

摘要: 生物基因组计划的相继启动对系统发育学的发展产生了深远影响。将基因组学和系统发育学结合起来的“系统发育基因组学”在这一背景下应运而生, 这门崭新交叉学科的出现成为重建“生命之树”的一条迷人途径。文章重点围绕系统发育基因组学的概念以及分析方法进行扼要介绍, 并探讨了它目前潜在的问题以及将来有待发展和完善的地方。

关键词: 系统发育学; 基因组学; 系统发育基因组学; 稀有基因组改变; 全基因组改变

中图分类号: Q111

文献标识码: A

文章编号: 0253-9772(2006)11-1445-06

Phylogenomics——An Attractive Avenue to Reconstruct “Tree of Life”

YU Li^{1,2}, ZHANG Ya-Ping²

(1. *Laboratory of Conservation and Utilization of Bio-resource, Yunnan University, Kunming 650091, China*; 2. *Laboratory of Molecular Evolution and Genome Diversity, Kunming Institute of Zoology, the Chinese Academy of Sciences, Kunming 650023, China*)

Abstract: The successive start-up of genome project in recent years has far-reaching impact on present phylogenetic studies. A new synthesized discipline——“Phylogenomics”, combing genomics with phylogenetics together, is emerging, and becomes an attractive avenue to reconstruct “Tree of Life”. In the review, the conception and analytic method of phylogenomics are summarized. The potential problems are also suggested.

Key words: phylogenetics; genomics; phylogenomics; rare genomic change (RGCs); whole-genome features (WGFs)

尽管生物界生物多种多样, 但是它们都由于一个共同的进化历史而有着不同程度的关联, 这种关联也使研究物种之间进化关系的学科——系统发育学(Phylogenetics)变得非常有意义^[1]。追溯生物界不同生物类型的起源及进化关系, 即重建生物类群的系统发育树已经成为进化生物学领域中一个十分重要的内容, 并日益受到广泛的关注。美国和一些欧洲国家已率先启动了类似人类基因组计划的“生命之树”计划(Tree of Life, TOL), 将对生物学研究的发展

产生深远影响。就目前而言, 推断物种间系统发育关系已经为分析所有进化问题所必需^[2]。

最初的系统发育学家认为生物界不同生物的进化历史可直接从化石记录获知, 因为它是生物进化史唯一的直接证据, 但是化石记录往往零散且不完整, 总是缺少许多“中间环节”的重要化石^[2~4]。直到20世纪70年代, 系统发育重建已经转向比较形态学的方法, 将形态学上的差异程度作为物种间的进化距离。研究者可以根据化石和现存物种的形态

收稿日期: 2006-06-01; 修回日期: 2006-07-07

基金项目: 云南省科技厅项目和国家自然科学基金资助 [Supported by Scientific and Technical Fund Project of Yunnan and National Natural Science Foundation of China (NSFC)]

作者简介: 于黎(1976—), 女, 辽宁人, 博士, 助理研究员, 研究方向: 分子进化。E-mail: yuli1220@yahoo.com.cn

通讯作者: 张亚平(1965—), 男, 云南昭通人, 研究员, 中国科学院院士, 研究方向: 分子进化。Tel: 0871-5190761; Fax: 0871-5195430; E-mail: zhangyp@mail.kiz.ac.cn

学特征性状构建出物种进化历史的主要框架^[3,4]。然而,形态性状过于复杂的进化历史致使可靠的同源性状很难找到,往往物种之间性状状态的相似被揭示为趋同而不是同源,因此生命之树上很多分支都存在着争议,无法解决^[4]。

20世纪80年代分子,系统发育学研究发生突破性的进展,从以往的比较形态学研究开始转向核型、免疫学研究,然后发展到染色体显带、氨基酸序列研究,到现代利用物种中的大量遗传物质信息,即DNA序列来重建系统发育关系^[5]。DNA序列由于本身相对稳定的进化演变过程和巨大的信息量等优势,大大提高了对物种间系统发育关系的推断能力。而且,现代系统发育学研究还处在不断的深入和发展过程中,例如从最初的只能使用很少量的DNA数据到现在通过自动测序仪的广泛应用以及许多生物的全基因组测序的陆续启动可产生大量核苷酸和蛋白质序列进行进化研究;从简单的推断物种系统发育关系的应用深入到揭示进化机制和规律;从对DNA序列的常规分析策略扩展到用于处理基因组大量不同信息的新方法等等。因此,现代系统发育学已经开始逐步渗透到生物学的其他领域,它又同时与其他学科互相吸收与综合,如近年来新出现的将基因组学和系统发育学结合起来的交叉学科“系统发育基因组学(Phylogenomics)”。

自从Craig Venter和他的小组在1995年首先报道了第一个微生物(*Haemophilus influenzae*)的基因组序列,便标志着生物学研究以迅雷不及掩耳之势由前基因组时代进入了一个崭新的阶段——基因组时代^[6]。很显然,基因组测序将导致核苷酸序列的迅猛骤增,还将提供大量有关重复基因,大段DNA缺失/插入,基因易位,转座子插入等信息^[3,4],这些都会对系统发育学研究的发展产生深远影响。“系统发育基因组学(Phylogenomics)”正是在这一背景下产生,成为重建“生命之树”的一条迷人途径。从系统发育学家的角度而言,可以将其定义为利用大规模的基因组学方法进行的系统发育分析。虽然基因组数据的获得在很大程度上满足了以往系统发育学对分子数据的需求,可以进行多个基因或整个基因组水平上的分析。但是如此海量的分子性状却对现有的系统发育重建方法提出了挑战,系统学家的注意力又开始从寻找不同来源的数据转向改进传统和开发新的

构树方法。如此一来,用于系统发育基因组学研究的一些重建方法和分析策略也开始不断的呈现在人们面前,反过来又促进了系统发育基因组学的发展和运用。如用来评估直系同源基因可靠性的RIO方法(Resampled Inference of Orthologs)^[7],快速构建基因组树的SHOT方法(Shared Orthologue and gene-order Tree)^[8]以及筛选用于基因组系统发育分析的保守氨基酸座位或域的Phylomat方法(A Protein Motif Analysis Tool for Phylogenomics)^[9]等等。

最近几年中,国际上许多研究小组已经开始深入开展系统发育基因组学的研究,希望能够利用这个崭新的工具澄清很多在以前因为只能使用很少量数据的研究中无法解决的诸多问题^[10-14]。当然,重建生物的进化历史只是系统发育基因组学应用的一个内容,它还可以用来阐述由基因复制导致的基因组进化问题^[15-18],构建各种生物化学路径的进化历史^[19,20]以及预测基因功能^[21-23]等等。“系统发育基因组学”一词开始在越来越多文章的标题,摘要和关键词中出现,也有越来越多的学者将自己的研究成果冠以是由“系统发育基因组学方法,(Phylogenomic approach)”分析得到的。《Nature》,《Science》和《PNAS》等非常有影响力的杂志都开始关注这个崭新的领域,通过具体典型的例子或总结它的进展来阐述系统发育基因组学对进化生物学的影响。“系统发育基因组学”作为一门崭新的交叉学科出现,涉及面非常广泛,而且很多相关分析方法还在探索过程中,因此在这里我们避免过多描述具体的统计学理论和算法,只是重点围绕系统发育基因组学的概念,分析方法以及在重建生命之树方面的应用进行宏观上的扼要介绍,旨在揭示其中所蕴藏的学科魅力。另外也探讨了它目前潜在的问题以及将来有待发展和完善的地方。希望能够对国内系统发育基因组学的推广和研究具有一定的指导作用,为迎接正在到来的国际系统发育基因组学领域研究的全面整合和发展开辟道路。

1 系统发育基因组学的概念

“系统发育基因组学”是进化生物学领域中一门崭新的交叉学科,最先由Eisen J A和Hanawalt P C^[19]正式全面提出。由于许多生物的全基因组测序的陆续启动而导致大量核酸和蛋白质序列的迅猛骤

增, 要求采用最新的概念或分子技术来促进对已获得的基因组序列的分析。“系统发育基因组学”正是在这一需求和背景下产生的。因此, 它来源于生命科学中两个主要领域的结合: “基因组学(Genomics)”, 即基因和基因组的结构和功能研究, 和“系统发育学(Phylogenetics)”, 即生物间的进化关系研究。在实践中, 这两个方面紧密相关, 它们之间的互相渗透, 互相影响已经日益显现: 一方面, 进化分析作为一个来源丰富的方法手段可以促进基因组研究以提高我们对基因组序列, 包括结构和功能的了解。反过来, 基因组研究也可以促进进化分析。大量的基因组数据将为重建生物间进化关系的研究提供海量的分子性状, 使很多在以前因为只使用很少量数据的研究中无法解决的问题有望得以澄清。总而言之, 基因组分析和进化研究之间在很多情况下不可分割而且能够结合起来促进共同发展^[19,24], 因而出现了系统发育基因组学。

一般认为, 系统发育基因组学是在系统发育分析中使用由高强度的测序和/或已有的基因组数据库中获得的规模或多个基因座位数据信息(通常只有较少的分类群代表)的一种分析方法, 目的是构建可靠的物种系统发育树或了解个别基因树的进化过程^[25]。而系统发育基因组学的一个复杂应用就是提高对未被描述过的基因的功能预测。这时系统发育基因组学也可以定义为利用系统发育信息对高生产量的基因组进行功能注释^[26]。

2 系统发育基因组学分析方法

目前, 系统发育基因组学研究所用的系统重建方法主要分为两类: 基于序列分析和基于“全基因组特征(whole-genome features, WGFs)”分析的方法, 有少量研究还使用“罕有的基因组改变(rare genomic changes, RGCs)”进行系统发育重建。基于全基因组特征和罕有的基因组改变这两种方法也被称为非序列分析方法^[27]。在以往仅仅基于DNA序列核苷酸差异的系统发育研究中, 核苷酸碱基的趋同进化, 位点和谱系间不同的替换速率以及位点间非独立替换等方面常常会影响到系统发育重建, 使系统发育结果很难解释。相比之下, 系统发育基因组学的一个明显优势就是可以提供除DNA序列以外的其他信息进行进化关系分析, 对传统的序列比对分析方法进行补充和验证。

2.1 序列分析方法

系统发育基因组学方法可以通过比较原始序列, 由序列比对来推断谱系关系。因为其理论基础和方法学背景被大多数学者熟知, 因此这种分析方法在目前使用最广泛。

一般来说, 当同源DNA序列比对确定以后, 有两种途径可以进行基因组数据的系统发育重建: 第一种是先把多组序列数据合并为一组, 然后在这一合并的序列数据基础上使用标准的构树方法(如最大简约法, 最大似然法, 距离法和贝叶斯法等)构建系统树, 即SuperMatrix途径。这种方法在一般情况下都能得到比较好的结果, 但是存在两个明显的缺陷, 一是合并的序列比对中必须包括一组相同的物种; 二是只能使用一个核苷酸替换模型对合并的序列比对进行分析^[27,28]。第二种途径是先根据各个基因序列分别构建各自的系统发育树, 然后将这些系统树通过一定的优化技术, 如Matrix Representation using Parsimony(MRP)方法, 合并成为一个“Super-Tree”^[27,29,30]。这个途径是目前唯一可以构建包含较多物种的生物类群的完整系统发育树的方法^[31], 有的学者甚至认为它是解决生命之树的一条必要途径^[32-34]。

目前这两种方法已经在一些原核生物和真核生物类群的系统发育重建方面有所应用, 例如通过鉴定很少发生水平基因转移的保守基因区, 并使用这些保守基因来推断原核生物系统发育关系, 即只结合基因组中的大量“信息”基因(与大分子有相互作用的基因), 通过SuperMatrix途径构建系统树^[10,35-38], 或采用SuperTree途径对基于所有基因序列构建的系统树拓扑结构一致性进行研究, 将那些支持相似物种系统发育关系的基因定义为保守基因区来构建系统树^[10]。研究结果证实尽管水平基因转移在原核生物进化中频繁发生, 但是还没有完全擦除系统发育信号, 在基因组水平上还是很有可能重建真实系统发育关系的^[10]; 另外, SuperMatrix途径还被用来探讨真核生物中陆地植物^[39-41]以及有胎盘类哺乳动物18个目之间的系统发育关系^[42,43]等等。

2.2 非序列分析方法

2.2.1 全基因组特征(whole-genome features, WGFs)分析方法

除了比较序列差异以外, 系统发育基因组学研

究还可以基于基因含量(gene content or repertoire), 基因顺序(gene order)以及“DNA strings(基因组中寡核苷酸分布)”构建系统发育树^[27]。

比较不同物种间的基因组最明显办法就是分析它们的基因含量, 关系较近的物种会含有大量共同的基因, 关系较远的物种含有的共同基因较少。使用基因含量进行系统发育分析主要涉及两个方面: 确定基因间的同源/直源关系以及将基因出现或缺失的数据转变成树结构^[28]。基因顺序方法则是将基因组之间断裂点数量最小化, 或先挑选那些在基因组中物理邻近很保守的蛋白对, 然后对这些蛋白对的出现或缺失矩阵进行分析^[27]等等。

基因含量和基因顺序方法的优点在于: 可以基于大量信息进行同源基因的甄别, 准确性提高; 数据“捕捉”的是整个基因组信息, 避免了基因树/物种树的问题; 不用进行多重序列比对; 不易出现趋同或回复突变等“非同源相似”, 能够追溯古老类群的进化历史等等^[27,44,45]。然而, 数据和进化模型的缺乏以及较大的计算机强度使它们目前在系统发育基因组学中的应用还受一定的限制。另外, 一些研究还表明这两种方法对于分歧度中等的物种系统发育关系还不能很好的解决^[9]。它们也很容易受到基因丢失和水平基因转移的影响。

DNA strings即基因组中短寡核苷酸的分布(如结合频率)是基因组结构的一个特征印记, 是可以直接对非同源DNA 序列进行比较得到系统发育信息的一种方法, 能够真正从整个基因组水平构建基因组树。可以将核苷酸分布差异作为物种间进化距离, 然后使用标准的距离构树方法进行基因组系统发育分析^[27,46]。但这种方法在系统发育基因组学研究中的应用还不多。

2.2.2 罕有基因组改变(rare genomic changes, RGCs)分析方法

系统发育基因组学研究也可以利用一些罕有的基因组改变来构建系统发育树。“罕有的基因组改变”是指在特定进化枝的基因组中发生的大规模突变包括基因重复, 内含子的插入和缺失, 逆转座子整合(SINE 和 LINE)以及基因融合和分裂事件等等^[27]。大多数RGCs是由一个或几个突变事件引起的, 基因组大规模突变是罕有的, 不可能在同一系统分枝上发生两次, 也不可能在不同的系统发生分枝上

重复发生。由于它们在物种演化过程中发生的频率很低, 因而产生趋同或平行进化的概率也很低。这样, 既有效避免了那些会影响序列数据分析的潜在问题, 如非同源相似事件, 又可以当序列数据产生冲突或模糊不清的结果时, 作为一个独立的系统发育信息来源, 增强系统发育假说的分辨力^[47,48]。

到目前为止, “罕有的基因组改变”特征已经在一些类群的系统发育学研究中应用, 如Venkatesh B 等^[49]使用内含子插入/缺失探讨鱼类和昆虫类内部的系统发育关系; SINEs逆转座子整合提供强烈的证据支持鲸类和河马之间的姐妹群关系, 提出“whippo”假说^[50], 还支持有胎盘类哺乳动物中灵长目和啮齿目有很近的进化关系^[51]等等。另外, 一些其他类型的大规模突变, 如基因重复, 染色体重排等等^[51,52], 也可以作为重要的基因组进化特征来重建物种间的系统发育关系。

3 系统发育基因组学潜在的问题

一条重建生命之树的迷人途径, 有着巨大的吸引力。但是, 它与其他新的学科和方法一样, 都要在观察和实验中接受检验。很显然, 系统发育基因组学可以明显削弱分析中的“随机或取样误差”, 但是它与以往的系统发育研究一样无法避免由系统发育重建方法带来的“系统误差”, 如在基因组序列分析中也会存在碱基组成倾向, 长枝吸引和位点间速率异质性等现象而将没有进化关系的物种聚在一起。因此, 大量基因组信息的产生在增强了系统发育信号的同时进化噪音也随之增加。如何避免上述的潜在问题, 发挥系统发育基因组学的优势对于系统发育学家和进化基因组学家都将是未来研究中的一个巨大挑战^[27,28]。

除了系统误差可以造成系统重建方法不一致以外, 巨大的计算机负荷也使系统发育基因组学的应用困难重重, 尽管传统的系统发育构树方法(最大简约法, 最大似然法, 距离法和贝叶斯法)仍然可以完成目前大多数基于基因组数据的系统发育分析, 但是随着将来挖掘出更多不同类型的基因组信息以及所分析的物种数量的增加, 必然需要新的分析方法和策略来处理。而现阶段的一些新的系统重建方法或缺乏坚实的理论基础或还没有在实际中检验, 因此加入新的基因组数据或使用更可靠的分析程序都可能会改变我们现在得到的基因组树。

可以看出, 虽然系统发育基因组学有实现“重建生命之树”的巨大优势和潜力, 但是在现阶段它还有很多方面有待发展和完善。我们不能将系统发育基因组学看作是能解决一切痼疾的仙丹。但是, 即使如此, 系统发育基因组学的意义仍然不言而喻。虽然系统发育基因组学这门交叉学科涉及面如此广泛, 难度如此之大, 在目前看来它仍然是高强度, 费时, 费力的一种解决问题的手段, 但是它所蕴藏的巨大潜能和将要在未来研究中发挥的作用不可忽视。目前国际上几个基因组研究协会已经分别开始进行不同生物类群的全基因组序列测定, 全世界科学家的通力合作无疑将会大大促进系统发育基因组学的发展和运用。总而言之, 系统发育基因组学将是一项艰难但是非常有意义的崭新领域等待我们去开拓。

参 考 文 献(References):

- [1] MENG Jin, WANG Xiao-Ming. Phylogenetics-understanding and exploration for modern biosystematics: different schools in modern biosystematics. *Vertebrata Palasiatica*, 1988, 26(4): 309~313.
孟 津, 王晓鸣. 系统发育系统学——对现代生物系统学的理解与探讨: 现代生物系统学中的不同学派. 古脊椎动物学报, 1988, 26(4): 309~313.
- [2] PENG Yi-Xin, HUANG Shi-Jian. *Evolutionary Biology*. Wuhan: Wuhan University Press, 1997.
彭奕欣, 黄诗笺主编. 进化生物学. 武汉: 武汉大学出版社, 1997.
- [3] Nei M, Kumar S. *Molecular Evolution and Phylogenetics*. New York: Oxford University, 2000.
- [4] LÜ Bao-Zhong, ZHONG-Yang, GAO Li-Ping. *Molecular Evolution and Phylogenetics*. Beijing: Higher Education Press, 2002.
吕宝忠, 钟 杨, 高莉萍等译. 分子进化与系统发育. 北京: 高等教育出版社, 2002.
- [5] Bininda-Emonds ORP. Factors influencing phylogenetic inference: A case study using the mammalian carnivores. *Mol Phylogenet Evol*, 2000, 16(1): 113~126. [DOI](#)
- [6] Fleischmann R D, Adams M D, White O, Clayton R A, Kirkness E F, Kerlavage A R, Bult C J, Tomb J F, Dougherty B A, Merrick J M, et al. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 1995, 269(5223): 496~512. [DOI](#)
- [7] Zmasek C M, Eddy S R. RIO: Analyzing proteomes by automated phylogenomics using resampled inference of orthologs. *BMC Bioinformatics*, 2002, 3(1): 14. [DOI](#)
- [8] Korbel J O, Snel B, Huynen M A, Bork P. SHOT: a web server for the construction of genome phylogenies. *Trends in Genetics*, 2002, 18(3): 158~162. [DOI](#)
- [9] Graham W V, Tchong D K, Shirk A L, Attene-Ramos M S, Welge M E, Gaskins H R. Phylomat: An automated protein motif analysis tool for Phylogenomics. *Journal of Proteome Research*, 2004, 3(6): 1289~1291. [DOI](#)
- [10] Brochier C, Bapteste E, Moreira D, Philippe H. Eubacterial phylogeny based on translational apparatus proteins. *Trends in Genetics*, 2002, 18(1): 1~5. [DOI](#)
- [11] Daubin V, Gouy M, Perriere G. A phylogenomic approach to bacterial phylogeny: evidence of a core of genes sharing a common history. *Genome Research*, 2002, 12(7): 1080~1090. [DOI](#)
- [12] Dopazo H, Santoyo J, Dopazo J. Phylogenomics and the number of characters required for obtaining an accurate phylogeny of eukaryote model species. *Bioinformatics*, 2004, 20(Suppl.1): 116~121. [DOI](#)
- [13] Thomas J W, Touchman J W, Blakesley R W, Bouffard G G, Beckstrom-Sternberg S M, Margulies E H, Blanchette M, Siepel A C, Thomas P J, McDowell J C, Maskeri B, Hansen N F, Schwartz M S, Weber R J, Kent W J, Karolchik D, Bruen T C, Bevan R, Cutler D J, Schwartz S, Elnitski L, Idol J R, Prasad A B, Lee-Lin S Q, Maduro V V, Summers T J, Portnoy M E, Dietrich N L, Akhter N, Ayele K, Benjamin B, Cariaga K, Brinkley C P, Brooks S Y, Granite S, Guan X, Gupta J, Haghighi P, Ho S L, Huang M C, Karlins E, Laric P L, Legaspi R, Lim M J, Maduro Q L, Masiello C A, Mastrian S D, McCloskey J C, Pearson R, Stantripp S, Tiangson E E, Tran J T, Tsurgeon C, Vogt J L, Walker M A, Wetherby K D, Wiggins L S, Young A C, Zhang L H, Osoegawa K, Zhu B, Zhao B, Shu C L, De Jong P J, Lawrence C E, Smit A F, Chakravarti A, Haussler D, Green P, Miller W, Green E D. Comparative analyses of multi-species sequences from targeted genomic regions. *Nature*, 2003, 424(6950): 788~793. [DOI](#)
- [14] Kirkness E F, Bafna V, Halpern A L, Levy S, Remington K, Rusch D B, Delcher A L, Pop M, Wang W, Fraser C M, Venter J C. The dog genome: survey sequence and analysis. *Science*, 2003, 301(5641): 1898~1903. [DOI](#)
- [15] Vienne A, Rasmussen J, Abi-Rached L, Pontarotti P, Gilles A. Systematic phylogenomic evidence of en Bloc duplication of the ancestral 8p11.21-8p21.3-like region. *Mol Biol Evol*, 2003, 20(8): 1290~1298. [DOI](#)
- [16] McPartland J M. Phylogenomic and chemotaxonomic analysis of the endocannabinoid system. *Brain Research Reviews*, 2004, 45(1): 18~29. [DOI](#)
- [17] Gadelle D, Filee J, Buhler C, Forterre P. Phylogenomics of type II DNA topoisomerases. *Bioessays*, 2003, 25(3): 232~242. [DOI](#)
- [18] Malik H S, Henikoff S. Phylogenomics of the nucleosome. *Nature Structural Biology*, 2003, 10(11): 882~891. [DOI](#)
- [19] Eisen J A, Hanawalt P C. A phylogenomic study of DNA repair genes, proteins, and processes. *Mutation Research*, 1999, 435(3): 171~213.
- [20] Bouzat J L, McNeil L K, Robertson H M, Solter L F, Nixon J E, Beever J E, Gaskins H R, Olsen G, Subramaniam S, Sogin M L, Lewin H A. Phylogenomic analysis of the alpha proteasome gene family from early diverging eukaryotes. *J Mol Evol*, 2000, 51(6): 532~543.
- [21] Eisen J A, Wu M. Phylogenetic analysis and gene functional predictions: phylogenomics in action. *Theoretical Population Biology*, 2002, 61(4): 481~487. [DOI](#)

- [22] Eisen J A, Nelson K E, Paulsen I T, Heidelberg, J F, Wu M, Dodson R J, Deboy R, Gwinn M L, Nelson W C, Haft D H, Hickey E K, Peterson J D, Durkin A S, Kolonay J L, Yang F, Holt I, Umayam L A, Mason T, Brenner M, Shea T P, Parksey D, Nierman W C, Feldblyum T V, Hansen C L, Craven M B, Radune D, Vamathevan J, Khouri H, White O, Gruber T M, Ketchum K A, Venter J C, Tettelin H, Bryant D A, Fraser C M. The complete genome sequence of *Chlorobium tepidum* TLS, A photosynthetic, anaerobic, green-sulfur bacterium. *Proc Natl Acad Sci USA*, 2002, 99(14): 9509~9514. [\[DOI\]](#)
- [23] Sjolander K. Phylogenomic inference of protein molecular function: advances and challenges. *Bioinformatics*, 2004, 20(2): 170~179. [\[DOI\]](#)
- [24] Eisen J A, Fraser C M. Phylogenomics: intersection of evolution and genomics. *Science*, 2003, 300(5626): 1706~1707. [\[DOI\]](#)
- [25] Chen W J, Orti G, Meyer A. Novel evolutionary relationship among four fish model systems. *Trends in Genetics*, 2004, 20(9): 424~431. [\[DOI\]](#)
- [26] Zmasek C M, Eddy S R. A simple algorithm to infer gene duplication and speciation events on a gene tree. *Bioinformatics*, 2001, 17(9): 821~828. [\[DOI\]](#)
- [27] Delsuc F, Brinkmann H, Philippe H. Phylogenomics and the reconstruction of the tree of life. *Nat Rev Genet*, 2005, 6(5): 361~375. [\[DOI\]](#)
- [28] Wolf Y I, Rogozin I B, Grishin N V, Koonin, E V. Genome trees and the tree of life. *Trends in Genetics*, 2002, 18(9): 472~479. [\[DOI\]](#)
- [29] Baum B R. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon*, 1992, 41(1): 3~10. [\[DOI\]](#)
- [30] Ragan M A. Phylogenetic inference based on matrix representation of trees. *Mol Phylogenet Evol*, 1992, 1(1): 53~58. [\[DOI\]](#)
- [31] Bininda-Emonds O R P. The evolution of supertrees. *Trends Ecol Evol*, 2004, 19(6): 315~322. [\[DOI\]](#)
- [32] Soltis P S, Soltis D E. Molecular systematics: assembling and using the Tree of Life. *Taxon*, 2001, 50(3): 663~677. [\[DOI\]](#)
- [33] Pennisi E. Modernizing the tree of life. *Science*, 2003, 300(5626): 1692~1697. [\[DOI\]](#)
- [34] Gatesy J, Springer M S. A critique of matrix representation with parsimony supertrees. In: Bininda-Emonds ORP ed. *Phylogenetic Supertrees: Combining Information to Reveal the Tree of Life*, 2004.
- [35] Klenk H P, Spitzer M, Ochsenreiter T, Fuellen G. Phylogenomics of hyperthermophilic archaea and bacteria. *Biochemical Society Transactions*, 2004, 32(2): 175~178. [\[DOI\]](#)
- [36] Jain R, Rivera M C, Lake J A. Horizontal gene transfer among genomes: The complexity hypothesis. *Proc Natl Acad Sci*, 1999, 96(7): 3801~3806. [\[DOI\]](#)
- [37] Wolf Y I, Rogozin I B, Grishin N V, Tatusov R L, Koonin E V. Genome trees constructed using five different approaches suggest new major bacterial clades. *BMC Evol Biol*, 2001, 1(1): 8~29. [\[DOI\]](#)
- [38] Brown J R. Genomic and phylogenetic perspectives on the evolution of prokaryotes. *Syst Biol*, 2001, 50(4): 497~512. [\[DOI\]](#)
- [39] Qiu Y L, Lee J, Bernasconi-Quadroni F, Soltis D E, Soltis P S, Zanis M, Zimmer E A, Chen Z, Savolainen V, Chase M W. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genomes. *Nature*, 1999, 402(6760): 404~407. [\[DOI\]](#)
- [40] Soltis P S, Soltis D E, Chase M W. Angiosperm phylogeny inferred from multiple genes as a tool for comparative biology. *Nature*, 1999, 402(6760): 402~404. [\[DOI\]](#)
- [41] Barkman T J, Chenery G, McNeal J R, Lyons-Weiler J, Ellisens W J, Moore G, Wolfe A D, dePamphilis C W. Independent and combined analyses of sequences from all three genomic compartments converge on the root of flowering plant phylogeny. *Proc Natl Acad Sci USA*, 2000, 97(24): 13166~13171. [\[DOI\]](#)
- [42] Murphy W J, Eizirik E, Johnson W E, Zhang Y P, Ryder O A, O'Brien S J. Molecular phylogenetics and the origins of placental mammals. *Nature*, 2001, 409(6820): 614~618. [\[DOI\]](#)
- [43] Madsen O, Madsen O, Scally M, Douady C J, Kao D J, DeBry R W, Adkins R, Amrine H M, Stanhope M J, de Jong W W, Springer M S. Parallel adaptive radiations in two major clades of placental mammals. *Nature*, 2001, 409(6820): 610~614. [\[DOI\]](#)
- [44] Gribaldo S, Philippe H. Ancient phylogenetic relationships. *Theor Popul Biol*, 2002, 61(4): 391~408. [\[DOI\]](#)
- [45] Moret B M E, Warnow T. Advances in phylogeny reconstruction from gene order and content data. Zimmer E A, Roalson E H eds. In: *Molecular Evolution: Producing the Biochemical Data*, Part B, 2005.
- [46] Lin J, Gerstein M. Whole-genome trees based on the occurrence of folds and orthologs: implications for comparing genomes on different levels. *Genome Res*, 2000, 10(6): 808~818. [\[DOI\]](#)
- [47] Rokas A, Holland P W H. Rare genomic changes as a tool for phylogenetics. *Tree*, 2000, 15(11): 454~459.
- [48] NIE Wen-Hui. Comparative molecular cytogenetics study among carnivores [Dissertation]. Chinese Academy of Science. Kuming Institute of Zoology, 2005.
何文惠. 食肉动物的比较分子细胞遗传学研究[学位论文]. 中国科学院昆明动物研究所. 2005.
- [49] Venkatesh B, Ning Y, Brenner S. Late changes in spliceosomal introns define clades in vertebrate evolution. *Proc Natl Acad Sci USA*, 1999, 96(18): 10267~10271. [\[DOI\]](#)
- [50] Nikaido M, Rooney A P, Okada N. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. *Proc Natl Acad Sci USA*, 1999, 96(18): 10261~10266. [\[DOI\]](#)
- [51] Murphy W J, Pevzner P A, O'Brien S J. Mammalian phylogenomics comes of age. *Trends in Genetics*, 2004, 20(12): 631~639. [\[DOI\]](#)
- [52] O'Brien S J, Stanyon R. Ancestral primate viewed. *Nature*, 1999, 402(6757): 365~366. [\[DOI\]](#)