

doi:10.3969/j.issn.1672-5565.2014.01.06

协同系统发生的研究方法

王 智

(陕西学前师范学院生物科学与技术系, 陕西 西安 710100)

摘要:协同系统发生研究生态上相关群体的系统发生间的关系,其研究方法主要分为两类:基于事件法和整体拟合法。基于事件法包括布鲁克斯简约分析、组分分析、协调树分析等,此类方法考虑进化情境,逻辑性强,但分析结果过多,较难选择;整体拟合法考虑了系统发生的误差,提供较好的关系一致性,但缺乏联系进化情境。本文详细比较了两类方法下各种分析软件的优劣,提出了整合两类方法的可能性,以便更好地理解协同系统发生。

关键词:协同系统发生;基于事件法;整体拟合法;布鲁克斯简约分析

中图分类号:Q7 **文献标志码:**A **文章编号:**1672-5565(2014)-01-033-05

An overview of cophylogeny research methods

WANG Zhi

(Department of life science & technology, Shaanxi Xueqian Normal University, Xi'an Shaanxi 710100, China)

Abstract: Cophylogeny is the study of the relationships between phylogenies of ecologically related groups. These methods are divided into two kinds: Event-Based Methods and Global-Fit Methods. The former includes Brooks Parsimony Analysis, Component, Reconciled Tree Analysis, etc, which thinks of evolution scenario and is much logical, but we make choice difficultly because of more analytic result. The latter provides congruence in a relationship as well as taking into account phylogenetic error, but it lacks of linking with relevant evolution scenario. The paper compares detailedly kinds of analytical softs and shows advantage and the disadvantage in the methods and puts forward the possibility of integrating them to understand better cophylogeny.

Keywords: Cophylogeny; Event-based methods; Global-fit methods; Brooks parsimony analysis

1 协同系统发生的概念

协同系统发生(Cophylogeny, or parallel cladogenesis or co-cladogenesis)是研究生态上相关群体(如类群、地理区域,基因等)系统发生间的关系,在此,“主”(如有机体、寄生物的宿主或一个地区)系统发生是独立的,而“辅”(如基因、寄生物或有机体)系统发生在某种程度依赖于“主”系统发生^[1]。两个物种的协同进化是协同系统发生研究的主要领域之一。使用协同进化这个术语时要区别协同趋异(Codivergence)和协同成种(Cospeciation)。协同进化(Coevolution)指的是两个物种内或相互作用的种群内彼此进化变化的一般过程。协同趋异指平行分

支的进化过程(当一个物种分支为两个或多个物种时)或者说一个生物体实体的成种导致与此相关的其它生物体实体的成种^[2]。协同趋异与协同成种常常交替使用,一些学者认为协同趋异是相关类群分成截然不同的群体的过程,注重过程;而协同成种指的是它们已经独立进化为不同的物种,注重结果^[3]。协同趋异或协同成种可以说是协同进化的一部分。以宿主——寄生物这种串联模式为例,寄生物和宿主能彼此协同进化,相互作用产生新的特征和性状,如果这些新性状的差异足够大,可以形成新种,那么协同趋异(协同成种)就会发生。协同系统发生的研究内容就包括分析和比对生物的系统发生以揭示协同趋异的模式^[2]。

在对宿主——寄生物关系的早期研究中产生了一

些主流的关系规则,如 Farenholz 规则 (Farenholz's Rule, 1913)、Szidat 规则 (Szidat's Rule, 1940) 等, Farenholz 规则宣称寄生物的系统发生反映了宿主的系统发生^[4]。实际上,几乎所有协同成种的实例中,除了宿主反常地依靠其协作者的例子外,很少有遵循 Farenholz 规则的。因为此规则将协同进化的可能性狭窄到只有协同趋异,而协同趋异仅仅是协同系统发生产生的众多事件中的一种可能。协同趋异的模式没有绝对的规则,因为两个物种不必遵循相同的进化路径。其它四种可能的进化情境 (Evolution scenario) (见图 1) 增加了协同系统发生研究模式的复杂性^[5-7]。

第一种情境称为分类事件 (Sorting events), 表明似乎是宿主分类时寄生物不跟随。这有三种生物学可能,一是寄生物出现但未被检测到 (称为 x 事件); 二是被认为失掉载体 (Missing the boat, MTB), 宿主形成新物种而寄生物只跟随趋异的宿主, 从而丢掉了其它的宿主种类。当宿主迁移到一新的没有寄生物的栖息地形成新物种就会发生这种情况; 三是寄生物在宿主谱系上灭绝 (Drowning on arrival, DOA), 宿主与寄生物不再发生关联。第二种情境称为宿主转换事件 (Host switching events), 寄生物寄生于不同于原先宿主谱系的宿主上, 继而开拓新宿主的近亲为宿主。第三种情境称为复制 (Duplication) 或宿主内成种 (Intrahost speciation)。寄生物不刺激宿主形成新种, 而在进化上发生分歧, 两个寄

生种群分隔进入不同的生态位——可能为避免竞争, 以形成新种而结束, 但仍具相同的宿主。第四种情境称为惰性 (Inertia), 即寄生物对宿主趋异不发生任何反应, 这种“惰性”导致寄生物具有多个宿主, 从而使得一种寄生物能在不同宿主上存活。不同于“丢失载体”的情形, 此种寄生物遵循两个以上的趋异宿主, 虽然本身并不发生趋异。

这几种生物学进化情境为研究宿主——寄生物的协同系统发生提供了基础, 使用这些生物类别, 通过对比系统发生就可以更好地理解协同进化事件, 而这些事件导致了现存的物种分类结构。将要讨论的协同系统发生方法就是使用这些类别来描述进化结果。关于协同系统发生研究必须要提的一点是假定取样充分和彻底, 除此之外, 分类系统的构建和每个物种的系统发生也假定正确。这些假定是值得注意的, 因为那可能会对分析协同系统发生有潜在的影响。先前描述的场景是按自然和生物发生规律发生的, 人类在取样和系统发生上的误差可能会增加其他的进化情境, 从而导致对协同系统发生解释的不正确或不完全。比如, 观察到的一个寄生物种离开宿主而灭绝可能是生物进化的合法结果或仅仅是不完全取样的结果。更重要的是, 必须做到物种系统发生的准确构建以确保随后对协同系统发生的准确分析。物种系统发生的细小差别都可能产生对发生的协同进化情境不正确的描述。

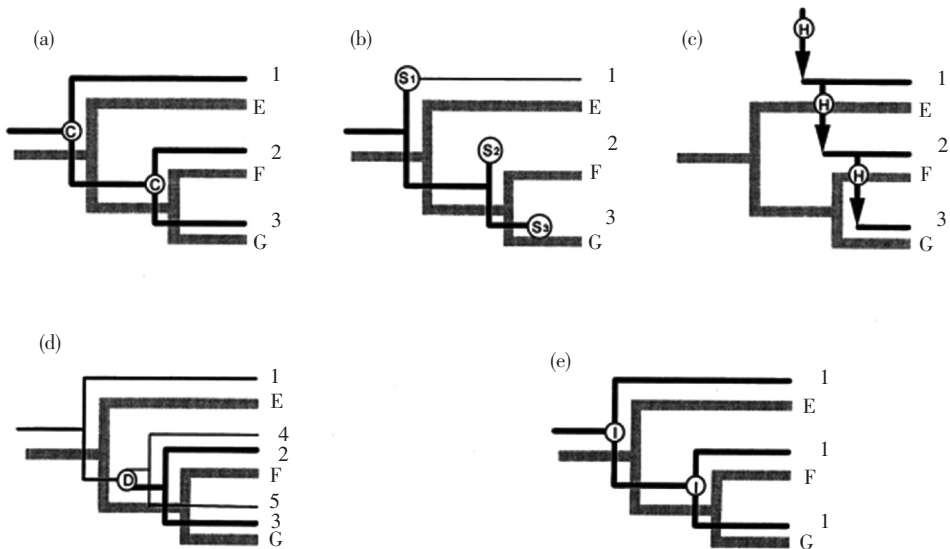


图 1 五种可能的进化情境

Fig.1 Five possible evolutionary scenarios

注: 此图引自参考文献[8]。(a) 协同成种事件('C'), (b) 分类事件('S'), (c) 寄主转换事件('H'), (d) 复制事件('D'), (e) 惰性('I')。图中 E-G 三个阴影线表明了三个宿主单元的系统发生, 1-5 线为寄生物的谱系。

Notes: The figure is cited from the 8th reference. (a) Cospeciation events ('C'). (b) Sorting events ('S'). (c) Host switching events ('H'). (d) Duplication event ('D'). (e) Inertia ('I'). Each figure shows a phylogeny of three host taxa (E-G; shaded line) with a parasite lineage (1-5; line) mapped upon it.

2 协同系统发生的研究方法

协同系统发生的研究方法大体分为两类:“基于事件法”(Event-based methods)和“整体拟合法”(Global-fit methods)^[9]。前者试图结合四种进化情境将寄生物的系统发生图做到宿主的系统发生作图上。这种作图可以通过“基于简约特征”(Parsimony-based character)的最优化(如布鲁克斯简约分析(Brooks parsimony analysis, BPA))或协调树分析(Reconciled tree analysis)(如TreeMap或TreeFitter)达到,在系统发生制图的不一致可能要归因于进化事件和进化情境,因为所提出的进化情境(通常是几个)首先是最优性标准为基础的。相对地,整体拟合法不提出进化情境,而是评估宿主和寄生物间的整体一致性,并识别有助于协同系统发生构建的个别的宿主——寄生物关联。此种方法中,一致性越高,协同成种就越普遍。

2.1 基于事件法

2.1.1 BPA

BPA是历史生物地理学的研究方法之一,最早由D.R.Brooks于1981年提出理论基础,后在1986年被E.O.Wiley正式命名为BPA^[10-12]。BPA是一种直接应用系统发育学原理,结合生物地理学知识,探讨“种历史”的方法,它以“替代假说”为框架,采用简约原则进行分析。BPA分析寄生物的系统发生时,将其转化成一套累积的二进制字符,然后经由简约法绘制到宿主树图上。这些二进制码通过代表各自身份的所有树节点定义每个分类单元。使用这些二进制码就可以构建主系统发生图,然后将宿主系统发生与独立证据比对,如独立识别的宿主关系或地理学和生物地理学证据。与基于事件法中的其他方法不同(如协调树分析)不同,BPA不是基于模型的,因此不会产生系统发生像模型那样的整齐的臆断,而且它并不试图将预先假设做最大化调整。但正如Page和其他学者所说^[13],BPA由于源于寄生物树的字符的非独立性,使得其过多估计了宿主转化的数目。

2.1.2 Component

Component即以计算机程序Component用于研究系统发生、树形分布、基因树(物种树)、宿主——寄生物的协同进化及生物地理学等,由Page发起^[14]。Component并非由数据推导得到树,而是将已存在的树导入程序中,通过几种树对比法,包括计算一致性树、树间的相似度及绘制一个树图到另一个上。也许对于用户来说,组分分析最令人感兴趣

的是它的作树图能力,它可以将寄生物树作图到宿主树上,物种树作图到区域树上,基因树作图到物种树上,使两个不一致树发生关联。Component运行快,界面友好,缺点是不对由于协同进化事件引起的宿主转换作解释^[15-16]。

2.1.3 协调树分析(Reconciled tree analysis)

协调树(Reconciled trees)的概念可追溯至Goodman等人试图协调公认的哺乳动物进化关系与血红蛋白基因进化关系的分歧^[17]。协调树分析通过对寄生物(或基因、有机体等“辅”)系统发生到宿主(或有机体、地区等“主”)系统发生作图来推断哪些进化关系是真实的,哪些是不适合的,任何不一致的结果都可以归于协同进化事件来协调。

协调树分析也是由Page开发,而由Component提供方法,它假定宿主的协同者有足够多的复制和缺失(灭绝、非样本类群)对于协同树完全适合宿主树的关系是必须的。宿主树要和协调过的协同树适合,可以通过三种不同方法达到:复制数目、增加的节点数目和独立缺失的数目。宿主树和其协同树的相似度可以通过Component的随机树发生器产生大量的随机树而统计测定。最初的协同树从而得以和每个随机树协调,对三种适合方法产生一个评估可能性分布^[18-19]。但Page承认,协调树分析经常产生一些非现实的协调协同树。

2.1.3.1 TreeFitter

对于宿主——寄生物系统图的协调经常产生多种可能的关系,这就使得很难确定哪种才代表真实的关系。TreeFitter是第一个把事件——代价(Events-cost)分析组合到基于简约的程序。它采取随意的代价分配使得复制、归类和宿主转换事件有零或正相关。协同趋异事件可能有正、负或零代价。这种事件——代价分配系统允许协调树分析法首次回复到更可能的代表实际的系统发生关系^[20]。

2.1.3.2 TreeMap

协调树分析中下一个主要的发展就是由Page开发的TreeMap,TreeMap是执行系统发生作图的一个Mac程序,由Component派生的,这个程序的主要更新是将宿主转换事件作为宿主——寄生物系统发生不一致的可能解释。当前通用的TreeMap2.02在程序中增加了许多特征,其中一个重要的就是Jungle运算法则,以解释以往程序分析时的中断。另一重要特征是可以指定宿主转换事件的最大数目,以消除某些不可能发生的生物学重建情境^[13,19]。

2.2 整体拟合法

如前所述,整体拟合法不像基于事件法那样提

出进化情境,实际上,它们使用统计方法评价宿主和寄生物间系统发生的一致性程度,确定有助于协同系统发生的特定关联。整体拟合法重要的一点是将最初系统发生重建时误差率考虑得较为充分,使用一个分隔的一致性检验计算了有关的不正确的系统发生概率。而基于事件法的臆断是建立在将要分析的系统发生树是完全而准确的基础上的。某些简单的一致性检验包括 K-H 检验 (Kishino-hasegawa test, KH test) 和不一致性长度差异检验 (Incongruence-length difference test, ILD test)。前者估算不同树形图概似值的方差,后者重点评价各种来源的 DNA 序列(所研究的宿主和寄生物)系统发生的一致性。

2.2.1 最大似然法(Maximum likelihood)

Huelsenbec 等人^[21]提出了两种不同的零假说,以验证检查是否宿主和寄生物系统发生是同一的。第一种方法即最大似然法,使用概率比验证,以评估同一和不同一的程度。第二种方法即最大事后概率法,使用贝叶斯推论(Bayesian inference)直接计算宿主和寄生物系统发生的事后概率。

2.2.2 ParaFit

ParaFit 是最新的协同系统发生数据检验,由 Legendre 开发^[22]。ParaFit 不是要将协同成种最大化,而是在观察协同成种时相当保守(第一类错误足够,第二类错误相当低),而且,它也不是要重建协同趋异情境。ParaFit 用一个置换检验协同趋异评价宿主——寄生物协同进化的整体假说,它使用三种类型的信息来描述矩阵式情形:寄生物系统发生、宿主系统发生和一套可观测到的宿主——寄生物关联。除了评价宿主——寄生物协同进化的一致性水平,ParaFit 还标记每个宿主——寄生物关联,可用于以后更多特定的研究。

3 结语与展望

协同系统发生方法不仅影响研究宿主——寄生物关系,而且影响许多类似关系,如病毒与人类的协同进化关系,可能会对医学研究提供某些思路,这些方法越过各物种对比基因的差异来理解蛋白质组的进化,可以更好理解所有有机体功能。协同系统发生方法也和生物地理领域相关,可以更好理解生物到达的不同地理区域的物种进化。

总之,每种方法论都有其利弊,因此必须经过权衡,选择最优方法。基于简约字符的最优化方法对于分析来说好而简单,虽然它忽视了某些协同进化场景情境;基于事件法似乎最具逻辑性,几乎把所有

可能的进化路径都考虑了,但是在一个给定分析中,过多的结果很难筛选,也很难发现真正的关系。这两种方法都没有考虑潜在的系统发生取样和重建误差。整体拟合法统计不仅考虑系统发生的误差,还提供较好的关系一致性,但缺乏联系协同进化情境。总的来说,改善每种方法都还有较大空间,如果能组合其优点,屏其弊端,那么研究协同系统发育关系将会受益更多。

参考文献(References)

- [1] CHARLESTON M A. Principles of cophylogenetic maps [C]// Lässig M, Valleriani A. Biological Evolution and Statistical Physics. Berlin; Springer-Verlag, 2002: 122-147.
- [2] CHARLESTON M A, PERKINS S L. Traversing the Tangle: Algorithms and Applications for Cophylogenetic Studies [J]. Journal of Biomedical Informatics, 2006, 39 (1): 62-71.
- [3] LIGHT J E, HAFNER M S. Codivergence in heteromyid rodents (Rodentia: heteromyidae) and their sucking lice of the genus Fahrenholzia (Phthiraptera: anoplura) [J]. Systematic Biology, 2008, 57(3): 449-465.
- [4] FAHRENHOLZ H. Ectoparasiten und Abstammungslehre [J]. Zoologischer Anzeiger, 1913(41): 371-374.
- [5] PAGE R D M. Parallel phylogenies: reconstructing the history of host-parasite assemblages [J]. Cladistics, 1994, 10(2): 155-173.
- [6] PATERSON A M, PALMA R L, GRAY R D. How frequently do avian Lice miss the boat Implications for coevolutionary studies [J]. Systematic Biology, 1999, 48 (1): 214-233.
- [7] PATERSON A M, GRAY R D. Host-parasite cospeciation, host switching and missing the boat [C]// Clayton-DH, Moore J. Host-Parasite Evolution: General Principles and Avian Models. Oxford: Oxford University Press, 1997: 236-50.
- [8] PATERSON A M, BANKS J. Analytical approaches to measuring cospeciation of host and parasites: through a glass, darkly [J]. International Journal for Parasitology, 2001, 31(9): 1012-1022.
- [9] DESDEVISES Y. Cophylogeny: insights from fish-parasite systems [J]. Parasitologia, 2007, 49(3): 125-128.
- [10] BROOKS D R. Hennig's Parasitological Method: A Proposed Solution. [J]. Systematics Biology, 1981, 30 (3): 229-249.
- [11] BROOKS D R. Historical ecology: a new approach to studying the evolution of ecological associations [J]. Annals of the Missouri Botanical Garden, 1985, 72(4): 660-680.
- [12] WILEY E O. Methods in vicariance biogeography [C]//

- Hovenkamp PH, et al. Systematics and Evolution: A Matter of Diversity. Utrecht: Institute of Systematic Botany, Utrecht University, 1986:83-306.
- [13] PAGE R D M. Parallel phylogenies: reconstructing the history of host-parasite assemblages [J]. Cladistics, 1994, 10(2): 155-173.
- [14] PAGE R D M. Component: Tree comparison software for Microsoft Windows, version 2.0. London: The Natural History Museum, 1993.
- [15] SLOWINSKI J B. Review of the computer program Component [J]. Cladistics, 1993, 9(3): 351-353.
- [16] CHARLESTON M A. Jungles: A new solution to the host/parasite phylogeny reconciliation Problem [J]. Mathematical Biosciences, 1998, 149(2): 191-223.
- [17] GOODMAN M, CZELUSNIAK J, MOORE G W, et al. Fitting the gene lineage into its species lineage: a parsimony strategy illustrated by cladograms constructed from globin sequences [J]. Systematic Biology, 1979, 28(2): 132-168.
- [18] PAGE R D M, CHARLESTON M A. From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem [J]. Molecular Phylogenetic Evolution, 1997, 7(2): 231-240.
- [19] PAGE R D M. Maps between trees and cladistic analysis of historical associations among genes, organisms, and areas [J]. Systematic Biology, 1994, 43(1): 58-77.
- [20] RONQUIST F. Reconstructing the history of host-parasite associations using generalised parsimony [J]. Cladistics, 1995, 11(1): 73-89.
- [21] HUELSENBECK J P, CRANDALL K A. Phylogeny Estimation and Hypothesis Testing Using Maximum Likelihood [J]. Annual Review of Ecology and Systematics, 1997, 28: 437-466.
- [22] LEGENDRE P, DESDEVISES Y, BAZIN E. A statistical Test for Host-Parasite Coevolution [J]. Systematic Biology, 2002, 51(2): 217-234.