

历史生物地理学中的 BPA 分析法

I: 初级 BPA^{*}

孟凯巴依尔^① 李枢强^{①**} 傅金钟^②

(^①中国科学院动物研究所 北京 100080; ^②Department of Zoology, University of Guelph, Canada)

摘要: BPA (brooks parsimony analysis) 是历史生物地理学研究的重要分析方法之一, 在过去的 20 年内得以发展和完善。除了用于生物地理分布格局的支序分析并解释其形成原因外, BPA 也可用于寄生物与宿主、物种与区系以及协同进化等方面的研究。本文详细介绍了 BPA 的原理以及初级 BPA 的分析方法。

关键词: BPA; 初级 BPA; 历史生物地理

中图分类号: Q332 **文献标识码:** A **文章编号:** 0250-3263(2003)05-64-05

The Application of BPA in Historical Biogeography

Part One: Primary BPA

MENG Kaibayier^① LI Shu-Qiang^① FU Jin-Zhong^②

(^①Institute of Zoology, Chinese Academy of Sciences, Beijing 100080, China;

^②Department of Zoology, University of Guelph, Canada)

Abstract: BPA (brooks parsimony analysis) is one of the most popular methods in the study of historical biogeography.

A brief introduction on the primary BPA is given in the present paper.

Key words: BPA; Primary BPA; Historical biogeography

BPA (brooks parsimony analysis, BPA) 是历史生物地理学的研究方法之一。它首先在 1981 年由 Brooks^[1,2] 提出理论基础, 后被 Wiley^[3] 在 1986 年正式命名为 BPA。

在过去的 20 多年中, BPA 的理论与方法得到进一步完善^[3-8]。Brooks 等^[9] 在 2001 年也对这种方法做了详细的解释与说明。鉴于国内应用 BPA 开展历史生物地理学方面的研究几乎空白, 而国外关于 BPA 的应用日益广泛, 相关讨论也日益深入。作者拟对 BPA 的原理和方法做一介绍。本文作为第一部分, 主要介绍 BPA 原理和初级 BPA 分析方法。二级 BPA 分析和综合实例等将在第二部分涉及。

1 BPA 的原理

BPA 是一种直接应用系统发育学 (phylogenetics) 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 (vicariance hypothes) 为框架, 采用简约原则进行分析。

假设 0 是 BPA 的基本假设之一, 由 Zandee 和 Roos^[10] 针对 Nelson 和 Plantick^[11] 的假设 1 和 2 提出。假设 0 认为, 祖先种的分布状况可以通过类群的系统发育关系和现生种的分布现状估计出来^[4]。当类群在某个分布区缺失时, 这种缺失可能是真实的, 也可能不是。在 BPA 分析中这种缺失数据 (missing data) 不提供任何信息, 而广布种 (widespread species) 在分布区内具有共近裔特征, 是历史生物地理学研究中具有意义的分析

* 国家自然科学基金资助项目 (No. NSFC-30270183, 39970102), 同时部分得到国家基础科学人才培养基金 (No. NSFC-J0030092), 中国科学院知识创新工程 (No. KSCX2-1-06A, KSCX3-10Z-01) 和中国科学院生物科学与技术研究特别支持费 (No. STZ-00-19) 的资助;

** 通讯作者, E-mail: liuq@panda.ioz.ac.cn;

第一作者介绍 孟凯巴依尔, 26 岁, 硕士, 实习研究员; 研究方向: 资源动物学。

收稿日期: 2003-04-18, 修回日期: 2003-07-20

依据。

BPA 承认种的个体性,因此种和种形成事件被认为相互独立。这类似于克氏辅助法则(kluge's auxiliary principle)中允许个别种或进化支序图的分支否定假设 0,即先验地假设数据的独立性。因此,区域支序图(area cladogram)中任何分支上任意一个种的分布都有可能假设 0 的潜在否定依据,如分支上的类群分布现状与替代事件无关等。

基于以上两点,进行 BPA 分析时每个原始数据都必须原封不动地使用,包括种的系统发育数据和地质历史数据。任何与假设 0 相违背的现象(包括多余种 redundant species、广布种以及对扩散现象的不同解释)都会促使分析产生偏差,甚至是错误。当某些数据需要人为修正时,必须有分析性解释。这是 BPA 与组分分析(component analysis, CA)、协调树分析(reconciled tree analysis, RTA)和三域综合分析(three area statement analysis, TAS)的区别。

关于种的形成, BPA 认为某一分布区内的种是通过不同机制形成的,且仅有一个唯一的形成机制和过程。某一个导致物种形成的事件不影响后续物种的形成。因此,不能武断地认定进化支序图中某分支是由一种机制形成的。另外, BPA 承认地理发育和生物进化是同时发生并有共同规律,地理发育和物种形成、种群以及生物区系的集合具有潜在的历史相关性。

2 初级 BPA 的研究方法

尽管 Brooks 本人现在认为 Hennig 简约分析更加适合作为 BPA 的分析基础(个人通讯, 2003),但目前的 BPA 在历史生物地理学研究中是通过 Wagner 简约分析寻求区域支序图进行的。BPA 的分析过程包括初级 BPA 分析(primary BPA)^[3,5]和二级 BPA 分析(secondary BPA)^[9]两个阶段。初级分析用于寻找共同格局(general pattern),二级分析用来分析具有网状进化历史的分布区,解决数据与假设的矛盾。

BPA 认为某一地理分布区内不同类群应具有共同的分布格局。根据生物类群系统发育树转换出的矩阵,利用 BPA 可以得出共同格局。这个分析过程被称为初级 BPA 分析,它可以用类似于 PAUP 等支序分析程序进行,其中外群设为零。

2.1 单个生物类群的 BPA 初级分析(图 1, 2) 在具有详细的地质历史资料的基础上,单个生物类群的 BPA 初级分析具有一定的研究意义,其结论同样具有一定的可信度。假设某个生物类群(如端足目钩虾科钩虾属)的种 1~5 分别在分布区 A~E 内分布,其 BPA 初级

分析包括以下步骤。

第一步:以性状分析为基础,建立这 5 个种的系统发育关系^[12],标记现生种和祖先种的位置(图 1)。

第二步:以物种对应的地理分布区代码取代系统发育树上的现生种的代码,根据变化后的分支图编制数据矩阵(表 1)。编制数据矩阵采用综合操作研究(inclusive ORing)编码方式^[6],即每个分布区的编码序列是系统发育树上现生种/祖先种的序列总合。系统发育分析中的分类单元(taxa)用分布区取代,系统发育分析中所指的特征序列用现生种与相关祖先种的序列取代。

BPA 应用二进制编码,如果某个种在相应分布区中存在,则编码为“1”;否则为“0”。如种 1、种 6 和种 9 在分布区 A 中的编码为“1”,其它为 0(表 1)。

第三步:通过简约分析产生初级分析的结果,即以种的系统发育关系为基础的区域支序图。如图 2 所示,它不仅体现了分布区的地理发育过程,还反映出种的进化关系。

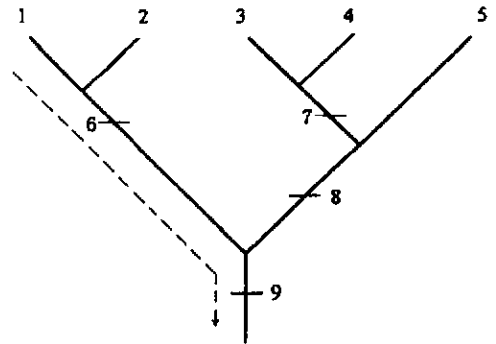


图 1 种 1~5 的系统发育树

虚线表示分布区 A 的编码路径,

数字表示现生种和假设祖先种

表 1 种 1~5 在分布区 A~E 的初级分析的数据矩阵

分类单元	特征序列(种序列)
A	100001001
B	010001001
C	001000111
D	000100111
E	000010011

BPA 认为,如果以生物学资料建立的区域支序图(图 2)和以地质学资料建立的区域支序图相同,那么种 1~5 的物种形成过程与分布区 A~E 的地理发育过程相同,也可以推测分布区 A~E 中其它生物类群具有这种分布格局。这种分布格局被认为是分布区 A~E 的共同分布格局。

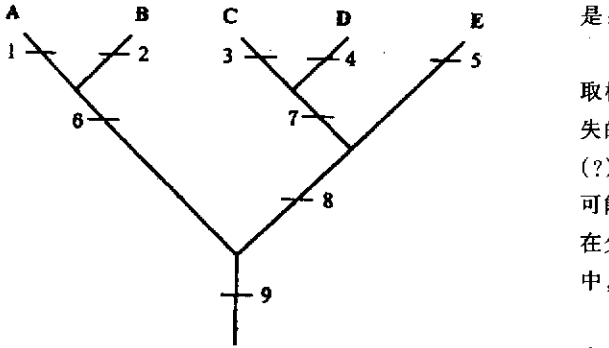


图2 单个生物类群 BPA 分析的区域支序图
数字表示特征序号

当引用多个类群进行 BPA 分析时,共同格局中常出现“缺失数据”现象。这可能是由于人为因素产生的,如取样不完全^[6];但也可能是客观现象。如果是后一种原因,则可能是由于扩散或灭绝现象造成的。

2.2 两个生物类群的 BPA 初级分析(图 3,4) 在具有详细的地质历史资料的基础上,两个生物类群的 BPA 初级分析结果的可信度超过单个类群。假设某一生物类群(如蜘蛛目跳蛛科跳蛛属)的种 1~5 和另外一个类群(如端足目钩虾科钩虾属)的种 10~13 在分布区 A~E 内不均匀分布,两个生物类群 BPA 初级分析的步骤

是:

第一步,编制数据矩阵。在某分布区内研究类群取样完全的前提下,BPA 认为种的不均匀分布是数据缺失的客观原因。在这种情况下,缺失数据编码为问号(?)^[4,5]。这种编码方式没有任何先验的假设,其缺失可能因为灭绝或扩散造成。如图 3 所示,种 1~5 分布在分布区 A~E 内,而种 10~13 仅分布在分布区 A~D 中,在分布区 E 中缺失。相应的数据矩阵如表 2 所示。

第二步,根据数据矩阵,应用支序分析软件^[12],产生区域支序图(图 4)。

第三步,结果分析。根据编码规则,分布区 E 中没有种 10~13 存在与否的假设。如前所述,这不符合 BPA 对数据的要求。在这种情况下,我们应提供假定的分析,即种 10~13 任意一个种都没有在 E 分布区存在过;或者某些种曾经在分布区 E 分布,但已经灭绝(此时忽视取样错误)。前一种假设认为物种形成形式与 Wiley 的异域物种形成^[13,14]相似,其中包括对种 5 的祖先种由分布区 D 扩散到分布区 E 的猜想。它暗示分布区 E 不是共同格局的一部分。后一种假设认为分布区 E 是共同格局的一部分,以替代物种形成机制解释种的缺失。在这种情况下,种 1~5 适合第一种解释,种 10~13 适合第二种解释。

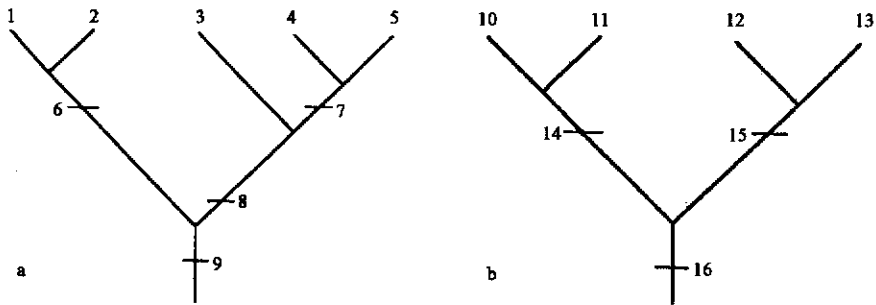


图3 种 1~5 (a)和种 10~13 (b)的系统发育树
数字表示现生种和假设祖先种

表 2 初级 BPA 分析的数据矩阵

分布区	分类单元	特征序列
A	1,10	1000010011000101
B	2,11	0100010010100101
C	3,12	0010000110010011
D	4,13	0001001110001011
E	5	000010111???????

“?”表示分布区内数据缺失

2.3 两个以上生物类群的 BPA 初级分析(图 5,6)

Brooks 等人^[15]认为寻找准确的共同格局至少需要 2 个以上的生物类群,并以其系统发育树作为“共近裔特征”。如果我们有三个生物类群(亲缘关系越远越好),比如在上一个例子(2.2)的基础上加上第 3 个支序图(图 5,种 17~21)。由这 3 个支序图编制的矩阵(表 3)可以得到新区域支序图(图 6)。新区域支序图表明分布区 E 中一个种发生绝灭的假设(第二种假设)比种 5 和种 21 的存在是由两个平行事件扩散形成的假设(第一种假设)更简约。

这个分析结果也说明灭绝事件不与替代理论的假设冲突。在证据缺乏的情况下它既不能证实假设,也不能否定假设。

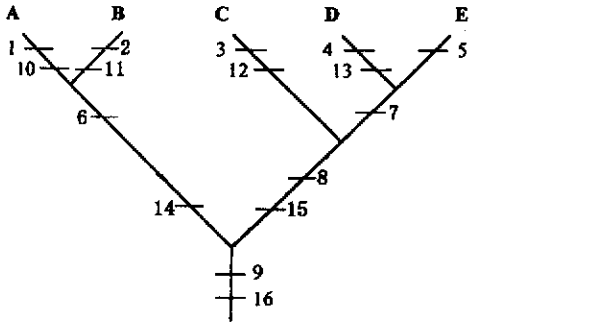


图4 两个生物类群 BPA 分析的区域支序图
数字表示“性状特征”序号

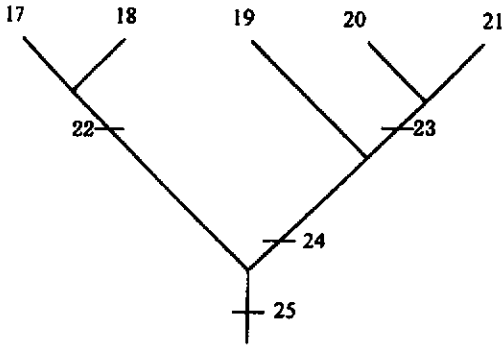


图5 种 17~21 的系统发育树
数字表示现生种和假设祖先种

表3 初级 BPA 分析的数据矩阵

分布区	分类单元	特征序列
A	1, 10, 17	1000010011000101100001001
B	2, 11, 18	0100010010100101010001001
C	3, 12, 19	0010000110010110001000011
D	4, 13, 20	0001001110001011000100111
E	5, 21	000010111???????000010111

“?”表示分布区内数据缺失

总之,BPA 的初级分析主要是为了寻找共同格局,并在一定条件下揭示种的形成/灭绝原因,如边缘隔离物种形成(peripheral isolates speciation)、平行演化种的第2次灭绝和同域物种的形成机理等。在 BPA 初级分析的结果中,这些现象或是表现为“谱系重叠”(lineage duplication),或是祖先种和后裔种出现在区域支序图的同一分支上。

但是在实际操作中,BPA 初级分析常常产生与 BPA 假设相矛盾的结果。研究两个以上的生物类群时,常

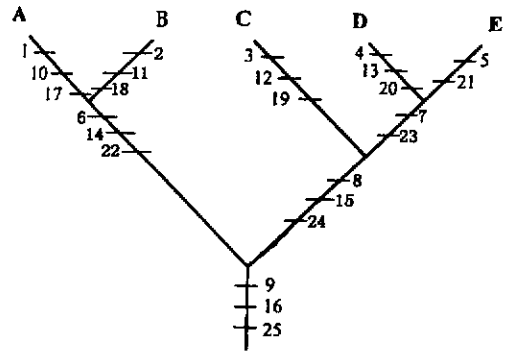


图6 三个生物类群 BPA 分析的区域支序图
数字表示特征序号

常不可避免地显示出网状进化格局。因此,大多数应用 BPA 进行的历史生物地理学研究都需要进行二级 BPA 分析。

有关二级 BPA 分析方法将在以后的文章中详细表述。

致谢 写作过程中,得到 BPA 创始人、加拿大多伦多大学 D.R. Brooks 教授的多方帮助,在此致谢。

参 考 文 献

[1] Brooks D R. Hennig's parasitological method: a proposed solution. *Syst Zool*, 1981, **30**: 229 ~ 249.

[2] Brooks D R. Historical ecology: a new approach to studying the evolution of ecological associations. *Ann Missouri Bot Garden*, 1985, **72**: 660 ~ 680.

[3] Wiley E O. Methods in vicariance biogeography. In: Hovenkamp P, et al. eds. *Systematics and Evolution: A Matter of Diversity*. Institute of Systematic Botany, Utrecht University, 1986. 83 ~ 306.

[4] Wiley E O. Vicariance biogeography. *Ann Rev Ecol Syst*, 1988a, **19**: 513 ~ 542.

[5] Wiley E O. Parsimony analysis and vicariance biogeography. *Syst Zool*, 1988b, **37**: 271 ~ 290.

[6] Cressey R F, Collette B, Russo J. Copepods and scombrid fishes: a study in host-parasite relationships. *Fish Bull*, 1983, **81**: 227 ~ 265.

[7] O'Grady R T, Deets G B. Coding multistate characters, with special reference to the use of parasites as characters of their hosts. *Syst Zool*, 1987, **28**: 1 ~ 21.

[8] Brooks D R. Parsimony analysis in historical biogeography and coevolution: methodological and theoretical update. *Syst Zool*, 1990, **39**: 14 ~ 30.

[9] Brooks D R, Veller M Van, McLennan D A. How to do BPA,

- really. *Journal of Biogeography*, 2001, **28**: 345 ~ 358.
- [10] Zandee M, Roos M C. Component-compatibility in historical biogeography. *Cladistics*, 1987, **3**: 305 ~ 332.
- [11] Nelson G, Platnick N I. *Systematics and Biogeography: Cladistics and Vicariance*. New York: Columbia Univ Press, 1981.
- [12] 孟凯巴依尔, 李枢强, 傅金钟. 联合使用 PAUP 和 MacClade 程序进行支序分析. *动物学杂志*, 2002, **37**(6): 59 ~ 62.
- [13] Wiley E O. *Phylogenetics, the theory and practice of phylogenetic systematic*. New York: Wiley-interscience, 1981.
- [14] Wiley E O, Mayden R L. Species and speciation in phylogenetic systematics, with examples from the North American fish fauna. *Ann Mo Bot Garden*, 1985, **71**: 596 ~ 635.
- [15] Brooks D R, McLennan D A. *Phylogeny, Ecology and Behavior: A Research Program in Comparative Biology*. Chicago: Univ Chicago Press, 1991.