

植物DNA条形码促进系统发育群落生态学发展

裴男才^{1,3*} 张金龙^{2,3} 米湘成² 葛学军¹

1 (中国科学院华南植物园, 中国科学院植物资源保护与可持续利用重点实验室, 广州 510650)

2 (中国科学院植物研究所植被与环境变化国家重点实验室, 北京 100093)

3 (中国科学院研究生院, 北京 100049)

摘要: 系统发育群落生态学是近年兴起的一个重要生态学研究分支, 它以群落生态学为基础并引入了系统发育的分析方法, 全面动态地反映了群落中物种内和物种间的相互作用关系, 揭示了群落格局形成的生态学过程, 研究了生物多样性的形成及维持机制。巴拿马BCI (Barro Colorado Island)样地的成功例子说明, 在固定样地进行长期的群落生态与系统发育研究切实可行且极具意义; DNA条形码的快速兴起对这一研究发挥着重要作用。本文先列举了群落生态与系统发育综合分析能解决的群落系统发育结构、群落生态位结构、生物地理学和性状进化等生态学问题; 接着介绍了标准植物DNA条形码以及利用片段组合(*rbcL*+*matK*+*trnH-psbA*)进行快速物种识别和近缘种区分、精确群落系统发育关系的构建以及群落生态学研究; 随后提出DNA条形码研究在类群水平上需注意两片段的条形码组合(*matK*+*rbcL*)在同属种鉴别能力上的不足, 而在较大尺度群落水平上需对实验设计进行优化。DNA条形码将为探讨物种多样性及其维持机理、系统发育beta多样性以及群落水平上功能性状进化研究提供新的思路。

关键词: 群落系统发育重建, 群落生态学, *rbcL*, *matK*, *trnH-psbA*, 大型固定样地, APG分类系统

Plant DNA barcodes promote the development of phylogenetic community ecology

Nancai Pei^{1,3*}, Jinlong Zhang^{2,3}, Xiangcheng Mi², Xuejun Ge¹

1 Key Laboratory of Plant Resource Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650

2 State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093

3 Graduate University of the Chinese Academy of Sciences, Beijing 100049

Abstract: There is a dynamic interplay between ecology and evolution within community ecology. Phylogenetic community ecology describes the intraspecific and interspecific relationships within a community, aiming to reveal the processes driving community assembly at multiple scales. Previous research has highlighted the role of phylogenetic and historical biogeographical data in explaining current patterns of global biodiversity. The success of using DNA barcoding in the construction of tropical forest community phylogenies highlights the usefulness and challenges of long-term research on community ecology and phylogenetics based on forest dynamic plots. In this paper, we illustrate the feasibility of a synthesis between community ecology and evolutionary biology in order to resolve particular ecological issues on community phylogenetic structure, community niche structure, biogeography, and trait evolution. We summarize progress on the development of a plant DNA barcoding system, and introduce the usage of a combination of DNA markers (*rbcL*+*matK*+*trnH-psbA*) for fast species discrimination and community phylogenetic reconstruction. We also explore the utilization of well-resolved phylogenies to understand community ecology. We discuss the limitations of core plant DNA barcodes (*rbcL*+*matK*) when identifying congeners, and propose an improved sequencing strategy suitable for studies at the community level. We expect that

收稿日期: 2010-10-14; 接受日期: 2011-03-08

基金项目: 国家重点基础研究发展计划(973)(2007CB411600)和中国科学院知识创新工程重要方向项目(KSCX2-YW-N-0807)

* 通讯作者 Author for correspondence. E-mail: nancai.pei@gmail.com

plant DNA barcodes will prove very useful for the study of species diversity, mechanisms of biodiversity maintenance, phylogenetic beta diversity and functional trait evolution.

Key words: community phylogenetic reconstruction, community ecology, *rbcL*, *matK*, *trnH-psbA*, large scale permanent plot, APG

系统发育群落生态学(phylogenetic community ecology)基于群落水平,运用系统发育分析方法来探讨局域范围(如以样地为尺度)内物种共存和生物多样性的维持机制,或者区域范围内(如以不同气候带为尺度)不同类型群落的构建规则(Webb *et al.*, 2002; Cavender-Bares *et al.*, 2009)。*Phylogenetics and Ecology*一书(Eggleton & Vane-Wright, 1994)的出版,表明系统发育学在20世纪90年代已引起生态学界关注,然而将分子系统发育方法真正融入生态学领域却是近几年的事。Campbell O. Webb提出将系统发育、性状进化和群落结构三方面结合起来,探讨群落结构与物种共存之间的联系,同时基本明确了系统发育群落生态学的主要任务和研究思路,取得了一系列成果(Webb, 2000; Webb *et al.*, 2002, 2006)。Johnson和Stinchcombe(2007)的研究也表明,群落内生态和进化两者存在着复杂的动态相互作用。以W. John Kress为代表的科研团队,在巴拿马Barro Colorado Island(BCI) 50 ha大样地上开展了植物DNA条形码与热带森林群落系统发育结构的研究,这标志着DNA条形码与植物系统发育群落生态学的成功结合,相关成果已发表在2009年的《美国科学院院刊》上(Kress *et al.*, 2009)。国际生物条形码联盟(Consortium for the Barcode of Life, CBOL) (<http://www.barcodeoflife.org/>)已经完成对植物标准DNA条形码的评价,并已进入大规模的标本收集、测序、物种鉴定等实质性工作阶段。在中国,从2004年开始,大样地平台经过多年的建设,已形成比较完善的网络,基本代表了中国各主要气候带的森林类型。今后可利用这个平台开展宏观和微观领域的科学研究,如全球变化与生物多样性、植物功能性状与系统演化的关系等(马克平, 2008, 2011)。

本文将从以下几个方面展开叙述:(1)群落生态与系统发育综合分析的重要性及其可解决的科学问题;(2)植物DNA条形码研究近况以及利用DNA条形码构建的分子系统在群落生态研究中的优势及进展;(3)DNA条形码研究存在的问题及其在植物生态学研究中的展望。通过本文的介绍,希望能给

对DNA条形码和大样地群落生态学研究感兴趣的读者提供一些思路。

1 群落生态学与系统发育生物学的综合分析

以往群落生态学和系统发育生物学这两个学科的研究方向和内容是相对分离的,在群落水平上探讨群落过程和模式对生物进化的影响尚未引起人们足够的重视(Johnson & Stinchcombe, 2007)。Webb等首次将系统发育、性状进化和群落结构三种分析相结合(Webb, 2000; Webb *et al.*, 2002),并列举了一系列可解决的生态学问题:

(1)群落的系统发育结构:在一个样方或生境内,物种间的系统发育关系是否为非随机的?或者说,这些物种在系统发育关系上是否具有关联性?Webb(2000)采用两个群落系统发育关系指数——净亲缘指数(net relatedness index, NRI)和最近亲缘指数(nearest taxon index, NTI)来描述群落的系统发育结构。有关这两个指数的推导及计算过程详见Webb等(2002)和Swenson等(2006)。当生境过滤作用是群落的主导过程时,群落主要由系统发育关系较近的物种构成,此时群落的NRI和NTI比较高;而当种间竞争是群落的主导过程时,群落主要由系统发育关系较远的物种构成,此时群落的NRI比较低,而NTI比较高。这两个指数从不同侧面反映了两个相反群落生态过程。

(2)群落生态位结构:群落中物种的生态位差异是否按某种系统发育关系排列?长期以来,生态学家们通过研究群落中物种的生态性状分布来了解群落组成。了解生态位差异的系统发育分布信息,有助于人们理解生态位结构是由生态因素、进化因素抑或是两者共同作用形成的(Ackerly *et al.*, 2006)。

(3)生物地理学、性状进化与群落分析:生态位格局是在何处以及何种群落条件下形成的?生态性状的积累本质上是随机的,且与物种形成相比,是一个漫长的过程(Harvey & Pagel, 1991)。现存植物群落中物种的生态性状可以反映这些物种及其最近祖先的生物地理历史。性状演化模式和生态相

互作用的不同组合可以形成相似的群落系统发育结构,此时需要加入性状数据、群落数据和系统发育关系等参数来探讨群落的演化和聚集方式。不同的森林群落聚集方式可通过竞争、生境过滤(habitat filtering),或者中性作用产生(Webb *et al.*, 2006)。当性状高度保守时,也即当多个性状对群落中的物种个体产生作用时,系统发育群落结构最为明显;而当具有进化性状的类群在群落中常见时,生境过滤作用就形成显著的系统发育结构(Kraft *et al.*, 2007)。

综合系统发育分析方法,将实验与假说验证相结合来分析群落结构、生态位分化、生物地理学和性状进化数据,对生态学将有极大的促进作用(Webb *et al.*, 2002, 2006; 钟扬等, 2004; 方精云, 2009)。如在解释局域和区域尺度上的生物多样性现状和维持模式时,或者在预测生态系统过程和全球变化对生物演化的影响等科学问题上,将多方面的数据与更精确的检测手段相结合,可弥补以往单一研究方法的不足,最终达到解释现象和探索未知的目的(Donoghue, 2008; Cavender-Bares *et al.*, 2009)。

2 标准植物DNA条形码与群落系统发育研究

DNA条形码(barcoding)这一概念由加拿大皇家学会会员、Guelph大学的Paul Hebert教授等人最先提出(Hebert *et al.*, 2003)。在动物界, DNA条形码是一段650 bp长的线粒体细胞色素c氧化酶亚基I(cytochrome c oxidase I, *COI*)基因序列,它在不同的分类水平上(门、目、种)都具有良好的识别能力,甚至在近缘物种上也能得到丰富的变异信息。由此科学家们将这段*COI*基因片段确定为动物DNA条形码,并广泛应用在昆虫、鱼类、鸟类等动物物种水平的识别和鉴定上(Hebert *et al.*, 2003; Hajibabaei *et al.*, 2006)。由于植物的线粒体基因组进化速率慢,遗传分化小,因此*COI*不适用于植物。

为了寻找合适的片段,世界各国植物学工作者进行了广泛的评价和筛选工作(Chase *et al.*, 2007; Kress & Erickson, 2007; Pennisi, 2007; Lahaye *et al.*, 2008)。国际生物条形码联盟于2009年初步确定并推荐使用的标准陆生植物核心条形码是叶绿体基因*rbcL*和*matK*(CBOL Plant Working Group, 2009)。在实践中,可根据需要补充一些变异性较大、能够分辨到更低分类阶元的候选DNA片段(如间隔区

trnH-psbA, 或者核基因片段ITS)(Gonzalez *et al.*, 2009; Kress *et al.*, 2009)。更多植物DNA条形码序列的选取标准,可参看宁淑萍等(2008)。

DNA条形码是分类学家识别物种的工具,也可用于检测生物产品的纯度以及甄别管制的物种(如外来入侵物种和濒危物种)(Dick & Kress, 2009; van de Wiel *et al.*, 2009),还可用于探索、检测以前没有被描述过的物种和疑难物种(Lahaye *et al.*, 2008)。更重要的是,应用标准植物DNA条形码片段,可以快速获得大量植物DNA序列,从而构建大尺度的分子系统树,用于更大分类阶元的分子系统学研究,甚至可以结合群落生态学进行植物群落系统发育研究(Kress *et al.*, 2009)。

分子系统学的发展趋势是依靠多个片段,而研究近缘类群则需要能提供较多变异信息的片段;即使当一些片段对某些特定类群不适用而缺乏完整的序列时,运用矩阵的分析方法也能得到比较准确的系统发育关系(Wiens, 2003)。APG分类系统(Angiosperm Phylogeny Group)(APG III, 2009)主要依照植物的3个基因组DNA的序列,以亲缘分枝的方法进行分类,更多地关注被子植物大类群间的系统发育关系。因为它选取的DNA片段分辨率有限,对大多数植物类群只能解决到目或者科的水平,不能满足更低分类阶元水平研究的需要。DNA条形码继承和发扬了原有的分类系统框架,利用其易扩增、低成本并能得到丰富变异信息序列的优势,进行大尺度的分子系统发育重建,并予以修正和完善(Hebert *et al.*, 2003; Lahaye *et al.*, 2008)。Kress等(2009)提出使用*rbcL*+*matK*+*trnH-psbA*的片段组合构建超级矩阵(supermatrix)(图1)来解决不同分类阶元的系统发育关系。

总之,当所研究的植物类群关系跨度大,且需要同时处理好高分类阶元的基本框架(科以上水平)和低分类阶元的准确位置(属,甚至种水平)这两种关系时,所选择的DNA片段必须兼顾保守性和快速进化这两方面。

3 DNA条形码在系统发育群落生态学研究中的优势及应用

准确的物种识别和精确的DNA条形码系统发育关系给群落生态学研究提供了新动力。为了探究群落内所有物种的系统发育关系, Campbell Webb

Species	<i>rbcL</i>	<i>matK</i>	<i>trnH-psbA</i> I	<i>trnH-psbA</i> II	<i>trnH-psbA</i> III
Sp 1	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	ACTTAAGGCCTTTACC	??????????????	????????????????????
Sp 2	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	ACTTAAGGCCTTTA??	??????????????	????????????????????
Sp 3	??TTCAGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	?CTTAAGGCCTTTACC	??????????????	????????????????????
Sp 4	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	ACTTAAG---TTTACC	??????????????	????????????????????
Sp 5	ACTTCCGGCCAACCTTTTA	CCTAAG---CTCTAGACAACA	???TAAGGCCTTTACC	??????????????	????????????????????
Sp 6	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	ACTTAAGGCCTTTACC	??????????????	????????????????????
Sp 7	ACTACCGGTCAACTTTTA	CCTAAGGTTCTCTAGACA???	????????????????	CTCTCAATTGGCA	????????????????????
Sp 8	ACTTCCGGTCAACTATTA	CCTAAGGTTCTCTAGACAACA	????????????????	CTCTCAA-TGGCA	????????????????????
Sp 9	ACTTCCGGTCAACTTTTA	???AAGGTTCTCTAGACAACA	????????????????	CTCTCAATTGTCA	????????????????????
Sp 10	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTC---ACAACA	????????????????	CTCT---TGGCA	????????????????????
Sp 11	ACTTCCTGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	????????????????	CT-TCAATCGGCA	????????????????????
Sp 12	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	????????????????	CTCTCAATTAGCA	????????????????????
Sp 13	??TTCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	????????????????	CTCTCACTTGGCA	????????????????????
Sp 14	ACTTCCGGTCAACTTTTA	????????????????????	????????????????	???????????????	TT---ATTAAACTCTTACA
Sp 15	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	????????????????	???????????????	TTCCGGATTAAACTCTTACA
Sp 16	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	????????????????	???????????????	??CCGGATT-----TACA
Sp 17	ACTTCCGGTCAACTATTA	CCTAAGGTTCTCTAGACAACA	????????????????	???????????????	TTCCGGATTAAACTCTTACA
Sp 18	ACTT---GTCAACTTTTA	CCTAAGGTTCTCTAGAC????	????????????????	???????????????	TTCCGGATTAAACTCTT???
Sp 19	ACTTCCGGTCCACTTTTA	?????GTTCT--GACAACA	????????????????	???????????????	????????????????????
Sp 20	ACTTCCGGTCAACTTTTA	CCTAAGGTTCTCTAGACAACA	????????????????	???????????????	????????????????????

图1 *rbcL*+*matK*+*trnH-psbA*的3位点组合片段的超级矩阵示意图(改自Kress *et al.*, 2009)。第一列为物种名称, 第二列之后的为各片段的序列排列情况, 其中*trnH-psbA* (也即*trnH-psbA* I, *trnH-psbA* II, *trnH-psbA* III等)片段以目为单位进行归类。3个片段共同组成一个矩阵。

Fig. 1 Illustration of supermatrix approach in sequence alignment for phylogenetic analysis (adopted from Kress *et al.*, 2009). The first column represents species name, while others represent sequence alignments of *rbcL*, *matK*, and *trnH-psbA*. In *trnH-psbA*, each order consists of a dataset.

编写了在线Phyloomatic程序(<http://www.phylodiversity.net/phyloomatic/phyloomatic.html>), 用户只需要输入相应格式的物种名录, 即可获得基于APG系统的物种系统发育关系。在早期的系统发育分析软件Phylocom中(<http://www.phylodiversity.net/phylocom/>), 还内置了“bladj”模块, 用户依据一定规则就可推算出进化树各节点距今的时间。当然, 这种估算存在许多弊端, 即当同一科属下分类单元较多时, 运算结果会出现多分枝结构(polytomy), 给群落系统发育结构分析带来较大偏差, 得出错误结果的几率升高(图2)。另外, Phyloomatic方法没有整合裸子植物的信息, 如松柏类植物的进化信息在进化树中得不到体现。而实际上, 针叶树也往往是建群种之一, 如我国中亚热带地区的浙江古田山大样地就以甜槠(*Castanopsis eyrei*)、木荷(*Schima superba*)以及马尾松(*Pinus massoniana*)等物种占优势, 而在温带地区的吉林长白山大样地则以阔叶红松林为主。Kress等(2009)的结果显示Phyloomatic倾向于错误接受零

假设(Type II error), 而利用DNA条形码建立的物种系统发育关系则增加了拒绝零假设的概率。

3.1 快速识别物种以及区分近缘物种

当专业人士缺乏的时候, 要快速准确地鉴定疑难物种, DNA条形码是个很好的方法(Pennisi, 2007; Lahaye *et al.*, 2008)。在热带雨林(综合巴拿马的BCI样地和法属圭亚那地区Nouragues研究站的结果), *rbcL*, *matK*和*trnH-psbA*三个条形码片段单独用于识别木本物种的准确率分别为62%、55%和74%; 如果使用国际条形码联盟推荐的两个片段组合(*rbcL*+*matK*), 其准确率约为76%; 如果使用三个片段组合(*rbcL*+*matK*+*trnH-psbA*), 准确率能提高至84%(其中在BCI样地可达98%)(Gonzalez *et al.*, 2009; Kress *et al.*, 2009)。在亚热带森林(综合鼎湖山、古田山和波多黎各Luquillo样地的结果), *rbcL*, *matK*和*trnH-psbA*三个条形码片段单独识别木本物种的准确率分别为70%、70%和78%; 如果使用国际条形码联盟推荐的两个片段组合, 其准确率约83%;

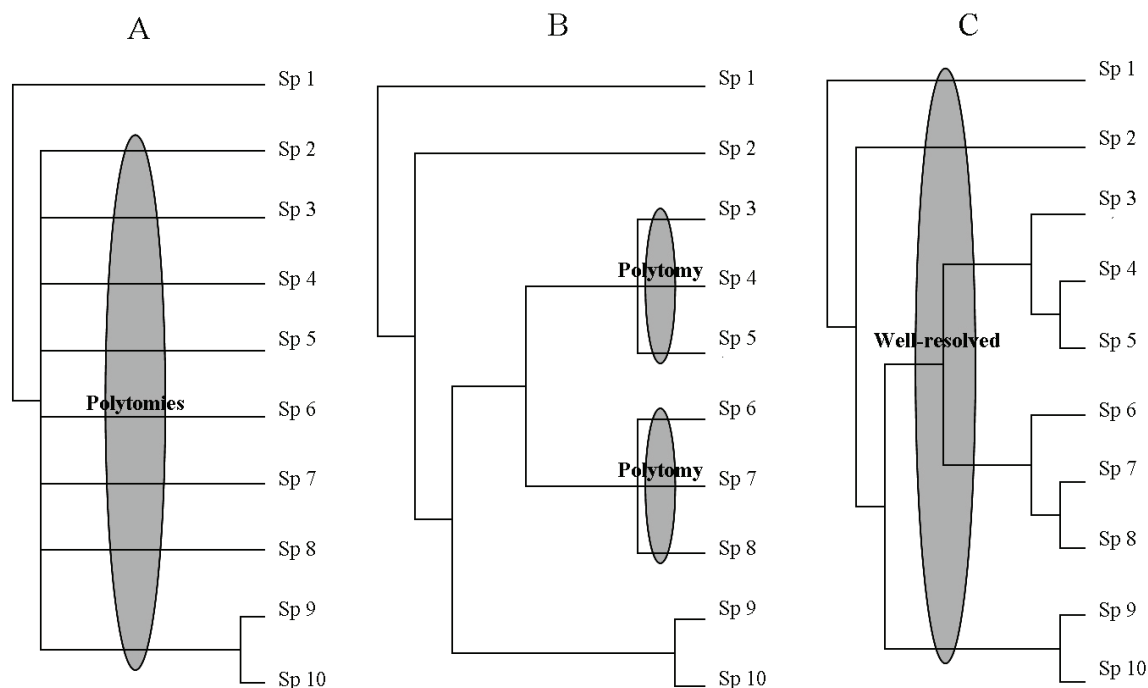


图2 10个物种的系统发育关系示意图。图A用PhyloMatic的方法获得, 图B和C用条形码(Barcode)的方法获得。Sp 1–10表示物种1–10, 以此类推。图A表示当使用在线软件PhyloMatic获取该10个物种的系统发育关系时, 物种2–10的位置未得到解决, 也就是常说的多分枝结构。图B表示, 当使用2个DNA条形码片段(如*rbcL*+*matK*)时, 物种3–5和物种6–8的系统发育关系尚不明确, 而其他物种的位置得到了较好解决。图C表示, 当使用3个DNA条形码片段构成的超级矩阵(如*rbcL*+*matK*+*trnH-psbA*)时, 全部物种的系统发育关系都得到了解决。

Fig. 2 A graphic representation of phylogenetic trees of ten species using PhyloMatic (A) and Barcode (B, C) approaches. Sp 1–10 means species 1–10. (A) Phylogenetic relationships of Sp 2–10 are not well-resolved, i.e., there are polytomies among these taxa when using an online query tool PhyloMatic; (B) Phylogenetic relationships of Sp 3–5, and Sp 6–8 are not well-resolved using a two-locus DNA barcode (*rbcL*+*matK*); (C) Phylogenetic relationships of Sp 1–10 are well-resolved using a supermatrix of three DNA barcodes (*rbcL*+*matK*+*trnH-psbA*).

如果使用三个片段组合, 准确率能提高至88%以上 (Kress *et al.*, 2010; Pei *et al.*, unpublished data)。

总体而言, 在热带和亚热带森林, 单个DNA条形码片段在物种水平上的识别率并不高(55–78%), 但是在属的水平上却可以达到90–100%; 而通过多个片段的组合, 在物种水平上的正确识别率可达80–98%。剩下的一些暂时未能正确识别到物种水平的植物类群, 可能是因为: (1)物种鉴定本身有偏差, 这种情况在野外即使是做专科专属的分类学家身上也会经常发生, 可以通过再次鉴定予以确认; (2)同属近缘种间的变异太小, 导致使用三个条形码片段也不能分辨开来, 此时可增加进化速率更快的片段(如*trnH-psbA*间隔区)予以区分(Kress & Erickson, 2007); (3)存在杂交的情况, 此时使用叶绿体基因不能达到分辨目的, 可以增加或者换用核基因ITS来

提高分辨率(Chen *et al.*, 2010)。由此我们认为, DNA条形码作为一种标准快速识别物种的方法是可行的, 即使在研究背景相对不充分的森林群落, 通过结合地理学或者形态学数据也能提高物种识别率(Gonzalez *et al.*, 2009)。

3.2 构建精确的群落系统发育关系

利用序列信息构建进化树的方法已经较为成熟。从早期的距离法(如neighbor-joining), 到后来的最大简约法(maximum parsimony)、极大似然法(maximum likelihood)及贝叶斯方法(Bayesian methods), 都已经历了较长时间的检验。贝叶斯方法虽然也是应用十分广泛的算法, 但是当运算量太大时, 其马尔可夫链极难收敛, 因而也就难以获得准确的进化树。为了快速运算超大规模数据的进化树并进行相应的Bootstrap检验, 新近开发出的软件

有PhyML(Guindon & Gascuel, 2003)、RA×ML (Stamatakis, 2006)等。一些科研机构也开发出相应的平台以支持用户远程提交运算任务, 例如超级计算机CIPRES(Cyberinfrastructure for Phylogenetic Research) (<http://www.phylo.org/>)。当前学术界有较为完备的分子钟理论和方法, 如非参数速率平滑法(non-parametric rate smoothing, NPRS)、似然罚分法(penalized likelihood)、贝叶斯法等, 分别可在“r8s”(Sanderson, 2003)和“ape” (<http://ape.mpl.ird.fr/>)等软件上运算。

用DNA条形码序列构建群落系统发育关系的优势在于处理系统树的末端分枝时能为分类单元提供准确的系统位置, 可根据所用的DNA片段计算整个群落内所有类群内部真实的分枝长度。相比由Phyloomatic软件产生系统树和采用Phylocom软件中的“bladj”程序模拟运算分枝长度, 基于DNA条形码的分子系统树在进化树的分辨率和进化时间的准确性上都更胜一筹。Kress等(2009)认为, 用3个植物DNA条形码组成的超级矩阵重建BCI大样地群落内物种的系统发育关系, 其结果要比直接使用APG III更好。具体表现在:

(1)在大的系统发育框架上, 由DNA条形码重建的系统发育关系与APG的拓扑结构基本一致, 样地中所包含的全部24个目都被分配在适当的系统位置上, 说明对于局域尺度下的植物群落而言, 利用多片段条形码组合所提供的信息可满足系统发育树的构建, 其效果已接近APG系统。

(2)在系统发育框架末端(例如科、属和物种水平上), 得益于3个DNA条形码组成的超级矩阵(包含进化速率快的片段*trnH-psbA*), 此时获得的末端节点二歧分枝结构比仅由Phyloomatic自动生成的多分枝效果好, 说明在具体的样地中, 如果有充分的序列信息, 加上科研人员的分类学背景知识, DNA条形码可发挥出巨大的优势。这进一步印证了前文的论点——对群落中每个物种进行测序所获得的分子证据, 可以用于探讨种间关系以及完善植物系统发育与演化进程的相关研究(Kress *et al.*, 2010)。

3.3 在系统发育群落生态学中的应用

将DNA条形码构建的进化树和群落数据一起输入到Phylocom和Picante(R软件自带的程序包)程序, 可分析群落结构的成因。如果样方在空间上是连续的, 在分析时首先需要去除空间自相关的作

用。常用的方法是用R软件的spdep程序包(<http://www.r-project.org/>)或者SAR模型(Spatial autoregression model)。然后, 根据每个400 m²小样方所属的生境类型, 用*t*检验来确定各种生境类型中小样方的聚集方式是倾向于系统发育集聚、散布还是随机分布。最后, 对于每个小样方, 用成对*t*检验(paired *t*-test)来比较用条形码和Phyloomatic获得的NRI和NTI值。

以往计算群落结构的相关指数时, 所使用的系统发育关系通常由研究人员自己编撰或者来自一些在线软件(如Phyloomatic)。当所研究群落内物种数较少, 物种关系简单时(如面积小于1 ha的小型样方), 采用这种粗略的系统发育关系也能获得较理想的结果; 而当取样面积大幅增加后(如CTFS的大型固定样地, 5 ha到50 ha), 群落内的物种数随之增多, 种间关系也更复杂, 此时需要一种更为精确的方法。DNA条形码的出现就能满足这一要求。在其他参数设置保持不变的前提下, 使用由DNA条形码构建的系统发育关系(Barcode Phylogeny), 可以得到群落结构指数的准确数值(例如NRI和NTI), 还可以对以往的研究进行校准和完善(简要的研究方案和操作指南见附录I)。例如在BCI样地, Kress等(2009)利用Barcode Phylogeny和Phyloomatic Phylogeny对样地内的7类生境进行比较, 得出不同的群落聚集方式, 并认为利用Barcode Phylogeny得到的结果更为可靠(Kembel & Hubbell, 2006)。Gonzalez等(2010)也认为, DNA条形码所构建的精确系统发育关系, 可以提高有关群落结构计算的统计力。

4 DNA条形码研究存在的问题

国际植物条形码联盟所建议的陆生植物标准DNA条形码片段是叶绿体基因的*matK*和*rbcL*片段。测试发现, 这两个片段的组合在550个物种(包含445种被子植物, 38种裸子植物和67种隐花植物)共计907个样本中, 扩增成功率和测序成功率可达90–98%, 物种识别准确率达72% (CBOL Plant Working Group, 2009)。但整个植物界物种约有40万种, 种间变异大, 不同生活型的植物(例如多年生木本、一年生草本)进化速率也极不相同, 因此这项新兴技术将面临严峻的挑战。

从类群水平上看, 两片段的条形码组合*matK*+

*rbcL*难以分辨为数众多的同属种,例如在桦木科桤木属(*Alnus*)的23个物种中,其准确率不到50%(Ren *et al.*, 2010);而在小檗科小檗属(*Berberis*)的16个物种中准确率只有23%(Roy *et al.*, 2010)。这是因为:(1)*matK*和*rbcL*片段本身所能提供的变异信息位点不够;(2)叶绿体基因组本身的信息量有限(汪小全和洪德元, 1997)。要想快速而准确地识别近缘类群并构建可信的系统发育关系,需要考虑增加基因区域作为标准的植物条形码片段。

从局域群落水平上看,如果在大样地或者保护区的尺度上开展条形码研究,必须考虑到成本问题,需要对实验设计进行优化。单个DNA条形码片段(不论是编码基因还是非编码基因片段)所能提供的有用信息较少,甚至有些引物本身的通用性较差。目前认为3个片段的组合是比较合适的,既可提供足够的信息位点,又可控制成本。具体方法是:有选择地对某些类群进行第三个片段测序。比如在鼎湖山样地,樟科、杜鹃花科、冬青科、桑科等包含较多的物种,有时需要进行第三甚至第四个片段的测序。具体科、属或者物种水平,由于其系统发育位置相对固定,可以着重选用进化速率较快的片段,减少甚至不用进化速率过慢的片段(比如跳过*rbcL*而直接使用*matK*片段来构建大的系统发育关系框架),最后根据实际需要,增加其他进化速率更快的片段(如*trnH-psbA*),或者使用核基因ITS(Chen *et al.*, 2010)。通过优化的片段组合和测序方略,就可以较好地解决植物DNA条形码研究经常遇到的问题。

另外,需要考虑不同生活型植物的进化速率问题。一般来说草本植物的进化速率会明显快于木本植物。当构建热带-亚热带森林全体物种的系统发育关系时,需要根据实际情况来选用DNA条形码片段。特别是在属或者物种水平,通常只能获得系统发育关系未解决的多歧分枝结构(polytomy),这可能是同一DNA条形码片段在不同生活型植物上的不同进化速率造成的结果。这种多歧分枝结构还会对后续的数据分析产生不利影响。

5 植物DNA条形码研究展望

5.1 探讨物种多样性的形成和维持机理

有关物种共存的理论主要有两类:生态位理论(niche theory)和中性理论(neutral theory)。当前,人

们逐渐淡化了对两种理论的争论,转向整合各自的合理成分,解释具体群落的构建方式,进而探讨物种多样性的形成和维持机理(Hubbell, 2001; Chase & Leibold, 2003; Silvertown, 2004; 牛克昌等, 2009)。

物种水平上的生态位差异在共存理论中扮演了重要的角色,而对于它在构建不同类型群落过程中的作用却有争议。虽有研究显示物种间的生态位差异对于成分复杂的热带森林多样性维持确实产生作用(Kraft *et al.*, 2008),这印证了生态位理论的观点,但也不能排除在此过程中中性作用的影响。在群落水平上,将DNA条形码所构建的精确系统发育关系与功能性状结合起来,有助于理解群落的构建机制。最近已有相关的研究结合分析性状、生境以及群落内的分类群,获取精确的系统发育关系之后发现,昆虫传粉特征和动物散布种子的方式对雨林下层林分的单子叶和系统位置比较原始的胡椒科植物的生境过滤作用极为重要;也就是说在特定的生境下,某些植物功能性状的确发挥着重要作用,促使一些类群的植物聚集在一起(Mayfield *et al.*, 2009)。当然,这种现象是否普遍,还需要更多的证据来支持。

5.2 群落水平上功能性状进化的意义及功能性状的选择

测度从个体到群落水平上的植物功能性状是生态学研究的一个重要方面(Tilman, 1988; Baraloto *et al.*, 2009)。从群落水平上描述不同气候梯度上的植被类型功能性状的进化,可以从更宽广的角度上理解气候因子、植被类型和生态系统服务之间的关联性(Wright *et al.*, 2007; Pavoine *et al.*, 2010)。动植物间的协同作用(如互利共生、竞争、捕食等关系)一直是进化生物学家们关注的热点。近来有研究(Lomáscolo *et al.*, 2010)发现,桑科榕属植物功能性状(比如果实大小和颜色)的进化方向,与食果动物(也即种子散布者)的形态、感知能力和行为状态差异是一致的。也就是说,食果动物的差异造成果实性状的进化方向不同,导致在榕属中出现许多果实大小、颜色、气味均有明显差异的物种,形成了比较高的物种多样性。当要筛选植物功能性状用于分析群落水平上研究物种关系、物种与环境或群落间的相互作用时,就需要考虑性状的代表性,所选取的性状及其变化信息既要能反映物种间的差异,也要有相对的保守性。

一个周全的植物功能性状选取策略, 首先要考虑该策略是否有利于回答特定的生态学问题; 有时还需要对某些性状在不同个体上进行重复采集 (Baraloto *et al.*, 2009)。更重要的是, 所选择的指标应有进化意义且尽量方便测量。具体来说, 可以从两大方面来选择功能性状: (1) 营养器官, 如根系(菌根、土壤有机物成分等)、茎秆(胸径大小、密度等)、叶片(大小、质地、叶片内营养元素等) (Ackerly & Cornwell, 2007; Swenson & Enquist, 2007; Kraft *et al.*, 2010; Wright *et al.*, 2010); (2) 繁殖器官, 如花(色泽、气味、对称性、开花物候等)、果实(质地类型、颜色、可食性等)、种子(大小、重量、散布方式等) (Cubas, 2004; Wright *et al.*, 2007; Kraft *et al.*, 2008; Lomáscolo *et al.*, 2010)。另外, 还有很多其他指标如生理生态方面的含水量、耐荫性等 (Hewitt, 1998; Niinemets & Valladares, 2006; Valladares & Niinemets, 2008)。需要根据实际情况选择、收集并分析相关指标, 综合功能性状信息和精确的DNA条形码系统关系, 进行群落水平上的性状系统发育分析。

5.3 系统发育beta多样性

Beta多样性所反映的是物种组成在不同时空尺度上的变化, 描述的是群落间的“不相似性”: 两个群落所处的环境差异越大, 物种组成越复杂, 则其beta多样性就越高; 环境差异越小, 物种组成越简单, 则其beta多样性就越低 (Ives & Helmus, 2010)。Beta多样性之所以成为生态学研究热点, 是因为如果沿着不同环境梯度的群落间差异大 (beta多样性高), 那么就表明有非随机过程在构建这些生态群落。通过解析这些影响因素和过程, 可以进一步探讨生物多样性格局形成和维持的机制, 这正是理论生态学研究的一个重要方向 (Whittaker, 1960; Condit *et al.*, 2002; Graham & Fine, 2008; 陈圣宾等, 2010)。

系统发育beta多样性指的是任何两个地点的生物个体间的系统发育距离 (也即系统树上的分枝长度), 可以被看作是包含 (演化) 时间维度的beta多样性 (Graham & Fine, 2008; Gómez *et al.*, 2010)。通常, 除非两群落间的物种数目和组成完全一样, 使用beta多样性指数与系统发育beta多样性指数会得到不同的结果。例如, 群落A、B、C的物种数均为100, 分别有50(10)、25(8)和25(2)个不同的属(科), 如果用

beta多样性指数进行计算时, 那么群落A与B, 群落A与C之间都将有比较高的beta多样性; 而当用系统发育beta多样性指数进行计算时, 那么群落A与B的beta多样性较低, 而群落A与C之间beta多样性则较高。这是因为运用系统发育beta多样性指数计算群落间的“不相似性”时, 考虑了不同物种间的系统发育关系, 而beta多样性指数则没有考虑。

生态学家们往往基于物种来归类群落内的植物。在野外, 近缘物种往往难以分辨。而基因信息 (通过测序获得研究对象的DNA序列) 就能弥补传统手段的不足, 真实反映现存植物的系统发育关系, 有助于推断随机和非随机过程在调控植物群落结构中的作用。通过比较系统发育beta多样性即不相似性, 就可以分析出不同群落所处的生境类型及其相互作用关系, 结合各自群落内的物种组成, 共同探讨物种构建机制、生物多样性格局等诸多重要生态学问题 (Faith, 1992; Graham & Fine, 2008; Losos, 2008; Faith *et al.*, 2009)。相应的指标有Phylogenetic Sorensen's Index (PhyloSor) (Bryant *et al.*, 2008), UniFrac Index (Lozupone *et al.*, 2006), Rao's Quadratic entropy (Rao, 1982)等。今后, 基于DNA条形码构建的精确群落系统发育关系, 可用于研究不同生物群落的演化历史和功能性状对beta多样性的影响, 不同时空尺度对beta多样性的影响, 还可应用于生物入侵、生物多样性保护等领域 (Faith & Baker, 2006; Faith *et al.*, 2010; 陈圣宾等, 2010)。

致谢: 中科院西双版纳热带植物园曹坤芳研究员、中科院植物研究所马克平研究员以及两位匿名审稿专家对本文初稿修改提供了具有建设性的意见, 特致谢忱!

参考文献

- Ackerly DD, Cornwell WK (2007) A trait-based approach to community assembly: partitioning of species trait values into within- and among-community components. *Ecology Letters*, **10**, 135–145.
- Ackerly DD, Schilck DW, Webb CO (2006) Niche evolution and adaptive radiation: testing the order of trait divergence. *Ecology*, **87**, S50–S61.
- APG III (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG. III. *Botanical Journal of the Linnean Society*, **161**, 105–121.
- Baraloto C, Paine CET, Patino S, Bonal D, Herault B, Chave J

- (2009) Functional trait variation and sampling strategies in species-rich plant communities. *Functional Ecology*, **24**, 208–216.
- Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL (2008) Microbes on mountainsides: contrasting elevational patterns of bacterial and plant diversity. *Proceedings of the National Academy of Sciences, USA*, **105**, 11505–11511.
- Cavender-Bares J, Kozak KH, Fine PVA, Kembel SW (2009) The merging of community ecology and phylogenetic biology. *Ecology Letters*, **12**, 693–715.
- CBOL Plant Working Group (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences, USA*, **106**, 12794–12797.
- Chase JM, Leibold MA (2003) *Ecological Niches: Linking Classical and Contemporary Approaches*. University of Chicago Press, Chicago.
- Chase MW, Cowan RS, Hollingsworth PM, van den Berg C, Madrinan S, Petersen G, Seberg O, Jorgensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M (2007) A proposal for a standardised protocol to barcode all land plants. *Taxon*, **56**, 295–299.
- Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C (2010) Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*, **5**, e8613.
- Chen SB (陈圣宾), Ouyang ZY (欧阳志云), Xu WH (徐卫华), Xiao Y (肖隰) (2010) A review of beta diversity studies. *Biodiversity Science (生物多样性)*, **18**, 323–335. (in Chinese with English abstract)
- Condit R, Pitman N Jr, Leigh EG, Chave J, Terborgh J, Foster RB, Núñez P, Aguilar S, Valencia R, Villa G, Muller-Landau HC, Losos E, Hubbell SP (2002) Beta-diversity in tropical forest trees. *Science*, **295**, 666–669.
- Cubas P (2004) Floral zygomorphy, the recurring evolution of a successful trait. *Bioessays*, **26**, 1175–1184.
- Dick CW, Kress WJ (2009) Dissecting tropical plant diversity with forest plots and a molecular toolkit. *BioScience*, **59**, 745–755.
- Donoghue MJ (2008) A phylogenetic perspective on the distribution of plant diversity. *Proceedings of the National Academy of Sciences, USA*, **105**, 11549–11555.
- Eggleton P, Vane-Wright RI (1994) *Phylogenetics and Ecology*. Academic Press, London.
- Faith DP (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, **61**, 1–10.
- Faith DP, Baker AM (2006) Phylogenetic diversity (PD) and biodiversity conservation: some bioinformatics challenges. *Evolutionary Bioinformatics Online*, **2**, 121–128.
- Faith DP, Ferrier S, Rosauer D (2010) Phylogenetic beta diversity as a tool in biodiversity conservation and monitoring. Powerpoint Presentation at Ecological Society of America Symposium. Pittsburgh, PA.
- Faith DP, Lozupone CA, Nipperess D, Knight R (2009) The cladistic basis for the phylogenetic diversity (PD) measure links evolutionary features to environmental gradients and supports broad applications of microbial ecology's "phylogenetic beta diversity" framework. *International Journal of Molecular Sciences*, **10**, 4723–4741.
- Fang JY (方精云) (2009) Community ecology comes to a new era. *Biodiversity Science (生物多样性)*, **17**, 531–532. (in Chinese)
- Gómez JP, Bravo GA, Brumfield RT, Tello JG, Cadena CD (2010) A phylogenetic approach to disentangling the role of competition and habitat filtering in community assembly of Neotropical forest birds. *Journal of Ecology*, **79**, 1181–1192.
- Gonzalez MA, Baraloto C, Engel J, Mori SA, Pétronelli P, Riéra B, Roger A, Thébaud C, Chave J (2009) Identification of Amazonian trees with DNA barcodes. *PLoS ONE*, **4**, e7483.
- Gonzalez MA, Roger A, Courtois EA, Jabot F, Norden N, Paine CET, Baraloto C, Thébaud C, Chave J (2010) Shifts in species and phylogenetic diversity between sapling and tree communities indicate negative density dependence in a lowland rain forest. *Journal of Ecology*, **98**, 137–146.
- Graham CH, Fine PVA (2008) Phylogenetic beta diversity: linking ecological and evolutionary processes across space in time. *Ecology Letters*, **11**, 1265–1277.
- Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.
- Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN (2006) DNA barcodes distinguish species of tropical Lepidoptera. *Proceedings of the National Academy of Sciences, USA*, **103**, 968–971.
- Harvey PH, Pagel MD (1991) *The Comparative Methods in Evolutionary Biology*. Oxford University Press, Oxford.
- Hebert PDN, Ratnasingham S, de Waard JR (2003) Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **270**, S96–S99.
- Hewitt N (1998) Seed size and shade-tolerance: a comparative analysis of North American temperate trees. *Oecologia*, **114**, 432–440.
- Hubbell SP (2001) *The Unified Neutral Theory of Biodiversity and Biogeography*. Princeton University Press, Princeton.
- Ives AR, Helmus MR (2010) Phylogenetic metrics of community similarity. *The American Naturalist*, **176**, E128–E142.
- Johnson MTJ, Stinchcombe JR (2007) An emerging synthesis between community ecology and evolutionary biology. *Trends in Ecology and Evolution*, **22**, 250–257.
- Kembel SW, Hubbell SP (2006) The phylogenetic structure of a neotropical forest tree community. *Ecology*, **87**, S86–S99.
- Kraft NJB, Cornwell WK, Webb CO, Ackerly DD (2007) Trait

- evolution, community assembly, and the phylogenetic structure of ecological communities. *The American Naturalist*, **170**, 271–283.
- Kraft NJB, Metz MR, Condit RS, Chave J (2010) The relationship between wood density and mortality in a global tropical forest data set. *New Phytologist*, **188**, 1124–1136.
- Kraft NJB, Valencia R, Ackerly DD (2008) Functional traits and niche-based tree community assembly in an Amazonian forest. *Science*, **322**, 580–582.
- Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*, **2**, e508.
- Kress WJ, Erickson DL, Jones FA, Swenson NG, Perez R, Sanjur O, Bermingham E (2009) Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proceedings of the National Academy of Science, USA*, **106**, 18621–18626.
- Kress WJ, Erickson DL, Swenson NG, Thompson J, Uriarte M, Zimmerman JK (2010) Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot. *PLoS ONE*, **5**, e15409.
- Lahaye R, van der Bank M, Bogarin D, Warner J, Pupulin F, Gigot G, Maurin O, Duthoit S, Barraclough TG, Savolainen V (2008) DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences, USA*, **105**, 2923–2928.
- Lomáscolo SB, Levey DJ, Kimball RT, Bolker BM, Albarn HT (2010) Dispersers shape fruit diversity in *Ficus* (Moraceae). *Proceedings of the National Academy of Sciences, USA*, **107**, 14668–14672.
- Losos JB (2008) Phylogenetic niche conservatism, phylogenetic signal and the relationship between phylogenetic relatedness and ecological similarity among species. *Ecology Letters*, **11**, 995–1003.
- Lozupone C, Hamady M, Knight R (2006) UniFrac—an online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics*, **7**, 371.
- Ma KP (马克平) (2008) Large scale permanent plots: important platform for long term research on biodiversity in forest ecosystem. *Journal of Plant Ecology (Chinese version)* (植物生态学报), **32**, 237. (in Chinese)
- Ma KP (马克平) (2011) Assessing progress of biodiversity conservation with monitoring approach. *Biodiversity Science* (生物多样性), **19**, 125–126. (in Chinese)
- Mayfield MM, Boni MF, Ackerly DD (2009) Traits, habitats, and clades: identifying traits of potential importance to environmental filtering. *The American Naturalist*, **174**, E1–E22.
- Niinemets Ü, Valladares F (2006) Tolerance to shade, drought, and waterlogging of temperate northern hemisphere trees and shrubs. *Ecological Monographs*, **76**, 521–547.
- Ning SP (宁淑萍), Yan HF (颜海飞), Hao G (郝刚), Ge XJ (葛学军) (2008) Current advances of DNA barcoding study in plants. *Biodiversity Science* (生物多样性), **16**, 417–425. (in Chinese with English abstract)
- Niu KC (牛克昌), Liu YN (刘怿宁), Shen ZH (沈泽昊), He FL (何芳良), Fang JY (方精云) (2009) Community assembly: the relative importance of neutral theory and niche theory. *Biodiversity Science* (生物多样性), **17**, 579–593. (in Chinese with English abstract)
- Pavoine S, Vela E, Gachet S, Bélair GD, Bonsall MB (2011) Linking patterns in phylogeny, traits, abiotic variables and space: a novel approach to linking environmental filtering and plant community assembly. *Journal of Ecology*, **99**, 165–175.
- Pennisi E (2007) Wanted: a barcode for plants. *Science*, **318**, 190–191.
- Rao CR (1982) Diversity and dissimilarity coefficients: a unified approach. *Theoretical Population Biology*, **21**, 24–43.
- Ren B-Q, Xiang X-G, Chen Z-D (2010) Species identification of *Alnus* (Betulaceae) using nrDNA and cpDNA genetic markers. *Molecular Ecology Resources*, **10**, 594–605.
- Roy S, Tyagi A, Shukla V, Kumar A, Singh UM, Chaudhary LB, Datt B, Bag SK, Singh PK, Nair NK, Husain T, Tuli R (2010) Universal plant DNA barcode loci may not work in complex groups: a case study with Indian *Berberis* species. *PLoS ONE*, **5**, e13674.
- Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics*, **19**, 301–302.
- Silvertown J (2004) Plant coexistence and the niche. *Trends in Ecology and Evolution*, **19**, 605–611.
- Stamatakis A (2006) RA×ML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, **22**, 2688–2690.
- Swenson NG, Enquist BJ (2007) Ecological and evolutionary determinants of a key plant functional trait: wood density and its community-wide variation across latitude and elevation. *American Journal of Botany*, **94**, 451–459.
- Swenson NG, Enquist BJ, Pither J, Thompson J, Zimmerman JK (2006) The problem and promise of scale dependency in community phylogenetics. *Ecology*, **87**, 2418–2424.
- Tilman D (1988) *Plant Strategies and the Dynamics and Structure of Plant Communities*. Princeton University Press, Princeton, NJ.
- Valladares F, Niinemets Ü (2008) Shade tolerance, a key plant feature of complex nature and consequences. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 237–257.
- van de Wiel CCM, van der Schoot J, van Valkenburg JLCH, Duistermaat H, Smulders MJM (2009) DNA barcoding discriminates the noxious invasive plant species, floating pennymort (*Hydrocotyle ranunculoides* L. f.), from non-invasive relatives. *Molecular Ecology Resources*, **9**, 1086–1091.
- Wang X-Q (汪小全), Hong D-Y (洪德元) (1997) Progress in molecular systematics of plants in recent five years. *Acta Phytotaxonomica Sinica* (植物分类学报), **35**, 465–480. (in Chinese with English abstract)

- Chinese with English abstract)
- Webb CO (2000) Exploring the phylogenetic structure of ecological communities: an example for rain forest trees. *The American Naturalist*, **156**, 145–155.
- Webb CO, Ackerly DD, McPeck MA, Donoghue MJ (2002) Phylogenies and community ecology. *Annual Review of Ecology and Systematics*, **33**, 475–505.
- Webb CO, Losos JB, Agrawal AA (2006) Integrating phylogenies into community ecology. *Ecology*, **87**, S1–S2.
- Whittaker RH (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, **30**, 279–338.
- Wiens JJ (2003) Missing data, incomplete taxa, and phylogenetic accuracy. *Systematic Biology*, **52**, 528–538.
- Wright IJ, Ackerly DD, Bongers F, Harms KE, Ibarra-Manriquez G, Martinez-Ramos M, Mazer SJ, Muller-Landau HC, Paz H, Pitman NCA, Poorter L, Silman MR, Vriesendorp CF, Webb CO, Westoby M, Wright SJ (2007) Relationships among ecologically important dimensions of plant trait variation in seven Neotropical forests. *Annals of Botany*, **99**, 1003–1015.
- Wright SJ, Kitajima K, Kraft N, Reich P, Wright I, Bunker D, Condit R, Dalling J, Davies S, Diaz S, Engelbrecht B, Harms K, Hubbell S, Marks C, Ruiz-Jaen M, Salvador C, Zanne A (2010) Functional traits and the growth-mortality tradeoff in tropical trees. *Ecology*, **91**, 3664–3674.
- Zhong Y (钟扬), Shi SH (施苏华), Ren WW (任文伟) (2004) Molecular phylogeny analysis. In: *Plant Life History Evolution and Reproductive Ecology* (植物生活史与繁殖生态学) (ed. Zhang DY (张大勇)). Science Press, Beijing. (in Chinese)

(责任编辑: 曹坤芳 责任编辑: 时意专)

附录I 植物DNA条形码应用于群落系统发育结构研究: 研究方案及操作指南

(<http://www.biodiversity-science.net/CN/article/downloadArticleFile.do?attachType=PDF&id=9528>)

植物DNA条形码应用于群落系统发育结构研究: 研究方案及操作指南

1 总DNA提取、PCR扩增及测序

用CTAB法或试剂盒提取植物叶片组织的DNA,除实验部分的DNA之外,其余的均保存在-80℃冰柜作为备用。使用国际条形码联盟公布的通用引物进行PCR扩增和测序(引物见附表1)。各引物的扩增程序为:(1)*rbcL*与*trnH-psbA*片段可共用一个程序:(95℃ 3 min, (95℃ 30 sec, 53℃ 45 sec, 72℃ 1 min) ×34 cycles, 72℃ 7 min); (2) *matK*片段为:(95℃ 3 min (95℃ 30 sec, 51℃ 45 sec, 72℃ 1.5 min) ×38 cycles, 72℃ 7 min)。提供PCR反应体系和测序引物,由测序公司完成全部序列工作。*rbcL*片段长约540 bp,一个反应测完;*matK*片段长约900 bp,需测序正反两个反应;*trnH-psbA*片段长度范围300–800 bp,一个反应可以测完(单独*trnH*,或者单独*psbA*端都可以),若存在poly A/T结构(也即出现连续重复多个碱基A或者T,数目可达8个甚至更多),则需要双向测序。

2 序列编辑、比对及BLAST

对于一个反应可以完成的个体,使用Chromas软件(<http://www.technelysium.com.au/chromas.html>)检验序列测序质量和长度,编辑后保存为Fasta格式的文档。对于两个反应的个体,使用DNASTar软件(<http://www.dnastar.com/>)的Seqman程序,选定一个方向作标准,双向校准序列,编辑后保存为Fasta格式的文档(注意小写字母的碱基,往往是错误的碱基插入,应去除;而重复插入若是3的倍数,则可不必要去除,因为不影响最终结果)。

用MEGA软件(<http://www.megasoftware.net/>)打开经过CLUSALx软件(<http://www.clustal.org/>)运行后得到的aln文件,转换成Fasta格式,随后就可以在SE-AL v2.0a 11 Carbon软件(<http://tree.bio.ed.ac.uk/software/seal/>)上分析序列。

(1) 对于 *matK* 序列,使用 transalign 软件(<http://life.anu.edu.au/molecular/software/transalign/>)通过反向翻译(Backtranslation,其作用是纠正可能的测序错误,并得到完全正确的*matK*序列)来执行比对步骤。反向翻译步骤是:(i)把核苷酸转化成氨基酸序列(通过NIH密码子表格11来获得质体基因组);(ii)在CLUSAL W软件(<http://www.clustal.org/>)上用氨基酸序列执行比对,然后反向翻译成核苷酸序列,以比对过的Fasta格式输出;(iii)全部的*matK*序列以同样的方式同时进行相互间比对。

(2)对于*trnH-psbA*序列,其Fasta文件按照科或目的模块进行分割。具体做法是:当样地中一个科只有一个物种,那么该个体就与同目的其他科放在一起进行比对;当一个目中只有一个物种时,这个物种的*trnH-psbA*序列不参加系统发育比对。将全部模块的序列集合在一起,使用Muscle软件(<http://www.drive5.com/muscle/>)的默认参数进行比对。比对过的*trnH-psbA*序列数据汇集起来,与*rbcL*序列结合,共同组成一个超级矩阵格式放到一个nexus文件。

(3)对于多基因nexus文件的数据,在MacClade4(<http://mesquiteproject.org>)中进行处理:先导入*rbcL*序列,随后在*rbcL*的队列末端,导入其他引物的文件(如*matK*序列)或者像*trnH-psbA*这种引物的巢式模块数集(图1)。此外,其他一些系统发育分析方法可参考最新版本的系统发育研究手册(Lemey *et al.*, 2009)。

为了保证测序所得序列不会发生误差(人为因素导致的顺序紊乱,或者测序污染而得到错误的序列),对每一条序列在NCBI(<http://blast.ncbi.nlm.nih.gov/Blast.cgi>)中进行BLAST搜索(Basic Local Alignment Search Tool)。可依次对3个片段的序列进行逐条搜索,并记录最大匹配率。

3 在CIPRES上运行系统发育树的重建

3.1 RA×ML树

(1) 打开 nex 文件，在 TextWrangler 软件 (<http://www.barebones.com/products/textwrangler/>) 中改成Windows所需要的文件格式。

(2)在DnaSP软件(<http://www.ub.edu/dnasp/>)上生成phy文件。

(3)在CIPRES上运行前，注意添加外类群，计算Bootstrap values时设置高级参数。

(4)得到besttree、parsimony以及100(200或500)次重复的Bootstrap trees。

3.2 MP树

(1)准备好nex文件，导入CIPRES data中。

(2)选择PAUP程序，且在高级设置中取200–250个ratchets，运行后输出结果，保存为nex格式。

(3)在TreeView软件([http://taxonomy.zoology.gla.](http://taxonomy.zoology.gla.ac.uk/rod/treeview.html)

[ac.uk/rod/treeview.html](http://taxonomy.zoology.gla.ac.uk/rod/treeview.html)) 中打开，保存为 phy 和 newick格式后，即可在MEGA中操作。

(4)将得到的ML和MP树，使用Mesquite软件 (<http://mesquiteproject.org/mesquite/mesquite.html>)与 APG III系统树比较拓扑结构。

4 在Phylocom软件上运算系统发育结构数据(以NRI/NTI为例)

(1)将准备好的phylo/sample/traits三个文件导入 phylocom的dos_version中。

(2)打开软件phylocom批处理文件。

(3)在dos界面输入phylocom，回车。

(4)输入phylocom comstruct -m 3>output.txt，得到计算结果。

(5)在dos_version中找到一个名为output的文本文件，打开后复制到Excel中，保存。所获得的NRI和NTI数值即可用于分析群落的系统发育结构。

附表1 本研究所使用的条形码片段引物对

Appendix 1 Primer pairs for barcode loci *rbcL*, *matK*, and *trnH-psbA*

标记 Marker	引物 Primer	序列 Sequence (5'– 3')	长度 Size (bp)	参考文献 Reference
<i>rbcL</i>	SI_F	ATGTCACCACAAACAGAGACTAAAGC	554	Kress & Erickson (2007)
	SI_R	GTYAAATCAAGTCCACCYCG		
<i>matK</i>	KIM_3F	CGTACAGTACTTTTGTGTTTACGAG	Average 850	Kim KJ, unpublished primers
	KIM_1R	ACCCAGTCCATCTGGAATCTTGGTTC		
	KIM_1F	AATATCCAAATACCAAATCC	Average 850	Kim KJ, unpublished primers
	KIM_1R	ACCCAGTCCATCTGGAATCTTGGTTC		
	2.1a-f	ATCCATCTGGAAATCTTAGTTC	Average 890	
5r	GTTCTAGCACAAAGAAAGTCG		www.kew.org/barcoding/protocols.html	
<i>trnH-psbA</i>	<i>psbA</i>	GTTATGCATGAACGTAATGCTC	280–750, aver- age 470	Sang <i>et al.</i> , 1997
	<i>trnH</i>	CGCGCATGGTGGATTCAATCC		Tate & Simpson (2003)

参考文献

- Kress WJ, Erickson DL (2007) A two-locus global DNA barcode for land plants: the coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*, 2: e508.
- Kim KJ. School of Life Sciences and Biotechnology, Korea University (http://www.korea.edu/academics/01co_04.php), Seoul, Korea, unpublished primers.
- Sang T, Crawford DJ, Stuessy TF (1997) Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany*, **84**, 1120–1136.
- Tate JA, Simpson BB (2003) Paraphyly of *Tarasa* (Malvaceae) and diverse origins of the polyploid species. *Systematic Biology*, **28**, 723–737.