



# 植物 DNA 条形码技术应用于大样地群落学研究

裴男才<sup>①</sup>, 米湘成<sup>②</sup>, 陈步峰<sup>①\*</sup>

① 中国林业科学研究院热带林业研究所, 国家林业局珠江三角洲森林生态系统定位研究站, 广州 510520;

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

\* 联系人, E-mail: zsjcsdwcbf@126.com

2014-04-04 收稿, 2014-06-25 接受, 2014-07-08 网络版发表

国家自然科学基金(31200471)、国家林业公益性行业科研专项(20140430105)和中国林业科学院热带林业研究基本科研业务费专项(RITFYWZX201208)资助

**摘要** DNA 条形码技术在生物分类学、进化生态学、分子系统学、生物多样性科学、医学检疫等学科和领域发挥着重要作用. 本文梳理了近些年国内外以森林大样地为平台, 利用植物 DNA 条形码(即叶绿体基因片段 *rbcL*, *matK* 和 *trnH-psbA*) 在群落水平研究获得的若干重要进展. 围绕植物 DNA 条形码技术如何应用于大样地群落学研究这一主线, 从基本原理出发, 以研究案例为基础进行了相应的评述, 提炼出以下 4 方面: (1) 鉴别群落水平上的植物材料时, 利用多个 DNA 条形码片段的组合能达到较高的物种鉴别率, 而单个片段效果较好的为 *trnH-psbA*, 其次是 *matK* 和 *rbcL*; (2) 构建大样地植物群落系统发育关系时, 搭配使用不同进化速率的条形码片段, 以超级矩阵(Supermatrix)方法排列 DNA 序列, 在系统发育树构建软件上可获取较好的分支精度并得到较高的节点支持率; (3) DNA 条形码系统发育关系与群落生态学研究结合时, 能增强对群落共存机制和群落构建方式的探索和阐释能力; (4) DNA 条形码系统发育关系应用于生物多样性指数计算时, 能增强多样性指数的分辨能力和评估手段. 植物 DNA 条形码技术已在类群和群落水平上应用并得到较好的效果; 建议与传统的 Phylomatic 等方法互补优势, 在原有基础上进行适当的完善, 如优化取样策略、增加具有更多信息量的核基因片段 ITS 等, 提升 DNA 条形码技术的整体性能, 使该技术在区域和全球水平上开展群落学研究时也能较好地应用.

## 关键词

生命条形码  
生物群落  
森林生物多样性  
监测网络  
整合生物学

DNA 条形码(DNA barcoding)类似于商品零售业中的物品自动扫描条码, 其主要理念是通过获取生物组织中一段或几段较短的 DNA 序列对物种进行快速而准确鉴定的技术. 随着分子生物学和生物信息学技术的进步, 基于 DNA 条形码技术进行生物鉴定和分类, 并以此为基础进行关联学科(如生命科学、法医学、流行病学、医药和食品质量控制等领域)的研究和应用得到了快速发展. 最具代表性的是国际生命条形码研究计划(iBOL: The International

Barcode of Life Project)的正式提出和实施, 被认为是继基因组计划之后, 又一个生物大科学计划而受到空前的关注<sup>[1-5]</sup>.

加拿大生物学家 Paul Hebert 等人<sup>[6,7]</sup>于 2003 年发表了第一批动物 DNA 条形码研究案例, 利用线粒体基因细胞色素 C 氧化酶(*COI*)对动物物种进行鉴定, 并提出了 DNA 条形码概念. 此后, DNA 条形码技术也拓展至植物界和微生物界的应用<sup>[8]</sup>. iBOL 是当前世界科技领域所启动的最大规模的生物多样性基因

**引用格式:** 裴男才, 米湘成, 陈步峰. 植物 DNA 条形码技术应用于大样地群落学研究. 科学通报, 2014, 59: 2333-2341

Pei N C, Mi X C, Chen B F. Integration of plant DNA barcoding technology into community studies on forest dynamics plots (in Chinese). Chin Sci Bull (Chin Ver), 2014, 59: 2333-2341, doi: 10.1360/N972014-00033

组项目. 来自 25 个国家成百上千的生物多样性研究科学家、基因组专家、技术人员和伦理观察者一起工作, 旨在构建一个信息量高度丰富的、标准化的 DNA 条形码参考文库, 为创立全部多细胞生命的 DNA 识别系统奠定基础. 在第一阶段(2010~2015 年), iBOL 各合作方将对 50 万物种、500 万份标本进行 DNA 条形码分析. 在构建条形码文库过程中, iBOL 参与者还将搭建各类基础设施, 将其用于物种保存、生态系统监测、法医鉴定、农业害虫和入侵物种的控制等实践工作(<http://www.barcodeoflife.org/content/about/what-ibol>; 查询于 2014 年 3 月 27 日). 国际生命条形码联盟(CBOL: Consortium for the Barcode of Life, <http://www.barcodeoflife.org/>)是全球生物 DNA 条形码研究的协调机构, 在加拿大、中国、欧盟和美国等主要生物多样性研究地区设有中心节点, 50 多个国家和地区的 200 多个组织和机构分布于欧洲、北美洲、南美洲、大洋洲、亚洲、非洲(<http://www.barcodeoflife.org/content/about/member-organizations>; 查询于 2014 年 3 月 27 日). 首届国际生命条形码大会(International Barcode of Life Conference)于 2005 年在英国伦敦召开, 此后国际生命条形码大会成为每两年举办一次的系列国际大会, 分别于 2007, 2009, 2011, 2013 年在中国台北、墨西哥首都墨西哥城、澳大利亚阿德莱德和我国云南省昆明成功举办, 累计吸引了全球近百个国家和地区近 2000 人次的专家学者参加, 显示了生命条形码研究在世界范围内的活跃程度.

植物 DNA 条形码研究在 2005 年才有相关学术论文发表<sup>[9,10]</sup>. 与动物相比, 植物线粒体基因组进化速率较慢, *COI* 片段不适用于植物, 而寻找或筛选植物通用 DNA 条形码片段过程较长, 难以筛选出一个像在动物界较为通用的条形码片段 *COI*<sup>[9]</sup>. 2009 年国际生命条形码联盟植物工作组推荐 *rbcL+matK* 组合作为陆地植物的核心条形码片段, 而建议 *trnH-psbA* 和核基因组 ITS 作为候选片段继续进行评估<sup>[11,12]</sup>, 以满足对包括近缘类群和其他疑难类群在内的植物材料进行鉴别的需求. ITS 引物通用性差、序列长度变异大、poly 结构较多、存在多拷贝等缺点, 降低了其作为通用植物 DNA 条形码片段的可应用性<sup>[10,12-14]</sup>. “中国植物条形码研究团队”对主要来自中国的种子植物 75 科 141 属 1757 种共约 6286 个样本进行综合评价, 发现 ITS 引物通用性在被子植物中可达 79%,

增加该片段后能将物种分辨率从 50%(*rbcL+matK*)提高至 70%~79%, 因此建议将 ITS 作为种子植物的核心 DNA 条形码<sup>[11]</sup>. 目前 *rbcL*, *matK* 和 *trnH-psbA* 已广泛用于植物大样地群落学研究中, ITS 也开始被使用. 植物 DNA 条形码研究已有比较成熟的工作流程、方案和科学议题, 在专科专属等类群水平上发展迅速, 涵盖了各大重要的陆生植物类群, 在植物物种水平上的分类和常规鉴定具有重要应用价值<sup>[11,15-21]</sup>, 其中具有代表性的如中国植物 DNA 条形码研究团队于 2011 年 5 月在 *Journal of Systematics and Evolution* 杂志上发表了“Plant DNA barcoding in China”专辑; *Molecular Ecology Resources* 杂志也刊登了较多的 DNA 条形码的研究结果. 基于植物 DNA 条形码数据、地理信息数据和计算机信息技术等元素的新一代植物志(iFlora), 有望带动国际国内植物学、生态学和生物多样性等学科的新一轮发展<sup>[22]</sup>. 此外, 利用高通量测序技术获得条形码基因扩增子序列的 DNA metabarcoding 技术也有望在环境监测、资源管理和生物多样性评估等方面得到应用<sup>[23-27]</sup>.

大型森林动态监测样地(forest dynamics plot, 即大样地)是生物多样性和群落学研究的重要平台<sup>[28]</sup>. 植物 DNA 条形码研究在群落维度上也紧跟发展步伐, 特别是在以大样地为平台开展的研究中有较大的现实需求, 但目前已有一些研究暂未较系统地叙述植物 DNA 条形码与群落学研究的有机结合. 本文以近些年国际和国内已开展的若干重要研究工作为基础, 结合作者的一些理解和思考, 着重从鉴别植物物种、构建群落系统发育关系、揭示群落系统发育结构以及评估生物多样性指数等 4 个方面展开, 探讨如何将植物 DNA 条形码技术应用于大样地群落学研究, 试图为该方面的研究和应用提供参考.

## 1 鉴别群落内植物物种

基本原理: DNA 序列包含着物种的基本遗传信息, 特别是编码基因决定着生物体内蛋白质的构成. 特定碱基组合体现出各生物组织的特异性; 而不同的碱基组合, 维持着大千世界生物物种的多样性<sup>[29,30]</sup>. 准确的物种鉴别对于群落生物学研究极为重要; DNA 条形码的出现, 在一定程度上可弥补分类学家的不足而造成的诸多不便, 为解决物种水平上的准确鉴定工作提供了理论支撑和技术保障. 生命条形码计划最初的设想是通过使用一段短的 DNA 序

列达到物种水平上的鉴别<sup>[7]</sup>。在标本不全或残缺的情况下, DNA 条形码技术可进行快速且相对准确的物种鉴定。用于物种鉴定的方法: (i) BLASTn 方法, 即通过 GenBank 在线数据库或者利用本地物种构建条形码参考数据库, 按不同条形码片段逐条或者打包搜索, 并记录到科、属和种水平上的最大匹配率<sup>[31]</sup>; (ii) TaxonDNA 方法, 需将所要鉴定的 DNA 条形码序列汇总为一个总的文库, 按照 TaxonDNA 软件的格式进行数据处理, 然后计算出研究对象在该文库中相应的匹配率<sup>[32]</sup>。此外, 还有基于 DNA 序列信息构建系统发育进化树的方法进行物种鉴定, 如距离法、最大简约法、极大似然法及贝斯方法等。

研究案例: 在热带森林区域, Kress 等人<sup>[33]</sup>率先利用 3 个条形码片段(*rbcL*, *matK* 和 *trnH-psbA*), 在本地条形码数据库中采用 BLASTn 方法对巴拿马 Barro Colorado Island (BCI) 50 hm<sup>2</sup> 热带湿润平原雨林样地上的 296 个植物物种进行搜索和鉴别。发现使用单个条形码片段时, *trnH-psbA* 的综合成功鉴别率(即测序成功率和物种识别成功率之乘积, 简称为“鉴别率”;下同)最高(90%), 其次是 *rbcL*(70%)和 *matK*(69%); 使用 *matK+rbcL+trnH-psbA* 的三片段组合在物种水平鉴别率高达 98%。Gonzalez 等人<sup>[34]</sup>在(法属)圭亚那 Nouragues 亚马孙热带雨林中的 2 个 1 hm<sup>2</sup> 样地, 检测了 8 个候选植物 DNA 片段(即 *rbcLa*, *rpoCl*, *rpoB*, *matK*, *ycf5*, *trnL*, *trnH-psbA* 和 ITS)在辨别 252 种幼树上的表现。结果显示, 当使用单个条形码片段时, *trnH-psbA* 的鉴别率最高(57%), 其次是 *rbcL*(53%)和 *matK*(41%); 使用三片段组合时鉴别率约 70%。在亚热带森林区域, Kress 等人<sup>[35]</sup>使用与 BCI 同样的方法, 对波多黎各 Luquillo 16 hm<sup>2</sup> 样地上 143 个物种进行鉴别。结果显示, 当使用单个条形码片段时, *trnH-psbA* 的鉴别率最高(90.4%), 其次是 *rbcL*(84.5%)和 *matK*(70.4%); 使用三片段组合时鉴别率达 93%。我们对我国鼎湖山 20 hm<sup>2</sup> 大样地 183 个植物物种进行物种水平上的鉴别, 发现使用单个条形码片段时, *trnH-psbA* 的鉴别率最高(75%), 其次是 *matK*(70%)和 *rbcL*(56%); 使用三片段组合时鉴别率在 87%以上<sup>[36]</sup>。卢孟孟等人<sup>[37]</sup>对云南哀牢山自然保护区内 51 科 111 属中 204 个树种的 525 个乔木个体开展了 DNA 条形码对鉴定亚热带森林树种有效性的评估工作, 发现在物种和属水平的鉴别成功率上, 单片段 ITS 分别为

68%~81%和 99%~100%, 核心条形码 *rbcL+matK* 组合的成功率是 53%~60%和 87%~91%; *rbcL+matK+trnH-psbA*+ITS 联合时可以成功鉴别 75%~80%的全部乔木物种。此外, 他们还发现 ITS 片段在亚热带森林部分重要树种类群(樟科和壳斗科等)中的测序成功率较差, 因此建议对这些植物类群采用 *trnH-psbA* 作为 DNA 条形码是一个更好的选择。

总的来看, 在热带地区和亚热带地区的森林样地中, 各片段准确识别植物物种的平均鉴别率分别为 *rbcL*(66%, 74.8%), *trnH-psbA*(79.5%, 87.5%)和 *matK*(80%, 89.5%), 说明植物 DNA 条形码片段在局部群落表现较为理想。亚热带森林中的鉴别率略高于热带森林, 这可能与亚热带森林局部群落内生长的植物同属种较少有关。3 个 DNA 条形码片段组合在热带和亚热带森林样地中鉴别率分别可达 84%和 90%<sup>[36]</sup>。增加 ITS 片段可能有助于鉴别成功率的提高, 但目前缺乏在大样地群落水平上的研究报道, 有必要对 ITS 片段在群落水平上作更多的案例分析。今后, 可依据现有 4 个 DNA 条形码片段的物种鉴别成功率, 确定在不同类型的森林样地使用相适合的条形码组合, 在达到物种分辨的同时, 尽量控制成本。将植物 DNA 条形码方法与形态分类特征、细胞学信息以及化石证据等相结合, 还能进一步提高未知植物样本的正确物种鉴别率。这在进化生态学和保护生物学领域体现得更为明显, 有利于增进了解现存生物多样性现状和分布格局, 可为制定更符合实际情况的生物保护策略提供重要参考依据<sup>[1,17,38-40]</sup>。

## 2 构建植物群落系统发育关系

基本原理: 群落内各物种间的差异, 可表现为 DNA 序列上的碱基变化。利用各物种 DNA 条形码的序列信息, 通过对不同进化速率的条形码片段进行序列排列以生成矩阵<sup>[41]</sup>, 采用分子系统学常用的系统发育树构建方法(如邻接法<sup>[42]</sup>、最大简约法<sup>[43]</sup>、最大似然法<sup>[44]</sup>等), 可快速且相对准确地获得特定群落的系统发育树。

研究案例: 在 BCI 大样地, Kress 等人<sup>[33]</sup>利用 DNA 条形码序列, 采取 3 个条形码片段组合的超级矩阵方法(Supermatrix; 具体为: 片段 *rbcL* 序列在前, 确保系统发育树的大框架正确, 能解决目和科分类阶元上的节点关系; 随后是 *matK* 和以目或者大科为单元的 *trnH-psbA* 序列, 试图在属和物种水平上将研

究材料定位到合适的系统发育位置), 构建出精确的植物群落系统发育关系(内含 23 目 57 科 181 属 296 个物种), 且有 97.5% 的分支节点支持率在 50% 以上; 相比来看, 仅用基于 Angiosperm Phylogeny Group (APG) 的 Phylomatic<sup>[45]</sup> 在线建树软件获得系统发育树上, 只有 48.4% 的分支节点支持率在 50% 以上. 而在 Luquillo 大样地, Kress 等人<sup>[35]</sup> 使用与 BCI 同样的方法, 但辅以最新版的 APG 在线数据库为框架, 以 DNA 条形码序列构建目与科分类阶元上的系统发育树, 同样构建出较为精确的植物群落系统发育关系(24 目 50 科 108 属 143 个物种), 且 90.8% 的分支节点支持率在 50% 以上; 相比来看, 仅用 APG 在线软件获得的 Phylomatic 系统发育树上, 只有 69.7% 的分支节点支持率在 50% 以上. 在鼎湖山大样地, 我们利用 DNA 条形码序列, 构建出了精确的植物群落系统发育关系(24 目 51 科 110 属 183 个物种), 有 99.4% 的分支节点支持率在 50% 以上; 相比来看, 仅用 APG 在线软件获得的 Phylomatic 系统发育树上, 只有 36.8% 的分支节点支持率在 50% 以上<sup>[46]</sup>.

BCI 大样地的例子说明, 在物种丰富度高、亲缘关系复杂的局部植物群落中进行充分样本采集时, 只需利用 DNA 条形码序列构建的群落系统发育关系与 APG 的拓扑结构较为一致, 各节点的支持率也高; 这种方法称为不受约束的建树方法(An unconstrained tree). 而 Luquillo 大样地的例子说明, 在物种丰富度较低、亲缘关系较简单的局部植物群落中对有限数量的物种进行样本采集时, 可借用 APG 系统的大框架, 然后在目水平的分支利用 DNA 条形码序列进行建树, 也能与 APG 较为一致; 这种方法称为受约束的建树方法(A constrained tree). 鼎湖山大样地的情况与 BCI 和 Luquillo 大样地有相同之处, 但也不完全一致: 该植物群落中物种丰富度较低、亲缘关系较简单, 但只需利用 DNA 条形码序列即可构建出较为精确的群落系统发育关系, 各主要被子植物分支与 APG 系统保持较好的一致性, 为植物群落系统发育关系构建提供了新的研究案例<sup>[46]</sup>.

生命之树的构建是一项长期而艰巨的工作, 需要不断创新研究方法, 吸取有益工作理念, 收集补充更多材料和信息, 开展整合的系统发育学研究<sup>[47]</sup>. DNA 条形码方法获得的系统发育进化树, 由于各片段受自身所含有的遗传信息量限制, 在辨别近缘类群(特别是物种数量较多的大属)时有一定的困难. 但

在群落水平上, 由于各植物材料的系统发育关系界线较为明显, 使得 DNA 条形码技术在构建群落系统发育关系时能发挥出优势, 便于跨区域多群落间的比较研究. 这是 DNA 条形码技术在该研究领域得以快速应用的关键所在; 今后可在近缘类群比重较高的植物群落(如热带-亚热带区域)中增加条形码片段的数量, 而在近缘类群比重较低的植物群落(如温带区域)中减少条形码片段的数量.

### 3 揭示群落系统发育结构分布格局

基本原理: 相对于较粗略的系统发育关系(如基于在线数据库获取的 Phylomatic 系统发育关系: <http://phylodiversity.net/phyloomatic/>), DNA 条形码系统发育关系(具有准确的二歧分支拓扑结构和高的节点支持率)在探讨群落共存机制和群落构建方式等方面能提供更为可靠的预测和评估<sup>[48,49]</sup>. 因此, 在相应的群落系统发育分析中应考虑增加物种水平上的系统发育信息, 而不仅限于属与科等更高分类阶元的粗略信息.

研究案例: 在 BCI 大样地, Kress 等人<sup>[33]</sup> 通过使用 DNA 条形码序列构建的精确分子系统发育关系, 发现样地内 7 种生境类型中的 5 种呈现出显著的系统发育结构, 这比此前使用基于 APG 在线数据库 Phylomatic 方法的研究更前进了一步, 表明精确的系统发育末端节点在探索群落系统发育结构时发挥了至关重要的作用, 且能减少发生统计误差的概率<sup>[33]</sup>. 在西双版纳 20 hm<sup>2</sup> 热带季节雨林大样地, Yang 等人<sup>[50]</sup> 获取了 400 多个树种 DNA 条形码系统发育关系, 在 3 个树龄径级、6 个空间尺度、6 种生境类型以及 10 个植物功能性状上探讨群落共存机制, 结果支持物种共存的确定性模型而非中性模型, 同时支持在较大的空间尺度上生境过滤机制作用较大, 而较小的空间尺度种间竞争作用较为显著. 在鼎湖山大样地, 我们分别用植物 DNA 条形码和 Phylomatic 方法构建了鼎湖山大样地植物群落系统发育关系, 两者均发现山谷和低坡生境为系统发育聚集分布格局, 表明近缘物种共存于这些低海拔生境, 生境过滤可能起主导作用; 两者均发现高坡和山脊生境为系统发育扩散分布, 表明远缘物种共存于这些高海拔生境, 竞争排斥可能起主导作用<sup>[51]</sup>. 对于高谷生境, Phylomatic 方法得到的结果为系统发育扩散分布; 而条形码方法得到的结果为系统发育随机分布, 这表明与系统发育有关的作用可能在这种生境类型下不

起作用或者作用不明显. 在古田山 24 hm<sup>2</sup> 大样地, 曹科等人<sup>[52]</sup>以 156 种木本植物为对象, 利用条形码方法获取各物种的 DNA 序列信息, 用  $K$  值法检验了叶片氮含量、磷含量、叶面积、木质密度、比叶面积和种子重量 6 种功能性状的系统发育信号, 并运用系统独立对比方法分析了性状与物种多度的关系. 结果表明, 6 种功能性状都表现出较强的系统发育信号, 可能受到该样地木本植物系统发育历史的影响; 前 5 种性状与多度显著相关, 但种子重量与物种多度未检测到显著的相关性, 表明群落内不同物种的资源获取方式可能影响群落的结构. 同在古田山大样地, Liu 等人<sup>[53]</sup>提取了比叶面积、叶面积、种子重量、树木密度、最大树高等 5 种植物性状数据, 与环境因子、空间因子、系统发育因子进行独立和交叉分析. 结果表明大部分的性状表现为聚集分布结构, 可能受到非随机过滤作用的影响. 生境过滤作用在形成群落中的重要性由方差分解的结果得到验证. 环境因子和空间因子在解释性状分布中有较大作用, 说明扩散限制是古田山群落共存的重要决定因子. 但该研究未能检测到系统发育进化关系在性状分布中的显著作用, 推测单纯依赖系统发育结构难以解释基于性状分析的群落物种共存方式. 对于植物性状空间分布中未被解释的方差, 作者认为可能是由随机过程、或是性状和环境因子选择的偏差而造成.

Swenson 等人<sup>[48]</sup>通过模型模拟和运算推导, 认为对群落系统发育结构更容易受系统发育树大框架的精确程度影响, 其次也与末端进化分支是否得到理想的二歧结构有关. 条形码方法和 PhyloMatic 方法所构建的群落系统发育关系, 在大的框架水平上基本一致, 但前者在末端进化分支的解决程度上优于后者. BCI 和鼎湖山森林大样地的研究结果显示基于条形码方法的系统发育指数(如 NRI 和 NTD)数值倾向大于由 PhyloMatic 方法获得的数值, 这可能与使用 3 个 DNA 条形码序列构建的条形码系统发育关系比 PhyloMatic 系统发育关系更精确有关. DNA 条形码方法会在呈现系统发育聚集分布的群落或者生境中表现得更显著, 而在系统发育扩散分布的群落或者生境中表现得更不显著. 而古田山大样地的研究结果提醒我们, 除了系统发育因子, 还需要充分考虑功能性状、环境因子以及空间因素对群落系统发育结构的影响. 系统发育信息增进了我们对现存森林群落结构和功能的了解, 标准化的 DNA 条形码方法为区域

和全球尺度上的比较生态学研究提供了快捷、可行的途径.

#### 4 评估生物多样性指数

基本原理: 基于标准化的植物 DNA 条形码方法而获得的系统发育关系, 可便捷地用于计算生物多样性指数, 并可供不同区域多样性指数的比较分析和评估工作, 丰富当前生物多样性研究的内涵.

研究案例: 在 Nouragues 森林样地, Gonzalez 等人<sup>[54]</sup>获取不同精确程度的群落系统发育关系, 具体为: (1) 利用 *rbcL* 片段获取条形码群落系统发育关系; (2) 利用 APG 在线数据库获取该样地的群落系统发育关系; (3) 综合条形码和 APG 的方法获取该样地的群落系统发育关系. 结果显示, 描述系统发育多样性指数的 Faith's 指数(即系统发育丰富度)和 Rao's 指数(即系统发育均一度)在不同年龄径级层次上, 发生了不同程度的变化, 暗示着该森林样地群落格局可能不受系统发育约束. 在古田山大样地森林群落, 基于植物 DNA 条形码方法, Feng 等人<sup>[55]</sup>对具有不同科学意义的 9 个系统发育  $\beta$  多样性指数( $D_{pw}$ ,  $D_{nn}$ , PhyloSor, UniFrac, Rao's  $D$ , Rao's  $H$ ,  $\Pi_{ST}$ ,  $P_{ST}$  以及 PCD)进行比较后发现: (1) 多度是否加权对同一个指数的影响在小尺度上并不明显; 随着尺度的增大, 多度加权与否的形式之间的差别才逐渐显现出来; (2)  $P_{ST}$  指数和 Rao's  $H$  指数在较大尺度上比其他指数能够更好地反映生境和空间距离的衰减效应; 随着尺度的减小,  $AW-D_{nn}$  指数和  $D_{nn}$  指数能够更好地反映生境和空间距离的衰减效应; (3) 物种多度加权处理的  $D_{pw}$  指数与 Rao's 二次熵的 Rao's  $D$  定义相似; PhyloSor 指数与 UniFrac 指数生态学含义相似, 具有较高的相关性;  $D_{nn}$  指数与 PhyloSor 指数,  $D_{nn}$  指数与 UniFrac 指数的生态学意义虽有差别, 但其相关性较高. 饶米德等人<sup>[56]</sup>通过模型模拟不同生态学过程下的群落结构, 分析了古田山大样地不同空间尺度下物种和系统发育  $\beta$  多样性指数的距离衰减效应. 研究发现: (1) 该样地群落的物种及系统发育多样性, 存在明显的距离衰减格局; (2) 生境过滤和扩散限制共同作用, 可以更好地解释物种  $\beta$  多样性维持机制; (3)  $D_{nn}$  指数的结果表明生境过滤和扩散限制共同维持着古田山样地群落系统发育  $\beta$  多样性格局; 而  $D_{pw}$  指数在小尺度上(如 20 m×20 m)生境过滤和扩散限制共同作用于群落系统发育  $\beta$  多样性, 但在大尺度上(如 40 m×40 m

和 50 m×50 m), 生境过滤逐渐成为解释群落系统发育结构的主要过程. 古田山大样地群落的系统发育  $\beta$  多样性与物种  $\beta$  多样性的格局表明, 生境过滤和扩散限制起着重要的维持作用. 在热带至温带森林区域, 以分布于亚洲和美洲的 6 个森林大样地为研究对象 (即中国的古田山和鼎湖山, 波多黎各的 Luquillo, 巴拿马的 BCI, 美国的 Wabikon Lake 和 Smithsonian Conservation Biology Institute), Swenson 等人<sup>[57]</sup>利用植物 DNA 条形码(*rbcL*, *matK* 和 *trnH-psbA*)构建该 6 个大样地的群落系统发育关系, 将系统发育和功能  $\alpha$  多样性(测量指数有 NRI, NTI, SES PW 和 SES NN), 与系统发育和功能  $\beta$  多样性(测量指数有 S.E.S.  $D_{\text{m}}$ ' 和 S.E.S  $D_{\text{pw}}$ ')相互关联, 在 20 m×20 m, 40 m×40 m 以及 100 m×100 m 这 3 个空间尺度上计算各指数的数值. 研究发现: (1) 非生物环境过滤因素在构建局部群落和支配群落组合的空间转化中发挥作用; (2) 本研究的 6 个森林大样地中, 系统发育  $\alpha$  和  $\beta$  多样性计量结果与功能  $\alpha$  和  $\beta$  多样性没有很强的关联性.

系统发育多样性指数计算和评估是生物多样性研究工作的重要组成部分. 在群落水平上, 通过植物 DNA 条形码技术, 可获得更接近实际情况的群落系统发育关系, 能反映出群落间各物种较为真实的进化关系, 这比以往通过模拟或估算的进化关系更为有效. 受全球变化的影响, 生物多样性研究进入了新的阶段; 不同类型的多样性指数从不同角度和层次描述生物多样性的历史和现实状况, 而 DNA 条形码方法则为区域和全球尺度上不同森林群落的多样性指数计算和评估创造了可供比较的机会. 目前这方面的工作尚处于探索阶段, 但随着更多研究的加入, 将来可为特定区域的生物多样性评估和跨区域的比较工作提供新的思路.

## 5 小结与展望

DNA 条形码是一门实用性较强的物种鉴定技术, 需与理论结合, 解决实际科学问题. 一般来说, 一份完整的标准植物 DNA 条形码材料清单通常包括物种的凭证标本和野外采集记录、供提取基因组 DNA 的植物材料样品、植物局部和生境特征、图片影像信息、DNA 序列信息、储备 DNA, 以及提交给相应数据库的编号等. 当然, 这种技术还有待完善, 在实践过程中需要处理好几个关系: (1) 与标本馆或博物馆的关系: 植物标本的采集和收录工作是标本馆的主

要工作之一, DNA 条形码的研究工作不是简单的标本重复采集和辨认. 已有的馆藏标本可为当前的 DNA 条形码野外采集提供参考信息来源, 而正在进行的 DNA 条形码信息采集工作将极大地丰富现存的标本储存数量和质量, 从而使得两者实现互补和完善. (2) 与 GenBank 的关系: GenBank (<http://www.ncbi.nlm.nih.gov/genbank/>)拥有目前最为全面和丰富的 DNA 序列库, 可供全世界研究人员共享使用; DNA 条形码的研究工作不是简单的重复测序, 能在部分重要类群有新的补充和发现. DNA 条形码的标准化研究方法在从事此类研究的国际同行中得到充分执行, 所构建的条形码数据文库包含大量生物学信息, 可以快速地在区域甚至全球尺度上开展比较性研究. DNA 条形码研究借力于 GenBank 中的海量序列信息, 同时也在一定程度上丰富了可用的 DNA 序列信息库, 发展了生命科学的研究框架. (3) 与生命之树的关系: 生命之树(Tree of Life)<sup>[58]</sup>和被子植物系统发育研究组(Angiosperm Phylogeny Group)<sup>[59,60]</sup>主要关注探索生命世界的起源和进化历史、构建各生物类群之间系统发育关系. DNA 条形码的研究工作不局限于某一类群或群落的系统发育关系构建, 更关注物种水平上的鉴定和辨别, 特别是在较低分类阶元的系统位置归类上具有较高的解决能力, 能在部分难解的进化分支上提高分辨率和支持率, 是生命之树大框架的重要组成部分<sup>[61,62]</sup>.

植物 DNA 条形码已逐渐成为现代生物学研究的一个重要工具, 正逐步与关联学科融合, 可在微观和宏观 2 个层面上促进各自的发展, 这将是值得期待的研究新方向. 在微观层面, 借力于新一代高通量测序技术, 可将转录组学(transcriptomics, 也即 RNA 测序)方面的研究与 DNA 条形码、系统发育以及功能性状研究相融合, 在种内和种间水平进行比较基因组学研究, 深度挖掘基因功能, 精确定位关键位点, 获取生物系统发育进化关系等方面更为详尽的信息<sup>[63-65]</sup>. 在宏观层面, 我们可充分利用现有的 DNA 条形码资源库, 吸取整合生物学、生物地理学等学科理论和技术优势, 如新一代智能植物志 iFlora 共享和开放数据平台的构思与推广应用, 进行大区域、跨学科、多层次、长周期的合作研究<sup>[66,67]</sup>, 探索大尺度生物多样性格局、生物资源保存和利用、生态环境变迁等关系人类生存和长远发展等难题的解决之道, 为决策者提供备选预案. DNA 条形码技术在生态学研究领域中的应用,

涉及群落生态(community ecology)、入侵(species invasion)、大进化(macroevolution)、性状进化(trait evolution)、食物网络和营养结构(food webs and trophic interactions)、集合群落(metacommunities)和空间生态学(spatial ecology)等方面. 尽管存在种下单元研究样本采集量较少和居群水平比较研究变异信息不足等缺陷, DNA 条形码技术仍有望在未知生物新类群的探索、海量数据的管理应用、下一代测序技术

等方面获得新的发展<sup>[68]</sup>. 此外, 近些年以 CTFS 和 CForBio 网络大样地为平台开展生态学研究获得了长足进展. 目前, 在具有植物 DNA 条形码研究基础的大样地, 有必要在局部、区域和全球尺度上尽快开展多群落水平的 DNA 条形码物种识别、群落系统发育多样性指数评价、地上地下动植物交互作用等工作, 为探讨生物多样性格局和维持机制提供更先进的技术支撑.

**致谢** 第七届海峡两岸森林动态样区研讨会上得到了与会专家的指点, 外审专家提供了中肯的修改意见, 特致谢忱.

## 参考文献

- 1 Pennisi E. Wanted: A barcode for plants. *Science*, 2007, 318: 190–191
- 2 Moritz C, Cicero C. DNA barcoding: Promise and pitfalls. *PLoS Biol*, 2004, 2: e354
- 3 Janzen D H, Hallwachs W, Blandin P, et al. Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Mol Ecol Resour*, 2009, 9: 1–26
- 4 裴男才, 陈步峰. 生物 DNA 条形码: 十年发展历程、研究尺度和功能. *生物多样性*, 2013, 21: 616–627
- 5 陈士林, 庞晓慧, 罗焜, 等. 生物资源的 DNA 条形码技术. *生命科学*, 2013, 25: 458–466
- 6 Hebert P D N, Cywinska A, Ball S L, et al. Biological identifications through DNA barcodes. *Proc R Soc London Ser B Biol Sci*, 2003, 270: 313–321
- 7 Hebert P D N, Ratnasingham S, deWaard J R. Barcoding animal life: Cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc R Soc London Ser B Biol Sci*, 2003, 270: S96–S99
- 8 Chase M W, Fay M F. Barcoding of plants and fungi. *Science*, 2009, 325: 682–683
- 9 Chase M W, Salamin N, Wilkinson M, et al. Land plants and DNA barcodes: Short-term and long-term goals. *Phil Transac R Soc B Biol Sci*, 2005, 360: 1889–1895
- 10 Kress W J, Wurdack K J, Zimmer E A, et al. Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci USA*, 2005, 102: 8369–8374
- 11 China-Plant-BOL-Group. Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be incorporated into the core barcode for seed plants. *Proc Natl Acad Sci USA*, 2011, 108: 19641–19646
- 12 Kress W J, Erickson D L. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS One*, 2007, 2: e508
- 13 Cowan R S, Chase M W, Kress W J, et al. 300000 species to identify: Problems, progress, and prospects in DNA barcoding of land plants. *Taxon*, 2006, 55: 611–616
- 14 宁淑萍, 颜海飞, 郝刚, 等. 植物 DNA 条形码研究进展. *生物多样性*, 2008, 16: 417–425
- 15 Kress W J, Erickson D L. *DNA Barcodes: Methods and Protocols*. Berlin: Humana Press, 2012
- 16 Hollingsworth P M, Graham S W, Little D P. Choosing and using a plant DNA barcode. *PLoS One*, 2011, 6: e19254
- 17 Lahaye R, van der Bank M, Bogarin D, et al. DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci USA*, 2008, 105: 2923–2928
- 18 Vere N D, Rich T C G, Ford C R, et al. DNA barcoding the native flowering plants and conifers of Wales. *PLoS One*, 2012, 7: e37945
- 19 CBOL-Plant-Working-Group. A DNA barcode for land plants. *Proc Natl Acad Sci USA*, 2009, 106: 12794–12797
- 20 高连明, 刘杰, 蔡杰, 等. 关于植物 DNA 条形码研究技术规范. *植物分类与资源学报*, 2012, 34: 592–606
- 21 张金梅, 王建秀, 夏涛, 等. 基于系统发育分析的 DNA 条形码技术在澄清芍药属牡丹组物种问题中的应用. *中国科学 C 辑: 生命科学*, 2008, 38: 1166–1176
- 22 李德铤, 王雨华, 伊廷双, 等. 新一代植物志: iFlora. *植物分类与资源学报*, 2012, 34: 525–531
- 23 罗亚皇, 刘杰, 高连明, 等. DNA 条形码在生态学研究中的应用与展望. *植物分类与资源学报*, 2013, 35: 761–768
- 24 Yu D W, Ji Y, Emerson B C, et al. Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol Evol*, 2012, 3: 613–623

- 25 唐敏, 伊廷双, 王欣, 等. Metabarcoding 技术在植物鉴定和多样性研究中的应用. 植物分类与资源学报, 2013, 35: 769–773
- 26 Ji Y, Ashton L, Pedley S M, et al. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecol Lett*, 2013, 16: 1245–1257
- 27 Taberlet P, Coissac E, Pompanon F, et al. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol Ecol*, 2012, 21: 2045–2050
- 28 马克平. 大型固定样地: 森林生物多样性定位研究的平台. 植物生态学报, 2008, 32: 237
- 29 Simon A L. *Encyclopedia of Biodiversity*, 2nd ed. Oxford: Academic Press, 2013
- 30 Kimura M. *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press, 1983
- 31 Altschul S F, Madden T L, Schäffer A A, et al. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*, 1997, 25: 3389–3402
- 32 Meier R, Shiyang K, Vaidya G, et al. DNA barcoding and taxonomy in Diptera: A tale of high intraspecific variability and low identification success. *Syst Biol*, 2006, 55: 715–728
- 33 Kress W J, Erickson D L, Jones F A, et al. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proc Natl Acad Sci USA*, 2009, 106: 18621–18626
- 34 Gonzalez M A, Baraloto C, Engel J, et al. Identification of Amazonian trees with DNA barcodes. *PLoS One*, 2009, 4: e7483
- 35 Kress W J, Erickson D L, Swenson N G, et al. Advances in the use of DNA barcodes to build a community phylogeny for tropical trees in a Puerto Rican forest dynamics plot. *PLoS One*, 2010, 5: e15409
- 36 裴男才. 利用 DNA 条形码技术识别植物物种. 应用生态学报, 2012, 23: 1240–1246
- 37 卢孟孟, 慈秀芹, 杨国平, 等. 亚热带森林乔木树种 DNA 条形码研究——以哀牢山自然保护区为例. 植物分类与资源学报, 2013, 35: 733–741
- 38 Webb C O, Slik J W F, Triono T. Biodiversity inventory and informatics in Southeast Asia. *Biodiv Conserv*, 2010, 19: 955–972
- 39 Whittaker R. Evolution and measurement of species diversity. *Taxon*, 1972, 21: 213–251
- 40 Faith D P. Conservation evaluation and phylogenetic diversity. *Biol Conserv*, 1992, 61: 1–10
- 41 Zhang J, Pei N, Mi X. *phylotools: Phylogenetic tools for Eco-phylogenetics*. R package version 0.1.2. (<http://CRAN.R-project.org/package=phylotools>), 2012
- 42 Saitou N, Nei M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol Biol Evol*, 1987, 4: 406–425
- 43 Swofford D L. *PAUP\* Phylogenetic Analysis Using Parsimony (\*and Other Methods) Version 4*. Sunderland, MA(USA): Sinaur Associates, Inc, 2003
- 44 Stamatakis A. RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 2006, 22: 2688–2690
- 45 Webb C O, Donoghue M J. *PhyloMatic: Tree assembly for applied phylogenetics*. *Mol Ecol Notes*, 2005, 5: 181–183
- 46 裴男才. 利用植物 DNA 条形码构建亚热带森林群落系统发育关系——以鼎湖山样地为例. 植物分类与资源学报, 2012, 34: 263–270
- 47 Beaulieu J M, Ree R H, Cavender-Bares J, et al. Synthesizing phylogenetic knowledge for ecological research. *Ecology*, 2012, 93: S4–S13
- 48 Swenson N G. Phylogenetic resolution and quantifying the phylogenetic diversity and dispersion of communities. *PLoS One*, 2009, 4: e4390
- 49 裴男才, 张金龙, 米湘成, 等. 植物 DNA 条形码促进系统发育群落生态学发展. 生物多样性, 2011, 19: 284–294
- 50 Yang J, Zhang G, Ci X, et al. Functional and phylogenetic assembly in a Chinese tropical tree community across size classes, spatial scales and habitats. *Funct Ecol*, 2014, 28: 520–529
- 51 Pei N, Lian J-Y, Erickson D L, et al. Exploring tree-habitat associations in a Chinese subtropical forest plot using a molecular phylogeny generated from DNA barcode loci. *PLoS One*, 2011, 6: e21273
- 52 曹科, 饶米德, 刘晓娟, 等. 古田山木本植物功能性状的系统发育信号及其对群落结构的影响. 生物多样性, 2013, 21: 564–571
- 53 Liu X, Swenson N G, Zhang J, et al. The environment and space, not phylogeny, determine trait dispersion in a subtropical forest. *Funct Ecol*, 2013, 27: 264–272
- 54 Gonzalez M A, Roger A, Courtois E A, et al. Shifts in species and phylogenetic diversity between sapling and tree communities indicate negative density dependence in a lowland rain forest. *J Ecol*, 2010, 98: 137–146
- 55 Feng G, Zhang J, Pei N, et al. Comparison of phylobetadiversity indices based on community data from Gutianshan forest plot. *Chin Sci Bull*, 2012, 57: 623–630
- 56 饶米德, 冯刚, 张金龙, 等. 生境过滤和扩散限制作用对古田山森林物种和系统发育  $\beta$  多样性的影响. 科学通报, 2013, 58: 1204–1212
- 57 Swenson N G, Erickson D L, Mi X, et al. Phylogenetic and functional alpha and beta diversity in temperate and tropical tree communities. *Ecology*, 2012, 93: S112–S125
- 58 Maddison D R, Schulz K S. *The Tree of Life Project*. Tucson: University of Arizona, 2007

- 59 Stevens P F. Angiosperm Phylogeny Website. Version 12, July 2012 (<http://www.mobot.org/MOBOT/research/APweb/>), 2014
- 60 APGIII. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Bot J Linnean Soc*, 2009, 161: 105–121
- 61 陈之端, 李德铎. 生命条形码与生命之树. *植物分类与资源学报*, 2013, 35: 675–681
- 62 Dong W, Cheng T, Li C, et al. Discriminating plants using the DNA barcode *rbcLb*: An appraisal based on a large data set. *Mol Ecol Resour*, 2014, 14: 336–343
- 63 Swenson N G. The functional ecology and diversity of tropical tree assemblages through space and time: from local to regional and from traits to transcriptomes. *ISRN Forestry*, 2012, 2012: Article ID 743617
- 64 Farnsworth E J, Chu M, Kress W J, et al. Next-generation field guides. *BioScience*, 2013, 63: 891–899
- 65 米湘成, 裴男才, 马克平. 群落系统发育学研究进展. 见: 《新生物学年鉴 2013》编委会, 编. *新生物学年鉴*. 北京: 科学出版社, 2013. 266–289
- 66 Pyron R A, Burbrink F T. Phylogenetic estimates of speciation and extinction rates for testing ecological and evolutionary hypotheses. *Trends Ecol Evol*, 2013, 28: 729–736
- 67 Mouquet N, Devictor V, Meynard C N, et al. Ecophylogenetics: Advances and perspectives. *Biol Rev*, 2012, 87: 769–785
- 68 Joly S, Davies T J, Archambault A, et al. Ecology in the age of DNA barcoding: The resource, the promise and the challenges ahead. *Mol Ecol Resour*, 2014, 14: 221–232

## Integration of plant DNA barcoding technology into community studies on forest dynamics plots

PEI NanCai<sup>1</sup>, MI XiangCheng<sup>2</sup> & CHEN BuFeng<sup>1</sup>

<sup>1</sup> Forest Ecosystem Station of the Pearl River Delta, State Forestry Administration, Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangzhou 510520, China;

<sup>2</sup> State Key Laboratory of Vegetation and Environmental Change, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China

Plant DNA barcoding technology has important applications in biodiversity, taxonomy, evolutionary ecology, molecular systematics, and forensics studies. This paper summarizes important progress in the application of plant DNA barcodes (specifically, *rbcL*, *matK*, and *trnH-psbA*) at the community level, presenting the rationale and case studies in the context of global forest dynamics plots. The results showed that the single DNA barcode region *trnH-psbA* achieved the highest rate of correct species identification, followed by *matK* and *rbcL*. Combining the three barcode regions resulted in excellent discrimination of plant species and reconstruction of community phylogenies (with improved topological resolution and high bootstrap support values for nodes) both in the tropics and subtropics. Furthermore, integration of plant DNA barcodes with community ecology may not only strengthen the capability to explore and elucidate mechanisms of community assembly, but also improve the efficiency of comparative evaluation of biodiversity indexes. Plant DNA barcoding is proven to be effective in research at the clade and community levels. The toolkit is also expected to show promise at the regional and global scales with refinements, such as optimization of sampling strategy and addition of the nuclear internal transcribed spacer region to increase the genetic information content, and to improve the technology's performance.

**barcode of life, biological community, forest biodiversity monitoring network, integrative biology**

doi: 10.1360/N972014-00033