Check for updates

Advances of Community-Level Plant DNA Barcoding in China

Nancai Pei[1]*, Bufeng Chen[1] and W. J. Kress[2]

[1] Key Laboratory of State Forestry Administration on Tropical Forestry Research, Research Institute of Tropical Forestry, Chinese Academy of Forestry, Guangzhou, China, [2] Department of Botany, MRC-166, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

DNA barcoding is a commonly used bio-technology in multiple disciplines including biology, environmental science, forensics and inspection, etc. Forest dynamic plots provide a unique opportunity to carry out large-scale, comparative, and multidisciplinary research for plant DNA barcoding. The paper concisely reviewed four previous progresses in China; specifically, species discrimination, community phylogenetic reconstruction, phylogenetic community structure exploration, and biodiversity index evaluation. Further, we demonstrated three major challenges; specifically, building the impetus to generate DNA barcodes using multiple plant DNA markers for all woody species at forest community levels, analyzing massive DNA barcoding sequence data, and promoting theoretical innovation. Lastly, we raised five possible directions; specifically, proposing a "purpose-driven barcode" fit for multi-level applications, developing new integrative sequencing strategies, pushing DNA barcoding beyond terrestrial ecosystem, constructing national-level DNA barcode sequence libraries for special plant groups, and establishing intelligent identification systems or online server platforms. These efforts will be potentially valuable to explore large-scale biodiversity patterns, the origin and evolution of life, and will also facilitate preservation and utilization of biodiversity resources.

Keywords: DNA barcode, community phylogeny, reproductive trait, forest biology, subtropical forest

INTRODUCTION

DNA barcoding, a bio-technology characterized by standardization, universality and efficiency (Hebert et al., 2003), is widely used in multiple disciplines including biology, environmental science, forensics and cross-boarder inspection, etc. DNA barcoding technology may be most promising to achieve the goals of rapid and accurate species identification and sustainable utilization of biological resources (CBOL Plant Working Group, 2009; Janzen et al., 2009; Harris and Bellino, 2013; Hollingsworth et al., 2016). Meanwhile, many drawbacks and problems cannot be neglected with use of DNA barcodes, including failure of amplification or sequencing, difficulties in finding universal primers, lack of barcoding gap, hybridization and introgression in some plant groups (Collins and Cruickshank, 2013; Zinger and Philippe, 2016). Concerns of plant DNA barcoding may differ in biologists; specifically, taxonomists mainly focus on clades, systematists on phylogenies, while ecologists on communities.

China initiated its national-level project of DNA barcoding (including animals, plants, and microbes) in 2008, and was invited as one of the four iBOL (International Barcode of Life[1]) central

---

[1]http://ibol.org/

node nations in the world (Che et al., 2010). Vegetation types in China are extremely diverse involving tropics, subtropics and temperate zones. Forest dynamic plots (FDPs), according to the standard Smithsonian CTFS/ForestGEO protocol (Condit, 1998) and also conducted based on Forestry Standards for "Observation Methodology for Long-term Forest Ecosystem Research" of the People's Republic of China (LY/T 1952-2011), provide us an opportunity to study large-scale and multidisciplinary research on forests including DNA barcoding. Presently, 15 FDPs were set up across the Chinese mainland (13 belonging to CForBio monitoring network[2]; plus Heishiding FDP in Guangdong province, and Jianfengling FDP in Hainan province). Additionally, Tai-Po-Kau FDP in Hong Kong and three FDPs (i.e., Lienhuachih, Fushan, and Nanjenshan) in Taiwan were also set up successively. Nowadays, there are four, nine, and six FDPs in tropics, subtropics and temperate zone, respectively. At least seven FDPs (i.e., Dinghushan, Gutianshan, Changbaishan, Xishuangbanna, and three FDPs in Taiwan) are available with plant DNA barcoding sequences (mostly *rbc*L, *mat*K, *trn*H-*psb*A, and occasionally ITS2). Additional two FDPs are underway to release or publish DNA barcoding sequence data (e.g., Badagongshan in Hunan province, Jianfengling in Hainan Island, etc.).

Based on Dinghushan FDP (a lower subtropical forest in south China), a Chinese research team cooperating with international scientists, initiated a plant DNA barcoding project (i.e., community-level) in 2008. The first research article was published in 2011, utilizing a well-resolved DNA barcode phylogeny to explore tree-habitat associations. Subsequently, approximately a dozen international and domestic publications of closely related or comparative topics were available in China (Pei et al., 2014). Besides of contributions from Chinese forests, studies on plant DNA barcoding involving forest communities are published worldwide across tropics, subtropics and temperate zones (Gonzalez et al., 2009; Kress et al., 2009, 2010; Burgess et al., 2011; Parmentier et al., 2013; Saarela et al., 2013; de Boer et al., 2014). Generally, applications of plant DNA barcoding can be divided into clade and community levels, and this paper focuses on the latter. We developed selection criteria that all investigations should be dominated by or participated with Chinese researcher(s) and included publications be carried out in one or more Chinese forest communities. Here, we briefly demonstrated recent progress, major challenges, and possible future directions of the community-level plant DNA barcoding in China.

PREVIOUS ADVANCES

(1) Species discrimination had become easy and rapid by means of DNA barcoding technology. Generally, the single DNA barcoding marker *trn*H-*psb*A raised relatively high rates of species discrimination, followed by *mat*K and *rbc*L. The combination of *rbc*L+*mat*K (a core barcode for land plants recommended by CBOL Plant Working

Group) averagely discriminated 88.6, 83.8, and 72.5% at the local, regional and global scales, respectively (Pei et al., 2015a). An additional intergenic spacer, either *trn*H-*psb*A or ITS2, had also proven to be useful in different taxonomic groups (China Plant BOL Group et al., 2011; Liu et al., 2015; Pei et al., 2015a). Rates of species discrimination varied along a latitudinal gradient and were negatively correlated with ratios of closely related taxa and generally depended on geographic scales in global FDPs (Pei et al., 2015a).

(2) Community phylogenetic reconstruction, via a super-matrix approach with the three-locus barcode combination (*rbc*L+*mat*K+*psb*A-*trn*H) provided a well-resolved phylogenetic framework. Thus, a DNA barcoding-based phylogeny could assign almost all species to a proper evolutionary position in a systematic classification (Pei, 2012; Erickson et al., 2014), when compared to "Phylomatic phylogenies," which were usually accompanied by more polytomies and resulted in a larger bias of phylogenetic community structure. In addition, a powerful R-package named "phylotools" (Zhang et al., 2012) had been developed to build a super-matrix for multi-locus DNA barcodes, and easily calculated the inequality among lineages and phylogenetic similarity for large datasets.

(3) The exploration of phylogenetic community structure benefited greatly from well-resolved phylogenies generated from plant DNA barcodes. Phylogenetic signal has been detected in plant-habitat associations (i.e., closely related species tend to prefer similar habitats; Pei et al., 2011). In addition, patterns of co-occurrence within habitats and functional traits across spatial and size scales were typically non-random with respect to community phylogenies and gave strong support for a deterministic model rather than for a neutral model (Pei et al., 2011; Yang et al., 2014).

(4) Biodiversity index evaluation proved to be more effective and comparable with the aid of standardized plant DNA barcoding procedures. Phylobetadiversity might not be significantly affected by species abundance when scales were relatively small, which might result from reductions in evenness in communities as scales increased. Moreover, *phylogenetic* measures of alpha and beta diversity were not strong predictors of *functional* alpha and beta diversity, but partitioning the variation in phylogenetic and functional beta diversity showed that environmental distance was generally a better predictor of beta diversity in diverse forests worldwide (Feng et al., 2012; Swenson et al., 2012; Yang et al., 2015). Overall, in view of cost-effectiveness and the trade-off between sequence recovery and species resolution, we suggested the combination of markers *rbc*L+*mat*K+*trn*H-*psb*A as a priority for DNA-based studies on forest communities (Pei et al., 2015a).

MAJOR CHALLENGES

Though numerous important progress had been made in the past 10 years, we raised that the following three challenges might be tough but remain promising: (1) Building the impetus to

---

[2]http://www.cfbiodiv.org/

generate DNA barcodes using multiple plant DNA markers for all woody species at forest community levels, which requires significant research resources and investment; (2) Analyzing massive DNA barcoding sequence data, which needs powerful computational systems and critical infrastructures to perform large-scale and multidisciplinary research projects; and (3) Promoting theoretical innovation, which calls for raising novel scientific hypotheses and publishing a series of influential papers in top academic journals.

FUTURE DIRECTIONS

Comparative analyses of community phylogenies from forest dynamics plots or natural reserves are feasible owing to the universal and standard working routine of plant DNA barcoding (Kress and Erickson, 2012). When combined with conservative plant traits (e.g., flowering phenology; Pei et al., 2015b), effects of individuals on assembly patterns within communities, dissimilarity of diverse communities along an environmental gradient, and non-random processes therein should be more thoroughly explored. Possible directions are: (1) Proposing a "purpose-driven barcode" (e.g., metabarcoding and mini-barcode) fit for multi-level applications such as identifying living organisms, reconstructing community phylogenies, detecting environmental biodiversity information, and exploring ecological network structure (Little, 2014; Kress et al., 2015; Evans et al., 2016); (2) Developing new integrative sequencing strategies (e.g., genome skimming; Hollingsworth et al., 2016) to generate mega-phylogenies in face of the post-genomic era; (3) Pushing DNA barcoding beyond terrestrial ecosystem, to include aquatic ecosystem (including mangroves; Richards and Friess, 2016) with complicated biotic interactions and abiotic extreme conditions to explore mechanisms of formation, maintenance, and evolution; (4) Constructing national-level DNA barcode sequence libraries of economically valuable tree species for commercial authentication and endangered plant taxa against illegal international trade (Xu et al., 2015; Zhao et al., 2015); and (5) Establishing intelligent identification systems (e.g., Leafsnap) or online server platforms (e.g., i-Flora and e-Flora; Kumar et al., 2012; Li et al., 2012; Zeng et al., 2014; Pei, 2016) for land plants integrating genetic, morphological and environmental information, which will make DNA-based plant identification more precise, more convenient, and more interesting. These pursuits will be valuable to explore large-scale biodiversity patterns, the origin and evolution of life, and will also facilitate preservation and utilization of biological resources. We hope that this paper will be an addition to the field.

AUTHOR CONTRIBUTIONS

NP conceived and wrote the draft, NP, BC, and WJK revised the paper and approved it for publication.

FUNDING

REFERENCES

Burgess, K. S., Fazekas, A. J., Kesanakurti, P. R., Graham, S. W., Husband, B. C., Newmaster, S. G., et al. (2011). Discriminating plant species in a local temperate flora using the rbcL+matK DNA barcode. *Methods Ecol. Evol.* 2, 333–340. doi: 10.1111/j.2041-210X.2011.00092.x

CBOL Plant Working Group (2009). A DNA barcode for land plants. *Proc. Natl. Acad. Sci. U.S.A.* 106, 12794–12797. doi: 10.1073/pnas.0905845106

Che, J., Huang, D., Li, D., Ma, J., and Zhang, Y. (2010). DNA barcoding and the international barcode of life project in China. *Bull. Chin. Acad. Sci.* 24, 257–260.

China Plant BOL Group, Li, D. Z., Gao, L. M., Li, H. T., Wang, H., Ge, X. J., Liu, J. Q., et al. (2011). 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. U.S.A.* 108, 19641–19646. doi: 10.1073/pnas.1104551108

Collins, R. A., and Cruickshank, R. H. (2013). The seven deadly sins of DNA barcoding. *Mol. Ecol. Resour.* 13, 969–975. doi: 10.1111/1755-0998.12046

Condit, R. (1998). *Tropical Forest Census Plots*. Berlin: Springer-Verlag. doi: 10.1007/978-3-662-03664-8

de Boer, H. J., Ouarghidi, A., Martin, G., Abbad, A., and Kool, A. (2014). DNA barcoding reveals limited accuracy of identifications based on folk taxonomy. *PLoS ONE* 9:e84291. doi: 10.1371/journal.pone.0084291

Erickson, D. L., Jones, F. A., Swenson, N. G., Pei, N., Bourg, N. A., Chen, W. N., et al. (2014). Comparative evolutionary diversity and phylogenetic structure across multiple forest dynamics plots: a mega-phylogeny approach. *Front. Genet.* 5:358. doi: 10.3389/fgene.2014.00358

Evans, D. M., Kitson, J. J. N., Lunt, D. H., Straw, N. A., and Pocock, M. J. O. (2016). Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. *Funct. Ecol.* 30, 1904–1916. doi: 10.1111/1365-2435.12659

Feng, G., Zhang, J. L., Pei, N., Rao, M. D., Mi, X. C., Ren, H. B., et al. (2012). Comparison of phylobetadiversity indices based on community data from Gutianshan forest plot. *Chin. Sci. Bull.* 57, 623–630. doi: 10.1007/s11434-011-4869-1

Gonzalez, M. A., Baraloto, C., Engel, J., Mori, S. A., Pétronelli, P., Riéra, B., et al. (2009). Identification of Amazonian trees with DNA barcodes. *PLoS ONE* 4:e7483. doi: 10.1371/journal.pone.0007483

Harris, S. E., and Bellino, M. (2013). DNA barcoding from NYC to Belize. *Science* 342, 1462–1463. doi: 10.1126/science.1230006

Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B* 270, 313–321. doi: 10.1098/rspb.2002.2218

Hollingsworth, P. M., Li, D. Z., van der Bank, M., and Twyford, A. D. (2016). Telling plant species apart with DNA: from barcodes to genomes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 371:20150338. doi: 10.1098/rstb.2015.0338

Janzen, D. H., Hallwachs, W., Blandin, P., Burns, J. M., Cadiou, J. M., Chacon, I., et al. (2009). Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Mol. Ecol. Resour.* 9, 1–26. doi: 10.1111/j.1755-0998.2009.02628.x

Kress, W. J., and Erickson, D. L. (eds). (2012). *DNA Barcodes: Methods and Protocols*. Berlin: Humana Press. doi: 10.1007/978-1-61779-591-6

Kress, W. J., Erickson, D. L., Jones, F. A., Swenson, N. G., Perez, R., Sanjur, O., et al. (2009). Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18621–18626. doi: 10.1073/pnas.0909820106

Kress, W. J., Erickson, D. L., Swenson, N. G., Thompson, J., Uriarte, M., and Zimmerman, J. K. (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. doi: 10.1371/journal.pone.0015409

Kress, W. J., García-Robledo, C., Uriarte, M., and Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends Ecol. Evol.* 30, 25–35. doi: 10.1016/j.tree.2014.10.008

Kumar, N., Belhumeur, P., Biswas, A., Jacobs, D., Kress, W. J., Lopez, I., and Soares, J. B. (2012). "Leafsnap: a computer vision system for automatic plant species identification," in *Computer Vision – ECCV 2012*, eds A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, C. Schmid et al. (Berlin: Springer), 502–516.

Li, D. Z., Wang, Y. H., Yi, T. S., Wang, H., Gao, L. M., and Yang, J. B. (2012). The next-generation flora: iFlora. *Plant Diversity Resour.* 34, 525–531. doi: 10.3724/SP.J.1143.2012.12135

Little, D. P. (2014). A DNA mini-barcode for land plants. *Mol. Ecol. Resour.* 14, 437–446. doi: 10.1111/1755-0998.12194

Liu, J., Yan, H. F., Newmaster, S. G., Pei, N., Ragupathy, S., and Ge, X. J. (2015). The use of DNA barcoding as a tool for the conservation biogeography of subtropical forests in China. *Divers. Distrib.* 21, 188–199. doi: 10.1111/ddi.12276

Parmentier, I., Duminil, J., Kuzmina, M., Philippe, M., Thomas, D. W., Kenfack, D., et al. (2013). How effective are DNA barcodes in the identification of African rainforest trees? *PLoS ONE* 8:e54921. doi: 10.1371/journal.pone.0054921

Pei, N. (2012). Building a subtropical forest community phylogeny based on plant DNA barcodes from Dinghushan plot. *Plant Divers. Resour.* 34, 263–270. doi: 10.3724/SP.J.1143.2012.11173

Pei, N. (2016). Intelligent plant identification via integration of genetic, morphological and environmental characteristics. *World For. Res.* 29, 29–32.

Pei, N., Erickson, D. L., Chen, B. F., Ge, X. J., Mi, X. C., Swenson, N. G., et al. (2015a). Closely-related taxa influence woody species discrimination via DNA barcoding: evidence from global forest dynamics plots. *Sci. Rep.* 5:15127. doi: 10.1038/srep15127

Pei, N., Kress, W. J., Chen, B. F., Erickson, D. L., Wong, K. M., Zhang, J. L., et al. (2015b). Phylogenetic and climatic constraints drive flowering phenological patterns in a subtropical nature reserve. *J. Plant Ecol.* 8, 187–196. doi: 10.1093/jpe/rtv009

Pei, N., Lian, J. Y., Erickson, D. L., Swenson, N. G., Kress, W. J., Ye, W. H., et al. (2011). Exploring tree-habitat associations in a Chinese subtropical forest plot using a molecular phylogeny generated from DNA barcode loci. *PLoS ONE* 6:e21273. doi: 10.1371/journal.pone.0021273

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

Richards, D. R., and Friess, D. A. (2016). Rates and drivers of mangrove deforestation in Southeast Asia, 2000-2012. *Proc. Natl. Acad. Sci. U.S.A.* 113, 344–349. doi: 10.1073/pnas.1510272113

Saarela, J. M., Sokoloff, P. C., Gillespie, L. J., Consaul, L. L., and Bull, R. D. (2013). DNA barcoding the Canadian arctic flora: core plastid barcodes (*rbcL* + *matK*) for 490 vascular plant species. *PLoS ONE* 8:e77982. doi: 10.1371/journal.pone.0077982

Swenson, N. G., Erickson, D. L., Mi, X., Bourg, N. A., Montana-Forero, J., Ge, X., et al. (2012). Phylogenetic and functional alpha and beta diversity in temperate and tropical tree communities. *Ecology* 93, S112–S125. doi: 10.1890/11-1180.1

Xu, C., Dong, W., Shi, S., Cheng, T., Li, C., Liu, Y., et al. (2015). Accelerating plant DNA barcode reference library construction using herbarium specimens: improved experimental techniques. *Mol. Ecol. Resour.* 15, 1366–1374. doi: 10.1111/1755-0998.12413

Yang, J., Swenson, N. G., Zhang, G., Ci, X., Cao, M., Sh, A L., et al. (2015). Local-scale partitioning of functional and phylogenetic beta diversity in a tropical tree assemblage. *Sci. Rep.* 5:12731. doi: 10.1038/srep12731

Yang, J., Zhang, G., Ci, X., Swenson, N. G., Cao, M., Sha, L., et al. (2014). Functional and phylogenetic assembly in a Chinese tropical tree community across size classes, spatial scales and habitats. *Funct. Ecol.* 28, 520–529. doi: 10.1111/1365-2435.12176

Zeng, C. X., Wang, Y. N., Wang, Y. H., and Wang, H. (2014). Plant DNA barcoding and biodiversity server platform. *Biodivers. Sci.* 22, 285–292. doi: 10.3724/SP.J.1003.2014.13268

Zhang, J. L., Pei, N., and Mi, X. C. (2012). *Phylotools: Phylogenetic Tools for Eco-Phylogenetics.* Available at: http://cran.r-project.org/web/packages/phylotools/index.html

Zhao, S., Chen, X., Song, J., Pang, X., and Chen, S. (2015). Internal transcribed spacer 2 barcode: a good tool for identifying Acanthopanacis cortex. *Front. Plant Sci.* 6:840. doi: 10.3389/fpls.2015.00840

Zinger, L., and Philippe, H. (2016). Coalescing molecular evolution and DNA barcoding. *Mol. Ecol.* 25, 1908–1910. doi: 10.1111/mec.13639