Sabbatical Report, Fall 2010
Amy L. Denton, Ph.D.

EVOLUTION AND DIVERSIFICATION IN THE BAMBOO FAMILY USING EMERGING MOLECULAR AND ANALYTICAL TOOLS

I. Summary of Purpose and Goals

The purpose of this sabbatical leave was to provide faculty retraining in emerging technology and analytical methodology necessary for research, publication, and instructional improvement.

For fifteen years I have maintained an interest in the evolution and diversification of land plants using the techniques of comparative DNA sequencing (the study of variations in the genetic material among related species) and molecular phylogenetics (the use of statistical analyses to determine the rates and patterns of change occurring in DNA sequences to reconstruct evolutionary relationships among organisms). This work requires massive amounts of DNA sequence data and is collaborative, involving scientists in many laboratories. Continuing innovations in high-throughput sequencing technologies and associated computational data analysis methods have changed dramatically the scope of work required for peer-reviewed publication in molecular evolutionary biology. Since I began my appointment at CSUCI, maintaining my currency in this type of research has been challenging because my work requires research infrastructure, instrumentation and time not available due to limitations in laboratory facilities and my focus on teaching and service. The field of phylogenetics develops rapidly and has a steep learning curve. My sabbatical leave allowed me to spend four months in the laboratory of Dr. Scot Kelchner at Idaho State University as part of an international, NSF-funded research group[1] studying the evolution and diversification of bamboos, learning and mastering new techniques for the rapid screening of genetic diversity and the associated methods of data analysis. Dr. Kelchner is an expert in the field of bamboo biology and applied phylogenetics, including the development of DNA sequence alignment protocols, improvement of phylogenetic models, and method testing via computer simulation experiments. He has published extensively in the highest-impact journals in our field and has given invited talks around the world. We have been collaborators since my appointment as an Affiliate Professor in the Department of Biological Sciences at ISU in 2005.

The approximately 1400 described species of bamboo are members of the grass family associated with tropical to temperate forests worldwide. They are the fastest growing plants on earth and some species exceed 30 cm in diameter and 100 m in height. Bamboos have economically important uses as building materials, textiles, food, and ornamental plants, and are ecologically significant as a key component of many remnant habitats that preserve threatened and endangered plant and animal species. Despite this importance, the evolutionary history and patterns of diversity in bamboos are poorly known relative to other grasses and bamboo classification remains problematic. These issues create difficulties for biologists, conservationists, and those interested in developing bamboo as a sustainable resource. A

[1] Bamboo Phylogeny Group (BPG); http://www.eeob.iastate.edu/research/bamboo/collaborators.html

principal goal of this project is to generate a phylogeny for the bamboos using DNA sequence data using innovative analytical and statistical methods.

In evolutionary biology, similarity among organisms or species is attributable to inheritance from a common ancestor. A phylogeny, or evolutionary tree, is a hypothesis that reconstructs the historical relationships, and in some cases the timing, of events that occur as species diversify. Phylogenies are generated using specialized computer applications that infer process from pattern: the process of organismal evolution is deduced from patterns of DNA variation. It is important to note that a single, all-purpose formula does not exist for phylogenetic analysis of DNA data. Although many algorithms, procedures, and programs continue to be developed and improved, their reliability is dependent upon the size and composition of the particular molecular data set under analysis. The merits and shortcomings of these various methods of phylogenetic inference and interpretation are subject to much scientific debate, because the danger of generating incorrect results is greater in computational molecular phylogenetics than in many other scientific disciplines. It is critical, therefore, that the user understand what the analytic method is actually doing with the data. This is why it was so important for me to update my analytical toolbox using real experimental data, under the direction of an expert well-versed in both the biology of the species under investigation as well as in the theoretical underpinnings of phylogenetic methodology.

II. Summary of Outcomes

a) Nuclear DNA phylogenetics in bamboos:

In order to produce a statistically robust phylogeny of the bamboo family, DNA sequences from the both the chloroplast (plant cellular organelles in which photosynthesis occurs) and the nuclear genome must be analyzed. Both types of DNA present unique obstacles to phylogenetic analyses due to unusual modes of inheritance, gene duplication patterns, and mutation rates. While Dr. Kelchner's lab has considerable expertise in chloroplast DNA amplification and sequencing in bamboos, no nuclear DNA marker sequences have been designed or tested. Nuclear marker development has been my research specialty since graduate school and my role in this collaborative effort was to use my expertise in the evolution of the cellular housekeeping genes RNA polymerase II (*RPB2*) and RNA polymerase III (*RPC1*) for the generation and troubleshooting of appropriate nuclear gene regions for DNA amplification and sequencing work. I trained Dr. Kelchner and his students (one Ph.D. candidate, three undergraduates) in the methods of nuclear gene sequencing, and together we generated *RPB2* and *RPC1* sequence data sets for a preliminary survey of bamboo species. These data were analyzed within the context of the collaborative bamboo phylogeny project described above, and were used as preliminary data for the generation of a sampling plan to study diversification in ecologically important bamboo species of South America. I am continuing to collaborate with Dr. Kelchner and his graduate student Dr. Amanda Fisher (Ph.D. May 2011) on this project at CSUCI. This work will result in the publication of at least one peer-reviewed article in a plant molecular evolution or phylogenetics journal in 2012.

b) Phylogenetic utility of plastid introns:
Working with Dr. Kelchner's students, I generated DNA sequence data for three chloroplast intron (non-coding) regions in fourteen problematic bamboo and related graminoid species to test the performance of these molecular markers in phylogenetic analysis. We used these data to independently corroborate a much larger phylogenetic reconstruction of the bamboos, demonstrating that these regions of DNA can be used to infer evolutionary relationships among problematic taxa quickly and inexpensively. The results of this work will be presented in poster format at two national meetings this summer: Botany 2011 (American Society of Plant Taxonomists, Society for Economic Botany, American Fern Society, Botanical Society of America), St. Louis MO, July 9-13, and Evolution 2011 (Society for the Study of Evolution, Society of Systematic Biologists, American Society of Naturalists, University of Oklahoma, Norman OK, June 17-21. The abstracts and poster are appended to this report.

c) Nuclear DNA phylogenetics in *Myoporum* (Scrophulariaceae):
I used my knowledge of nuclear DNA evolution in plants to evaluate confusing data generated by Chang Liu, an M.S. student (2008) in the Kelchner lab. This work involved using *RPB2* intron sequence data to estimate the copy number of *RPB2* in the endemic Australian genus *Myoporum* in the absence of accurate chromosome counts. In collaboration with Dr. Kelchner, I am continuing the extensive computer-based re-alignment of these sequences at CSUCI and expect this work to form the basis of a peer-reviewed article in 2012 or 2013.

d) Evolution of land plants:
Since graduate school I have maintained and interest in deep evolution of land plants. This work also takes the form of comparative genomics, but requires the development of molecular markers that evolve at a very slow rate, suitable for estimations of divergence times of deep lineages (ca. 400 million years). This work requires massive amounts of sequence data and is collaborative, involving many labs. Dr. Sean Graham, an associate professor of botany at the University of British Columbia and the Research Director at the UBC Botanic Garden, joined Dr. Kelchner and me at ISU for two weeks in December to develop a research plan using the *RPB2* markers to study the early divergence of monocot lineages. We have already done some preliminary work using these markers in ferns and mosses and have a draft manuscript in the early stages of preparation.

e) Training in emerging phylogenetic methods:
Through direct analysis of data generated during my sabbatical, as well as in weekly journal club discussions, I received training in the application of several new tests and software packages for phylogenetic analysis (jModeltest[2], CONSEL[3], SplitsTree[4]) which I will continue to use in my research and in teaching (BIOL 431-Bioinformatics, BIOL 494-Independent Research, BIOL 506-Molecular Evolution, BINF 500-DNA and Protein Sequence Analysis).

In sum, my sabbatical leave at Idaho State University provided me with an experience not

---

[2] Posada, D. 2008. jModelTest: Phylogenetic Model Averaging. *Molecular Biology and Evolution* 25: 1253-1256

[3] Shimodaira, H. & Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17,** 1246-1247

[4] Huson, D. and D. Bryant. 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution* 23:254-267

possible during my normal workload assignment due to constraints in time, instrumentation, and the availability of local expertise. During my leave I acquired new technical and analytical proficiency required to maintain currency in my field and generated data for future publication, teaching, and extramural funding. It also allowed me to continue work-in-progress with current collaborators (Kelchner and Graham), applying new methodologies learned at ISU. Tangible outcomes of this sabbatical are: 1) *at least* one collaborative scientific publication on the phylogeny of the bamboo family, 2) data generated and analytical methods learned for one future collaborative publication on deep evolution in land plants, 3) appropriate methods learned for future research and teaching, and 4) real data sets converted into learning exercises for students in BIOL 303 (Evolution), BIOL 311 (Plant Biology), BIOL 431 (Bioinformatics), BIOL 506 (Molecular Evolution), and BINF 500 (DNA and Protein Sequence Analysis. The skills I learned are relevant far beyond my applications to plant evolution – these techniques are used across the biological subdisciplines of biomedicine, molecular biology, conservation biology, bioinformatics, and genomics. My leave also provided an extended opportunity for me to interact with my scientific peers on a level that I could not otherwise. This informs my teaching in an important way by allowing me to share with my students cutting-edge research ideas and the newest data and results.

Abstract for: Botany 2011 (American Society of Plant Taxonomists, Society for Economic Botany, American Fern Society, Botanical Society of America), Annual Meeting, St. Louis MO, July 9-13, 2011

Evolution 2011 (Society for the Study of Evolution, Society of Systematic Biologists, American Society of Naturalists), Annual Meeting, University of Oklahoma, Norman, OK, June 17-21, 2011

Title

Three introns corroborate the chloroplast phylogeny of bamboos.

Authors

Jill R. Carskaddon[1], Sara B. Spangler[1], Tamara C. Richards[1], Amanda E. Fisher[1], Amy L. Denton[2], Scot A. Kelchner[1]

[1]Department of Biological Sciences, Idaho State University, Pocatello, Idaho

[2]Department of Biology, California State University Channel Islands, Camarillo, California

Abstract (196 words)

Bamboos (Poaceae: Bambusoideae) include between 1,200 and 1,400 species, and are found in tropical to temperate areas worldwide. The evolutionary history of bamboos has been difficult to infer with comparative sequence analysis: problems include rapid radiation events, relatively slow rates of sequence evolution, limited variation in trialed nuclear loci, and lack of convincing resolution at key nodes in gene topologies. Although some researchers advocate a phylogenomic approach for problematic taxa like bamboos, a reasonable (and inexpensive) alternative might be to target loci that are strong phylogenetic performers. Three such introns were recently identified by the authors using multilocus comparisons with a corroborated chloroplast topology of 14 bamboo species. In the present study, we resample the taxon set of the Bamboo Phylogeny Group (BPG) and estimate a chloroplast phylogeny using sequence data from the *atpF, trnG,* and *petB* introns. Although our combined data set is only one third the size of the BPG alignment, it recovers all but two of the supported branches in the BPG estimation. Our analysis provides the first independent corroboration of BPG's chloroplast phylogeny for bamboos, and suggests that these three introns could be more widely used for phylogenetic inference in the grasses.

Three introns corroborate a first estimate of bamboo phylogeny

Jill R. Carskaddon[1], Sara B. Spangler[1], Tamara C. Richards[1], Amanda E. Fisher[1], Amy L. Denton[2], and Scot A. Kelchner[1]

[1]Department of Biological Sciences, Idaho State University; [2]Department of Biology, California State University Channel Islands

Kelchner Lab

Background

- Bamboos, of which nearly 1,400 species are recognized, occur worldwide in temperate to tropical regions
- Bamboo phylogeny has been hindered by long branch attraction, short internodes, and heterogeneous rates of sequence evolution
- Recently, using nearly 10kb of sequence data, the Bamboo Phylogeny Group (BPG) has estimated the first chloroplast phylogeny to include all bamboo subtribes [1]
- Chloroplast group II introns have now been shown to provide high quality, rapidly evolving DNA sequences for accurate phylogeny estimation in bamboos [2]

Aim

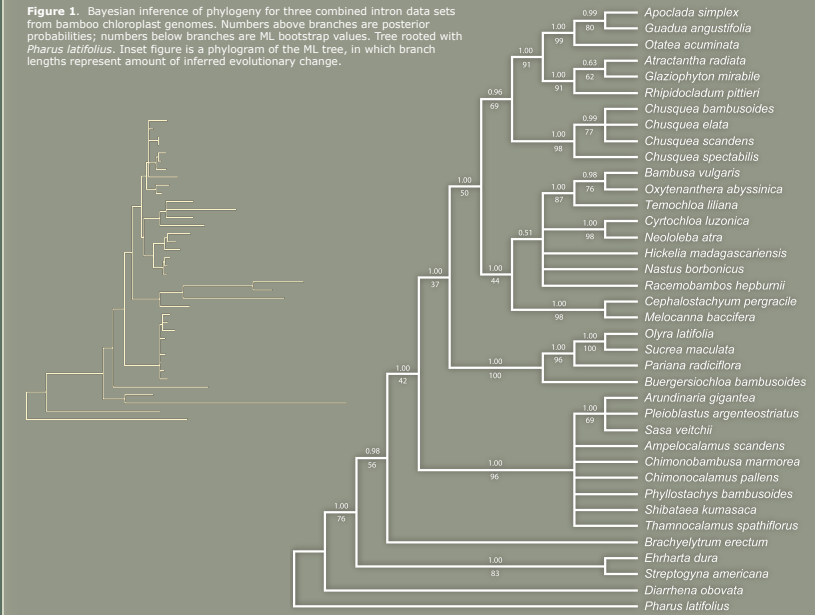
- Investigate whether 3 short plastome introns can provide sufficient DNA characters to corroborate evolutionary relationships among bamboo chloroplast genomes

Methods

- DNA was isolated from vouchered BPG leaf tissue of 33 bamboos and 5 outgroup taxa worldwide, representing each taxonomic tribe and subtribe
- Primer design, followed by PCR amplification of the group II introns in *atpF*, *petB,* and *trnG* genes of the chloroplast genome
- DNA sequencing, editing, and criterion based manual alignment with structure and mutation mechanisms [3]
- Model selection by dynamical LRT and AIC differences (delta ≤ 2) using jModelTest [4]
- Maximum likelihood (ML) analysis using PAUP* [5] with a best fit model; partitioned Bayesian inference using MrBayes [6] with unlinked parameters
- ML bootstrap analysis (1,000 pseudoreplications) using PhyML 3.0.1 [7]

Results

Figure 1. Bayesian inference of phylogeny for three combined intron data sets from bamboo chloroplast genomes. Numbers above branches are posterior probabilities; numbers below branches are ML bootstrap values. Tree rooted with *Pharus latifolius*. Inset figure is a phylogram of the ML tree, in which branch lengths represent amount of inferred evolutionary change.

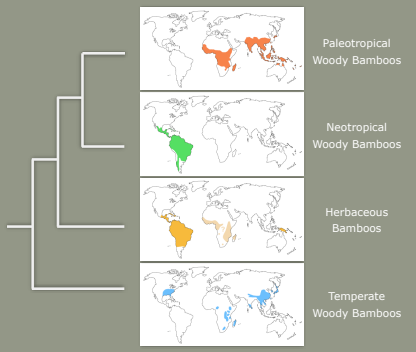


Rh pidocladum gem natum Photo by Lynn C ark | *Arundinar a g gantea* Photo rom USDA plants | *Raddia distichophy a* Photo by C eo Ca deron | *Phy ostachys v vax* Photo by J S Peterson | *Chusquea (Neuro epis) mo s* Photo by X mena Londoño

Table 1. Characteristics of each intron.

	Alignment Length	Variable Nucleotides	Potentially Informative	Scored Indels	Total Informative
atpF intron	756	119	49	5	51
petB intron	784	173	77	5	82
trnG intron	630	124	45	2	47

Figure 2. Geographic ranges of bamboo plastome lineages.



Conclusions

- 2kb of DNA data from chloroplast introns was able to resolve all major lineages of bamboos
- Our estimation is the first to replicate the BPG bamboo plastome phylogeny [1] using an independent data set
- *TrnG* intron was easiest to amplify and sequence; more difficult was the *atpF* intron; the *petB* intron is not recommended due to multiple internal mononucleotide repeats that complicate the sequencing efforts
- Corroboration of the BPG tree by our new data suggests that sources of bias (*e.g.* long branch attraction, heterotachy) might not significantly affect the plastome phylogeny of bamboos

Acknowledgements

Range maps of major bamboo lineages are from the Bamboo Biodiversity website (www.eeob.iastate.edu/research/bamboo/). Sequencing was performed by the Molecular Research Core Facility at Idaho State University. Lynn Clark and Jimmy Triplett provided leaf material for many of the DNA isolations in this study. Shannalee Hansen contributed DNA sequences to the alignment matrix. The work was supported by an NSF Research Experience for Undergraduates supplement to Jill Carskaddon and NSF Award 0515828 to Scot A. Kelchner.

Literature Cited

[1] Bamboo Phylogeny Group (in prep.); [2] Hansen (2010) *M.Sc. Thesis, Idaho State University;* [3] Kelchner (2000) *Ann Missouri Bot Gard* 87:482; [4] Posada (2008) *Mol Biol Evol* 25:1253; [5] Swofford (2002) *Sinauer*; [6] Huelsenbeck & Ronquist (2001) *Biometrics* 17:754; [7] Guindon & Gascuel (2003) *Syst Biol* 52:696