Molecular Phylogenetics and Evolution 67 (2013) 404–413

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

# Higher level phylogenetic relationships within the bamboos (Poaceae: Bambusoideae) based on five plastid markers

Scot A. Kelchner *, Bamboo Phylogeny Group [1]

*Department of Biological Sciences, Campus Stop 8007, Idaho State University, Pocatello, ID 83209-8007, USA*

ARTICLE INFO

*Article history:*
Received 2 July 2012
Revised 14 November 2012
Accepted 6 February 2013
Available online 20 February 2013

*Keywords:*
Bamboo
Plastid phylogeny
Exploratory data analysis
Noncoding DNA
*ndhF*
Microstructural characters

ABSTRACT

Bamboos are large perennial grasses of temperate and tropical forests worldwide. Two general growth forms exist: the economically and ecologically important woody bamboos (tribes Arundinarieae and Bambuseae), and the understory herbaceous bamboos (tribe Olyreae). Evolutionary relationships among the 1400 + described species have been difficult to resolve with confidence. Comparative analysis of bamboo plastid (chloroplast) DNA has revealed three to five major lineages that show distinct biogeographic distributions. Taxon sampling across tribes and subtribes has been incomplete and most published data sets include a relatively small number of nucleotide characters. Branching order among lineages is often poorly supported, and in more than one study herbaceous bamboos form a clade within the woody bamboos. In this paper, the Bamboo Phylogeny Group presents the most complete phylogeny estimation to date of bamboo tribes and subtribes using 6.7 kb of coding and noncoding sequence data and 37 microstructural characters from the chloroplast genome. Quality of data is assessed, as is the possibility of long branch attraction, the degree of character conflict at key nodes in the tree, and the legitimacy of three alternative hypotheses of relationship. Four major plastid lineages are recognized: temperate woody, paleotropical woody, neotropical woody, and herbaceous bamboos. Woody bamboos are resolved as paraphyletic with respect to Olyreae but SH tests cannot reject monophyly of woody species (Arundinarieae + Bambuseae).

Published by Elsevier Inc.

## 1. Introduction

Bamboos are remarkably robust forest grasses that number more than 1400 described species in 115 genera (Bamboo Phylogeny Group [BPG], 2012). Most familiar and useful are those with "woody" (lignified) stems that belong to the tribes Arundinarieae (temperate woody bamboos) and Bambuseae (tropical woody bamboos). The roughly 1300 woody species often play critical roles in the ecology of their forest habitats and have long been of economic importance to humans (see McClure, 1966; Judziewicz et al., 1999). Less popularly known are the herbaceous bamboos (tribe Olyreae), a lineage of about 120 non-woody understory species found in tropical forests, predominantly in the New World. Together the three tribes constitute the grass subfamily Bambusoideae, members of which share the unique character of asymmetrically invaginated arm cells in the chlorenchyma (Zhang and Clark, 2000). Most species also possess fusoid cells in the mesophyll, papillae on at least the abaxial leaf epidermis, and a broad leaf blade with a basal constriction called a pseudopetiole (GPWG, 2001; Judziewicz and Clark, 2007; BPG, 2012).

Members of Bambusoideae are found from sea level to over 4000 m in mainly forest and high montane grassland habitats worldwide, except for Europe and Antarctica (Fig. 1). At least 40% of species (including many awaiting formal description) are endemic to the Americas. In both the Americas and Asia, several taxa can reach gigantic proportions: plant height can be to 40 m and clonal individuals can occupy many hectares (Judziewicz et al., 1999). Famously, woody bamboos can take between 7 and 120 years to flower, often in gregarious events that lead to mass seed production and an accompanying dieback of parent plants. Because of this rarity of flowering, woody bamboos are usually encountered in their vegetative stage and identification of species can be difficult for a non-specialist.

Bambusoideae is one of three subfamilies in the BEP clade of grass family Poaceae (GPWG, 2001; GPWG II, 2012). Although each is monophyletic in molecular analyses, branching order of the BEP subfamilies has been difficult to resolve. Several recent analyses (e.g., Bouchenak-Khelladi et al., 2008; Davis and Soreng, 2010; GPWG II, 2012; Wu and Ge, 2012) support a sister relationship between Bambusoideae (B) and Pooideae (P; the bluegrasses), with

* Fax: +1 208 282 4570.
*E-mail address:* kelchner@isu.edu

[1] Bamboo Phylogeny Group members: Lynn Clark, Gilberto Cortés, Reyjane P. Oliveira, Soejatmi Dransfield, Tarciso Filgueiras, Amanda E. Fisher, Gerald F. Guala, Trevor Hodkinson, Emmet Judziewicz, M. Kumar, De-Zhu Li, Ximena Londoño, M. Teresa Mejia, Ana Paula Santos-Gonçalves, Chris Stapleton, Sarawood Sungkaew, Jimmy Triplett, Elizabeth Widjaja, Khoon Meng Wong, Nian-He Xia.

1055-7903/$ - see front matter Published by Elsevier Inc.
http://dx.doi.org/10.1016/j.ympev.2013.02.005

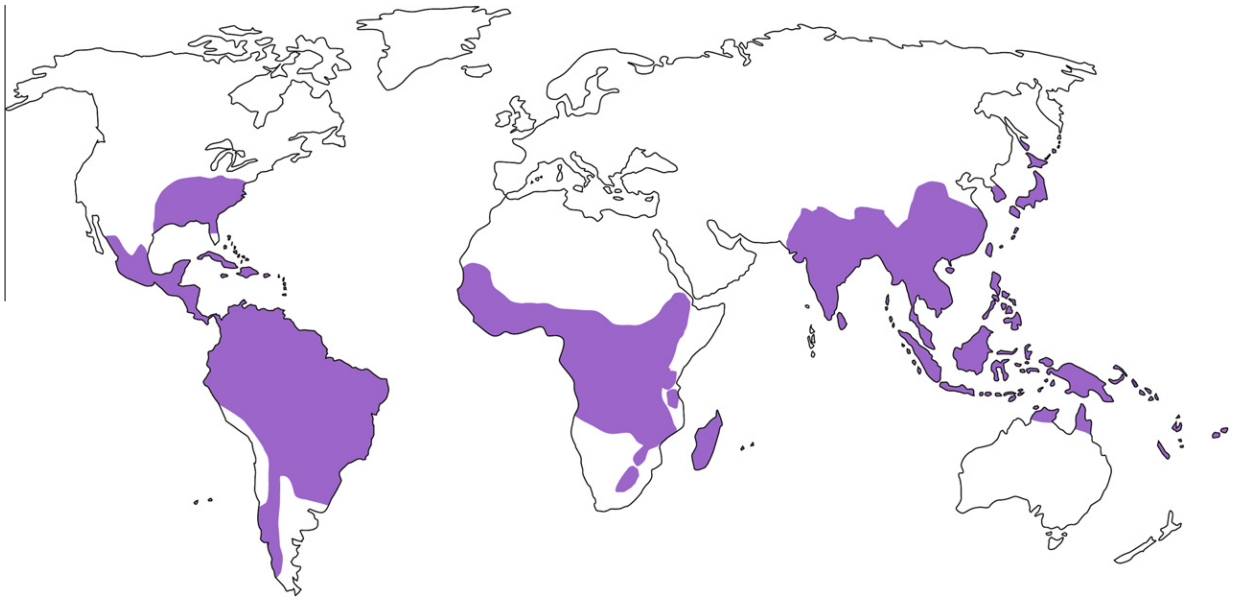

**Fig. 1.** World distribution of bamboos (Poaceae: Bambusoideae).

the B + P clade being sister to Ehrhartoideae (E; the rices). Branching order of the three clades varies across studies, possibly as a function of outgroup sampling.

Within Bambusoideae, three to five major lineages are recovered by comparative DNA sequence analyses (Clark et al., 1995, 2007; Kelchner and Clark, 1997; Zhang and Clark, 2000; Bouchenak-Khelladi et al., 2008; Sungkaew et al., 2009) despite the often marked dissimilarity of taxa and genetic loci sampled. Temperate woody bamboos form the most robust of these lineages, appearing as an uninterrupted branch in molecular phylogenies that suggests a long evolutionary history followed by recent and rapid radiation (Hodkinson et al., 2010). Herbaceous bamboos are also clearly monophyletic and at many loci show rates of sequence evolution much higher than those of other bamboos and more similar to grasses in general (Gaut et al., 1997). The tropical woody bamboos form between one and three main lineages; they are most commonly resolved as two (paleotropical woody; neotropical woody) or three (paleotropical woody; Arthrostylidiinae + Guaduinae; Chusqueinae). Rate of sequence evolution in subtribe Chusqueinae also appears to be accelerated, though to a lesser degree than the herbaceous lineage.

Since the first comparative DNA sequence analysis of bamboos by Kelchner and Clark (1997), more than 20 molecular phylogenetic analyses have been published on taxonomic subsets of Bambusoideae (reviewed by BPG (2012)). In that time, several phenomena that complicate phylogeny estimation have been observed in most bamboo nuclear and plastid (chloroplast) DNA data sets. These include strong heterogeneity of sequence evolution rates among lineages, poor support for internodes that separate major clades, long branch attraction that can affect both ingroup and outgroup topology, lack of strict congruence among trees from different studies (or data sets from different genomic compartments), and problems with attaining supported resolution in certain lineages such as the temperate woody and paleotropical woody clades.

Molecular phylogeny estimation in bamboos is thus widely considered to be a difficult problem. Maximum parsimony has been the most common optimization criterion used for tree selection in bamboo molecular systematics, yet it is well known that parsimony can mislead phylogeny estimation when the above phenomena are a feature of the data set and its analysis. Recent studies have used Bayesian inference together with parsimony (e.g., Yang et al., 2007; Peng et al., 2008; Sungkaew et al., 2009; Triplett and Clark, 2010), which at least provides an opportunity to test robustness of a phylogeny estimation to certain changes in the model of character evolution employed. Frequently the two analytical frameworks do not produce the same topology (e.g., Yang et al., 2007; Peng et al., 2008), suggesting that the problematic issues listed here for bamboo molecular studies can indeed affect the estimate of bamboo phylogeny.

In response to the challenge of generating a well sampled, data rich, and carefully tested molecular phylogeny of Bambusoideae, the Bamboo Phylogeny Group was formed to better coordinate efforts among many of the world's bamboo systematists (BPG, 2006). Molecular studies had already suggested that a taxonomic reorganization of bamboos was inevitable; BPG members sought to align that reclassification to a strongly supported estimate of phylogeny. The endeavor was to begin with a phylogeny estimation of chloroplast genomes (plastomes) in bamboos because the task could be readily addressed with existing collections of silica-gel-dried leaf tissue from which sequence data of chloroplast loci are easily obtained. Taxon sampling was designed to include all subtribes of bamboos, and DNA sequences would be taken from five plastome loci. Phylogenetic analyses were to employ several available tests that determine data quality, tree stability, and whether bias and error were negatively affecting the phylogeny estimation.

In this paper, the Bamboo Phylogeny Group presents its first rigorously tested plastome phylogeny estimation. The estimate represents the breadth of taxonomic, geographic, and morphological diversity in the subfamily Bambusoideae. Aims of this study include (i) improvement of resolution and support for commonly recognized lineages, (ii) identification of previously unknown relationships, (iii) establishment of branching order among major clades, (iv) stabilization of the plastid phylogeny estimation, and (v) identification of remaining areas of concern in the topology that might hinder further progress in deciphering the biogeography, history, and evolution of bamboos.

## 2. Material and methods

### 2.1. Taxon sampling

A total of 40 species, including 33 ingroup (Bambusoideae) and seven outgroup (Ehrhartoideae, Pharoideae, Pooideae) taxa, was

analyzed (Table 1). Ingroup species were chosen to represent the three major lineages of bamboos now recognized as tribes (Arundinarieae, Bambuseae and Olyreae), and each of the currently or, in the case of Arundinarieae, previously recognized, subtribes of Bambusoideae (BPG, 2012). With the exceptions of Parianinae, Racemobambosinae and the monotypic Buergersiochloinae, each subtribe was represented by two to four species. *Neurolepis* taxa are treated here as *Chusquea* after Fisher et al. (2009). In addition to representing taxonomic diversity, the particular species sampled in this study were also chosen to represent as much morphological diversity as possible, given samples available at the time. The seven outgroup taxa consisted of three species each from BEP clade subfamilies Pooideae and Ehrhartoideae (including a sister taxon to the main lineage of each subfamily), and one species from the Pharoideae, an early diverging Poaceae lineage that is sister to most of the remaining grasses (GPWG, 2001; GPWG II, 2012). Leaf material was collected in the field or from greenhouse-grown specimens and silica-gel dried tissue (Chase and Hills, 1991) except for *Oryza sativa* and *Triticum sativum* for which sequences were downloaded from GenBank (Table 1).

### 2.2. DNA sequencing and alignment

Five plastid DNA regions were used in this analysis: one gene (*ndhF*), two group II introns (*rpl16* intron, *rps16* intron), and two intergenic spacers (*trnD–trnT*, *trnT–trnL*). Total genomic DNA extractions were performed with Qiagen DNeasy Plant Mini Kits (Qiagen, Valencia, USA) on vouchered silica gel-dried leaf tissue. Isolations were cleaned using Qiagen PCR Purification Columns and then quantified with fluorometry. Target regions were PCR-amplified in the laboratories of Clark and Kelchner. Each automated sequence file was reviewed for base-calling errors, poor sequence reads, and multiple peaks. Unusual sequences were verified by conducting a second DNA isolation, PCR amplification, and sequence from the original vouchered leaf materials. BLAST searches were used to confirm probable homology of each target sequence.

Regions were aligned individually by process partition using a criterion-based manual methodology that infers secondary structures and mutational mechanisms to then inform the positioning of gaps (Kelchner and Clark, 1997; Kelchner, 2000). Specific criteria for choosing insertions, deletions, and hairpin inversions (Kelchner and Wendel, 1996) as scored microstructural characters followed Kelchner (2000) and Löhne and Borsch (2005). Gap placements in the combined alignment are staggered to reflect hypothesized independent insertion and deletion events (Morrison, 2006, 2009). This approach allowed us to recover nucleotide substitution characters within length mutations that had reasonable biological evidence of homology. Insertions and deletions that involved mononucleotide repeats were not coded as microstructural characters.

### 2.3. Data quality and model selection

Ambiguous characters in each alignment were included for pairwise characterization of sequence structure (base frequency and sequence length comparisons) but were excluded prior to phylogenetic analyses of individual data sets following Morgan and Kelchner (2010). In the combined data set, ambiguous regions were deleted entirely from the matrix prior to phylogeny estimation for the purposes of decreasing the overall quantity of missing data and for facilitating maximum likelihood and network analyses across multiple software platforms.

Base composition homogeneity was assessed for individual and combined data sets by the chi-squared test of base frequency inequalities implemented in PAUP* 4.0b10 (Swofford, 2002). PAUP* was also used to measure variation in uncorrected pairwise

**Table 1**
Taxon samples used in this study, with classification and voucher information provided. Subtribe nomenclature, when previously contradictory, follows that of BPG (2012).

| Ingroup taxa | |
|---|---|
| Tribe Arundinarieae | |
| *Ampelocalamus scandens* J.R. Xue and W.D. Li | (LC 1291) |
| *Arundinaria gigantea* (Walter) Muhl. | (JT 197) |
| *Chimonobambusa marmorea* (Mitford) Makino | (JT 69) |
| *Chimonocalamus pallens* J.R. Xue and T.P. Yi | (JT238) |
| *Phyllostachys bambusoides* Siebold and Zucc. | (LC 1289, JT 121) |
| *Pleioblastus argenteo-striatus* (Regel) Nakai | (JT 66) |
| *Sasa veitchii* (Carriere) Rehder | (JT 126, LC 1325) |
| *Shibataea kumasaca* (Zoll. ex Steud.) Makino | (LC 1290) |
| *Thamnocalamus spathiflorus* (Trin.) Munro | (LC 1319) |
| Tribe Bambuseae | |
| Subtribe Arthrostylidiinae | |
| *Atractantha radiata* McClure | (ASG 599, AMC 4362) |
| *Glaziophyton mirabile* Franch. | (LSS 1066) |
| *Rhipidocladum pittieri* (Hack.) McClure | (LC 1349, LC and WZ 1349) |
| Subtribe Bambusinae | |
| *Bambusa vulgaris* Schrad. ex. J.C. Wendl. | (GSK 666) |
| *Cyrtochloa luzonica* (Gamble) S. Dransf. | (SD 1323) |
| *Neololeba atra* (Lindl.) Widjaja | (LC and JT 1663) |
| *Oxytenanthera abyssinica* (A. Rich.) Munro | (LC and JT1664) |
| *Temochloa liliana* S. Dransf. | (SD 1494) |
| Subtribe Chusqueinae | |
| *Chusquea bambusoides* (Raddi) Hack. | (LC 1029) |
| *Chusquea elata* (Kunth) L.G. Clark (=*Neurolepis elata*) | (LC and PA1409) |
| *Chusquea scandens* Kunth | (LC and XL 1235) |
| *Chusquea spectabilis* L.G. Clark (=*Neurolepis aperta*) | (LC 919) |
| Subtribe Guadinae | |
| *Apoclada simplex* McClure and L.B. Sm. | (LC and WO 1027) |
| *Guadua angustifolia* Kunth | (LC and XL 931) |
| *Otatea acuminata* (Munro) C. Calderon and Soderstr. | (LC and WZ 1348, LC et al. 1312) |
| Subtribe Hickeliinae | |
| *Hickelia madagascariensis A.* Camus | (SD 1290, SD 1292) |
| *Nastus borbonicus* J.F. Gmel. | (LC and SD 1656) |
| Subtribe Melocanninae | |
| *Cephalostachyum pergracile* (Munro) R.B. Majumdar | (WZ 8400635, SD 1435) |
| *Melocanna baccifera* (Roxb.) Kurz | (XL and LC 930) |
| Subtribe Racemobambosinae | |
| *Racemobambos hepburnii* S. Dransf. | (WKM2891) |
| Tribe Olyreae | |
| Subtribe Buergersiochloinae | |
| *Buergersiochloa bambusoides* Pilg. | (SD 1365, SD 1382) |
| Subtribe Olyrinae | |
| *Olyra latifolia* L. | (XL and LC911) |
| *Sucrea maculata* Soderstr. | (LC and WZ 1345) |
| Subtribe Parianineae | |
| *Pariana radiciflora* Doell in Martius | (LC and WZ 1344) |
| Outgroup taxa | |
| Ehrhartoideae | |
| *Ehrharta dura* Nees ex Trin. | (NB 1118) |
| *Oryza sativa* L. | (GenBankNC 0013201) |
| *Streptogyna americana* C.E. Hubb. | (RP and GD 12310, GSK 657) |
| Pharoideae | |
| *Pharus latifolius* L. | (LC 1302) |
| Pooideae | |
| *Brachyelytrum erectum* (Schreb.) P. Beauv. | (JT 199) |
| *Diarrhena obovata* (Gleason) Brandenberg | (LC and WZ 1216, JT 290) |
| *Triticum aestivum* L. | (GenBankNC 0027621) |

*Voucher abbreviations:* AMC – Andre Mauricio Carvalho; ASG – Ana Paula Santos-Gonçalves; GD –Gerrit Davidse; GSK – Gabriel Sanchez-Ken; JT – Jimmy Triplett; LC – Lynn Clark; LSS – Luis Sergio Sarahyba; NB – Nigel Barker; PA – Patricio Asimbaya; RP – Richard Pohl; SD – Soejatmi Dransfield; WKM – Wong Khoon Meng; WO – Walter de Oliveira; WZ – Weiping Zhang; XL – Ximena Londoño.

distances among sequences and the number of potentially informative substitutions observed in each alignment matrix. Character saturation was assessed by plotting uncorrected pairwise distances against model-corrected distances (Philippe et al., 1994), a method that can be use for noncoding DNA because it does not rely on codon position. Model-corrected distances were produced using a best-fit model for the combined data set (discussed below).

For phylogeny estimations using probability frameworks (Section 2.4), a model of character evolution had to be specified. Both the hierarchical likelihood ratio test (hLRT) and the Akaike Information Criterion (AIC) were used to ascertain model adequacy among the 56 models available for comparison in ModelTest (Posada and Crandall, 1998). The hLRT method might be biased due to its hierarchical approach (Kelchner and Thomas, 2007), but all hLRT results were compared to models that show reasonable evidence of being "best fit" by AIC (i.e., those models that have AIC differences of $\Delta \leqslant 2$; the rationale follows Posada and Buckley (2004) and Kelchner (2009)).

Conflict among character state changes in the data was evaluated in two ways. First, the data were visualized with a neighbor-net analysis using the program SplitsTree 4 (Huson and Bryant, 2006). Networks of uncorrected and model-corrected distances were surveyed to identify sequences involved in large reticulations that indicate character conflict (Morrison, 2005, 2010). Second, nonparametric bootstrap analyses were performed on each data set using maximum likelihood in PhyML 3 (Guindon and Gascuel, 2003). The best-fit model was set for each data partition with all parameter values estimated in 1000 pseudoreplicates. Results were surveyed for bootstrap-supported incongruence among trees recovered for each sequence region. In one case (*rpl16* intron), the Shimodaira–Hasegawa test (SH test, Shimodaira and Hasegawa (1999)) was used to determine if an observed incongruence with bootstrap support of 80 was significantly different to the combined nucleotide phylogeny estimation when given the *rpl16* intron data and three different models of character evolution. The SH test was performed with PAUP* using RELL and 10,000 pseudoreplicates.

### 2.4. Tree estimation

Three analytical frameworks were applied to establish whether the phylogeny estimation for Bambusoideae plastomes was robust to reasonable changes in assumptions about character evolution (Penny et al., 1992; Kelchner and Thomas, 2007). Two of the frameworks, maximum parsimony (MP) and Bayesian inference (BI), could be applied to both nucleotide and microstructural characters. The third framework, maximum likelihood (ML), is most efficiently used for phylogeny estimation from nucleotide data. Therefore, only nucleotide characters were used for topology comparisons that evaluated robustness across frameworks. In the estimations produced by MP and BI from combined data, scored microstructural changes were included at the end of the alignment matrix as 0/1 standard characters.

Equal weight MP analyses (sensu Fitch, 1971) were conducted in PAUP* using a rapid, suboptimal search strategy in which 1000 replicates of random taxon-addition starting trees produced a pool of 100,000 topologies from which a secondary TBR search was swapped to completion and a strict consensus of the most parsimonious trees was produced. This strategy provided a computationally tractable way to perform nonparametric MP bootstrap analyses using PAUP* (1000 pseudoreplicates). The most parsimonious topology was not sought for or required by our assessment of robust signal across analytical frameworks.

ML analysis was performed in PAUP* when the aim was to obtain an optimal (i.e., maximized likelihood estimate, or MLE) topology for the individual and combined data sets. The program PhyML was used for nonparametric ML bootstrapping (10,000 pseudoreplicates). In both programs, the model was designated to be one of the best fit candidate models for the data set, with model parameter values estimated during analysis and microstructural characters excluded. Each PAUP* MLE search was conducted twice, and MLE topologies were used primarily for SH testing of alternative hypotheses. Although PhyML might not be an optimal search algorithm for MLE topologies, it is computationally inexpensive and is likely to be conservative when used for nonparametric bootstrapping: if some pseudoreplicates fail to produce truly optimal estimates, then the spurious topological conflict over ten thousand repetitions would slightly lower bootstrap support values.

BI was conducted with MrBayes 3.1 (Huelsenbeck and Ronquist, 2001) using a partitioned GTR + I + G model for reasons outlined by Huelsenbeck and Rannala (2004) with all parameter values estimated during analysis. A Dirichlet prior was used for base frequencies and the rate matrix. A uniform prior was used for the shape parameter ($\alpha$), proportion of invariable sites (I), and topology. Branch lengths were unconstrained. Partitions were designated for each data set and for the microstructural characters and all parameters were unlinked across partitions. Four separate analyses were performed with four runs of four chains each that continued for 20 million generations. Posterior probabilities were calculated using the final 10 million generations in each run, with sampling every 100 generations; this cut-off value was computationally efficient and it conservatively satisfied a rough convergence criterion that average standard deviation of split frequencies should stabilize below 0.01.

### 2.5. Rooting and potential long branch attraction

Estimated phylogenies were rooted with *Pharus latifolius*. *P. latifolius* is, however, one of many long branches in this study that attaches to a short internode in each tree. To survey for potential error caused by long branch attraction (LBA) in our analyses, taxon removal experiments were performed in PAUP* with combined nucleotide data using ML (model GTR + I + G). The following taxon removals were tested individually: (i) Olyreae; (ii) all outgroup taxa except *Streptogyna americana*; (iii) Pooideae; (iv) *Triticum aestivum*; (v) Ehrhartoideae; (vi) *Oryza sativa*; (vii) Chusqueineae; and (viii) *Streptogyna americana*. The MLE topology from each analysis was checked for changes in relationship among remaining taxa when rooted with *P. latifolius*. The removal of *Streptogyna americana* also included a nonparametric ML bootstrap analysis (1000 pseudoreplicates) because of its high level of character conflict at key nodes in the SplitsTree neighbor-net graph.

### 2.6. Testing of alternative hypotheses

Three ML constraint analyses of combined nucleotide data were conducted with PAUP*. In the first analysis, all woody bamboos (Arundinarieae and Bambuseae) were required to form a single clade that was sister to herbaceous bamboos (Olyreae). In the second analysis, Ehrhartoideae was constrained to be sister to Bambusoideae. In the third analysis, Chusqueinae was forced to be sister to a clade of all other tropical woody bamboos (Paleotropical + Arthrostylidiinae + Guaduinae). A fourth analysis involved the incongruent topology from the *rpl16* intron data set. Constraint MLE topologies were tested individually against the unconstrained MLE topology of combined nucleotide data in a two–tree SH test (sensu Buckley et al., 2001) that maintains excellent type 1 error control yet avoids the conservative behavior the test shows in multiple tree comparisons (Shimodaira, 2002); in this form, the test reduces to a KH test with correction for a *posteriori* topologies (Shimodaira and Hasegawa, 1999; Goldman et al., 2000). All SH tests were performed with PAUP* and used 10,000 RELL replicates.

Because SH test results can vary with model choice (Buckley et al., 2001), each test was performed three times using the models JC, HKY + G, and GTR + I + G.

## 3. Results

### 3.1. Sequence alignments and data quality

Sequence recovery from the five plastid regions was generally straightforward. PCR amplifications resulted in single bands and the sequence files showed no problematic double peaks. Of the GenBank sequences downloaded for the analysis, the *rps16* intron in *T. aestivum* appeared to be inaccurately labeled (see below, this section) and was scored as missing data in the combined analyses. GenBank accession numbers for sequences generated by this study are KC020491–KC020545, and KC020547–KC020602; a table of GenBank accession numbers by sequence region and taxon is included in the online Supplementary material.

Alignment length (number of characters) by region was 2120 for *ndhF*, 1364 for *rpl16* intron, 1006 for *rps16* intron, 1465 for *trnD–trnT* spacer, and 1066 for *trnT–trnL* spacer. Percent informative characters for all regions was between 10% and 11% except for the *rps16* intron (6.96%). The combined alignment with ambiguous regions removed was 6657 nucleotide characters with 10.03% informative characters. Microstructural changes for which reasonable estimation of homology could be made totaled 37: two in *ndhF* (including one hairpin inversion), 13 in *rpl16* intron (including one hairpin inversion), three in *rps16* intron, nine in *trnD–trnT* spacer (including one hairpin inversion, with an additional hairpin inversion located in a deleted ambiguous region), and ten in *trnT–trnL*. A list of character types and scoring across taxa is included in Supplementary material.

None of the alignments showed significant differences in base composition among taxa. Saturation plots indicated slight deviation when model correction was applied, but the deviation was not strong (Supplementary material) and there was no obvious plateau to indicate problematic saturation of the noncoding data (e.g., Simon et al., 2009). Neighbor-net analysis of uncorrected and model-corrected pairwise distances showed considerable character conflict among outgroup sequences, particularly in relation to *S. americana* and the herbaceous bamboo lineage Olyreae. The network also allowed for the identification of an incorrect sequence for the *rps16* intron in *T. aestivum*, which was subsequently removed from the alignments. A table of data set characteristics is included in Supplementary material.

Phylogeny estimations of individual data sets showed no well supported incongruence among region signal except in the *rpl16* intron for the clade *Rhipidocladum pittieri* + *Glaziophyton mirabile*. However, an SH test showed no significant difference between this arrangement and the combined data topology when given *rpl16* intron data and three models of character evolution (Supplementary materials). The clade is also not present in the MLE topology of *rpl16* intron data produced by PAUP*. These observations, as well as the general lack of character conflict among ingroup sequences in the neighbor-net analysis and the relatively low bootstrap values on incongruent nodes from each region, suggested that the same tree-like signal was present among data sets and that combination of data was acceptable.

### 3.2. Phylogeny estimations

Models selected for each data set, together with relevant AIC information, are presented in Supplementary material. Each region had GTR + G or GTR + I + G as one of its best-fit models from the model candidate pool according to the criterion of AIC differences $\Delta \leqslant 2$. Hierarchical likelihood ratio test results were either TVM + G or TVM + I + G for all data sets. The combined matrix had both GTR + I + G and TVM + I + G as potentially best-fit using the same AIC and hLRT criteria. ML analyses of each data set, therefore, used the following models: GTR + I + G for *ndhF*, GTR + G for *rpl16* intron, GTR + G for *rps16* intron, GTR + G for *trnD–trnT* spacer, GTR + G for *trnT–trnL* spacer, and GTR + I + G for combined nucleotide data.

ML bootstrap topologies for each region, with minor incongruences identified, are presented in Supplementary material. Although the percent of informative characters is nearly equal for all regions except *rps16* intron, *ndhF* gave good resolution throughout its tree whereas noncoding regions provided improved resolution at mid-levels of divergence (e.g., among the four major lineages of Bambusoideae). The position of Olyreae is dependent on noncoding DNA data in this study.

The *ndhF* phylogeny estimation was congruent with the topology of the combined nucleotide analysis. Each noncoding region, however, showed between two and five minor incongruences. All but one of these had low bootstrap support (between 50 and 69) on generally very short branches. The case of the *rpl16* intron incongruence with a bootstrap value of 80 is discussed above and was ruled out as being significant conflicting signal. *Rps16* intron showed the most incongruence (five nodes) but also had the least number of potentially informative characters and the least number of resolved nodes with bootstrap support of $\geqslant 80$. Similarly supported nodes in other data sets totaled 15 in *ndhF*, 13 in *rpl16* intron, 15 in *trnD–trnT*, and only nine in *trnT–trnL* despite its 10.88% of potentially informative characters. Microstructural characters resolved 14 nodes in the MP consensus tree of 641 most parsimonious topologies (Supplementary material) with the single incongruence of *Atractantha radiata* + *R. pittieri*; the tree has a length of 83, a consistency index of 0.446, a homoplasy index of 0.554, and a retention index of 0.711.

The independent phylogeny estimations of *rps16* intron, *trnD–trnT*, and *trnT–trnL* data place Olyreae within Bambusoideae, but *ndhF* and *rpl16* intron data recover only a polytomy. *Rpl16* intron and *trnT–trnL* place Ehrhartoideae sister to Bambusoideae with low bootstrap support (69 and 57, respectively). This differs from *ndhF*, *trnD–trnT* and the combined data which place Pooideae as sister to Bambusoideae with ML bootstrap support below 67. Bootstrapping of *rps16* intron data produced a polytomy involving these three lineages.

Phylogenetic analysis of combined nucleotide characters produced a congruent topology across MP, ML and BI frameworks (Fig. 2a). The topology was robust to all models of character evolution trialed. Four main lineages are present within Bambusoideae: the Paleotropical Woody clade (subtribes Bambusinae, Racemobambosinae, Hickeliinae, Melocanninae), the Neotropical Woody clade (subtribes Chusqueinae, Arthrostylidiinae, Guaduinae), the Herbaceous clade (subtribes Buergersiochloinae, Olyrinae, Parianinae), and the Temperate Woody clade (tribe Arundinarieae). Pooideae is weakly supported as sister to Bambusoideae (ML bootstrap of 67, posterior probability of 0.94). Twenty-seven nodes were recovered with ML bootstrap support of $\geqslant 80$ although five nodes within the Paleotropical Woody clade and the Temperate Woody clade show either very low bootstrap support or low posterior probabilities. MP and BI analyses that included microstructural characters with the combined nucleotide data gave the same topology as Fig. 2a (see Supplementary materials).

### 3.3. LBA and alternative hypothesis testing

Taxon removal experiments to survey for long branch attraction resulted in no changes to ingroup or outgroup topologies. MLE topology estimates from each taxon removal data set show only one minor difference: the position of *Racemobambos hepburnii*

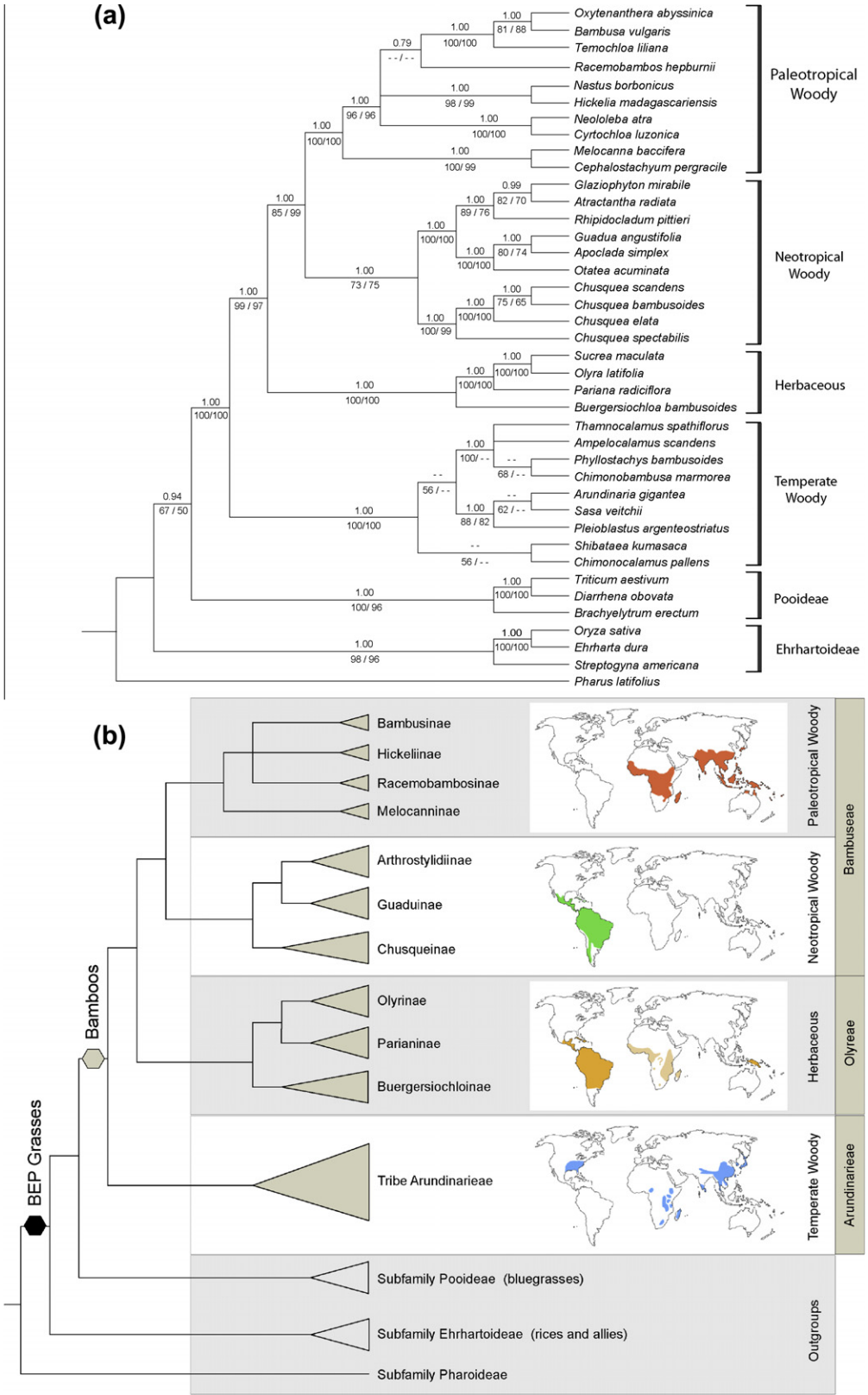


**Fig. 2.** Phylogeny estimation and geographic distribution of major bamboo lineages based upon plastid (chloroplast) DNA sequences from five loci: *ndhF*, *rpl16* intron, *rps16* intron, *trnD–trnT* intergenic spacer, and *trnL–trnF* intergenic spacer. (A) Consensus posterior topology of the partitioned BI analysis of combined nucleotide and microstructural characters. The topology is congruent with the MP strict consensus topology of the same data set and the ML topology of nucleotide data alone. Posterior probabilities are shown above branches. ML bootstrap values are listed first below each branch, followed by MP bootstrap values. Paired dashes indicate support values below 50 (bootstraps) or 0.5 (posterior probabilities). The ML estimate of topology shows additional resolution (Fig. 4). (B) Subtribe relationships and geographic distributions of the four major lineages of bamboos recovered in this study. Clade order follows that of (A). Question mark in the Herbaceous diagram represents uncertainty about the native status of *Olyra latifolia* in Africa and Madagascar.

usually collapses into a polytomy with other members of the Paleotropical Woody clade (see Supplementary materials). Pooideae remains sister to Bambusoideae in all taxon removal experiments which included members of that lineage, when *P. latifolius* is used to root each topology.

Two–tree SH tests of three alternative hypotheses of bamboo and outgroup relationships failed to reject any alternative constraint MLE topology for each of three character evolution models (attained significance $p = 0.05$). A table of SH test results can be found in Supplementary materials.

## 4. Discussion

Bamboos have long been a problematic plant group to examine with molecular phylogenetic techniques. Rate heterogeneity among lineages, lack of sequence variation in certain genera and subtribes, and short internodes deep within the topology have prevented consistent and well supported resolution of branching order among and within major bamboo clades. Despite the difficulty of comparative sequence analysis in Bambusoideae, we have been able to use 6.7 kb of high quality nucleotide data and 37 microstructural changes to produce a robust and rigorously tested phylogeny estimation for bamboo chloroplasts. Nuclear data will be required before an organismal phylogeny of bamboos can be established with greatest confidence, but the generation of a reliable chloroplast phylogeny is an essential step in that process. We therefore present an evaluation of our phylogeny estimation, its strengths and weaknesses, and corresponding evidence from published literature that allow us to assess whether this topology should be considered a best estimate of higher level relationships among bamboo plastid genomes.

### 4.1. Strengths of the phylogeny estimation

A worldwide effort of Bamboo Phylogeny Group members to attain leaf tissue samples for this study has improved upon previous efforts of taxon sampling and vouchering of materials for bamboo molecular analysis. The phylogeny estimation represents the greatest breadth to date of bamboo taxonomic, morphological and geographic diversity to be included in a single analysis. The estimate is derived from data representing multiple process partitions in the plastid genome, including one fast-evolving gene, two group II introns, and two intergenic spacers. Congruence of supported tree resolution among each of those partitions indicates that the data contains one dominant signal. Phylogeny estimation from microstructural characters mirrors the estimation derived from nucleotide characters. These observations suggest that contradictions in the data to this dominant signal are minor and are likely due to conflict among homoplastic characters.

Potential bias in the phylogeny estimation was not detected with available tests for error. There is no evidence of base composition inequalities among sequences and lineages, and no problematic site saturation. The topology is robust to moderate changes in the model of character evolution (sensu Kelchner and Thomas, 2007) including the analytical framework (MP, ML, BI). Taxon removal experiments cause no alteration of ingroup topology that would indicate the misleading effect of long branch attraction.

With nearly 6.7 kb of nucleotide data showing 10–11% sequence variation among taxa, posterior probabilities and bootstrap values are high on most nodes of the Bambusoideae phylogeny estimation. This observation indicates that more sequence data from the chloroplast genome in these specimens is likely to show the same branching order among well supported clades. Further sequencing of chloroplast regions, if bias and poor data quality are minimal, is therefore unlikely to alter our estimation of higher-level plastid relationships among the bamboos. An exception might be in the resolution of monophyly of neotropical bamboos (Section 4.2). Additional independent sequencing efforts have, encouragingly, provided corroborative evidence of these results (Kelchner et al., in prep.).

### 4.2. Weaker nodes of the tree

Despite many positive signs that our topology is robust and well supported, there remain a few areas of concern in the chloroplast phylogeny estimation that warrant further consideration. A result that justifies caution is the failure to reject alternative hypotheses of relationships in the three cases tested. Although the SH test can be unnecessarily conservative when several topologies are compared at once, the test is proper to use when the number of candidate topologies is very small and it does give better type 1 error control than the AU test (Shimodaira, 2002). In our case only two trees were compared in each test, conditions under which the SH test reduces to a KH test (Kishino and Hasegawa, 1989) with correction for the comparison of *a posteriori* topologies.

One possible explanation for the failure to reject alternative hypotheses can be found in the neighbor-net network of gap-free, model-corrected sequence data (Fig. 3) and in the branch lengths of the MLE phylogram of combined nucleotide data (Fig. 4). Substantial character conflict exists in the network, represented in the diagram as boxes outlining alternative edges of character support. The most important ingroup relationship affected by this conflict is the edge between temperate woody bamboos and herbaceous bamboos, a part of the network to which all outgroup sequences join. Although this character conflict might arise simply from stochastic parallelism and reversal of nucleotide changes, it is problematic enough that bootstrap-supported resolution still represents indecisive signal in SH tests. The conflict allows for a possibility that outgroup sequences are misrooting to the temperate clade sequences, which produces an apparent paraphyly of the woody bamboos that is due to analytical error and not evolutionary history.

A second problem spot is the monophyly of Neotropical woody bamboos, a relationship that is likely to be unstable due to a paucity of unique character variation at the node. Again the issue is present despite bootstrap values of 73 (ML) and 75 (MP), and posterior probability (1.00) for the clade (Fig. 2). The branch is comparatively short in the MLE phylogram of combined nucleotide data (Fig. 4) and involves a clade (Chusqueinae) with faster than normal rates of sequence evolution for woody bamboo chloroplasts. An SH test failed to reject the possibility that Neotropical woody bamboos are paraphyletic with respect to the Paleotropical Woody clade, which allows for a reasonable alternative hypothesis that Chusqueinae is sister to a lineage that includes all remaining tropical woody bamboos.

Although Melocanninae is supported as sister to the remaining paleotropical woody bamboos (bootstraps 96 ML, 96 MP; posterior probability 1.00), the network analysis (Fig. 3) shows character conflict from *R. hepburnii* that likely prevents supported resolution of branching order within the paleotropical woody lineage. The final problem area is already well known to bamboo phylogeneticists: a lack of sufficient resolution to clarify relationships within the temperate woody clade. Encouragingly, the absence of a clear bifurcation pattern here is not due to character conflict (Fig. 3) but rather an absence of observed character state variation in the sequence data. Adding further high quality sequences from additional chloroplast loci therefore holds considerable promise in the eventual resolution of major lineages within this clade (e.g., Triplett and Clark, 2010; Zheng et al., 2010; Zhang et al., 2011), although if the radiation is recent enough the bamboos in the temperate clade might still fail to exhibit clear signs of lineage sorting.

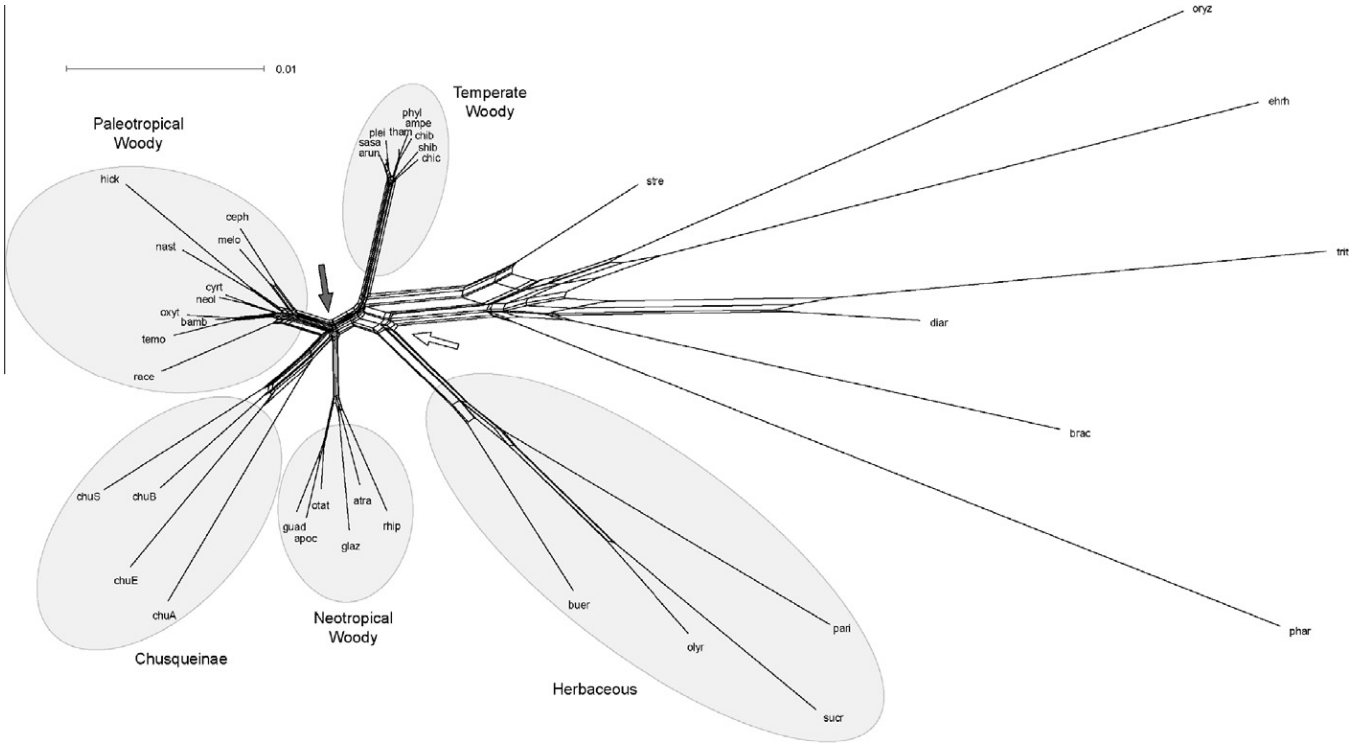


**Fig. 3.** A data-display network of combined nucleotide sequences. The neighbor-net method is used here to explore data conflict and not to estimate phylogeny. For visual clarity, taxon names have been abbreviated as the first four letters of the genus name except in the following cases: chuB (*Chusquea bambusoides*), chuS (*Chusquea scandens*), chuA (*Chusquea spectabilis*), chuE (*Chusquea elata*), chib (*Chimonobambusa marmorea*) and chic (*Chimonocalamus pallens*). The shaded arrow shows a network location that corresponds to the question of monophyly in neotropical woody bamboos. The white arrow indicates the joining location of outgroup sequences to the long edges between the herbaceous and temperate woody bamboo groups.

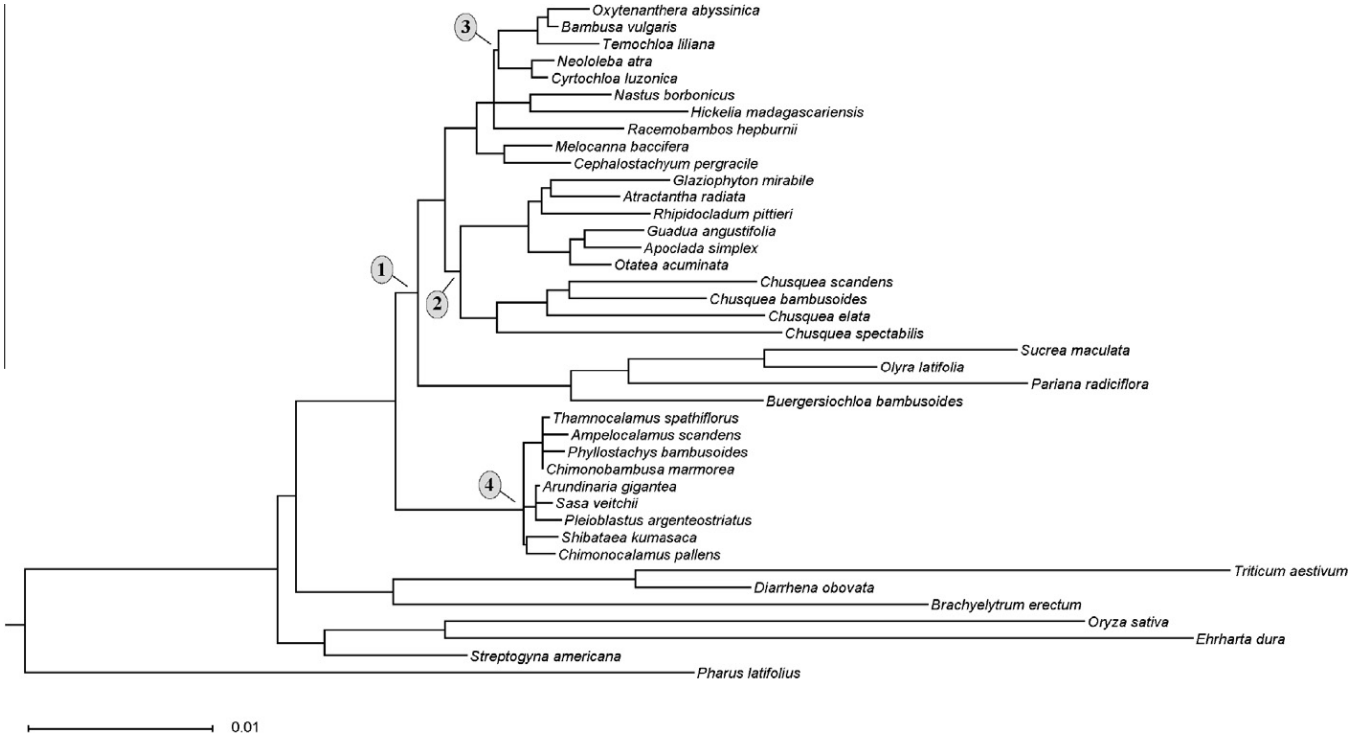


**Fig. 4.** Maximized likelihood estimate of topology based upon combined nucleotide data and a best-fit model (GTR + I + G). Nodes of persistent difficulty are (1) paraphyly of woody bamboos with respect to the herbaceous lineage, (2) monophyly of neotropical woody bamboos, (3) monophyly of Bambusinae *sensu stricto*, and (4) the long branch leading to a probable rapid diversification of temperate woody bamboos.

Despite the robustness of the ingroup topology to long branch attraction experiments, several long branches still exist in the phylogeny estimation that might not improve with additional taxon sampling. The branch leading to the radiation of Temperate

Woody bamboos, for example, could remain uninterrupted: newly sequenced members continue to show strong sequence similarity to all other species in the clade (Triplett and Clark, 2010; Zheng et al., 2010; Hodkinson et al., 2010). Similarly, the sister relationship of *Buergersiochloa bambusoides* to all other herbaceous bamboos will probably remain the earliest split in that lineage: the remaining unsampled species taxonomically fall within the Parianinae and Olyrinae clades further up that branch. So far, outgroup taxa have proven to be relatively distant in sequence similarity and will continue to cause problematic character conflict when joining the long branch between herbaceous and temperate woody clades in unrooted topologies and networks.

### 4.3. Congruence with other studies

Several features of our topology are worth comparison with published literature on bamboo phylogenetics. Presented here is a short discussion of congruence and inconsistencies between our phylogeny estimation and those published in other studies for the purpose of evaluating our topology. A more complete review of bamboo molecular phylogenetics literature is presented in an accompanying classification paper by the Bamboo Phylogeny Group (BPG, 2012).

Four major bamboo lineages were recovered that show biogeographic and morphological distinction (Fig. 2). The finding is consistent with two other recent studies that have comparable taxon and marker sampling (Bouchenak-Khelladi et al., 2008; Sungkaew et al., 2009). The Temperate Woody clade has long been distinct in molecular analyses. Early evidence came from plastid RFLPs (Watanabe et al., 1994; Kobayashi, 1997), followed by comparative DNA sequence analysis (Kelchner and Clark, 1997; Zhang and Clark, 2000; Bouchenak-Khelladi et al., 2008; Peng et al., 2008 and several subsequent studies). Lack of resolution within the Temperate Woody clade is common to all chloroplast DNA studies, including the present one.

The Herbaceous clade has also been recognized for more than a decade, both in chloroplast and nuclear DNA studies (e.g., Clark et al., 1995; Kelchner and Clark, 1997; Mathews et al., 2000; Bouchenak-Khelladi et al., 2008; Sungkaew et al., 2009). A Paleotropical Woody clade first appeared in Kelchner and Clark (1997) and was supported with additional taxon sampling in Zhang and Clark (2000). This clade has since been recovered in the well sampled studies of Bouchenak-Khelladi et al. (2008) and Sungkaew et al. (2009). A novel result of the present study is the well-supported inclusion of Hickeliinae, represented by two Malagasy species, in the Paleotropical Woody clade. The subtribe forms a trichotomy with Racemobambosinae and the Bambusinae. Although Clark et al. (2007) included a more extensive sampling of Hickeliinae in their analysis, the subtribe was only weakly associated with the other paleotropical woody bamboos.

A tropical woody clade (Paleotropical + Neotropical Woody bamboos) is well supported in both Bouchenak-Khelladi et al. (2008) and Sungkaew et al. (2009), with only slightly less support in our analysis. More problematic has been a confirmation of the Neotropical Woody clade. Previous to this study, the only published molecular evidence of a Neotropical Woody clade was a preliminary *rpl16* intron analysis by Kelchner and Clark (1997) and a more taxon and data rich analysis by Sungkaew et al. (2009). We note that bootstrap support for this clade in both Sungkaew et al. (2009) and the present study is moderate at best, that problematic character conflict exists at these nodes in the network analyses, and that possible paraphyly of Neotropical woody bamboos with respect to the Paleotropical Woody clade cannot be rejected with SH tests.

As in Bouchenak-Khelladi et al. (2008), Sungkaew et al. (2009), GPWG II (2012) and Kelchner et al. (in prep) our tree suggests paraphyly of plastids in woody bamboos. The topology, however, is incongruent with Mathews et al.'s (2000) phylogeny estimation of nuclear *phyB* sequences in grasses, the only published nuclear data set to include taxa from all three woody and herbaceous bamboo tribes. Although the *phyB* analysis was focused on tribal relationships in Poaceae, a bamboo clade was recovered with a woody lineage sister to an herbaceous lineage, as predicted by the traditional classification scheme of two Bambusoideae tribes. Only the herbaceous clade had bootstrap support. SH tests of our plastid data fail to reject this alternative arrangement of taxa, which suggests that the potential incongruence of nuclear signal with most plastid studies could be due to analytical error from the accumulation of homoplasious characters in bamboo plastid lineages. Alternatively, it is also possible that the incongruence is historically accurate and represents an echo of early hybridization events among ancestral populations of diverging herbaceous and woody bamboos. Until more data is available from the nuclear genome and directed morphological analyses, commitment to a stance of paraphyly in woody bamboos is inadvisable.

### 4.4. Summary and future directions

In this paper, the Bamboo Phylogeny Group presents a rigorously tested plastid phylogeny estimation of branching order among the world's major bamboo lineages. The tree is well resolved, most branches have high support values, and all tribes and subtribes of bamboos are adequately positioned within the topology. Corroboration of the tree by an independent study (Kelchner et al., in prep.) and congruence of major lineages with previous molecular phylogeny estimations suggests that we have converged upon a predominant plastid signal for Bambusoideae.

Four main plastid lineages are recognized that correspond to distinct biogeographic categories: paleotropical woody, neotropical woody, temperate woody, and herbaceous bamboos. The subtribe Hickeliinae is robustly placed in a trichotomy with Racemobambosinae and Bambusinae that is sister to Melocanninae in the paleotropical woody clade. Chusqueinae is sister to the remaining neotropical woody bamboos, and Buergersiochloinae is sister to the remaining herbaceous subtribes. Resolution within tribe Arundinarieae is weak because of limited sequence variation among sampled taxa.

Monophyly of the woody bamboos remains plausible because of character conflict in the plastid data and the possible incongruence of plastid signal with a phylogeny estimation based on nuclear DNA sequences. What is evident, however, is that a deep split exists in the woody bamboo plastid phylogeny between tropical species and temperate ones. By classifying the two woody lineages as separate tribes (Arundinarieae and Bambuseae), the Bamboo Phylogeny Group can provide a stable system for bamboo taxonomy that is resilient to the possibility of monophyly of woody bamboo nuclear genomes (BPG, 2012).

A thorough reconstruction of organismal phylogeny for bamboos will require morphological and nuclear data to be evaluated in addition to the chloroplast sequences presented here. A detailed morphological analysis of bamboo species is currently being prepared by members of the BPG. Ongoing efforts in bamboo plastid comparisons are focusing on the addition of many more taxa in most subtribe lineages (BPG, in prep.) and on whole plastome sequences (Zhang et al., 2011) to generate additional resolution within each of the four major bamboo clades. This is likely to reinforce and enhance our knowledge of subtribal plastid relationships over the next few years. Sequencing of low-copy nuclear loci is also underway that will offer critical evidence for an accurate reconstruction of bamboo evolutionary history. These data sets, when available, should provide a robust framework for advanced testing of morphological and molecular character evolution, biogeographic

hypotheses, and discernment of evolutionary steps in the development of modern day Bambusoideae.

## Acknowledgments

Chloroplast DNA sequences were produced in the laboratories of Lynn Clark and Scot Kelchner with the assistance of Jimmy Triplett, Amanda Fisher, Chi-Sing Ho, Idaho State University's Molecular Research Core Facility, and the DNA Facility of the Iowa State University Office of Biotechnology. Funding for the Project was provided by National Science Foundation awards DEB-0515828 to Scot Kelchner and DEB-0515712 and DEB-9806877 to Lynn Clark. Field work for temperate bamboo collections was supported in part by National Geographic Society Grant 7336-02 to Lynn Clark and De-Zhu Li.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2013.02.005.

## References

Bamboo Phylogeny Group, 2006. The bamboo phylogeny project. Bamboo 27, 11–14.

Bamboo Phylogeny Group, 2012. An updated tribal and subtribal classification of the bamboos (Poaceae: Bambusoideae). In: Proceedings of the 9th World Bamboo Congress, 10–12 April 2012, Antwerp, Belgium, pp. 3–27.

Bouchenak-Khelladi, Y., Salamin, N., Savolainen, V., Forest, F., van der Bank, M., Chase, M.W., Hodkinson, T.R., 2008. Large multi-gene phylogenetic trees of the grasses (Poaceae): progress towards complete tribal and generic level sampling. Mol. Phylogenet. Evol. 47, 488–505.

Buckley, T.R., Simon, C., Shimodaira, H., Chambers, G.K., 2001. Evaluating hypotheses on the origin and evolution of the New Zealand alpine cicadas (Maoricicada) using multiple-comparison tests of tree topology. Mol. Biol. Evol. 18, 223–234.

Chase, M.W., Hills, H.H., 1991. Silica gel: an ideal material for field preservation of leaf samples for DNA studies. Taxon 40, 215–220.

Clark, L.G., Dransfield, S., Triplett, J., Sánchez-Ken, J.G., 2007. Phylogenetic relationships among the one-flowered, determinate genera of Bambuseae (Poaceae: Bambusoideae). Aliso 23, 315–332.

Clark, L.G., Zhang, W., Wendel, J.F., 1995. A phylogeny of the grass family (Poaceae) based on *ndhF* sequence data. Syst. Bot. 20, 436–460.

Davis, J.I., Soreng, R.J., 2010. Migration of endpoints of two genes relative to boundaries between regions of the plastid genome in the grass family (Poaceae). Am. J. Bot. 97, 874–892.

Fisher, A.E., Triplett, J.K., Ho, C.S., Schiller, A.D., Oltrogge, K.A., Schroder, E.S., Kelchner, S.A., Clark, L.G., 2009. Paraphyly in the bamboo subtribe Chusqueinae (Poaceae: Bambusoideae) and a revised infrageneric classification for *Chusquea*. Syst. Bot. 34, 673–683.

Fitch, W.M., 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. 20, 406–416.

Gaut, B.S., Clark, L.G., Wendel, J.F., Muse, S.V., 1997. Comparisons of the molecular evolutionary process at *rbcL* and *ndhF* in the grass family (Poaceae). Mol. Biol. Evol. 14, 769–777.

Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49, 652–670.

Grass Phylogeny Working Group [GPWG], 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). Ann. Missouri Bot. Gard. 88, 373–457.

Grass Phylogeny Working Group II [GPWG II], 2012. New grass phylogeny resolves deep evolutionary relationships and discovers $C_4$ origins. New Phytol. 193, 304–312.

Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

Hodkinson, T.R., Chonghaile, G.N., Sungkaew, S., Chase, M.W., Salamin, N., Stapleton, C.M.A., 2010. Phylogenetic analyses of plastid and nuclear DNA sequences indicate a rapid late Miocene radiation of the temperate bamboo tribe Arundinarieae (Poaceae, Bambusoideae). Plant Ecol. Divers. 3, 109–120.

Huelsenbeck, J.P., Rannala, B., 2004. Frequentist properties of Bayesian posterior probabilities of phylogenetic trees under simple and complex substitution models. Syst. Biol. 53, 904–913.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. Biometrics 17, 754–755.

Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. Mol. Biol. Evol. 23, 254–267.

Judziewicz, E.J., Clark, L.G., 2007. Classification and biogeography of New World grasses: Anomochlooideae, Pharoideae, Ehrhartoideae, and Bambusoideae. Aliso 23, 303–314.

Judziewicz, E.J., Clark, L.G., Londoño, X., Stern, M.J., 1999. American Bamboos. Smithsonian Institution Press, Washington, DC.

Kelchner, S.A., 2000. The evolution of non-coding chloroplast DNA and its application in plant systematics. Ann. Missouri Bot. Gard. 87, 482–498.

Kelchner, S.A., 2009. Phylogenetic models and model selection for noncoding DNA. Plant Syst. Evol. 282, 109–126.

Kelchner, S.A., Clark, L.G., 1997. Molecular evolution and phylogenetic utility of the *rpl16* intron in *Chusquea* and the Bambusoideae (Poaceae). Mol. Phylogenet. Evol. 8, 385–397.

Kelchner, S.A., Thomas, M.A., 2007. Model use in phylogenetics: nine key questions. Trends Ecol. Evol. 22, 87–94.

Kelchner, S.A., Wendel, J.F., 1996. Hairpins create minute inversions in non-coding regions of chloroplast DNA. Curr. Genet. 30, 259–262.

Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. J. Mol. Evol. 29, 170–179.

Kobayashi, M., 1997. Phylogeny of world bamboos analysed by restriction fragment length polymorphisms of chloroplast DNA. In: Chapman, G.P. (Ed.), The Bamboos. Academic Press, London, England, pp. 227–234.

Löhne, C., Borsch, T., 2005. Molecular evolution and phylogenetic utility of the *petD* group II intron: a case study in basal angiosperms. Mol. Biol. Evol. 22, 317–332.

Mathews, S., Tsai, R.C., Kellogg, E.A., 2000. Phylogenetic structure in the grass family (Poaceae): evidence from the nuclear gene phytochrome B. Am. J. Bot. 87, 96–107.

McClure, F.M., 1966. The Bamboos: A Fresh Perspective. Harvard University Press, Cambridge, Massachusetts.

Morgan, M.J., Kelchner, S.A., 2010. Inference of molecular homology and sequence alignment by direct optimization. Mol. Phylogenet. Evol. 56, 305–311.

Morrison, D.A., 2005. Networks in phylogenetic analysis: new tools for population biology. Int. J. Parasitol. 35, 567–582.

Morrison, D.A., 2006. Multiple sequence alignment for phylogenetic purposes. Aust. Syst. Bot. 19, 479–539.

Morrison, D.A., 2009. A framework for phylogenetic sequence alignment. Plant Syst. Evol. 282, 127–149.

Morrison, D.A., 2010. Using data-display networks for exploratory data analysis in phylogenetic studies. Mol. Biol. Evol. 27, 1044–1057.

Peng, S., Yang, H.-Q., Li, D.-Z., 2008. Highly heterogeneous generic delimitation within the temperate bamboo clade (Poaceae: Bambusoideae): evidence from GBSSI and ITS sequences. Taxon 57, 799–810.

Penny, D., Hendy, M.D., Steel, M.A., 1992. Progress with methods for constructing evolutionary trees. Trends Ecol. Evol. 7, 73–79.

Philippe, H., Sorhannus, U., Baroin, A., Perasso, R., Gasse, F., Adoutt, A., 1994. Comparison of molecular and paleontological data in diatoms suggests a major gap in the fossil record. J. Evol. Biol. 7, 247–265.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Posada, D., Crandall, K.A., 1998. ModelTest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Simon, D.M., Kelchner, S.A., Zimmerly, S., 2009. A broadscale phylogenetic analysis of group II intron RNAs and intron-encoded reverse transcriptases. Mol. Biol. Evol. 26, 2795–2808.

Sungkaew, S., Stapleton, S.M.A., Salamin, N., Hodkinson, T.R., 2009. Non-monophyly of the woody bamboos (Bambuseae; Poaceae): a multi-gene region phylogenetic analysis of Bambusoideae s.s. J. Plant Res. 122, 95–108.

Swofford, D.L., 2002. PAUP* Phylogenetic Analysis using Parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts.

Triplett, J.K., Clark, L.G., 2010. Phylogeny of the temperate woody bamboos (Poaceae: Bambusoideae) with an emphasis on *Arundinaria* and allies. Syst. Bot. 35, 102–120.

Watanabe, M., Ito, M., Kurita, S., 1994. Chloroplast DNA phylogeny of Asian bamboos (Bambusoideae, Poaceae) and its systematic implication. J. Plant Res. 107, 253–261.

Wu, Z.-Q., Ge, S., 2012. The phylogeny of the BEP clade in grasses revisited: evidence from the whole-genome sequences of chloroplasts. Mol. Phylogenet. Evol. 62, 573–578.

Yang, H.-Q., Peng, S., Li, D.-Z., 2007. Generic delimitations of *Schizostachyum* and its allies (Gramineae: Bambusoideae) inferred from GBSSI and *trnL-F* sequence phylogenies. Taxon 56, 45–54.

Zhang, W.P., Clark, L.G., 2000. Phylogeny and classification of the Bambusoideae (Poaceae). In: Jacobs, S.W.L., Everett, J. (Eds.), Grasses: Systematics and Evolution. CSIRO, Melbourne, Australia, pp. 35–42.

Zhang, Y.-J., Ma, P.-F., Li, D.-Z., 2011. High-throughput sequencing of six bamboo chloroplast genomes: phylogenetic implications for temperate woody bamboos (Poaceae: Bambusoideae). PLoS ONE 6, e20596.

Zheng, C.-Z., Zhang, Y.-X., Triplett, J.K., Yang, J.-B., Li, D.-Z., 2010. Large multi-locus plastid phylogeny of the tribe Arundinarieae (Poaceae: Bambusoideae) reveals ten major lineages and low rate of molecular divergence. Mol. Phylogenet. Evol. 56, 821–839.