RESEARCH ARTICLE

The evolution and utility of ribosomal ITS sequences in Bambusinae and related species: divergence, pseudogenes, and implications for phylogeny

HUI-XING SONG, SU-PING GAO, MING-YAN JIANG, GUANG-LI LIU, XIAO-FANG YU and QI-BING CHEN*

School of Landscape Architecture, Sichuan Agricultural University, Wenjiang, Chengdu 611130, Sichuan, People's Republic of China

Abstract

Ribosomal internal transcribed spacer (ITS) sequences are commonly used for phylogenetic reconstruction because they are highly reiterated as components of rDNA repeats, and hence are often subject to rapid homogenization through concerted evolution. Concerted evolution leads to intragenomic uniformity of repeats even between loci on nonhomologous chromosomes. However, a number of studies have shown that the ITS polymorphism within individuals is quite common. The molecular systematics of Bambusinae and related species were recently assessed by different teams using independently generated ITS sequences, and the results disagreed in some remarkable features. Here we compared the ITS sequences of the members of *Bambusa* s. l., the genera *Dendrocalamus*, *Dinochloa*, *Gigantochloa*, *Guadua*, *Melocalamus*, *Monocladus*, *Oxytenanthera*, *Thyrsostachys*, *Pleioblastus*, *Pseudosasa* and *Schizostachyum*. We have reanalysed the ITS sequences used by different research teams to reveal the underlying patterns of their different results. After excluding the sequences suspected to represent paralogous loci, a phylogenetic analysis of the subtribe Bambusinae species were performed using maximum parsimony and maximum-likelihood methods. The implications of the findings are discussed. The risk of incorporating ITS paralogues in plant evolutionary studies that can distort the phylogenetic signal should caution molecular systematists.

[Song H.-X., Gao S.-P., Jiang M.-Y., Liu G.-L., Yu X.-F. and Chen Q.-B. 2012 The evolution and utility of ribosomal ITS sequences in Bambusinae and related species: divergence, pseudogenes, and implications for phylogeny. *J. Genet.* **91**, 129–139]

Introduction

The ribosomal internal transcribed spacer (ITS) sequences have been used successfully in studying phylogenetic and genomic relationships of plants at lower taxonomic levels (Hamby and Zimmer 1992; Baldwin 1993; Baldwin *et al.* 1995; Hsiao *et al.* 1995; Sang *et al.* 1995; Wendel *et al.* 1995; Zhang and Sang 1999; Volkov *et al.* 1999; Feliner *et al.* 2004; Liu *et al.* 2006). There are multiple copies of this ribosomal array in the genome, but they appear to undergo rapid concerted evolution and all copies appear to be virtually identical (Wendel *et al.* 1995; Baldwin *et al.* 1995; Álvarez and Wendel 2003; Nieto-Feliner and Rossselló 2007). Therefore, in many studies, the rDNA is treated as a single gene, and both direct sequencing and cloning of PCR products were used for ITS analysis, with only one clone for each individual sequenced even if PCR products were cloned (Hsiao *et al.* 1995; Liu *et al.* 2006; Yang *et al.* 2008; Silva *et al.* 2012). However, in recent years a number of studies have shown that the concerted evolution of rDNA is not completely as expected, and ITS polymorphism within individuals is quite common (Baldwin *et al.* 1995; Wendel *et al.* 1995; Muir and Schlötterer 1999; Denduangboripant and Cronk 2000; Mayol and Rosselló 2001; Rosselló *et al.* 2006, 2007; Nieto-Feliner and Rossselló 2007; Zhang and Ge 2007; Kim *et al.* 2008; Göer and Grimm 2008; Grimm and Denk 2008; Pilotti *et al.* 2009). In that case, if only one cloned sequence is used as the representative of the ITS sequence of one individual, some polymorphism will be absent and the resulting phylogenetic analysis will be suspect. These results underline that the use of ITS as a universal marker should be evaluated on a case-by-case basis.

Bambusinae is a woody bamboo subtribe placed in the tribe Bambuseae subfamily Bambusoideae of grass family Poaceae (Ohrnberger 1999). The subtribe comprises 10 to 13 genera, found mostly in the tropical and subtropical regions of the Old World, many of which are economically exploited by the communities in South and South-East Asia. Because of the long time period before flowering in woody bamboos, vegetative characters such as culm sheath, ligule, blade, branching and rhizome have often been used to classify these species instead of flowers. It has been difficult to relate these

*For correspondence. E-mail: landscapesicau@hotmail.com.

Keywords. Bambusinae; internal transcribed spacer; phylogenetic analysis; pseudogenes; evolution.

characteristics to floral distinctions, which is one cause of controversy in the classification of bamboo genera. A sound taxonomy is needed for conservation and management of the woody bamboos belonging to this subtribe.

The recent availability of molecular data enabled taxonomists to review the phylogeny of the Bambusinae. Some studies on taxa of Bambusinae based on ITS nuclear rDNA sequences have been undertaken such as for Bambusa (Sun *et al.* 2005) and major groups of the paleotropical woody bamboos (Yang *et al.* 2008). However, only limited species were sampled in previous molecular studies, and only one ITS sequence was typically cloned to represent the multiple copies of each plant ribosomal array. As a result, the phylogenetic relationships of the genera were not well resolved.

The present paper examines the ITS polymorphism within and among the species of the subtribe Bambusinae and the related species, assessing nucleotide diversity among homologous ITS repeats, and infer the molecular phylogeny of the related genera based on all of the published ITS sequences available. Specifically, the objectives of this study were to (i) evaluate the utility of ITS sequence in the subtribe Bambusinae; and (ii) reveal relationships of the Bambusinae species.

Materials and methods

The ITS sequences of the members of broad concept *Bambusa* and the related species of *Dendrocalamus*, *Dinochloa*, *Gigantochloa*, *Guadua*, *Melocalamus*, *Monocladus*, *Oxytenanthera*, *Thyrsostachys*, *Pleioblastus*, *Pseudosasa* and *Schizostachyum* by different research teams were obtained from GenBank (http://www.ncbi.nlm.nih.gov/genbank/). A species of temperate woody bamboo, *Phyllostachys heteroclada*, was used as outgroup. The species name, clone number, GenBank accession number and the authors are listed in table 1.

Sequences were aligned with CLUSTALW (Thompson *et al.* 1994) and the alignment was manually refined using Genedoc (Nicholas *et al.* 1997). The aligned matrix was subjected to a neighbour-joining (NJ) analysis using the Kimura two parameter distance estimates (Kimura 1980). Constant sites were subsequently removed from the alignments. Then the matrix was first subjected to a maximum-likelihood (ML) analysis, searching for the appropriate evolutionary model using Modeltest (Posada and Crandall 1998), and then inputting the appropriate model(s) obtained from the likelihood ratio tests and the Akaike information criterion (AIC), into the ML analysis. All analyses were carried out using PAUP 4.0 (Swofford 1998). Because some sequences from the same species were diverged into different clades, we compared the GC content, the sequence length, and the nucleotide difference among them. Based on the methods of Mayol and Rosselló (2001) we identified the rDNA sequences that appeared to represent paralogous loci. These sequences were excluded in subsequent phylogenetic analysis. The consensus sequence for each group was created with BioEdit (Hall 1999) and aligned with CLUSTAL W (Thompson *et al.* 1994) to show the major difference of the types.

The phylogenetic analyses based on the reduced sequence matrix were conducted using parsimony and distance methods with PAUP* 4.0 (Swofford 1998). We first carried out a NJ analysis using the Kimura two parameter distance estimates (Kimura 1980). Maximum parsimony (MP) analyses were performed by heuristic search, tree bisection reconnection (TBR) branch swapping, and RANDOM stepwise addition with 100 replicates. To find the evolutionary model that fits best for given data, we subjected the alignment of the sequences to Modeltest (Posada and Crandall 1998) v3.7. The parameters of the best fit model were then used as input for a ML analysis. Topological robustness was assessed by bootstrap analysis with 500 replicates (Felsenstein 1985). In all analyses the functional outgroup (FOG) (Watrous and Wheeler 1981) used was *P. heteroclada.*

Some sequences in Bambusinae were closely related. To detect the lineage relationships among these closely related sequences in subtribe Bambusinae, a phylogenetic network reconstruction method was used to study relationships among the 44 sequences which belong to 34 species within Bambusinae group. We conducted clone genealogy by coalescent simulations using the Median-Joining model as implemented in the Network v4.0 software (Bandelt *et al.* 1999).

Results

Polymorphism of ITS sequences and phylogenetic analysis

A total of 99 sequences belonging to 72 species were included in this study. The aligned sequences revealed three distinctly different sequence types, which have a high level of variation in length and in sequence composition. Based on the alignment result combining with the sequence information in GenBank, we ascertained the uniform boundary for each sequence type respectively. The sequences ranged from 541 to 667 bp. The NJ tree (figure 1) showed three well supported clades also recovered from phylogenetic analysis, labelled as I, II and III, respectively. In clade I, the sequences were clustered matching the genera in general, different clones belonging to one or two genera were grouped together as subclades. However, species were mixed in clades II and III. Unexpectedly, the sequences from same species (i.e. *Bambusa bambos*, *B. multiplex*, *B. vulgaris*, *Dendrocalamus strictus*, etc.) were scattered into different main clades.

The three different types and the deduced paralogous loci

To compare the differences among sequence types, the sequence length and the G+C contents were counted (table 1). The length of the sequences in clade I ranged from 583 to 667 bp, and the length of sequences in clades II and

Table 1. Taxa included in this study and the related ITS sequence information of each taxon.

Genus	Species	Synonym	Abbreviations	GenBank accession no.	G+C (%)	Length (bp)	ITS type
Arundinaria	*Aru. alpina*		Aru_alp_1	AF454508	69	593	I
	Aru. faberi		Aru_fab_1	EU847129	68.2	592	I
	Aru. gigantea		Aru_gig_1	AY004759	70.4	588	I
	Aru. gigantea		Aru_gig_2	AF305726	70	588	I
Bambusa	*Bam. balcooa*		Bam_bal_1	EU244594	58.5	563	II
	Bam. bambos		Bam_bam_1	FJ410319	58.8	565	II
	Bam. bambos		Bam_bam_3	DQ915808	56.2	564	II
	Bam. bambos		Bam_bam_4	DQ131514	69.9	664	I
	Bam. bambos var. *gigantea*		Bam_bam_2	GQ464805	55.6	552	III
	Bam. beecheyana	*Sinocalamus beecheyana*	Bam_bee_1	AY839720	69.4	625	I
	Bam. blumeana		Bam_blu_1	DQ270125	69.6	643	I
	Bam. burmanica		Bam_bur_1	FJ410312	57.6	566	II
	Bam. chungii		Bam_chu_1	DQ270129	68.2	647	I
	Bam. chungii		Bam_chu_2	AY839709	68.5	649	I
	Bam. chungii		Bam_chu_3	GQ464806	56.5	551	III
	Bam. contracta		Bam_con_1	AY839703	68.7	645	I
	Bam. emeiensis	*Neosinocalamus affinis;*	Bam_eme_1	AY839711	68.8	640	I
	Bam. emeiensis	*Sinocalamus affinis*	Bam_eme_2	DQ270121	70.7	617	I
	Bam. emeiensis		Bam_eme_3	GQ464830	55.9	545	III
	Bam. flexuosa		Bam_fle_1	AY839701	69	666	I
	Bam. grandis		Bam_gra_1	AY839721	69.6	615	I
	Bam. hainanensis		Bam_hai_1	AY839702	69.4	639	I
	Bam. intermedia		Bam_int_1	DQ270127	69.8	666	I
	Bam. intermedia		Bam_int_2	AY839718	68.5	652	I
	Bam. multiplex		Bam_mul_1	FJ410317	58	564	II
	Bam. multiplex var. *riviereorum*		Bam_mul_2	EF450229	56.7	564	III
	Bam. multiplex		Bam_mul_3	AY839710	69.1	637	I
	Bam. multiplex		Bam_mul_4	DQ270126	69.7	637	I
	Bam. multiplex		Bam_mul_5	GQ464807	55.4	548	III
	Bam. nutans		Bam_nut_1	FI410315	51.9	566	II
	Bam. nutans		Bam_nut_2	AY839706	68.6	636	I
	Bam. oldhamii	*Dendrocalamopsis oldhamii;*	Bam_old_1	AY839707	69	625	I
	Bam. oldhamii	*Sinocalamus oldhamii*	Bam_old_2	DQ270124	68.3	660	I
	Bam. oldhamii		Bam_old_3	GQ464815	56.7	541	III
	Bam. oliveriana		Bam_oli_1	EU368961	52.3	564	II
	Bam. polymorpha		Bam_pol_1	EU244595	53.8	590	II
	Bam. sinospinosa		Bam_sin_1	DQ131515	69.7	656	I
	Bam. sinospinosa		Bam_sin_2	AY839714	68.8	652	I
	Bam. striata		Bam_str_1	EU523115	55.2	563	II
	Bam. subaequalis		Bam_sub_1	AY839712	69.3	636	I
	Bam. surrecta		Bam_sur_1	DQ270128	68.8	667	I
	Bam. surrecta		Bam_sur_2	AY839715	68.8	654	I
	Bam. teres		Bam_ter_1	FJ410316	52.3	566	II
	Bam. textilis		Bam_tex_1	AY839717	69.1	631	I
	Bam. tuldoides		Bam_tul_1	AY839708	68.8	654	I
	Bam. tulda		Bam_tulda	EF540854	48.6	565	II
	Bam. valida		Bam_val_1	AY839716	68.6	653	I
	Bam. vulgaris		Bam_vul_1	FJ410314	49.3	566	II
	Bam. vulgaris		Bam_vul_2	AY839705	69.2	652	I
	Bam. wamin		Bam_wam_1	FJ410318	52.6	567	II
Dendrocalamus	*Den. bambusoides*		Den_bam_1	DQ270136	69.4	617	I
	Den. brandisii		Den_bra_1	DQ270132	69.1	621	I
	Den. giganteus		Den_gig_1	DQ270133	69.7	622	I
	Den. giganteus		Den_gig_2	EU244593	56.5	565	II
	Den. latiflorus	*Sinocalamus latiflorus*	Den_lat_1	DQ270134	69.5	620	I
	Den. latiflorus		Den_lat_2	AY839713	69.8	621	I
	Den. membranacea	*Bam. membranacea*	Den_mem_1	DQ270138	70.5	650	I
	Den. membranacea		Den_mem_2	AY839704	70.1	649	I

Table 1 (*contd*).

Genus	Species	Synonym	Abbreviations	GenBank accession no.	G+C (%)	Length (bp)	ITS type
	Den. minor		Den_min_1	GQ464816	55.9	559	III
	Den. sinicus		Den_sin_1	DQ270135	69.4	621	I
	Den. strictus		Den_str_1	EU191757	56.2	562	II
	Den. strictus		Den_str_2	DQ270137	69.4	644	I
	Den. strictus		Den_str_3	AY839719	70.2	641	I
Dinochloa	*Din. malayana*		Din_mal_1	DQ131505	73.7	593	I
	Din. scandens		Din_sca_1	DQ131506	73.5	595	I
Gigantochloa	*Gig. albociliata*		Gig_alb_1	DQ270130	69	620	I
	Gig. atroviolacea		Gig_atr_1	EU543214	56.5	562	II
	Gig. levis		Gig_lev_1	GQ464820	55.2	554	III
	Gig. verticillata		Gig_ver_1	DQ270131	68.9	621	I
Guadua	*Gua. angustifolia*		Gua_ang_1	AY993946	71.7	607	I
	Gua. angustifolia		Gua_ang_2	GQ464821	55.6	551	III
Melocalamus	*Mel. arrectus*		Mel_arr_1	DQ131518	72	633	I
	Mel. arrectus		Mel_arr_2	GQ464826	56.3	559	III
	Mel. compactiflorus var. *fimbriatus*		Mel_com_1	DQ131517	72.2	633	I
	Mel. scandens		Mel_sca_1	DQ131516	71.2	641	I
Monocladus	*Mon. amplexicaulis*	*Bonia amplexicaulis*	Mon_amp_1	DQ131509	73	586	I
	Mon. levigatus	*Bonia levigatus*	Mon_lev_1	DQ131510	71.6	585	I
	Mon. saxatilis	*Bonia saxatilis*	Mon_sax_1	DQ131508	72.5	586	I
	Mon. saxatilis var. *solidus*	*Bonia saxatilis* var. *solidus*	Mon_sax_2	DQ131507	73	587	I
	Mon. saxatilis	*Bonia saxatilis*	Mon_sax_3	GQ464829	56.5	560	III
Oxytenanthera	*Oxy. abyssinica*		Oxy_aby_1	DQ131513	71.3	611	I
Pleioblastus	*Ple. simonii*		Ple_sim_1	EU847135	70.5	596	I
Pseudosasa	*Pse. japonica*		Pse_jap_1	EU847136	70.5	588	I
Schizostachyum	*Sch. blumei*		Sch_blu_1	DQ131536	72.8	586	I
	Sch. brachycladum		Sch_bra_1	DQ131535	73.3	588	I
	Sch. dumetorum		Sch_dum_1	DQ131530	72.6	590	I
	Sch. funghomii		Sch_fun_1	DQ131528	73.2	586	I
	Sch. funghomii		Sch_fun_2	GQ464842	56.6	554	III
	Sch. gracile		Sch_gra_1	DQ131538	73.2	583	I
	Sch. hainanense		Sch_hai_1	DQ131537	73.4	590	I
	Sch. jaculans		Sch_jac_1	DQ131531	72.9	587	I
	Sch. pseudolima		Sch_pse_1	DQ131529	73.1	585	I
	Sch. sanguineum		Sch_san_1	DQ131533	72.5	588	I
	Sch. xinwuense		Sch_xin_1	DQ131532	73.5	585	I
	Sch. zollingeri		Sch_zol_1	DQ131534	73.2	589	I
Thyrsostachys	*Thy. oliveri*		Thy_oli_1	DQ131512	68.2	623	I
	Thy. oliveri		Thy_oli_2	GQ464846	56.9	559	III
	Thy. siamensis		Thy_sia_1	DQ131511	68.3	619	I
Phyllostachys	*Phy. heteroclada*		Phy_het_1	EU847118	70.7	597	I

III were 541–564 bp and 562–590 bp, respectively. The average lengths for clades I, II and III were 629 ± 26.9 bp, 566 ± 6.5 bp and 554 ± 6.6 bp, respectively. The G+C content of all sequences is 65.96 ± 7.16%, which is 68.2–73.7% in clade I, 48.6–58.8% in clade II and 55.2–56.9% in clade III, with average content of 70.37 ± 1.74%, 56.21 ± 0.51%, 54.23 ± 3.43%, respectively. The consensus sequence for the three distinct groups and the alignment result (figure 2) showed the divergent sequence composition in many sites. The three consensus sequences can not be better aligned except for the 5.8S rRNA gene region. The sequences in clades II and III have a shorter length and lower GC content, therefore, they were suspected to represent paralogous loci according to the descriptions in materials and methods. These 29 sequences were excluded and there was a total of 70 sequences obtained from 54 species used in the following phylogenetic analysis.

Phylogenetic analysis based on the ITS matrix excluding pseudogenes

The aligned matrix of 730 character length included 393 constant characters, 127 variable parsimony-uninformative characters and 210 parsimony-informative characters. The MP trees had a minimal length of 800 steps; a consistency index (CI), excluding uninformative characters, of 0.6988; and a retention index (RI) of 0.7992.

Figure 1. Neighbour-joining gene tree inferred from all known ITS sequences. Numbers above branches indicate bootstrap values above 50%.

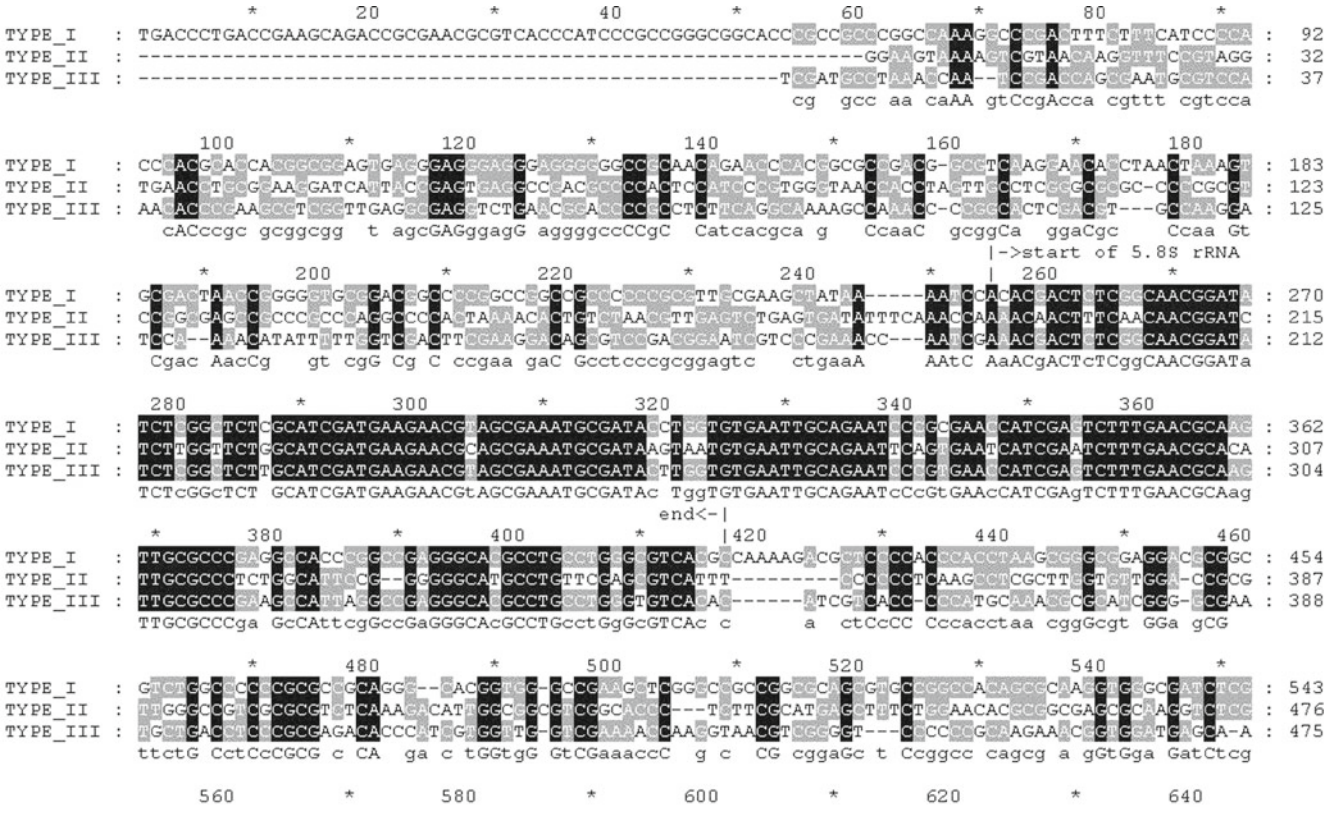

Figure 2. Alignment of the consensus sequences of the three ITS sequence types in this study. Black shading, 100% identity; dark grey, 75% identity; light grey, 50% identity; white, <50% identity.

Modeltest yielded the following two models: TrN+I+G best fit model under the likelihood ratio test (LRT) criterion and GTR+I+G under the AIC. ML analysis, using the TrN+I+G evolutionary model with the following parameters: gamma shape (alpha) = 0.6642; proportion of invariable sites = 0.233. The best tree in the ML analysis using the model showed a very similar topology to that of the strict consensus MP tree, with some minor differences of bootstrap support values (figure 3). The topology of both MP and ML trees showed that *Arundinaria*, *Pseudosasa* and *Pleioblastus* diverged from *Schizostachyum* showing the broad concept subtribe Bambusinae. The tree also showed that *Guadua* diverged from *Schizostachyum* and the remaining species of subtribe Bambusinae. The genera *Schizostachyum*, *Monocladus*, *Dinochloa*, *Melocalamus* and *Thyrsostachys* formed monophyletic groups with high (>90%) bootstrap values respectively (figure 3 dark lines). The topologies of the MP and the ML trees also were similar in regard to the remaining ingroup taxa. On the ML tree (figure 3), the remaining ingroups of species belong to *Bambusa*, *Dendrocalamus*, *Gigantochloa*, *Melocalamus* and *Thyrsostachys* which clustered as three large clades, labelled here as clades 1, 2 and 3. However, *Oxytenanthera abyssinica* diverged from these three clades. In clade 1, *B. surrecta*, *B. tuldoides*, *B. sinospinosa*, *B. valida*, *B. intermedia*, *B. flexuosa* and *B. bambos* formed a group with a bootstrap value of 95%, which was a sister taxa to a group formed by *B. chungii*, *B. emeiensis* and species of *Melocalamus*, in which *B. chungii* and *B. emeiensis* formed a clade with a bootstrap value of 86% and were sister species to the clade of *Melocalamus*. *B. hainanensis* diverged from these species as an outer clade. *B. vulgaris*, *Dendrocalamus strictus* and *D. membranacea* formed a clade with a bootstrap value of 100%, which was a sister taxon to *B. contracta* and a group formed by *B. blumeana* and *B. nutans*. The species of *Thyrsostachys* formed a clade with a bootstrap value of 100%, and had a close relationship with the clade formed by *Dendrocalamus bambusoides* and *Gigantochloa albociliata*. In clade 2, *B. grandis*, *B. beecheyana*, *B. oldhamii*, *Dendrocalamus latiflorus*, *D. sinicus* and *D. giganteus* formed a group, which was a sister taxon to the clade formed by *D. brandisii* and *Gigantochloa verticillata*. In clade 3, *B. subaequalis*, *B. multiplex* and *B. textiles* formed a clade with a bootstrap value of 84%.

Network analysis of Bambusinae

Most of the Bambusinae sequences are closely related, especially among the sequences of the speies of *Bambusa*, *Melocalamus*, *Dendrocalamus*, *Gigantochloa* and *Thyrsostachys*. Taking the potential for reticulation in closely related lineages into consideration, a phylogenetic network reconstruction method was used. A network of these 44 sequences

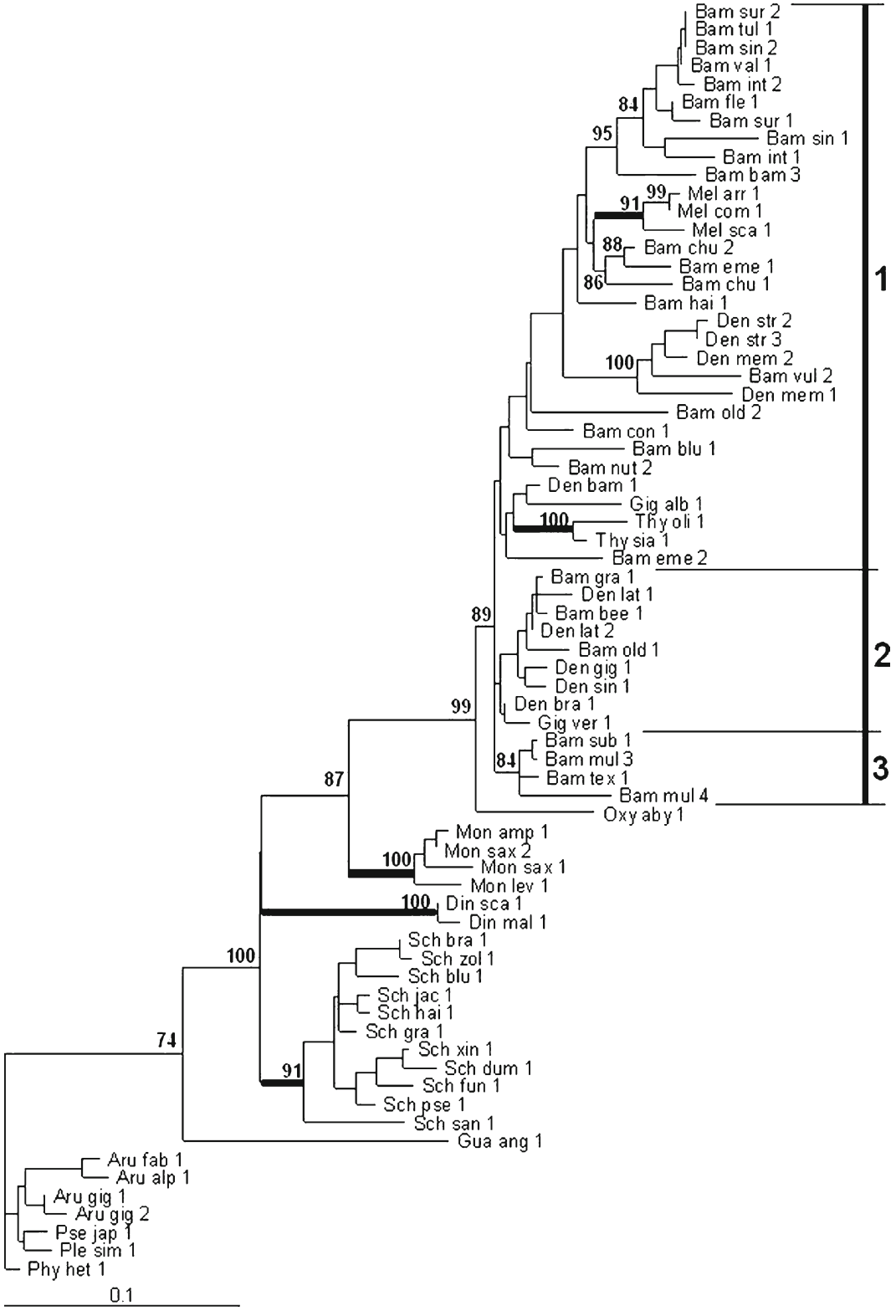

Figure 3. Phylogenetic reconstructions from the ITS matrix after excluding pseudogenes. The tree shown is maximum-likelihood tree, numbers above branches indicate bootstrap values above 50%. Five monophyletic clade were marked with dark lines (see text).

belonging to the five closely related species is shown in figure 4. These 44 sequences were grouped into a complex network. The Mel/Bam_hai group included all *Melocalamus* clones and a sequence of *B. hainanensis*. The Bam_chu/eme group indicated the close relationships between *B. chungii* and *B. emeiensis*. *Dendrocalamus strictus* and *B. vulgaris*

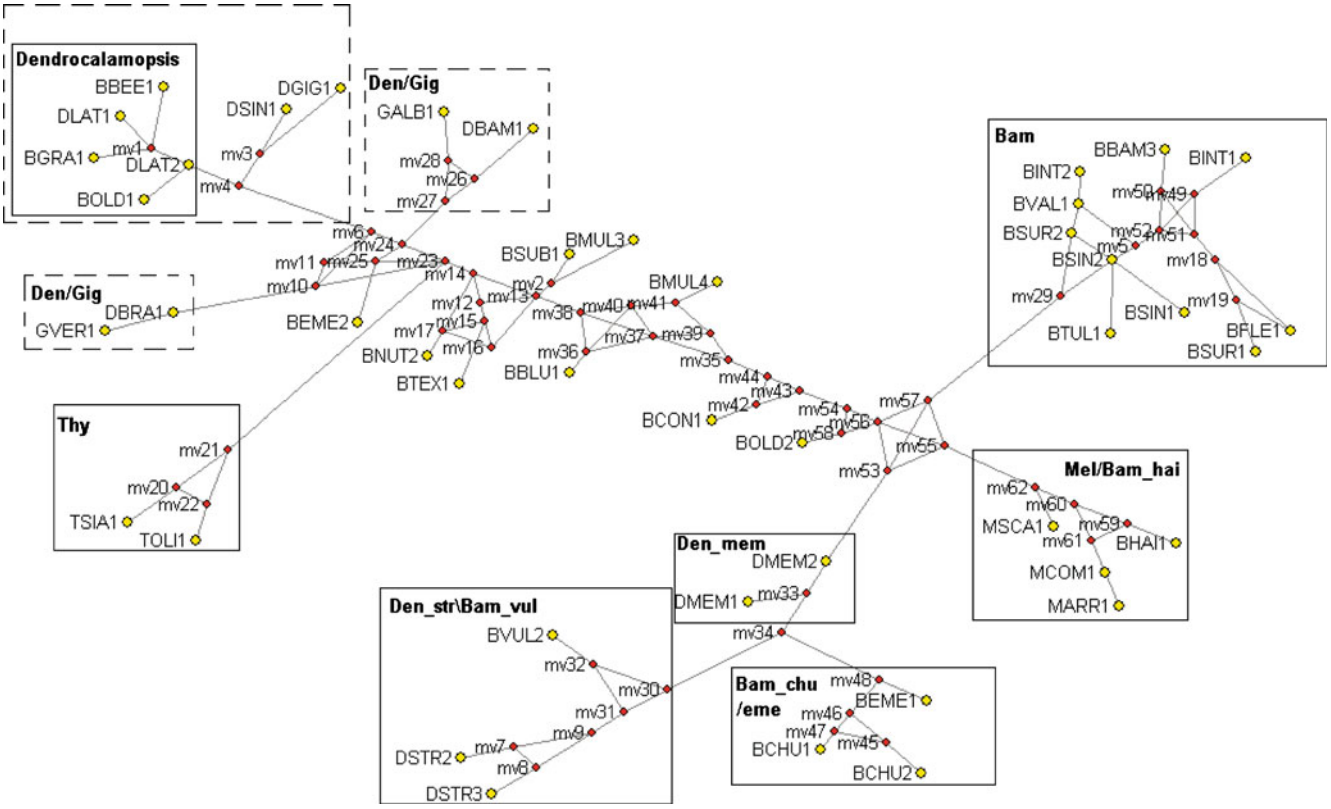

Figure 4. Network based on the 44 ITS haplotypes of 34 closely related Bambusinae species.

formed the group named Den_str/Bam_vul. The two sequences of *Thyrsostachys* formed the group of Thy, while each of the two sequences of *Gigantochloa* showed closer relationship with sequence of *Dendrocalamus* and clustered two Den/Gig groups. *Dendrocalamus latiflorus*, *B. grandis*, *B. beecheyana*, and *B. oldhamii* formed the group of *Dendrocalamopsis* and showed a close relationship with *D. sinicus* and *D. giganteus*. The Bam group included most of *Bambusa* speices: *B. surrecta*, *B. tuldoides*, *B. sinospinosa*, *B. valida*, *B. intermedia*, *B. flexuosa* and *B. bambos*.

Discussion

Ribosomal ITS sequences are commonly used for phylogenetic reconstruction because they are ubiquitous and present in high copy number. Ribosomal rDNA repeats often undergo rapid concerted evolution within and among arrays (Baldwin *et al.* 1995; Álvarez and Wendel 2003; Nieto-Feliner and Rossselló 2007). In recent years, a number of studies have shown that ITS polymorphism within individuals is quite common (Baldwin *et al.* 1995; Wendel *et al.* 1995; Buckler-IV *et al.* 1997; Denduangboripant and Cronk 2000; Mayol and Rosselló 2001; Bailey *et al.* 2003; Rosselló *et al.* 2006, 2007; Nieto-Feliner and Rossselló 2007; Zhang and Ge 2007; Kim *et al.* 2008; Göer and Grimm 2008; Grimm and Denk 2008; Pilotti *et al.* 2009). Mayol and Rosselló (2001) analysed the ITS sequences in *Quercus* generated independently by two different teams, and suggested that the divergent ITS alleles reported by one team were pseudogenes, i.e., nonfunctional paralogous loci. Paralogous ITS loci have been reported in many vascular plant groups (Buckler-IV *et al.* 1997; Yang *et al.* 1999; Kita and Ito 2000; Hartmann *et al.* 2001) and are associated with hybridization, introgression, and/or polyploidization events. Pseudogenes have been reported to be AT-rich since the G→A and C→T transitions are frequent mutation patterns that occurr at these nonfunctional DNA regions (Li and Graur 1991) and, accordingly, lower the G+C content. In the present study, the ITS sequences diverged into clade I were longer and had higher G+C content than those in clades II and III. In addition, sequences in clades II and III were found to have high rate of substitutions and deletions involving conserved regions (these sequences were diverged into two independent clades in the NJ analysis tree based on 5.8S sequence, analysis not shown). In this case, we excluded those sequences that form clades II and III from further phylogenetic analysis.

The phylogenetic analysis showed that *Arundinaria*, *Pseudosasa* and *Pleioblastus* genera were found to diverge from the species of subtribe Bambusniae. This is is supported by Li (1997) and Soderstrom and Ellis (1987) who were of the opinion that the three genera belong to the subtribe Arundinariinae. *Schizostachyum* was placed under the subtribe Melocanninae (Soderstrom and Ellis 1987; Li 1997),

but had a closer relationship with Bambusniae species than with *Guadua*. The *Dinochloa*, *Melocalamus*, *Thyrsostachys* and *Monocladus* genera were supported to be monophyletic, respectively, in the present analysis. *Dinochloa* was treated as a member of subtribe Bambusinae based on the similarity in ovary characters (Holttum 1956), and the result of combined analysis of the nuclear and chloroplast sequences (Yang *et al.* 2008). Based on these previous results, it was reasonable to find *Dinochloa* clustering as a subclade belonging to the main Bambusinae clade in the present investigation. *Dinochloa* was also recovered as a sister lineage to the clades containing *Dendrocalamus* by Pattanaik and Hall (2009) based on AFLP Fingerprints. Although Clayton and Renvoize (1986) and Soderstrom and Ellis (1987) merged *Monocladus* into *Bambusa*, our results agreed with the opinion of Xia (1996) and Sun *et al.* (2005) that *Monocladus* was a separate genus.

Gigantochloa species occur mainly in Malaysia, and they can be distinguished upon flowering by the termination of their spikelets in a long, empty lemma. *Gigantochloa* are very similar to *Bambusa* and *Dendrocalamus* in morphology, and it is difficult to separate the three genera without flowers. Watanabe *et al.* (1994), the first to study phylogenetic relationships among Asian bamboos using RFLP of chloroplast DNA recovered a clade representing subtribe Bambusinae sensu Ohrnberger (1999), containing *Bambusa*, *Dendrocalamus*, and *Gigantochloa*, suggesting close relationships among these genera. The study of Loh *et al.* (2000) and Ramanayake *et al.* (2007), using AFLP and RAPD respectively, also indicated a close relationship between *Bambusa* and *Gigantochloa*. The combined evidence from these earlier molecular studies and the present investigation suggest that taxa belonging to *Bambusa*, *Dendrocalamus* and *Gigantochloa* form a close complex but are relatively distant from *Dinochloa* and *Monocladus*. Holttum (1956) moved the Asiatic species of the genus *Oxytenanthera* to either *Dendrocalamus* or *Gigantochloa* on the basis of the absence of ovary appendage. Soderstrom and Ellis (1987) transferred *Oxytenanthera* and *Gigantochloa* into *Dendrocalamus*. Our results supported the close relationship between *Gigantochloa* and *Dendrocalamus*, but *Oxytenanthera abyssinica* was divergent from the *Bambusa* and *Dendrocalamus* species. In fact, *Oxytenanthera* is vegetatively similar to *Dendrocalamus* except for the structure of ovaries and florets. *Oxytenanthera* has a unique ovary (Holttum 1956) and three stigmas while *Dendrocalamus* has only one stigma.

When McClure (1940) published the new genus *Sinocalamus*, the four species he listed were *Sinocalamus latiflorus*, *S. beecheyana*, *S. oldhamii* and *S. affinis*. Now, *Dendrocalamus latiflorus*, *Bambusa beecheyana*, *B. oldhamii* and *Neosinocalamus affinis* (or *B. emeiensis*) are synonyms of these four species respectively. After the examination of inflorescence, Lin (1989) believed that *D. latiflorus* evolved from subgenus *Dendrocalamopsis* of *Bambusa*, which possesses many-flowered spikelets. This is congruent with the observation that *D. latiflorus* was recovered in a clade shared with *B. grandis*, *B. beecheyana*, and *B. oldhamii* and had a close relationship with *B. emeiensis* in the present investigation. In this paper, *B. emeiensis*, a synonym of *Neosinocalamus affinis*, was still a sister species to *B. chungii* with a highly supportive bootstrap value, while another sequence from *B. emeiensis* had a close relationship with *Dendrocalamus*, which is supported by the result that *N. affinis* grouped with the *Dendrocalamus* subclade based on nuclear and chloroplast sequences (Yang *et al.* 2008). *N. affinis* may be an intermediate species between *Bambusa* and *Dendrocalamus* as suggested by Li and Hsueh (1988) and Li (1997).

Type species of subgenus *Leleba* of *Bambusa*, *B. multiplex*, was closely related to *B. textilis*, *B. contracta*, *B. subaequalis*, and *B. nutans*. This supports the opinion that *B. nutans* and *B. contracta* species as members of subgenus *Leleba*, but contradicts the treatment of *B. subaequalis* as a member of *Bambusa* (sensu stricto) by Chia and Fung (1996). In addition, *B. vulgaris* is found as sister to *D. strictus* and *D. membranaceus* in the present paper, not to *B. nutans* as found by Sun *et al.* (2006).

As for the *Bambusa* species involved in this study, *B. hainanensis*, *B. chungii*, *B. intermedia* and *B. surrecta* were thought belong to subgenus *Linganania* by Xia *et al.* (2006), *B. hainanensis* had a close relationship with *Melocalamus* species in Network analysis. The inflorescence of this species is unknown and requires further investigation. The *B. intermedia* and *B. surrecta* species were embedded in the thorny species *B. sinospinosa* and *B. flexuosa*, which belong to the subgenus *Bambusa*; and *B. chungii* was divided into another clade. All these raised doubts about the monophyly of the subgenera *Linganania*, and confused the delineation of *Bambusa* (sensu stricto) and *Linganania*.

From the foregoing results, it appears that the *Dinochloa* and *Monocladus* were separate genera, while the species of *Dendrocalamus* were closely related to and nested in a polyphyletic *Bambusa*. Taxa belonging to *Bambusa*, *Dendrocalamus* and *Gigantochloa* form a close complex but are relatively distant from *Dinochloa* and *Monocladus*. *B. emeiensis* may be an intermediate species between *Bambusa* and *Dendrocalamus*. The phylogenetic trees generated in the present study are plausible hypothesis for relationships within subtribe Bambusinae, but need validation from other sources. This study confirms that the current taxonomic treatment of *Bambusa*, *Dendrocalamus* and related genera is unsatisfactory and needs revision. A broader study encompassing a wider selection of taxa from *Bambusa*, *Dendrocalamus*, *Oxytenanthera* and *Gigantochloa*, and inclusion of evidence from multiple data source such as morphologic, cytology, molecular makers and more sequences might be expected to produce a robust phylogenetic tree for this suite of closely related taxa. In addition, there are multiple copies of the ribosomal array in the genome, and they were thought to undergo rapid concerted evolution and it was believed that all copies appear to be virtually identical (Baldwin *et al.* 1995; Wendel *et al.* 1995; Álvarez and Wendel 2003;

Nieto-Feliner and Rossselló 2007). Therefore, the belief was that the rDNA needs to be treated as a single gene, and that only one clone for each individual needed to be sequenced. However, in recent years a number of studies together with our results have shown that the concerted evolution of rDNA is not entirely as was expected. Polymorphism of the ITS within individuals is quite common. Therefore, if only one clone is selected as the representative of one individual, polymorphism will be lost and the resulting phylogenetic analysis will be suspect. These results underline that the use of the ITS as a universal marker should be evaluated on a case-by-case basis. The risk of incorporating ITS paralogues in plant evolutionary studies which can distort the phylogenetic signal should caution molecular systematists. Without a detailed inspection of some basic features of the sequence, errors in evolutionary inferences might easily occur.

Acknowledgements

This work was funded by the Agricultural Science and Technology Achievements Transformation Funds Project of the Ministry of Science and Technology of China (2010GB24320621) and Key Technology Research and Development Program of Sichuan Province (no. 10ZC1136). Hui-Xing Song was supported by Foundation of Sichuan Educational Committee (no. 11ZA083). We wish to thank two anonymous reviewers for helpful comments that improved the manuscript. This paper benefited from the careful review of Bernard R. Baum, Agriculture and Agri-Food Canada, Ottawa.

References

Álvarez I. and Wendel J. F. 2003 Ribosomal ITS sequences and plant phylogenetic inference. *Mol. Phylogenet. Evol.* **29**, 417–434.

Bailey C. D., Carr T. G., Harris S. A. and Hughes C. E. 2003 Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. *Mol. Phylogenet. Evol.* **29**, 435–455.

Baldwin B. G. 1993 Molecular phylogenetics of *Calycadenia* (Compomtae) based on ITS sequences of nuclear ribosomal DNA: choromosomal and morphological evolution reexmained. *Am. J. Bot.* **80**, 220–238.

Baldwin B. G., Sanderson M. J. and Porter J. M. 1995 The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Ann. Mol. Bot. Gard.* **82**, 247–277.

Bandelt H. J., Fortser P. and Rohl A. 1999 Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48.

Buckler E. S. IV, Ippolito A. and Holtsford T. P. 1997 The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* **145**, 821–832.

Chia L. C. and Fung H. L. 1996 Bambusa and Leleba. In *Flora reipublicae popularis sinicae* (ed. P. C. Keng and Z. P. Wang), vol. 9, pp. 48–114. Science Press, Beijing, China (in Chinese).

Clayton W. D. and Renvoize S. A. 1986 *Genera graminum-grasses of the world*. Royal Botanic Gardens, Kew, UK.

Denduangboripant J. and Cronk Q. C. 2000 High intraindividual variation in internal transcribed spacer sequences in *Aeschynanthus* (Gesneriaceae): implications for phylogenetics. *Proc. Biol. Sci.* **267**, 1407–1415.

Feliner G. N., Larena B. G. and Aguilar J. F. 2004 Fine-scale geographical structure, intra-individual polymorphism and recombination in nuclear ribosomal internal transcribed spacers in Armefia (Plumbaginaceae). *Ann. Bot.* **93**, 189–200.

Felsenstein J. 1985 Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791.

Göer M. and Grimm G. W. 2008 General functions to transform associate data to host data, and their use in phylogenetic inference from sequences with intraindividual variability. *BMC Evol. Biol.* **8**, 86.

Grimm G. W. and Denk T. 2008 ITS evolution in Platanus: homoeologues, pseudogenes, and ancient hybridization. *Ann. Bot.* **101**, 403–419.

Hall T. A. 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids. Symp. Ser.* **41**, 95–98.

Hamby R. K. and Zimmer E. A. 1992 Ribosomal RNA as a phylogenetic tool in plant systematics. In *Molecular systematics of plants* (ed. P. S. Soltis, D. E. Soltis and J. J. Doyle), pp. 50–101. Chapman and Hall, New York, USA.

Hartmann S., Nason J. D. and Bhattacharya D. 2001 Extensive ribosomal DNA genic variation in the columnar cactus Lophocereus. *J. Mol. Evol.* **53**, 124–134.

Holttum R. E. 1956 The classification of bamboos. *Phytomorphology* **6**, 73–90.

Hsiao C., Chatterton N. J. and Asay K. H. 1995 Molecular phylogeny of the Pooideae (Poaceae) based on nuclear rDNA (ITS) sequences. *Theor. Appl. Genet.* **90**, 389–398.

Kim S. T., Sultan S. E. and Donoghue M. J. 2008 Allopolyploid speciation in Persicaria (Polygonaceae): insights from a low-copy nuclear region. *Proc. Natl. Acad. Sci. USA* **105**, 12370–12375.

Kimura M. 1980 A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120.

Kita Y. and Ito M. 2000 Nuclear ribosomal ITS sequences and phylogeny in East Asian Aconitum subgenus *Aconitum* (Ranunculaceae), with special reference to extensive polymorphism in individual plants. *Plant Syst. Evol.* **225**, 1–13.

Li D. Z. 1997 The flora of China Bambusoideae project: problems and current understanding of bamboo taxonomy in China. In *The bamboos* (ed. G. P. Chapman), pp. 61–81. Academic Press, London, UK.

Li D. Z. and Hsueh C. J. 1988 A study on the genus *Dendrocalamus* Nees from China. *J. Bamboo Res.* **7**, 1–19 (in Chinese).

Li W. H. and Graur D. 1991 *Fundamentals of molecular evolution*. Sinauer, Sunderland, USA.

Lin W. T. 1989 Comments on the genus *Dendrocalamus* Nees from China. *J. Bamboo Res.* **8**, 30–35 (in Chinese).

Liu Q., Ge S. and Tang H. 2006 Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast *trnL-F* sequences. *New Phytol.* **170**, 411–420.

Loh J. P., Kiew R., Set O., Gan L. H. and Gan Y. Y. 2000 A study of genetic variation and relationships within the bamboo subtribe Bambusinae using amplified fragment length polymorphism. *Ann. Bot.* **85**, 607–612.

Mayol M. and Rosselló J. A. 2001 Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. *Mol. Phylogenet. Evol.* **19**, 167–176.

McClure F. A. 1940 New genera and species of Bambusaceae from Eastern Asia. *Lingnan Univ. Sci. Bull.* **9**, 66–67.

Muir G. and Schlötterer C. 1999 Limitations to the phylogenetic use of ITS sequences in closely related species and populations - a case study in *Quercus petraea* (Matt.) Liebl, Chapter 11. In *Which DNA marker for which purpose*? Final Compendium of the Research Project: Development, optimization and validation of molecular tools for assessment of biodiversity in forest trees in European Union DGXII Biotechnology FWIV Research Program Molecular Tools for Biodiversity (ed. E. M. Gillet) (http://webdoc.sub.gwdg.de/ebook/y/1999/whichmarker/index.htm).

Nicholas K. B., Nicholas Jr. H. B. and Deerfield D. W. II. 1997 GeneDoc: analysis and visualization of genetic variation. *Embnet. News* **4**, 1–4.

Nieto-Feliner G. and Rossselló J. A. 2007 Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Mol. Phylogenet. Evol.* **44**, 911–919.

Ohrnberger D. 1999 *The Bamboos of the World*, pp. 250–280, Elsevier, Amsterdam, Holland.

Pattanaik S. and Hall J. B. 2009 Species relationships in *Dendrocalamus* inferred from AFLP Fingerprints, volume 5, pp. 27–40. VIII World Bamboo Congress Proceedings, Thailand.

Pilotti M., Brunetti B. and Tizzani L. 2009 Platanus×acerifolia genotypes surviving to inoculation with *Ceratocystis platani* (the agent of canker stain): first screening and molecular characterization. *Euphytica* **169**, 1–7.

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

Ramanayake S. M. S. D., Meemaduma V. N. and Weerawardene T. E. 2007 Genetic diversity and relationships between nine species of bamboo in Sri Lanka, using Random Amplified Polymorphic DNA. *Plant Syst. Evol.* **269**, 55–61.

Rosselló J. A., Cosín R. and Boscaiu M. 2006 Intragenomic diversity and phylogenetic systematics of wild rosemaries (*Rosmarinus officinalis* L. s.l., Lamiaceae) assessed by nuclear ribosomal DNA sequences (ITS). *Plant Syst. Evol.* **262**, 1–12.

Rosselló J. A., Lázaro A. and Cosín R. 2007 A phylogeographic split in *Buxus balearica* (Buxaceae) as evidenced by nuclear ribosomal markers: when ITS paralogues are welcome. *J. Mol. Evol.* **64**, 143–157.

Sang T., Crawford D. J. and Sutessy T. F. 1995 Documentation of reticulate evolution in *Paeonies* (Paeonia) using internal transcribed spacer sequences of nuclear ribosomal DNA: implications for biogeography and concerted evolution. *Proc. Natl. Acad. Sci. USA* **92**, 6813–6817.

Silva U. C. S., Rapini A., Liede-Schumann S. and Ribeiro P. L. 2012 Taxonomic considerations on Metastelmatinae (Apocynaceae) based on plastid and nuclear DNA. *Syst. Bot.* **37**, 795–806.

Soderstrom T. R. and Ellis R. P. 1987 The position of bamboo genera and allies in a system of grass classification. In *Grass systematics and evolution* (ed. T. R. Soderstrom, K. W. Hilu, C. S. Campbell and M. E. Barkworth). Smithsonian Institution Press, Washington DC and London.

Sun Y., Xia N. H. and Lin R. S. 2005 Phylogenetic analysis of *Bambusa* (Poaceae: Bambusoideae) based on internal transcribed spacer sequences of nuclear ribosomal DNA. *Biochem. Genet.* **43**, 603–612.

Sun Y., Xia N. H. and Stapleton C. M. A. 2006 Relationships between *Bambusa* species (Poaceae, Bambusoideae) revealed by random amplified polymorphic DNA. *Biochem. Syst. Ecol.* **34**, 417–423.

Swofford D. L. 1998 PAUP* *phylogenetic analysis using parsimony* (*and other methods), version 4.0b10. Sinauer, Sunderland, USA.

Thompson J. D., Higgins D. G. and Gibson T. J. 1994 CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680.

Volkov R. A., Borisjuk N. V. and Panchuk II. 1999 Elimination and rearrangement of parental rDNA in the alloteraploid nicotiana tabacum. *Mol. Biol. Evol.* **16**, 311–320.

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

Watrous L. E. and Wheeler Q. D. 1981 The outgroup comparison method of character analysis. *Syst. Zool.* **30**, 1–11.

Wendel J. F., Schnabel A. and Seelanan T. 1995 Bi-directional interlocus concerted evolution following allopolyploid speciation in cotton (*Gossypium*). *Proc. Natl. Acad. Sci. USA* **92**, 280–284.

Xia N. H. 1996 A study on the genus *Bonia* (Gramineae: Bambusoideae). *Kew Bull.* **51**, 565–569.

Xia N. H., Jia L. Z., Li D. Z. and Stapleton C. 2006 Bambusa. In *Flora of China* (ed. Z. Y. Wu and P. H. Raven), vol. 22, pp. 9–38. Science Press, St Louis, USA.

Yang H. Q., Yang J. B., Peng Z. H., Gao J., Yang Y. M., Peng S. and Li D. Z. 2008 A molecular phylogenetic and fruit evolutionary analysis of the major groups of the paleotropical woody bamboos (Gramineae: Bambusoideae) based on nuclear ITS, GBSSI gene and plastid trnL-F DNA sequences. *Mol. Phylogenet. Evol.* **48**, 809–824.

Yang Y. W., Lai K. N. and Tai P. Y. 1999 Molecular phylogenetics studies of *Brassica*, *Rorippa*, *Arabidopsis*, and allied genera based on the internal transcribed spacer region of 18s-25s rDNA. *Mol. Phylogenet. Evol.* **13**, 455–462.

Zhang D. and Sang T. 1999. Physical mapping of ribosomal RNA genes in Peonies (Paeonia, Paeoniaceae) by fluorescent in stiu hybridization: implications for phylogeny and concerted evolution. *Am. J. Bot.* **86**, 735–735.

Zhang L. B. and Ge S. 2007 Multilocus analysis of nucleotide variation and speciation in Oryza officinalis and its close relatives. *Mol. Biol. Evol.* **24**, 769–783.

Received 3 November 2010, in final revised form 15 February 2012; accepted 24 February 2012
Published on the Web: 9 August 2012