Molecular Phylogenetics and Evolution 54 (2010) 266–277

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Phylogeny and biogeography of the rice tribe (Oryzeae): Evidence from combined analysis of 20 chloroplast fragments

Liang Tang [a,b], Xin-hui Zou [a], Gaston Achoundong [c], Christina Potgieter [d], Gerard Second [e], Da-yong Zhang [b], Song Ge [a,f,*]

[a] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China
[b] College of Life Sciences, Beijing Normal University, Beijing 100875, China
[c] Herbier National du Cameroun, BP 1601 Yaoundé, Cameroon
[d] L' Institut de recherche pour le développement (IRD, UMR DIAPC), Montpellier, France
[e] Bews Herbarium, University of Kwazulu-Natal, South Africa
[f] Graduate University of Chinese Academy of Sciences, Beijing 100039, China

ARTICLE INFO

Article history:
Received 30 April 2009
Revised 7 August 2009
Accepted 10 August 2009
Available online 14 August 2009

Keywords:
Biogeography
Chloroplast fragments
Divergence time
Long-distance dispersal
Oryzeae
Phylogeny

ABSTRACT

Based on sequences of 20 chloroplast fragments, we generated a fully resolved phylogeny of Oryzeae and estimated divergence times of its major lineages as well as explored the historical biogeography of the tribe. Our results (1) confirmed the monophyly of Oryzeae and two-subtribe subdivision; (2) indicated that *Maltebrunia*, *Potamophila* and *Prosphytochloa* were genetically distinct enough to deserve generic status but *Maltebrunia* and *Prosphytochloa* were sister groups in the subtribe Oryzinae while *Potamophila* was a member in the subtribe Zizaniinae; (3) suggested that the previously unresolved phylogeny of the subtribe Zizaniinae was most likely explained by insufficient data due to rapid radiation; (4) provided the first well-corroborated timescale for the origin and divergence of Oryzeae, with the crown node of Oryzeae and the deepest split of *Oryza* at ~24 and 15 MYA, respectively; (5) developed a biogeographic history of the tribe and substantiated the important roles of long-distance dispersal in the origin and diversification of the tribe Oryzeae.

1. Introduction

The tribe Oryzeae consists of approximately 11 genera and ~70 species and is distributed throughout the tropical and temperate regions of the world (Clayton and Renvoize, 1986; Vaughan, 1994). This tribe is an important plant group in the grass family that includes the Asian cultivated rice (*Oryza sativa*) and several other economically important species, such as the salt tolerant *Oryza coarctata*, the *Zizania* species that are a part of the cuisine of China and North America (wild rice), and forage species in *Leersia* (Vaughan, 1994; Vaughan and Morishima, 2003; Lu and Ge, 2005). Beyond this, *Oryza*, along with its relatives in Oryzeae, has become an increasingly attractive system for biological studies at the genetic and genomic levels given the completion of the whole genome sequencing of two rice cultivars and well established genomics resources for *Oryza* (Kim et al., 2008) and other grasses (Bennetzen, 2009; Kellogg and Buell, 2009). To take full advantage of rice genetic and genome resources relies significantly on a clear understanding of the phylogenetic relationships of rice and its relatives (Ge et al., 1999; Kellogg, 2009).

As economically and theoretically important groups, rice and closely related genera in the tribe Oryzeae have been studied using various approaches, including morphology (Weatherwax, 1929; Kellogg and Watson, 1993; Terrell and Robinson, 1974; Terrell et al., 2001; Martínez-y-Pérez et al., 2006, 2008), cytology (Nayar, 1973), and molecular markers (Zhang and Second, 1989; Duvall et al., 1993; Ge et al., 2002; Guo and Ge, 2005). In particular, Ge et al. (2002) and Guo and Ge (2005) sampled all extant genera of Oryzeae for the first time in a phylogenetic context except for *Maltebrunia* that was considered to be a synonymy under *Potamophila* by Duistermaat (1987). Based on multigene sequences from chloroplast, mitochondrial and nuclear genomes, Guo and Ge (2005) established the most updated phylogeny of Oryzeae thus far and estimated the divergence time of major lineages in the tribe. They confirmed the monophyly of Oryzeae and its division of two subtribes (Oryzinae and Zizaniinae) and questioned the recognition of three monotypic genera (*Hydrochloa*, *Porteresia*, and *Prosphytochloa*). A recent phylogenomic investigation using sequences of 142 nuclear genes successfully reconstructed the

* Corresponding author. Address: State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, 20 Nanxincun, Xiangshan, Beijing 100093, China. Fax: +86 010 62590843.
E-mail address: gesong@ibcas.ac.cn (S. Ge).

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.08.007

phylogeny of all diploid *Oryza* genomes, which has been debated for a long time, and demonstrated the advantages of genomic information and large data sets in phylogenetic reconstruction (Zou et al., 2008).

Despite these studies, there remain a few uncertainties and unanswered questions with respect to the phylogeny of the tribe Oryzeae. For instances, phylogenetic relationships of genera within the subtribe Zizaniinae, particularly involving the basal lineages and the systematic position of *Hygroryza aristata*, were inconsistent among different studies and among data sets (Zhang and Second, 1989; Duvall et al., 1993; Ge et al., 2002; Guo and Ge, 2005). Taxonomic treatment of three genera, *Maltebrunia*, *Potamophila* and *Prosphytochloa*, has also been debated for centuries (Bentham, 1881; Hubbard, 1967; Clayton, 1970; Duistermaat, 1987) while phylogenetic reconstruction has never been attempted using molecular data due to lack of samples. In addition, different subdivision systems of the tribe (Hubbard, 1959; Terrell and Robinson, 1974; Guo and Ge, 2005) remain to be justified using molecular data with the inclusion of all extant genera.

The tribe Oryzeae is currently distributed in all continents except for Antarctica, with wide ecological amplitude (Clayton and Renvoize, 1986; Vaughan, 1994; Watson and Dallwitz, 1999). In addition to *Oryza* and *Leersia* with pantropical distributions and *Zizania* that is disjunctively distributed in eastern Asia and North America, all other genera in Oryzeae are confined to a specific continent of Asia, Africa, Australia, North and South America (Vaughan, 1994; Watson and Dallwitz, 1999) (Table 1). Such a biogeographic pattern throughout the continents raises an interesting topic regarding the origin and historical biogeographic connection of the Oryzeae lineages. Several studies have speculated the biogeographic history of Oryzeae species (Second, 1985b; Kellogg, 2009), particularly for the genus *Oryza* (Chang, 1976, 1985; Second, 1985a; Wang et al., 1992; Vaughan et al., 2005). Earlier speculations that either *Oryza* (Chang, 1976, 1985) or Oryzeae (Second, 1985b) evolved from a common ancestor in the Gondwana supercontinent before its fracture seem unlikely because these explanations are obviously in conflict with the known facts about the evolution of grasses and monocots (Kellogg, 2001; Gaut, 2002; Vaughan et al., 2005; Vicentini et al., 2008). With the assumption that ancestors of maize and rice separated at ~50 million years ago (MYA), Guo and Ge (2005) used sequences of the chloroplast *matK* and nuclear *GPA1* genes to estimate the divergence times for major Oryzeae lineages under a molecular clock. They suggested that two subtribes of Oryzeae split at ~20 MYA, *Oryza* branched off from the closely related *Leersia* at ~14 MYA, and the crown node age of *Oryza* was ~9 MYA. Based on these estimates, Guo and Ge (2005) discussed the biogeographic implications for the origin and divergence of Oryzeae. Nevertheless, such a molecular clock method based on single or a few genes invoked a number of limitations, including substitution rate heterogeneity among lineages, uncertainties of clock calibration, and no assessment of confidence intervals on dates (Gaut, 2002; Kellogg, 2009). More

Table 1
List of the Oryzeae species and outgroups used in this study.

Taxa	Accession No.[d]	Origin
Chikusichloa aquatica Koidz.	106186	Japan
Chikusichloa mutica Keng	GS0601	China
Hygroryza aristata (Retz.) Nees ex Wright and Arn.	105460	Sri Lanka
Leersia hexandra Sw. (4×)	105252	Philippine
Leersia oryzoides (L.) Sw. (4×)	GS0203	Guangdong, China
Leersia perrieri (A. Camus) Launert	105164	Madagascar
Leersia tisserantii (A. Chev.) Launert	105610	Cameroon
Luziola fluitans (Michx.) Terrell and H. Rob.	L.E. Urbatsh 8434	USA
Luziola leiocarpa Lindm.	82043	Argentina
Maltebrunia letestui (Koechlin.) Koechlin.	GS0801	Cameroun
Oryza sativa L. (A)[a]	Nipponbare	X15901 (GenBank)
Oryza rufipogon Griff. (A)	105480	India
Oryza glaberrima Steud. (A)	102236	Liberia
Oryza meridionalis N.Q. Ng (A)	105282	Australia
Oryza punctata Kotschy ex Steud. (B)	103903	Tanzania
Oryza punctata Kotschy ex Steud. (BC)	100125	India
Oryza malampuzhaensis Kishn. and Chandras (BC)	80768	India
Oryza officinalis Wall. ex G. Watt. (C)	104972	China
Oryza rhizomatis Vaughan (C)	105440	Sri Lanka
Oryza latifolia Desv. (CD)	100167	Costa Rica
Oryza australiensis Domin. (E)	105263	Australia
Oryza brachyantha A. Chev. and Roehr. (F)	105151	Sierra Leone
Oryza granulata Nees et Arn. ex G. Watt. (G)	M8-15	Hainan, China
Oryza neocaledonica Morat (G)	SG0901	New Caledonia
Oryza coarctata Roxb. (HK)	104502	Bangladesh
Oryza ridleyi Hook. f. (HJ)	105366	Thailand
Potamophila parviflora R. Br. (1)	85424	Australia
Potamophila parviflora R. Br. (2)	GS0803	Australia
Prosphytochloa prehensilis (Nees) Schweick. (1) (*Potamophila parviflora*)[b]	—	South Africa
Prosphytochloa prehensilis (Nees) Schweick. (2)	GS0802	South Africa
Rhynchoryza subulata (Nees) Baill.	100913	Argentina
Zizania aquatica L.	J. Alexander 200301	Massachusetts, USA
Zizania latifolia (Griseb.) Turcz. ex Stapf.	GS0202	Beijing, China
Zizaniopsis villanensis Quarin	85425	Argentina
Ehrharta erecta Lam.[c]	B. Bartholomeal 9130	USA
Phyllostachys aurea Carrière ex Riviére and C. Riviére[c]	GS0204	China

[a] Capital letters in parentheses represent the genome type.
[b] This accession was misidentified as *Prosphytochloa prehensilis* and should be *Potamophila parviflora* (see the text).
[c] Outgroups.
[d] All accessions coded by six numbers are provided by the Genetic Resources Center of the International Rice Research Institute (IRRI) at Los Banos, Philippines, and others were collected by the authors. --- unknown origin.

importantly, as pointed out by Kellogg (2009), biogeographic analyses were sensitive to taxon sampling and tree topology, because inclusion of more species and/or minor changes in the phylogeny would affect the inference. Consequently, a reasonable explanation of biogeography relies not only on an accurate estimate of divergence times but also on a reliable and resolved phylogeny.

In the present study, we included all extant genera of Oryzeae, and in particular the genus *Maltebrunia* that was missing in all previous molecular studies. By sequencing 20 chloroplast fragments, we generated a robust phylogeny of Oryzeae and found that the genera *Maltebrunia* and *Prosphytochloa* were most closely related and sister to *Leersia* in the subtribe Oryzinae, which is in striking contrast to previous studies that showed *Prosphytochloa* was a member of the subtribe Zizaniinae. In particular, we fully resolved sequential short interior branches on the tribal tree involving many Asian and American genera, demonstrating that these genera represented a scenario of rapid radiation in the early Miocene. With a robust phylogeny available, we were able to estimate the origin and divergence time of Oryzeae and its major lineages. Using two relaxed molecular clock approaches, we dated the divergence times of major lineages in Oryzeae, and provided the first well-corroborated timescale for the origin and divergence of Oryzeae. On this basis, we investigated the biogeographic history of the tribe and substantiate the important roles of long-distance dispersal in origin and diversification of the tribe Oryzeae.

2. Materials and methods

2.1. Taxon sampling

We included all 11 genera of Oryzeae in the present study by adding to our previous samples (Guo and Ge, 2005) a few of important collections. In the tribe, four monotypic genera are endemic to a specific region, i.e., *Potamophila* (Australia), *Prosphytochloa* (South Africa), *Rhynchoryza* (South America) and *Hygroryza* (Asia). Of them, *Potamophila* and *Prosphytochloa* are most closely related to *Maltebrunia* according to previous morphological studies (Duistermaat, 1987), and the three genera were considered as either a single genus or independent genera by different authors (Bentham, 1881; Clayton, 1970; Duistermaat, 1987; Vaughan, 1994; Guo and Ge, 2005). To clarify this uncertainty, we sampled one species from the genus *Maltebrunia*. *O. coarctata* was previously treated and widely used as a monotypic genus, *Porteresia* Tateoka (Tateoka, 1965). However, molecular phylogenies based on multiple genes from chloroplast, mitochondrial and nuclear genomes all indicated that this species is derived from within *Oryza* (Ge et al., 1999, 2002; Guo and Ge, 2005) and thus was moved back to *Oryza* (Lu and Ge, 2005). A total of 15 *Oryza* species representing all six diploid and four tetraploid genomes were sampled, including *Oryza neocaledonica*, a species that is endemic to New Caledonia (Morat et al., 1994) and has not been used in any previous phylogenetic study. Four representative species from *Leersia* were used, including two diploid and two tetraploid species. The genus *Zizania* consisting of four species is distributed in Asia and North America, and one species from North America (*Zizania aquatica*) and one from Asia (*Z. latifolia*) were included. We sampled two species, one each from Japan and China, from the genus *Chikusichloa* that is distributed exclusively in Eastern Asian with three species in total. *Zizaniopsis* and *Luziola* consist of five and 11 species, respectively, and have almost the same geographic distributions in Central and South America (Vaughan, 1994; Watson and Dallwitz, 1999). We sampled one *Zizaniopsis* species and two *Luziola* species as representatives since morphological and anatomical studies suggested that both genera were good monophyletic groups (Martínez-y-Pérez et al., 2006, 2008). A comprehensive study of subfamilial relationship of Poaceae indicated that subfamily Ehrhartoideae contained three tribes: Ehrharteae, Oryzeae, and Phyllorachideae, and Ehrharteae is sister to Oryzeae (GPWG, 2001; Bouchenak-Khelladi et al., 2008). Therefore, we chose one *Ehrharta* species as an outgroup. One *Phyllostachys* species of the closely related subfamily Bambusoideae was also included as an additional outgroup in the phylogenetic analysis. Detailed information on the sampled species including their scientific names, geographical distribution, accession numbers or vouchers, and countries of origins are listed in Table 1.

2.2. DNA extraction, amplification and sequencing

Total genomic DNA was isolated from silica-gel dried or fresh leaves using cetyltrimethylammonium bromide (CTAB) as described by Ge et al. (1999). In addition to *matK* and *trnL* sequences that were used in our previous studies (Ge et al., 2002; Guo and Ge, 2005), we further obtained 18 chloroplast fragments that were evenly distributed in the large and small single region of the chloroplast genome to improve the phylogenetic resolution. These fragments include protein coding genes, intergenic regions and introns. Primers were designed according to the conserved sequences among rice, maize, and wheat and their sequences are listed in Table S1. Schematic diagrams of the 20 fragments are provided in Fig. S1 (see supplementary materials).

Polymerase chain reactions (PCR) was performed in a total volume of 25 μL containing 5–50 ng of genomic DNA, 5.0 pM of each primer, 0.2 mM of each dNTP, 2.0 mM $MgCl_2$ and 0.75 U ex *Taq* DNA polymerase (TaKaRa, Shiga, Japan). Amplification was carried out in a T gradient 96 U thermocycler (Biomtre, Göttingen, Germany) as follows: 2 min at 94 °C followed by 32 cycles of 30 s at 94 °C, 30 s at 52–58 °C, 90–130 s at 72 °C (depending on the annealing temperature of specific primers and length of the amplified regions) and a final extension at 72 °C for 10 min. Products were examined by gel electrophoresis, and single products of expected size were directly purified using a DNA Purification kit (Amersham Pharmacia Biotech, Piscataway, USA). Multiple bands products were first separated by 1.5% agarose gel stained with ethidium bromide, and then the expected size bands were incised and gel-purified with a DNA Purification kit (Amersham Pharmacia Biotech, Piscataway, USA). Both strands of the resulting products were sequenced by using an ABI 3730 DNA sequencer (Applied Biosystems, Forster City, CA, USA).

2.3. Phylogenetic analysis

Sequences were aligned using ClustalX version 1.83 (Thompson et al., 1997) with additional manual refinements. The homogeneity across 20 fragments was tested using the incongruence length difference (ILD) test (Farris et al., 1994), as implemented in PAUP 4.0b10 (Swofford, 2002). Heuristic searches were performed with 10 random addition replicates and bisection–reconnection (TBR) branch swapping.

Single fragments and concatenated data set were analyzed separately by maximum parsimony (MP) and maximum likelihood (ML) criterions using PAUP 4.0b10 (Swofford, 2002). For MP analysis, heuristic searches were performed with 1000 random addition replicates followed by tree bisection–reconnection branch swapping. All characters were unordered and equally weighted with gaps treated as missing data. Topological robustness was assessed by 1000 bootstrap replicates. ML heuristic tree searches were carried out with the selected substitution model, and random taxon addition was repeated for 10 times followed by tree bisection–reconnection branch swapping. As ML tree search was time consuming using PAUP, we adopted PHYML2.4.5 (Guindon and Gascuel, 2003) to carry out 500 bootstrap replicate analyses.

Bayesian inference (BI) was conducted under MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Two independent runs of Metropolis-coupled Markov chain Monte Carlo were conducted simultaneously, with each run being one cold chain and three incrementally heated chains and all started randomly in the parameters space. Five million generations were run and every 1000 generations were sampled with the first 25% of samples discarded as burn-in. Tracer 1.4 was further used to check whether the chains have been converged (Rambaut and Drummond, 2007). In ML and BI analyses the best nucleotide substitution models for each data set were selected using Modeltest 3.7 by corrected Akaike information criterion (Posada and Buckley, 2004).

The approximately unbiased (AU) test (Shimodaira, 2002) was used for inferring the confidence of phylogenetic tree selection. This method adopts a multiscale bootstrap technique to reduce tree selection bias stemming from simultaneous comparison of many trees (Goldman et al., 2000). We performed the AU test, as implemented in the program CONSEL (Shimodaira and Hasegawa, 2001), to test whether sequential short interior branches on the trees would collapse and be recognized as hard polytomies. Baseml in PAML4.1 was used to calculate the sites likelihood for the concatenated data set (Yang, 2007), and the general time reversible substitution model plus five categories of discrete gamma distribution was used for likelihood calculation. Default scale parameters and numbers of replicates were used for the *P* value assessment.

2.4. Divergence time estimation

We used two relaxed-clock approaches, the Bayesian (Thorne et al., 1998; Thorne and Kishino, 2002) and penalized likelihood (PL) (Sanderson, 2002) methods, for dating approximate divergence times within Oryzeae, because rate constancy was rejected for the combined chloroplast data. These two approaches rely on appropriate calibrations. Many previous studies (i.e., Bremer, 2002; Gaut, 2002; Prasad et al., 2005; Vicentini et al., 2008) have estimated the age of the grass family and divergence time of major lineages within grasses but obtained different age estimations due to differences in the approaches used, different strategies in taxon and gene sampling, as well as different calibration points. Recently, Vicentini et al. (2008) used six fossils as calibration points and several calibration schemes to have obtained the most robust estimate to date for divergence times of major grass lineages. In their study, a divergence time of 34.5 ± 6.8 MYA between the tribes Oryzeae and Ehrharteae was obtained and thus used for calibrating the age of the stem node of Oryzeae. Two macrofossil records belonging to tribe Oryzeae have been reported to date. One silicified anthoecia (fertile lemmas and paleas) was found at Nebraska (North America) in the Miocene deposits (Thomasson, 1980) and described as *Archaeoleersia nebraskensis* Thomasson, which has features in common with the living *Leersia ligularis* Trin. The other was the spikelets found in a Miocene excavation in Germany (Heer, 1855) and was identified as *Oryza exasperata* (A. Braun) Heer. (Heer, 1855), which appears to be close to the extant *Oryza granulata* on the basis of morphology (G. Second pers. obs.). Because the absolute time of the two Miocene fossils was not indicated precisely (Heer, 1855; Thomasson, 1980), we used 5 MYA conservatively as the minimum age constraint for the crown nodes of *Oryza* and *Leersia* in our divergence time estimation. Collectively, these three calibration points were used to estimate dates of the main divergence events within Oryzeae.

We first performed a Bayesian relaxed-clock approach as implemented in Multidivtime program (Thorne et al., 1998; Thorne and Kishino, 2002). This method can deal with multiple loci and multiple fossil calibrations and adopts lower and upper bounds for node age constraints. It relaxes the rate constant hypothesis by imposing a prior autocorrelated rate variation model for substitution rate variation of nodes (Thorne et al., 1998; Thorne and Kishino, 2002). The combined chloroplast data was partitioned into five data partitions (three codon positions, intergenic and intron sequences) to account for their heterogeneous evolutionary rates, while a common set of divergence times was assumed to be shared by all data partitions (Thorne and Kishino, 2002). First, model parameters of the F84+Γ model were estimated for each partition by using Baseml in PAML 4.1 package (Yang, 2007). Then, maximum likelihood estimations of branch lengths and their variance–covariance matrix were obtained by Estbranches in Multidivtime program (Thorne et al., 1998). Lastly, Multidivtime was used to conduct MCMC simulation to estimate posterior distribution of divergence times as well as 95% credibility intervals, and the difference in substitution processes among partitions was account for by estimating branch lengths for individual partitions. Multidivtime takes advantages of a multivariate normal distribution to approximate the likelihood surface for the sake of saving MCMC computation (Thorne et al., 1998). The prior distributions for the age and the evolutionary rate of the ingroup root, as well as the rate variation parameter should be specified before running MCMC procedure. We used 34.5 MYA as the mean age of the ingroup root (rttm) as estimated by Vicentini et al. (2008) and indicated above. By dividing the median genetic distance between the ingroup root and tips by rttm as recommended in the Multidivtime manual, we obtained the mean rate of the ingroup root (rtrate) to be 0.1. The rate variation parameter (brown) determines the amount of rate variation per unit time (Thorne et al., 1998) and the mean of brown was calculated to be 6 when the product of brownmean and rttm was set to 2 as required. The standard deviations of the three priors were set to the same as their means to account for the high degree of uncertainty embedded in these parameters (Thorne and Kishino, 2002). When implementing MCMC runs, the first one hundred thousand generations were discarded as burn-in, and then every one hundred generations were sampled until a total of ten thousands samples were collected.

As an independent cross-check of divergence time estimations, we also used a penalized likelihood (PL) method (Sanderson, 2002), as implemented in program r8s (Sanderson, 2003). PL is a semiparametric smoothing method that assumes an autocorrelation in substitution rates by minimizing rate variation across branches on a tree. In this analysis, branches are allowed to change evolutionary rates but penalized when rates change across branches. A smoothing parameter is designed to determine the relative contribution of the penalty function, the optimal value of which is specified by a data-driven cross-validation criterion (Sanderson, 2002). The ML tree topology and its branch lengths from the concatenated data set were subjected for PL analysis. We used additive penalty function and tested the optimal value for the smoothing parameter on a log scales ranging from −5 to 5 with each step increasing by one order of magnitude. The level of optimal smoothing was estimated to be 0.1, allowing substantial rate variation among lineages of the concatenated chloroplast data. One hundred bootstrapped data sets were generated by Seqboot (Felsenstein, 2004) and branch lengths for each bootstrapped data set were estimated by PHYML2.4.5 with topology fixed as the ML tree. Divergence times of the 100 bootstrapped trees were obtained by using r8s, and the standard deviation (SD) of estimated node ages were calculated by *profile* command in r8s. Confidence intervals of node ages were estimated to be the node ages plus and minus twice the SD.

2.5. Ancestral area reconstruction

To infer historical scenarios for the biogeography of the rice tribe, dispersal–vicariance (DIVA) analysis, as implemented by DIVA 1.1 (Ronquist, 1996, 1997), was used to reconstruct ancestral distributions on the ML tree generated from the combined 20

chloroplast fragments. Four areas of endemism were defined based on the present distributions of Oryzeae species: Asia, Australia, America, and Africa. DIVA treats vicariance as a default model of speciation, and reconstructs ancestral areas by minimizing the number of dispersal and extinction events required to explain the observed distribution pattern. Species on the ML tree were replaced with their distribution areas, and then the ancestral area optimization was carried out with the number of ancestral areas either not constrained or constrained to be not more than two (Ronquist, 1996, 1997; Donoghue et al., 2001).

3. Results

3.1. Sequence characteristics

We successfully obtained 20 chloroplast fragments from all 34 species. The aligned length ranged from 586 base pair (bp) (*trnL*^{UAA} intron) to 2006 bp (*ndhF*), with total length of the 20 fragments being 21,149 bp. As shown in Table 2, the proportion of variable sites range from 10.6% (*ycf3* intron) to 22.3% (*rps16-trnQ*^{UUG}) and that of informative sites range from 6.17% (*rpl20-clpP*) to 13.04% (*rps16-trnQ*^{UUG}). As expected, the combined intergenic regions have the highest proportion of variable and informative sites (18.63% and 10.29%) followed by the combined introns (13.24% and 7.93%), and the combined coding regions have the lowest polymorphism (12.33% and 6.52%). The GC content of the 20 fragments varied between 29.5% (*rps16-trnQ*^{UUG}) and 41.1% (*ycf3* intron), but no significant difference in base frequency was found for any fragment among the Oryzeae species ($P > 0.98$). All sequences generated in this study have been deposited in GenBank, and their accession numbers are FJ908096–FJ908705.

3.2. Phylogenetic analysis

Phylogenetic trees generated by single chloroplast fragments were essentially similar but their resolutions were low as expected. High values of CI and RI index for all single fragments (CI = 0.788–0.880; RI = 0.863–0.940) suggest a low level of homoplasy in these fragments. To evaluate the suitability of combining the 20 fragments, we performed the incongruence length difference (ILD) test and found no significant incongruence among 20 fragments as a whole ($P = 0.228$). The result is consistent with the hypothesis that all chloroplast fragments evolved along the same underlying topology as a single gene and can be combined in phylogenetic analysis (Doyle, 1992; Ickert-Bond and Wen, 2006).

Phylogenetic analyses of the concatenated sequences using maximum parsimony (MP), maximum likelihood (ML) and Bayesian inference (BI) all resulted in a single tree with high bootstrap support or Bayesian posterior probability (PP) for all internal branches. Different rooting strategies using either *Phyllostachys* or *Ehrharta* or both as outgroups did not change the ingroup topologies. Fig. 1 shows the ML tree with bootstrap support values from ML and MP analyses. The posterior probabilities for all interior branches are 1.0 on the BI tree (see Fig. S2 in supplementary materials). The MP and BI trees are provided in Fig. S2 (see supplementary materials). As shown in Fig. 1, all 33 internal branches on the tree are resolved with 24 nodes having 100% support values for both MP and ML methods. The genera with multiple species are all monophyletic with 100% bootstrap supports. The tribe divides clearly into two main clades (I and II). The first clade (I) includes four genera, with *Oryza* sister to the other three genera (*Leersia*, *Maltebrunia*, *Prosphytochloa*). In this clade, *Maltebrunia letestui* and *Prosphytochloa prehensilis*-2 are most closely related, and sister to *Leersia*. The second clade (II) consists of the remaining seven genera and *Chikusichloa* is at the basal position, followed by *Potamophila* lineage (plus *Prosphytochloa prehensilis*-1) that is sister to the other five genera. It should be noted that two accessions of *Prosphytochloa prehensilis* are nested into two separate clades, with *P. prehensilis*-1 in the clade II and *P. prehensilis*-2 in the clade I. Particularly, the sequence identity between *Prosphytochloa prehensilis*-1 and *Potamophila parviflora* (99.4%) is much higher than that between *Prosphytochloa prehensilis*-1 and *Prosphytochloa prehensilis*-2 (93.7%). Repeated DNA extraction, PCR and sequencing for two *Prosphytochloa* accessions obtained the same results. A careful examination of the morphology and herbarium record of this species indicated that *Prosphytochloa prehensilis*-1 used here and in

Table 2
Characteristics of 20 single and combined chloroplast fragments (excluding outgroups).

cpDNA region	Aligned length (bp)		Number of variable sites (%)	Number of informative sites (%)	GC%	Model selected by AICc
	All	Coding				
atpB-rbcL	978	204	133 (13.60)	77 (7.87)	34.6	TVM+G
atpF intron	870	51	105 (12.07)	62 (7.13)	32.8	TIM+G
atpI-atpH	923	105	145 (15.71)	70 (7.58)	34.2	TVM+G
rps16-trnQ^{UUG}	1112	0	248 (22.30)	145 (13.04)	29.5	TVM+G
matK	1574	1574	261 (16.58)	135 (8.58)	34.3	TVM+I+G
ndhA intron	1001	18	149 (14.89)	93 (9.29)	32.5	TVM+G
ndhC-trnV^{UAC}	1161	300	202 (17.40)	116 (9.99)	35.7	GTR+G
ndhF	2006	2006	279 (13.91)	136 (6.78)	34.0	TVM+I+G
petB intron	1106	183	136 (12.30)	89 (8.05)	35.7	TIM+G
rpl20-clpP	1751	933	193 (11.02)	108 (6.17)	39.5	TVM+I+G
rps3-rps19	1259	1146	160 (12.71)	85 (6.75)	34.6	TVM+I+G
rps16 intron	910	24	135 (14.84)	70 (7.69)	34.4	TIM+G
trnC^{GCA}*-rpoB*	1061	0	196 (18.47)	100 (9.43)	32.3	TVM+G
trnG^{UCC} intron	728	0	88 (12.09)	53 (7.28)	35.0	HKY+I+G
trnL^{UAA} intron	586	0	89 (15.19)	52 (8.87)	35.0	HKY+G
trnP^{UGG}*-rpl33*	1234	306	169 (13.70)	86 (6.97)	32.8	GTR+G
trnS^{UGA}*-trnfM*^{CAU}	1305	186	235 (18.01)	123 (9.43)	40.1	TVM+G
trnT^{GGU}*-trnD*^{GUC}	1009	0	194 (19.23)	120 (11.89)	37.6	K81uf+G
trnV^{UAC} intron	744	0	83 (11.16)	49 (6.59)	37.6	TVM+G
ycf3 intron	857	123	91 (10.62)	54 (6.30)	41.1	K81uf+G
coding combined	7152	7152	882 (12.33)	466 (6.52)	36.3	TVM+I+G
intergenic combined	8636	0	1609 (18.63)	889 (10.29)	33.3	GTR+I+G
introns combined	5361	0	710 (13.24)	425 (7.93)	35.2	GTR+I+G
Total	21,149	7152	3201 (15.14)	1780 (8.42)	35.2	TVM+I+G

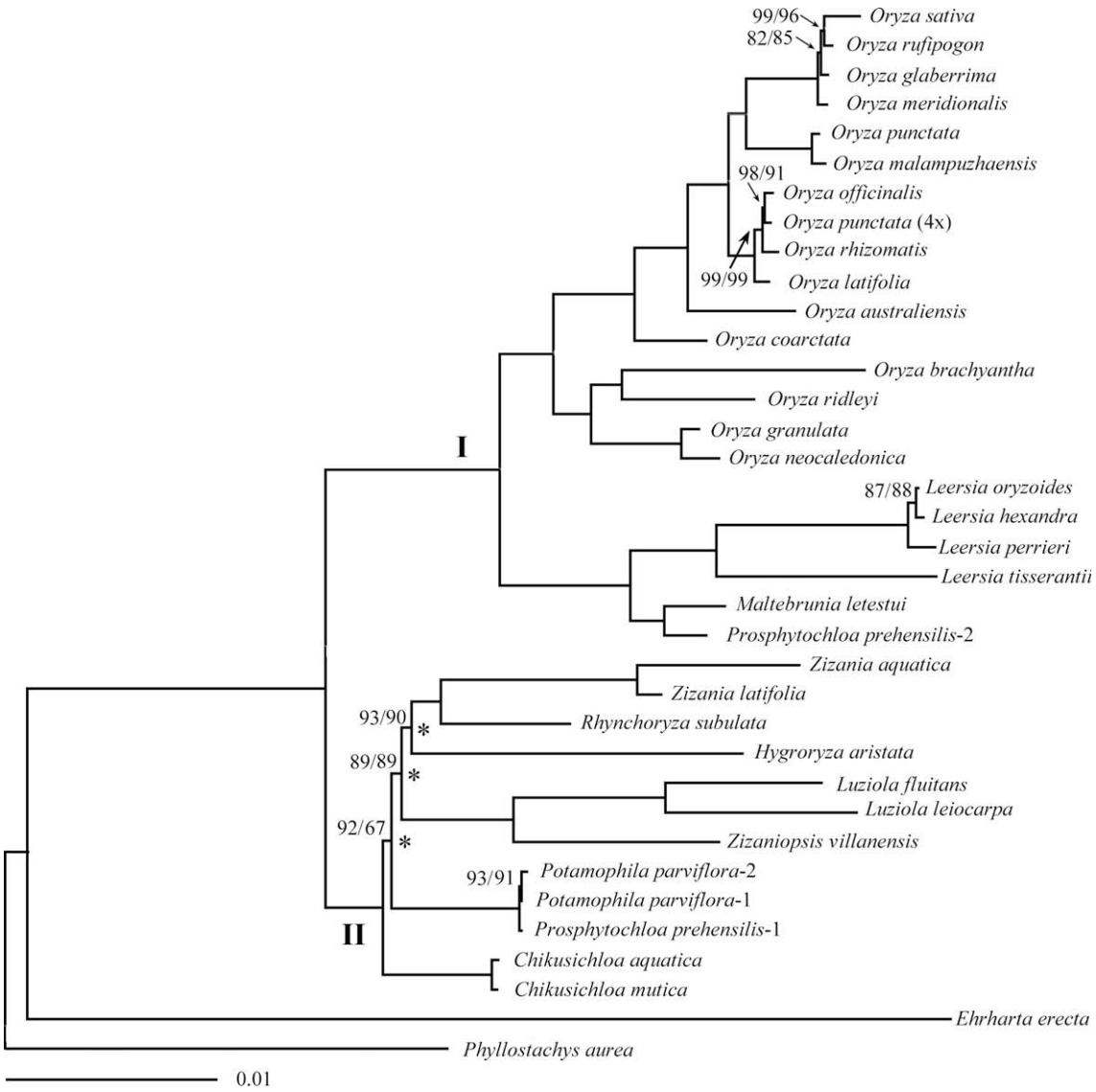

Fig. 1. ML phylogeny of Oryzeae inferred from the concatenated 20 chloroplast fragments under the TVM+I+G model. MP and BI inferences generated the same topology (Fig. S2 in supplementary materials). Numbers near branches are bootstrap percentages of ML and MP, respectively. The branches without numbers indicate 100% bootstrap supports. Stars indicate three successive short interior branches in clade II.

previously studies (Ge et al., 2002; Guo and Ge, 2005), was misidentified and should be *Potamophila parviflora*. A striking result is that *Maltebrunia* and *Prosphytochloa* formed a clade with sequence identity of 95.3% and are sister to *Leersia* in clade I (Fig. 1).

Within clade I, there are two apparent lineages, i.e., one including all *Oryza* species and the other consisting of three genera (*Leersia*, *Maltebrunia*, *Prosphytochloa*). The phylogenetic relationships of major *Oryza* lineages are nearly similar to that found in a recent phylogenomic study (Zou et al., 2008) except for F- and G-genomes, which were grouped into sister clade in this study. Such a sister relationship between F- and G-genomes was observed in previous phylogenetic studies using cytoplasmic gene sequences (Ge et al., 2002; Guo and Ge, 2005). *Oryza neocaledonica* is sister to *O. granulata*, confirming a close affinity of the two species (Vaughan and Morishima, 2003). Within clade II, all species-level relationships are fully resolved and all internal branches obtain bootstrap supports over 89% in both MP and ML trees except for one branch with 67% bootstrap support in MP tree (Fig. 1). It is interesting to note that three short but resolved internal branches occur at the basal positions of clade II, with *Chikusichloa* being the earliest divergent lineage. Two other subclades in clade II are sister to each other, with one subclade consisting of three genera (*Hygroryza*, *Rhynchoryza*, and *Zizania*) and the other of *Zizaniopsis* and *Luziola*.

To test whether the three sequential short internal branches at the base of clade II could be recognized as a polytomy, we performed an AU test as follows. We first collapsed one branch at a time, which represents one time of simultaneous speciation events. Then, two of three short branches were randomly selected and collapsed to represent two simultaneous speciation events. Finally, all three short internal branches were collapsed to represent three simultaneous speciation events. The seven partially resolved topologies were compared with the fully resolved ML tree (Fig. 1) by an AU test. All the collapsed trees were poorly resolved in comparison to the fully resolved ML tree ($P < 0.005$), suggesting that the three short internal branches differed remarkably from zero and thus did not represent a hard polytomy on the tree.

3.3. Divergence time estimation

The likelihood ratio test showed that the strict molecular clock was rejected for the concatenated chloroplast data set ($P < 0.001$).

Table 3
Divergence time estimation of the major Oryzeae lineages in millions of years plus their 95% credibility intervals based on Bayesian and penalized likelihood approaches.

Node in Fig. 2[a]	Divergence	Multidivtime Posterior mean and 95% credibility intervals	Penalized likelihood Estimated age and 95% credibility intervals
1 (C1)	Ehrharteae *vs.* Oryzeae	35.61 (28.47, 41.02)	34.5 (non)
2	Crown node of Oryzeae	24.37 (18.79, 29.87)	23.86 (22.82, 24.90)
3	Crown node of clade I	17.48 (12.91, 22.72)	16.79 (15.63, 17.95)
4 (C2)	Crown node of genus *Oryza*	15.03 (10.84, 20.03)	14.05 (12.65, 15.45)
5	A-, B-, C-genomes *vs.* E-genome	7.02 (4.63, 10.26)	6.26 (5.22, 7.30)
6	A-, B-genomes *vs.* C-genome	4.83 (3.02, 7.44)	4.17 (3.33, 5.01)
7 (C3)	Crown node of *Leersia*	7.61 (5.22, 11.26)	7.59 (6.79, 8.39)
8	Crown node of clade II	21.15 (15.99, 26.55)	20.41 (19.17, 21.65)
9	*Chikusichloa vs. Potamophlia*	20.56 (15.45, 25.90)	19.85 (18.55, 21.15)
10	(*Hygroryza*, *Rhynchoryza*, *Zizania*) *vs.* (*Zizaniopsis*, *Luziola*)	19.62 (14.68, 24.87)	19.34 (18.04, 20.64)
11	*Hygroryza vs.* (*Rhynchoryza*, *Zizania*)	18.90 (14.06, 24.08)	18.76 (17.38, 20.14)

[a] C1, C2 and C3 are three fossil calibrations used in divergence time estimation (see Section 2). The node numbers correspond to those on the tree in Fig. 2.

Therefore, we used two relaxed-clock approaches for divergence time estimation of the Oryzeae lineages. Divergence time estimates of major clades and three calibration points are shown in Table 3 and Fig. 2. In the Bayesian estimation, the divergence of two major clades (clades I and II) (node 2 in Fig. 2) is estimated to be 24.4 (18.8–29.9) MYA (Table 3), roughly between the late Oligocene and the Early Miocene. Clade I split further into *Oryza* and the lineage consisting of *Leersia*, *Prosphytochloa* and *Maltebrunia* (node 3 in Fig. 2) at the Early Miocene (~17.5 MYA). The deepest split in the genus *Oryza* (node 4 in Fig. 2) is estimated to be at ~15

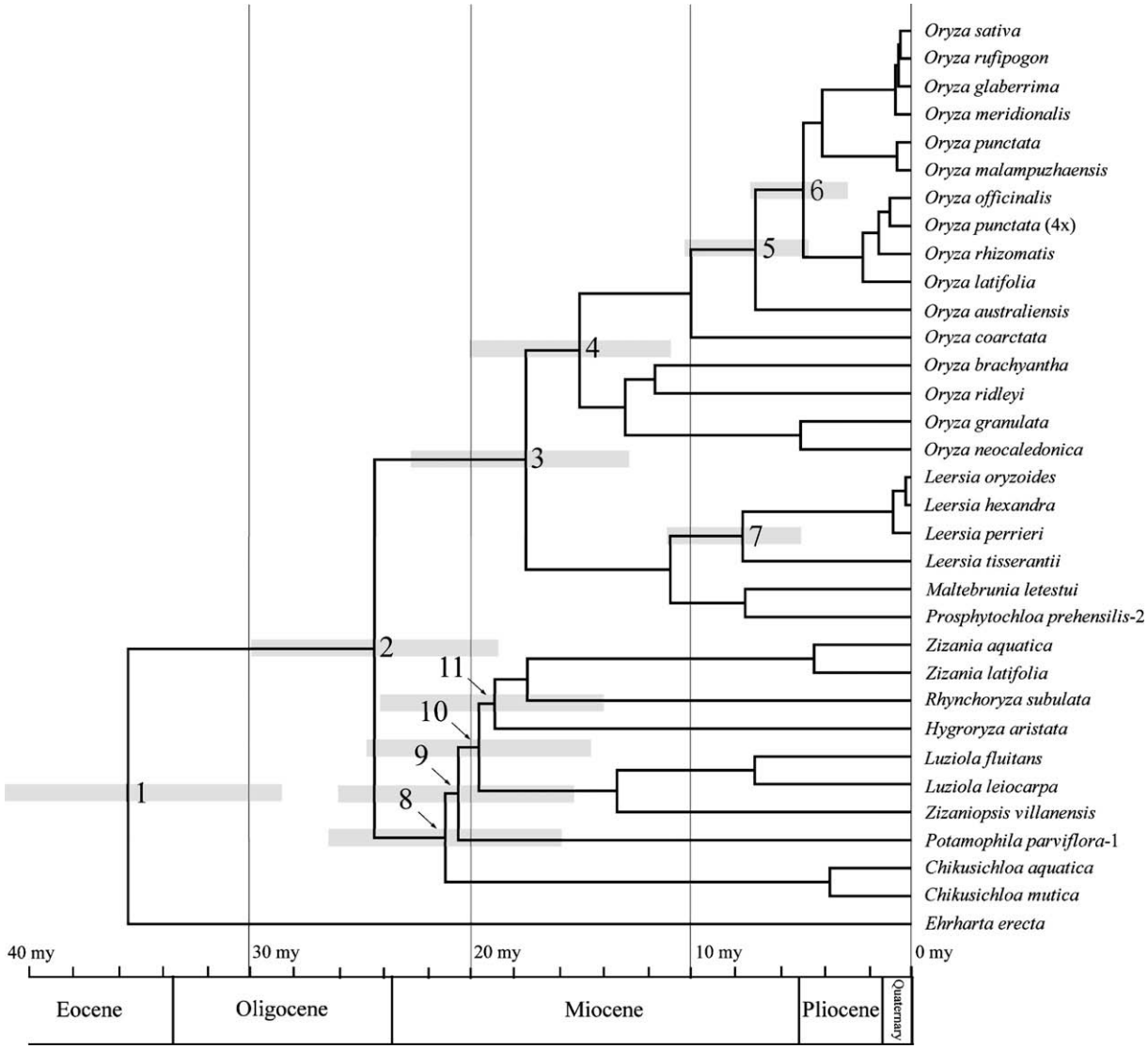

Fig. 2. Chronogram obtained for Oryzeae under a Bayesian relaxed-clock approach as implemented in Multidivtime based on the combined 20 chloroplast fragments. Numbers adjacent to the nodes correspond to node numbers in Table 3. Gray boxes indicate 95% confidence intervals on nodal ages.

(10.8–20.0) MYA (Table 3), the Middle Miocene. The radiation of A-, B-, and C-genomes (node 6 in Fig. 2) began at ~4.8 (3.0–7.4) MYA, in the Early Pliocene. The initial diversification of clade II appears to have occurred at ~21.2 (16.0–26.6) MYA, giving rise to four lineages within a short time interval. At the basal positions in clade II, three short internal branches (nodes 9 to 11 in Fig. 2) reflect three rapid divergence events that happened in the Early Miocene within less than 2.5 million years (21.15–18.90 = 2.25) (Table 3). The relaxed-clock penalized likelihood (PL) method generated slightly smaller but very similar estimates of divergence times for all Oryzeae lineages, with much smaller 95% confidence intervals (Table 3).

3.4. DIVA analyses

Dispersal–vicariance analyses were first conducted using DIVA without constraints. Nineteen dispersal events were required to explain the current geographical distribution of the Oryzeae lineages. Our results suggest that many nodes of the tree are widespread in all four areas, including the common ancestor of the tribes Oryzeae and Ehrharteae, the ancestor of Oryzeae, and the ancestor of clade II. Two other nodes were widespread in three areas (Fig. S3 in supplementary materials). Such uninterpretable results have been found in many previous analyses where multiple ancestral areas were present for some nodes (e.g., Davis et al., 2002; Yuan et al., 2005; Zhou et al., 2006; Inda et al., 2008), particularly at deeper nodes in the phylogeny (Kellogg, 2009). Fig. 3 shows the results of an analysis in which the number of ancestral areas was constrained to two. The two-area optimization required only one more dispersal event (20) and suggested that the common ancestor of Oryzeae and Ehrharteae was probably distributed either in Asia and Australia or in Asia and Africa (AB/AD). Due to the uncertainties regarding the biogeographic history of Ehrharteae, our analyses focused on biogeographic patterns within the tribe Oryzeae. The DIVA analysis revealed Asia (area "A" in Fig. 3) as the most likely region of origin of Oryzeae. At early diversification of the tribe, a few of the initial dispersals of the ancestor of Oryzeae were revealed, including from Asia (area "A") to Australia (area "B") (*Potamophila*) and/or to America (area "C") where vicariances and/or dispersal events gave rise to American genera (*Luziola*, *Rhynchoryza*, *Zizania*, and *Zizaniopsis*).

For clade I (the subtribe Oryzinae), DIVA specifies that the common ancestor of the genera *Leersia*, *Maltebrunia* and *Prosphytochloa* dispersed from Asia to Africa (area "D"). This single dispersal event resulted in a vicariance pattern between Asia and Africa and was followed by possible vicariance in the African lineage. The genus *Oryza* originated in Asia and subsequently expanded to Australia, Africa and America though multiple dispersals and recent vicariances, giving rise to the species diversity of this genus across these continents. The DIVA analyses indicated that at least eight dispersals were required to explain the observed biogeographic pattern of the *Oryza* species, with four from Asia to Africa and three from Asia to Australia and once from Asia to America (Fig. 3).

4. Discussion

4.1. A revised phylogeny of the rice tribe with special emphasis on the systematic position of Maltebrunia and Prosphytochloa

With a much larger data set of the combined 20 chloroplast fragments and the inclusion of additional species, the present study obtained an updated phylogeny of Oryzeae. Apart from one additional accession from both *Potamophila* and *Prosphytochloa*, we included, for the first time, a representative of *Maltebrunia* in our phylogenetic reconstruction. An important finding is that the newly sampled *Prosphytochloa prehensilis*-2, sister to *Maltebrunia*

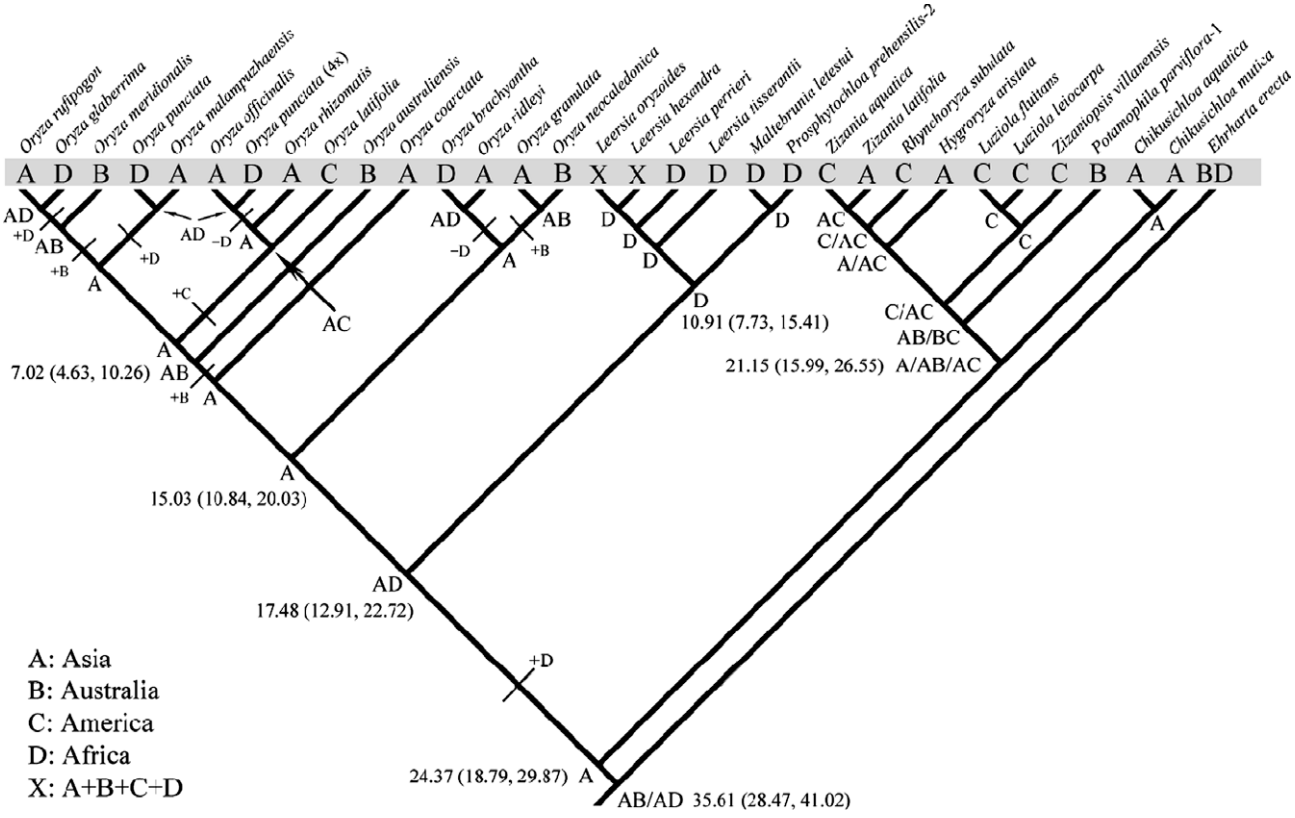

Fig. 3. Ancestral area reconstruction for internal nodes of the phylogeny shown in Fig. 1 using dispersal–vicariance analysis (DIVA) under two-area optimization. Capital letters (A–D) represent four areas of endemism. Double letters indicate two coexisting ancestral areas. A bar cross a clade reflects a dispersal event to a new area prefixed by "+".

letestui, formed a sister group to *Leersia* species in the subtribe Oryzinae (clade I), whereas *Prosphytochloa prehensilis*-1 was nested within the subtribe Zizaniinae (clade II). By carefully checking the gross morphology of the two accessions, along with examination of the herbarium record, we found that *Prosphytochloa prehensilis*-1 has characteristics of *Potamophila* and is a misidentified accession. Thus, previous phylogenies where *Prosphytochloa* formed a monophyletic group with *Potamophila* (Ge et al., 2002; Guo and Ge, 2005) were erroneous due to a misidentified sample of *Prosphytochloa* and the lack of *Maltebrunia*. Similar cases with misidentified accessions have happened frequently involving *Oryza* species in different genome groups (Wang et al., 1992; Ge et al., 2001; Zhu and Ge, 2005).

The genera *Potamophila* and *Maltebrunia* were originally described in Brown (1810) and Kunth (1829), respectively. Bentham (1881) suggested that *Maltebrunia* should be put into *Potamophila*. An additional closely related genus, *Prosphytochloa*, was established later on by Schweickerdt (1961). Disagreement regarding the taxonomic treatment of these three genera has remained since then. The subsequent morphological comparative studies by de Winter (1951) indicated significant morphological difference between *Potamophila prehensilis* (Nees) Benth and *Potamophila parviflora* R. Br., and therefore these two species should not coexist in the same genus. Other authors (Hubbard, 1967; Clayton, 1970) suggested to split *Potamophila* into three separate genera, including *Maltebrunia* and two monotypic genera, *Prosphytochloa* and *Potamophila*. Nevertheless, such a treatment was questioned by Duistermaat (1987) because these genera did not have fundamental differences in gross morphology, especially in the structure of spikelets. Recent molecular phylogenetic studies were not able to provide convincing evidence in the absence of *Maltebrunia* (Ge et al., 2002; Guo and Ge, 2005). Remarkably, the present study shows that *Potamophila* is highly differentiated genetically from both *Prosphytochloa* and *Maltebrunia* and the latter two also differ to some extent (Fig. 1). These findings provide convincing evidence that *Maltebrunia*, *Potamophila* and, to some extent, *Prosphytochloa* are genetically distinct enough to deserve generic status. *Prosphytochloa* and *Maltebrunia* are more closely related to the genus *Leersia* in the subtribe Oryzinae than to *Potamophila* in the subtribe Zizaniinae.

In general accordance with the results by Guo and Ge (2005), the present study supports the monophyly of Oryzeae and the subdivision of the tribe into two traditionally recognized subtribes, Oryzinae and Zizaniinae (clades I and II, respectively, in Fig. 1). A three-subtribe system (Terrell and Robinson, 1974) does not seem to be justified. Consequently, to account for this updated phylogeny, we suggest recognition of only two subtribes: (1) subtribe Oryzinae that consists of *Leersia*, *Maltebrunia*, *Oryza*, and *Prosphytochloa*; (2) subtribe Zizaniinae that includes seven genera, *Chikusichloa*, *Hygroryza*, *Luziola*, *Potamophila*, *Rhynchoryza*, *Zizania*, and *Zizaniopsis*.

4.2. Rapid radiation of the subtribe Zizaniinae and its implications for phylogenetic reconstruction

Despite numerous efforts to reconstruct the phylogeny of Oryzeae, only two studies (Ge et al., 2002; Guo and Ge, 2005) have sampled all the genera of this tribe except for *Maltebrunia* that was not recognized by some authors (e.g. Duistermaat, 1987). In the first study, Ge et al. (2002) generated a molecular phylogeny of Oryzeae based on sequences of chloroplast *matK* gene. Of the two major clades, the clade corresponding to the subtribe Zizaniinae was largely unresolved. Guo and Ge (2005) conducted a phylogenetic analysis of Oryzeae using sequences of two chloroplast genes (*matK* and *trnL*), one mitochondria gene (*nad1*) and two nuclear genes (*Adh1* and *GPA1*). Results showed that the combined data set of three cytoplasmic genes did not resolve the intergeneric relationships within the subtribe Zizaniinae and tree topologies were inconsistent between the two nuclear genes. In contrast, with respect to the subtribe Oryzinae, tree topologies were consistently and fully resolved with both the combined cytoplasmic genes and two single nuclear genes (Guo and Ge, 2005). Therefore, Guo and Ge (2005) suggested that such a phylogenetic uncertainty in the subtribe Zizaniinae was probably the result of rapid speciation, though other causes such as insufficient data, lineage sorting and hybridization could not be excluded entirely.

In this study, we used much larger data set to have obtained a phylogenetic tree, within which all genera in clade II (the subtribe Zizaniinae) were fully resolved. It is noteworthy that the internal branches of the basal lineages in clade II were very short relative to those in clade I (the subtribe Oryzinae) and involved six genera (Fig. 1). AU test showed that seven partially resolved trees were significantly worse than the fully resolved tree, indicating that the three short internal branches were reliable and did not represent hard polytomies. These observations indicate that the basal lineages in clade II radiated in rapid succession following the early split of the tribe Zizaniinae. Divergence time estimates showed that this radiation occurred within a time span of 2.25 million years (Fig. 2 and Table 3). Such a closely spaced series of speciation events is the most likely explanation for the incongruence among gene trees and unsolved topologies in previous investigations on the phylogeny of the subtribe Zizaniinae (Zhang and Second, 1989; Ge et al., 2002; Guo and Ge, 2005).

Rapid speciation or radiation is often featured by short internal branches in phylogenetic trees and is increasingly appreciated as molecular data have accumulated in recent decades (e.g., Fishbein et al., 2001; Verboom et al., 2003; Wortley et al., 2005; Rokas and Carroll, 2006; Whitfield and Lockhart, 2007; Zou et al., 2008). A growing body of evidence indicates that many previously unresolved phylogenies or polytomies attributed to rapid evolutionary radiations could be resolved with additional data and proper phylogenetic methods (Walsh et al., 1999; Wortley et al., 2005; Jian et al., 2008; Zou et al., 2008). In a recent study on phylogenetic reconstruction of the diploid genomes of *Oryza*, Zou et al. (2008) used 142 single-copy genes to fully resolve the relationships among all diploid genome types of *Oryza* and demonstrated that rapid speciation in an angiosperm genus can be resolved as long as a sufficient number of unlinked genes are sampled. A fully resolved phylogeny of the subtribe Zizaniinae suggests that inconsistent or unresolved phylogenetic relationships of this subtribe in previous studies can be explained by insufficient data as a result of rapid speciation involving these lineages.

4.3. Divergence time and biogeographic history of major Oryzeae lineages

Although the origin and divergence times of major grass lineages have been extensively investigated and remain actively debated (e.g., Kellogg, 2001; Bremer, 2002; Gaut, 2002; Verboom et al., 2003; Inda et al., 2008), relatively few empirical studies have been conducted on the divergence times of Oryzeae and their relatives (Guo and Ge, 2005; Kellogg, 2009). Guo and Ge (2005) was the single study so far to have obtained preliminary estimates of divergence times for major Oryzeae lineages using sequence data. They placed the crown node (divergence between clades I and II in Fig. 1) at ~20 MYA, the divergence of *Oryza* from *Leersia* at ~14 MYA and the crown node of *Oryza* at ~9 MYA. The present study used two relaxed-clock approaches to have obtained somewhat older divergence times (Table 3 and Fig. 2) than those estimated by Guo and Ge (2005). We estimate that the crown node of Oryzeae, the divergence between *Oryza* from *Leersia* and the deepest split of *Oryza* at ~24, 17, and 15 MYA, respectively. The

elevated estimates of divergence times of the Oryzeae lineages in this study seem reasonable and mostly likely reflect the following factors: (1) two important genera (*Maltebrunia* and *Prosphytochloa*) were included in this analysis, which might help generate a more robust phylogeny of this tribe; (2) a larger chloroplast dataset with a total of 22,175 bp in length provided higher resolution and thus generated a more fully resolved phylogenetic tree; (3) relaxed-clock methods under multiple temporal calibrations would not have suffered from a few flaws in previous studies, including different substitution rates among lineages, errors arising from a single calibration point, and no assessments of confidence intervals on dates. However, our calibration of the molecular clock was conservative and corresponds to the widely accepted time frame for grass evolution (Vicentini et al., 2008; Kellogg and Buell, 2009). A more ancient origin of grasses has been proposed recently based on the observation of oryzoid phytoliths from India dated at ~75 MYA (Prasad et al., 2005) but these need to be substantiated with additional fossils from different regions.

Despite some ambiguity regarding the common ancestral areas of Oryzeae and Ehrharteae, our estimates on divergence times suggest that the two tribes diverged from each other at the end of Eocene, possibly through vicariance either between Asia and Australia or between Asia and Africa. Moreover, the DIVA analysis suggests that Asia is an unambiguous ancestral distribution area for Oryzeae and this tribe started to diversify at the end of the Oligocene (~24 MYA). Because Oryzeae arose well after the breakup of Gondwanaland, the widespread geographic distribution of Oryzeae lineages seems unlikely to be explained by continental drift following the fragmentation of Gondwana. Therefore, multiple long-distance dispersals would be invoked to achieve the establishment of the biogeographic pattern of Oryzeae, involving Asia, Australia, Africa and America. This implicates oceanic dispersal as a major factor in divergence patterns of the major lineages in subtribe Oryzinae. As shown in Fig. 3, ancestral area reconstruction provided a unique solution under two-area constraints for this clade. The genus *Oryza* separated from the clade comprising genera *Leersia*, *Maltebrunia* and *Prosphytochloa* in the early Miocene (~17 MYA). Since the ancestor of Oryzeae occurred in Asia, an initial dispersal from Asia to Africa is required for the distribution area expansion and the subsequent vicariance or dispersal within *Oryza* and the lineage leading to *Leersia–Maltebrunia–Prosphytochloa*. It is obvious that the common ancestor of *Leersia*, *Maltebrunia* and *Prosphytochloa* originated in Africa and differentiated in the late Miocene (~11 MYA) into three genera found exclusively in Africa (Fig. 2).

The genus *Oryza* originated in Asia and started to diversify earlier than its African counterpart in the middle Miocene (~14–15 MYA). Our area reconstruction showed that *Oryza* diversification was likely accomplished by both vicariance and dispersals between the continents and suggested multiple dispersal events at least three times from Asia to Australia, four times from Asia to Africa and once from Asia to America (Fig. 3). Such a timescale and biogeographic pattern for *Oryza* corresponds to the Miocene collision of Laurasia with the Australia/New Guinea shard of Gondwanaland, an event that allowed considerable land migration from Asia to Australasia (Second, 1985a). Asian origin and dispersal to other regions by *Oryza* have been proposed in previous studies (Second, 1985a; Wang et al., 1992; Vaughan et al., 2005). Based on time estimates using isozyme data and the observation that *Oryza* species are adapted to forest environments primarily found in Asia, Second (1985a) assumed an Eurasiatic distribution of *Oryza* during the Tertiary period and suggested multiple recent migrations from Asia to Africa through a relatively arid environment, and a very recent introduction to America. Using RFLP markers, Wang et al. (1992) also detected multiple introduction events of *Oryza* from Asia to Africa and from Asia to Australia. These speculations are in agreement with the observation that two out of four species complexes in *Oryza* (the *O. ridleyi* and *O. granulata* complexes) that were considered close to 'primitive' or 'ancestral' *Oryza* are distributed across Southeast Asia and New Guinea (Vaughan et al., 2005).

The subtribe Zizaniinae began to diverge in the early Miocene (~21 MYA). Although the inferred distribution areas of Zizaniinae have three outcomes (A/AB/AC), Asia is always included as one of its ancestral areas (Fig. 3). Our Oryzeae divergence estimates suggested that following the split of genus *Chikusichloa*, three major lineages diverged rapidly within a very short period of time. Such radiation might be associated with possible multiple long-distance dispersal events between Asia and America through the Bering land bridge (Wen, 1999; Donoghue et al., 2001) and between Asian and Australian through the coasts of the Indian Ocean after the collision of Asia and Australia plates in the Miocene (Raven and Axelrod, 1972). The trans-Pacific dispersal may also be an alternative way for migration of Zizaniinae between Australia and South America, as numerous studies have demonstrated the exchange of plant species between the two continents in different plant groups (e.g., Renner et al., 2000; Winkworth et al., 2002; Sanmartin and Ronquist, 2004; Sanmartin et al., 2007; Inda et al., 2008).

This investigation, along with previous studies, demonstrates clearly that the Gondwana hypothesis on the origins of either the tribe Oryzeae or the genus *Oryza* (Chang, 1976, 1985; Second, 1985b) should be rejected because the origin and divergence of the tribe Oryzeae are long after the breakup of Gondwanaland. As discussed above, multiple trans-continent dispersals have to be considered as the main causes to explain the present-day worldwide distribution of the Oryzeae. On the basis of current *Oryza* diversity and the distribution of grass relatives, Vaughan et al. (2005) speculated that the Austral-Asian zone was the most likely region where *Oryza* first evolved and spread to other tropical regions, and proposed a hypothesis that explained the distribution of *Oryza* species in relation to the movement of animals including birds and humans. They suggested that the distribution of *Oryza* species on landmasses separated by oceans might best be explained by their introduction by birds and humans. Zhang and Ge (2007) conducted a molecular population genetics study on a few of the *Oryza* species and suggested that a long-distance dispersal from West Africa to Sri Lanka was more likely to play a role in the disjunctive distribution of *Oryza eichingeri*. Recent decades have witnessed increasingly important roles of dispersal in historical biogeography studies (Zink et al., 2000; de Queiroz, 2005; Yuan et al., 2005; Ickert-Bond and Wen, 2006; Inda et al., 2008). Renner (2004) reviewed biogeographic studies on plant genera which had trans-Atlantic disjunction and suggested that long-distance dispersal by wind and sea current seemed more likely to bring about the observed tropical Atlantic disjunct distribution. The studies of Southern Hemisphere biogeography also supported the important role of dispersal in shaping the distribution pattern of many plant species (Winkworth et al., 2002; Sanmartin and Ronquist, 2004; Sanmartin et al., 2007). Migrations across the North Atlantic land bridge and Bering land bridge have been suggested as the main dispersal routes for the development of floristic disjunction of Asia and North America (Tiffney, 1985; Wen, 1999; Donoghue et al., 2001; Xiang and Soltis, 2001; Ickert-Bond and Wen, 2006). Along with these studies, the present investigation on the tribe Oryzeae further highlights the importance of long-distance dispersal in the origin and diversification of plant species.

Acknowledgments

We thank F.-M. Zhang, L.-B. Zhang, X.-W. Xu, H.-Z. Sun, L.-L. Zang, X.-M. Zheng and other members of the Ge laboratory (Institute of

Botany, CAS) for technical assistance and encouragements. We also thank S.-Q. Tang, Jun Wen, Bruce G. Baldwin, Curtis A. Jones, J.-Y. Li for help in collecting materials, and X.-G. Xiang, Z.-Y. Yang and Wei Wang for discussion about biogeography. We are grateful to Ming Feng and Ming Li for assistance in running Linux Cluster for data analysis. We acknowledge the International Rice Research Institute (Los Banos, Philippines) for providing seed and leaf samples. This work was supported by the National Basic Research Program of China (2007CB815704), the National Natural Science Foundation of China (30121003), and grants from the Chinese Academic of Sciences.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2009.08.007.

References

Bennetzen, J.L., 2009. Biology and the beasts: individual investigator-driven research in the megaprojects era. Trends Genet. 25, 57–59.
Bentham, G., 1881. Mr. G. Bentham on Gramineae. Biol. J. Linn. Soc. Lond. 19, 14–134.
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.
Bremer, K., 2002. Gondwanan evolution of the grass alliance of families (Poales). Evolution 56, 1374–1387.
Brown, R., 1810. Prodromus Flora Novae Hollandiae et insulae van Diemen. Richard Taylor, London.
Chang, T.T., 1976. Origin, evolution, cultivation, dissemination, and diversification of Asian and African rices. Euphytica 25, 425–441.
Chang, T.T., 1985. Crop history and genetic conservation: rice – a case-study. Iowa State J. Res. 59, 425–496.
Clayton, W.D., 1970. Flora of tropical East Africa. Grasses 1, 23–33.
Clayton, W.D., Renvoize, S.A., 1986. Genera Graminum. Kew Bull. Additional Series, vol. 8, pp. 1–389.
Davis, C.C., Bell, C.D., Mathews, S., Donoghue, M.J., 2002. Laurasian migration explains Gondwanan disjunctions: Evidence from Malpighiaceae. Proc. Natl. Acad. Sci. USA 99, 6833–6837.
de Queiroz, A., 2005. The resurrection of oceanic dispersal in historical biogeography. Trends Ecol. Evol. 20, 68–73.
de Winter, B., 1951. A morphological, anatomical and cytological study of *Potamophila prehensilis* (Nees) Benth. Bothalia 6, 117–137.
Donoghue, M.J., Bell, C.D., Li, J.H., 2001. Phylogenetic patterns in Northern Hemisphere plant geography. Int. J. Plant Sci. 162, S41–S52.
Doyle, J.J., 1992. Gene trees and species trees—molecular systematics as one-character taxonomy. Syst. Bot. 17, 144–163.
Duistermaat, H., 1987. A revision of *Oryza* (Gramineae) in Malesia and Australia. Blumea 32, 157–193.
Duvall, M.R., Peterson, P.M., Terrell, E.E., Christensen, A.H., 1993. Phylogeny of North American Oryzoid grasses as construed from maps of plastid DNA restriction sites. Am. J. Bot. 80, 83–88.
Farris, J.S., Kallersjo, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. Cladistics 10, 315–319.
Felsenstein, J., 2004. PHYLIP (phylogeny inference package). Version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.
Fishbein, M., Hibsch-Jetter, C., Soltis, D.E., Hufford, L., 2001. Phylogeny of saxifragales (angiosperms, eudicots): analysis of a rapid, ancient radiation. Syst. Biol. 50, 817–847.
Gaut, B.S., 2002. Evolutionary dynamics of grass genomes. New Phytol. 154, 15–28.
Ge, S., Li, A., Lu, B.R., Zhang, S.Z., Hong, D.Y., 2002. A phylogeny of the rice tribe Oryzeae (Poaceae) based on *matK* sequence data. Am. J. Bot. 89, 1967–1972.
Ge, S., Sang, T., Lu, B.R., Hong, D.Y., 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc. Natl. Acad. Sci. USA 96, 14400–14405.
Ge, S., Sang, T., Lu, B.R., Hong, D.Y., 2001. Rapid and reliable identification of rice genomes by RFLP analysis of PCR-amplified *Adh* genes. Genome 44, 1136–1142.
Goldman, N., Anderson, J.P., Rodrigo, A.G., 2000. Likelihood-based tests of topologies in phylogenetics. Syst. Biol. 49, 652–670.
GPWG, 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). Ann. MO Bot. Gard. 88, 373–457.
Guindon, S., Gascuel, O., 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.
Guo, Y.L., Ge, S., 2005. Molecular phylogeny of Oryzeae (Poaceae) based on DNA sequences from chloroplast, mitochondrial, and nuclear genomes. Am. J. Bot. 92, 1548–1558.
Hubbard, C.E., 1959. Gramineae. In: Hutchinson, J. (Ed.), Families of Flowering Plants. Clarendon Press, London.
Hubbard, C.E., 1967. Tabula 3647: *Snowdenia petitiana* (A. Rich.) C.E. Hubbard. Hooker's Icones plantarum, vol. 37, p. t3647.
Ickert-Bond, S.M., Wen, J., 2006. Phylogeny and biogeography of Altingiaceae: evidence from combined analysis of five non-coding chloroplast regions. Mol. Phylogenet. Evol. 39, 512–528.
Inda, L.A., Segarra-Moragues, J.G., Muller, J., Peterson, P.M., Catalan, P., 2008. Dated historical biogeography of the temperate Loliinae (Poaceae, Pooideae) grasses in the northern and southern hemispheres. Mol. Phylogenet. Evol. 46, 932–957.
Jian, S.G., Soltis, P.S., Gitzendanner, M.A., Moore, M.J., Li, R., Hendry, T.A., Qiu, Y.L., Dhingra, A., Bell, C.D., Soltis, D.E., 2008. Resolving an ancient, rapid radiation in Saxifragales. Syst. Biol. 57, 38–57.
Kellogg, E.A., 2001. Evolutionary history of the grasses. Plant Physiol. 125, 1198–1205.
Kellogg, E.A., 2009. The evolutionary history of Ehrhartoideae, Oryzeae, and *Oryza*. Rice 2, 1–14.
Kellogg, E.A., Buell, C.R., 2009. Splendor in the grasses. Plant Physiol. 149, 1–3.
Kellogg, E.A., Watson, L., 1993. Phylogenetic studies of a large data set. I. Bambusoideae, Andropogonodae, and Pooideae (Gramineae). Bot. Rev. 59, 273–343.
Kim, H., Hurwitz, B., Yu, Y., Collura, K., Gill, N., SanMiguel, P., Mullikin, J.C., Maher, C., Nelson, W., Wissotski, M., Braidotti, M., Kudrna, D., Goicoechea, J.L., Stein, L., Ware, D., Jackson, S.A., Soderlund, C., Wing, R.A., 2008. Construction, alignment and analysis of twelve framework physical maps that represent the ten genome types of the genus *Oryza*. Genome Biol. 9, R45.
Kunth, K.S., 1829. Révision des Graminées 1:6. Librairie-Gide, Paris.
Lu, B.-R., Ge, S., 2005. Taxonomic treatment of *Porteresia coarctata* (Poaceae: Oryzeae). Nord. J. Bot. 23, 555–558.
Martínez-y-Pérez, J.L., Mejía-Saulés, T., Sosa, V., 2006. Phylogenetic relationships of *Luziola* (Poaceae: Oryzeae) and related genera from aquatic habitats. Can. J. Bot. 84, 1753–1770.
Martínez-y-Pérez, J.L., Mejía-Saulés, T., Sosa, V., 2008. A taxonomic revision of *Luziola* (Poaceae: Oryzeae). Syst. Bot. 33, 702–718.
Morat, P., Deroin, T., Couderc, H., 1994. Presence en Nouvelll-Caledonie d'une espece endemique du genre *Oryza* L. (Graminae). Bull. Mus. Nat. Paris 4, ser. 16, Sec. B, Adansonia 1, 3–10.
Nayar, N.M., 1973. Origin and cytogenetics of rice. Adv. Genet. 17, 153–292.
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.
Prasad, V., Stromberg, C.A.E., Alimohammadian, H., Sahni, A., 2005. Dinosaur coprolites and the early evolution of grasses and grazers. Science 310, 1177–1180.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. Available from: <http://beast.bio.ed.ac.uk/Tracer/>.
Raven, P.H., Axelrod, D.I., 1972. Plate tectonics and Australasian Paleobiogeography. Science 176, 1379–1386.
Renner, S.S., 2004. Plant dispersal across the tropical Atlantic by wind and sea currents. Int. J. Plant Sci. 165, S23–S33.
Renner, S.S., Foreman, D.B., Murray, D., 2000. Timing transantarctic disjunctions in the Atherospermataceae (Laurales): evidence from coding and noncoding chloroplast sequences. Syst. Biol. 49, 579–591.
Rokas, A., Carroll, S.B., 2006. Bushes in the tree of life. PLoS Biol. 4, 1899–1904.
Ronquist, F., 1996. DIVA version 1.1. Computer program and manual. Available from: <http://www.ebc.uu.se/systzoo/research/diva/diva.html/>.
Ronquist, F., 1997. Dispersal–vicariance analysis: a new approach to the quantification of historical biogeography. Syst. Biol. 46, 195–203.
Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Sanderson, M.J., 2002. Estimating absolute rates of molecular evolution and divergence times: a penalized likelihood approach. Mol. Biol. Evol. 19, 101–109.
Sanderson, M.J., 2003. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics 19, 301–302.
Sanmartin, I., Ronquist, F., 2004. Southern Hemisphere biogeography inferred by event-based models: plant versus animal patterns. Syst. Biol. 53, 216–243.
Sanmartin, I., Wanntorp, L., Winkworth, R.C., 2007. West Wind Drift revisited: testing for directional dispersal in the Southern Hemisphere using event-based tree fitting. J. Biogeogr. 34, 398–416.
Schweickerdt, H.G., 1961. Graminearum species nova et genus novum. Zur Kenntnis einiger südafrikanischer gramineen. Züchter 31, 193–195.
Second, G., 1985a. Evolutionary relationships in the sativa group of *Oryza* based on isozyme data. Genet. Sel. Evol. 17, 89–114.
Second, G., 1985b. Geographic origins, genetic diversity and the molecular clock hypothesis in the Oryzeae. In: Jacquard, P., Heim, G., Antonovics, J. (Eds.), Genetic Differentiation and Dispersial in Plants. Springer-Verlag, Berlin, Germany.
Shimodaira, H., 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51, 492–508.
Shimodaira, H., Hasegawa, M., 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics 17, 1246–1247.
Swofford, D.L., 2002. PAUP*: Phylogenetic analysis using parsimony (* and other methods). Sinauer, Sunderland, MA.
Tateoka, T., 1965. *Porteresia*, a new genus of Graminae. Bull. Natl. Sci. Mus. Tokyo 8, 405–406.

Terrell, E.E., Peterson, P.M., Wergin, W.P., 2001. Epidermal features and spikelet micromorphology in *Oryza* and related genera (Poaceae: Oryzeae). Smithsonian Contr. Bot. 91, 1–50.

Terrell, E.E., Robinson, H., 1974. Luziolinae, a new subtribe of oryzoid grass. Bull. Torrey Bot. Club 101, 235–245.

Thomasson, J.R., 1980. Archaeolersia nebraskensis gen. et sp. nov. (Gramineae–Oryzeae), a new fossil grass from the late Tertiary of Nebraska. Am. J. Bot. 67, 876–882.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.

Thorne, J.L., Kishino, H., 2002. Divergence time and evolutionary rate estimation with multilocus data. Syst. Biol. 51, 689–702.

Thorne, J.L., Kishino, H., Painter, I.S., 1998. Estimating the rate of evolution of the rate of molecular evolution. Mol. Biol. Evol. 15, 1647–1657.

Tiffney, B.H., 1985. The Eocene North-Atlantic Land-Bridge: its importance in tertiary and modern phytogeography of the Northern Hemisphere. J. Arnold Arbor. 66, 243–273.

Vaughan, D.A., 1994. The wild relatives of rice: a genetic resources handbook. International Rice Research Institute, Manila, Philippines.

Vaughan, D.A., Kadowaki, K., Kaga, A., Tomooka, N., 2005. On the phylogeny and biogeography of the genus *Oryza*. Breed. Sci. 55, 113–122.

Vaughan, D.A., Morishima, H., 2003. Biosystematics of the genus *Oryza*. In: Smith, C.W. (Ed.), Rice: Origin, History, Technology, and Production. John Wiley & Sons, Inc., pp. 27–65.

Verboom, G.A., Linder, H.P., Stock, W.D., 2003. Phylogenetics of the grass genus Ehrharta: evidence for radiation in the summer-arid zone of the South African Cape. Evolution 57, 1008–1021.

Vicentini, A., Barber, J.C., Aliscioni, S.S., Giussani, L.M., Kellogg, E.A., 2008. The age of the grasses and clusters of origins of C_4 photosynthesis. Global Change Biol. 14, 2963–2977.

Walsh, H.E., Kidd, M.G., Moum, T., Friesen, V.L., 1999. Polytomies and the power of phylogenetic inference. Evolution 53, 932–937.

Wang, Z.Y., Second, G., Tanksley, S.D., 1992. Polymorphism and phylogenetic-relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. Theor. Appl. Genet. 83, 565–581.

Watson, L., Dallwitz, M.J., 1999. The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. Version: 25th November 2008. Available from: http://delta-intkey.com/grass/.

Weatherwax, P., 1929. The morphology of the spikelets of six genera of Oryzeae. Am. J. Bot. 16, 547–555.

Wen, J., 1999. Evolution of eastern Asian and eastern North American disjunct distributions in flowering plants. Annu. Rev. Ecol. Syst. 30, 421–455.

Whitfield, J.B., Lockhart, P.J., 2007. Deciphering ancient rapid radiations. Trends Ecol. Evol. 22, 258–265.

Winkworth, R.C., Wagstaff, S.J., Glenny, D., Lockhart, P.J., 2002. Plant dispersal N.E.W.S from New Zealand. Trends Ecol. Evol. 17, 514–520.

Wortley, A.H., Rudall, P.J., Harris, D.J., Scotland, R.W., 2005. How much data are needed to resolve a difficult phylogeny? Case study in Lamiales. Syst. Biol. 54, 697–709.

Xiang, Q.Y.J., Soltis, D.E., 2001. Dispersal–vicariance analyses of intercontinental disjuncts: historical biogeographical implications for angiosperms in the Northern Hemisphere. Int. J. Plant Sci. 162, S29–S39.

Yang, Z.H., 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24, 1586–1591.

Yuan, Y.M., Wohlhauser, S., Moller, M., Klackenberg, J., Callmander, M.W., Kupfer, P., 2005. Phylogeny and biogeography of Exacum (Gentianaceae): a disjunctive distribution in the Indian Ocean Basin resulted from long distance dispersal and extensive radiation. Syst. Biol. 54, 21–34.

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

Zhang, S.-H., Second, G., 1989. Phylogenetic analysis of the tribe Oryzeae: total chloroplast DNA restriction fragment analysis (A preliminary report). Rice Genet. Newsl. 6, 76–80.

Zhou, S.L., Renner, S.S., Wen, J., 2006. Molecular phylogeny and intra- and intercontinental biogeography of Calycanthaceae. Mol. Phylogenet. Evol. 39, 1–15.

Zhu, Q.-H., Ge, S., 2005. Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. New Phytol. 167, 249–265.

Zink, R.M., Blackwell-Rago, R.C., Ronquist, F., 2000. The shifting roles of dispersal and vicariance in biogeography. P. Roy. Soc. Lond. B Biol. 267, 497–503.

Zou, X.H., Zhang, F.M., Zhang, J.G., Zang, L.L., Tang, L., Wang, J., Sang, T., Ge, S., 2008. Analysis of 142 genes resolves the rapid diversification of the rice genus. Genome Biol. 9, R49.