Molecular Phylogenetics and Evolution 55 (2010) 136–142

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

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

Multiple patterns of rDNA evolution following polyploidy in *Oryza*

Ying Bao [a,b,*], Jonathan F. Wendel [c], Song Ge [b]

[a] College of Life Science, Qufu Normal University, Qufu, Shandong 273165, PR China
[b] State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, PR China
[c] Department of Ecology, Evolution, & Organismal Biology, Iowa State University, Ames, IA 50011, United States

ARTICLE INFO

Article history:
Received 26 June 2009
Revised 6 October 2009
Accepted 18 October 2009
Available online 24 October 2009

Keywords:
Oryza
Polyploidy
Concerted evolution
Ribosomal DNA
Internal transcribed spacer sequences

ABSTRACT

Ribosomal ITS sequences are commonly used for phylogenetic reconstruction because they are included in rDNA repeats, and hence are ubiquitous and present in high copy number. Ribosomal rDNA repeats often undergo rapid concerted evolution within and between arrays. Interspecific hybridization merges divergent repeat types in a single nucleus, setting in motion evolutionary processes leading to coexistence, maintenance of paralogs, origin of novel sequence variants, loss of arrays, or inter-array sequence homogenization via concerted evolution. Here we examined ITS polymorphism within and among six *Oryza* tetraploids of varying genomic composition to infer the extent and direction of concerted evolution following allopolyploid speciation. We demonstrate that different polyploids have experienced varying fates, including maintenance or homogenization of divergent arrays, even among allopolyploids having the same genomic origins but in different geographic locations. Bidirectional concerted evolution, in which arrays become homogenized to alternative progenitor diploid types in different allopolyploid derivatives, is evident among species in one clade. Our results exemplify the panoply of outcomes for ribosomal DNA evolution following allopolyploid speciation.

1. Introduction

Ribosomal ITS sequences are commonly used for phylogenetic reconstruction because they are included in rDNA repeats, and hence are ubiquitous and present in high copy number. Ribosomal rDNA repeats often undergo rapid concerted evolution within and between arrays (Baldwin et al., 1995; Álvarez and Wendel, 2003; Nieto Feliner and Rosselló, 2007). The result of concerted evolution is that individual copies are identical or nearly so, such that repeats within a lineage appear to evolve more or less in unison because inter-repeat sequence variation is reduced to a negligible level due to sequence homogenization. Nevertheless, 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; O'Kane et al., 1996; 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 Rosselló, 2007; Zhang and Ge, 2007; Kim et al., 2008; Göker and Grimm, 2008; Grimm and Denk, 2008; Pilotti et al., 2009). The presence of multiple paralogs, recombinant mosaic sequences or a mixture of both within a genome are phenomena of ITS evolution in plants that need to be considered when conducting phylogenetic analysis (reviewed in Álvarez and Wendel, 2003). This complexity may be increased following allopolyploidization, which merges divergent repeat types in a single nucleus, thereby giving rise to at least an evolutionarily ephemeral coexistence of divergent ITS repeats (Volkov et al., 2007). A number of studies demonstrated the diversity of outcomes of this process, including maintenance of biparental rDNA repeats (e.g., Soltis and Soltis, 1991; O'Kane et al., 1996; Popp and Oxelman, 2001; Zhang et al., 2002), to loss of one parental copy (e.g., Wendel et al., 1995; Volkov et al., 1999; Kotseruba et al., 2003; Kovarik et al., 2005, 2008; Matyasek et al., 2007; Kim et al., 2008), to the origin of chimeric repeats (e.g., Volkov et al., 1999; Nieto Feliner and Rosselló, 2007).

To better understand rDNA repeat interactions following genomic mergers it is important to study the process in different natural polyploid systems. The genus *Oryza* is particularly appropriate in this regards, as approximately half of the extant *Oryza* species are considered to be allopolyploids derived from interspecific hybridization (Gopalakrishnan et al., 1965; Vaughan, 1989,1994; Ge et al., 1999; Lu and Ge, 2005). Based on interspecific crossing, subsequent cytogenetic analysis (review in Nayar, 1973), total genomic DNA hybridization (Aggarwal et al., 1997), and comparisons of homoeologous DNA sequences (Ge et al., 1999), four different genomic constitutions have been recognized among allopolyploid species (i.e., BBCC, CCDD, HHJJ and HHKK). Species with the genomic compositions of BBCC and CCDD have been especially well-studied phylogenetically. Cladistic analyses of multiple gene sequences or microsatellite markers have clarified phylogenetic

* Corresponding author. Address: College of Life Science, Qufu Normal University, 57 Jingxuan West Road, Qufu, Shandong 273165, PR China. Fax: +86 537 4456999.
E-mail addresses: baoyingus@126.com, baoyingus@gmail.com (Y. Bao).

1055-7903/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2009.10.023

relationship among these *Oryza* polyploids and their diploids donors (Ge et al., 1999; Bao and Ge, 2004; Bao et al., 2006). These data show that BBCC species had multiple origins, with B genome maternal parents for Asian species and C genome for African species. All CCDD species derive from a maternal CC genome parent and a paternal DD genome parent. Although no diploid, DD genome parent has been discovered to date, sequence analysis has shown that EE genome species are closely related to the DD genome progenitor that gave rise to the CCDD species (Ge et al., 1999; Bao and Ge, 2004). These clear relationships facilitate an analysis of the pattern and direction of ITS homogenization following allopolyploid formation.

Based on rDNA-FISH analysis, Chung et al. (2008) recently localized rDNAs on chromosomes of *Oryza* species. They demonstrated that allopolyploid species had four to eight rDNA loci in their haploid genomes, and that these rDNA sites were not always additive of their parental diploid rDNA array numbers.

In the present study, using cloning and sequencing of multiple repeats per individual, we examined ITS polymorphism within and among six *Oryza* allopolyploids having BBCC or CCDD genomes, assessing nucleotide diversity among homologous ITS repeats and their origins. Our objective was to infer the evolutionary outcome of reuniting two divergent ITS repeat types in these allopolyploids, thereby providing some insight into the consequences of allopolyploidization on ITS evolution in *Oryza*.

2. Materials and methods

2.1. Plant materials

A total of 25 *Oryza* accessions were used in this study, representing six allopolyploids and five diploids (Table 1). The allopolyploids included three species with BBCC genomes and three with CCDD genomes. For the diploids, we selected one species each with BB and EE genomes, and three with CC genomes. Also studied were individual representatives of a diploid with a GG genome. The latter taxon was selected as the phylogenetic outgroup, based on earlier evidence that the species with GG genomes are sister to the remainder of the genus (Wang et al., 1992; Ge et al., 1999; Zou et al., 2008). In most cases, two accessions were selected per taxon; exceptions were *O. malampuzhaensis* (three accessions) and *O. latifolia* (four accessions).

Seeds were kindly provided by the Internal Rice Research Institute (IRRI, Manila, Philippines), except that the outgroup accession (*O. granulata*) was collected from China. Seeds from each accession were germinated following methods described previously by Bao and Ge (2004) and kept in the greenhouse for DNA-extraction. Further identities and genomic identification of these accessions are described in Bao et al. (2005).

2.2. DNA-extraction, PCR, cloning, sequencing and ITS copy searching

Total DNA was extracted from fresh leaves, following extraction methods described by Bao and Ge (2004). PCR amplification was performed in 25 μl volumes of 10 mM Tris buffer (pH 8.3) containing 2.5 mM $Mgcl_2$, 0.2 mM of each dNTP, 5 pM of each primer, 100 ng of template DNA and 0.75 U of Taq polymerase (Takaya). In addition, 8% dimethylsulfoxide (DMSO) was added to the reaction mixture to facilitate denaturation during PCR (Baldwin et al., 1995). Two primers were used to amplify the ITS region, i.e., ITS5 (5′AGAAGTCGTAACAAGGTTTCCGTA3′) near the 3′ end of the 18S rRNA gene, and ITS4 (5′TCCTCCGCTTATTGATATGC3′) near the 5′ end of the 26S rRNA gene. The procedure for 35 cycles of amplification was: 1 min denaturation at 94 °C, 1 min annealing at 52 °C,

Table 1
Materials used in the study, and nucleotide diversity of ITS sequences.

Taxa	Acc. no.	Genome	Origin	Seq. no.	ITS type	GenBank no.	π	*Hap.*
O. officinalis	105085 106159	CC	Philippines Papua New Guinea	6[b] 1[a]	C	AF479067, EU574633–EU574637 AF479063	0.0000 0.0005*	H_1 H_2
O. eichingeri	101422 105159	CC	Uganda	2[a]	C	AF479069, AY181995	0.0000	H_3
O. rhizomatis	103410 105448	CC	Sri Lanka	2[a]	C	AF479065, AF479066	0.0000	H_4
O. punctata	104071 105607	BB	Chad	10[b] 1[a]	B	AF479070, AF479071 EU574638–EU574646	0.0023 0.0022*	H_7–H_10
O. punctata	104059 105158	BBCC	Nigeria Kenya	2[b] 2[b]	B C	AF479074, AF479075 AF479077, AF479078	0.0017 0.0034	H_11, H_12 H_5, H_6
O. minuta	104674 101082	BBCC	Philippines	31[b]	B	AY188606–AY188636	0.0082	H_13– H_31
O. malampuzhaensis	105223 105328 80764	BBCC	India	27[b]	B	AY188637–AY188663	0.0035	H_32– H_38
O. alta	100161 105143	CCDD	Brazil Guyana	15[b]	C	AF513601–AF513606 AF514909–AF514917	0.0058	H_39– H_46
O. grandiglumis	105664 105669	CCDD	Brazil	8[b]	C	AF520780, AF520781 AY151363–AY151368	0.0061	H_46– H_48
O. latifolia	100914 100167 105145 105141	CCDD	Mexico Costa Rica Colombia Costa Rica	45[b]	D	AF513577–AF513600 AF513553–AF513564 AY126015 AY151355–AY151362	0.0044	H_49– H_63
O. australiensis	105263 101144	EE	Australia	2[a]	E	AF520777, AF520778	0.0050	
O. granulata	C0024	GG	Hainan, China	1[a]	G	AF520779	–	

π refers to nucleotide diversity, according to Nei and Li (1979). *Hap.* refers to different haplotypes (cf. Fig. 1).
[a] Refers to the sequences obtained by PCR sequencing.
[b] Denotes sequences obtained following cloning.
* Refers to the values obtained based on all sequences from species of *O. officinalis* or *O. punctata*.

and 1.5 min extension at 72 °C. PCR was performed in a PTC-200 (Perkin Elmer) thermocycler. PCR products were electrophoresed in 1.5% agarose gels, and DNA fragments corresponding to the expected size were cut from the gel and purified using a DNA Purification Kit (Pharmacia) following the manufacturer's instructions. Purified products were sequenced directly or inserted into a pGEM-T-easy vector (Promega) for cloning.

In general, PCR-purified products of diploids were sequenced directly. In order to evaluate possible polymorphism within diploids, PCR products from two accessions, representing the BB and CC genomes, were cloned and sequences were obtained from individual clones. This latter procedure was followed for all allopolyploids. Sequencing was done on an ABI377 automated DNA sequencer with a Dye Terminator Cycle Sequencing Reaction Kit (PE Applied Biosystems). Both strands of the samples were sequenced. To avoid the influence of higher GC content during ITS sequencing, 10% DMSO was added to the sequencing reactions (Baldwin et al., 1995).

For purposes of identifying the genomic origin of individual clones in the polyploid genomes, sequences of the diploids with BB and CC were aligned and analyzed for the presence of diagnostic restriction sites. This process revealed that the restriction enzyme SamI could distinguish ITS clones originating from the B genome from those originating from the C genome.

To evaluate whether ITS sequences in BBCC-genome polyploids have both B and C genomic types, 40 clones of each accession were digested with the restriction enzyme SamI and the digestion products were run in agarose gels. If the BBCC genome species displayed two kinds of profiles, two clones with different profiles from each accession were sequenced. In cases where the restriction enzyme could not detect different genomic types among clones, 27–31 randomly selected clones were sequenced. For CCDD genome species, because there is no DD diploid in nature, it was not possible to design a diagnostic restriction site-based screening procedure, so we simply sequenced random clones (range of $n = 8–45$).

2.3. Data analysis

Sequences were aligned using ClustalW (Thompson et al., 1994) and refined manually. To evaluate the consequences of homogenization, we estimated ITS polymorphism levels with nucleotide diversity (π) (Nei and Li, 1979) and haplotype numbers (H), using DnaSP v4.0 (Rozas and Rozas, 1999). In order to reduce false base substitutions resulting from PCR polymerase mismatch, polymorphism observed in only one clone was removed from the analysis. To detect the lineage relationships between ITS clones in polyploids and those from their parental diploids, we conducted clone genealogy by coalescent simulations using the Median-Joining model as implemented in the Network v4.0 software (Bandelt et al., 1999).

To further estimate phylogenetic "preference" for ITS repeats in the allopolyploids, one sequence of each accession representing different ITS copies from all polyploids and diploids were used for an additional analysis. A total of 25 sequences were selected randomly from each monophyletic clade of the clone genealogy, and 2 EE genome diploid sequences and 1 GG genome outgroup sequence were also included. Maximum parsimony (MP), maximum likelihood (ML), and Bayesian inference (BI) methods were all performed for phylogenetic construction. MP and ML trees were implemented with PAUP version 4.0 (Swofford, 2002), with the branch-and-bound algorithm used for tree searching. Gaps were treated as missing values and these sites were excluded from the data matrix. Topological robustness was assessed by bootstrap analysis with 1000 replicates (Felsenstein, 1985) for MP analysis and with 500 replicates for ML analysis. ML and BI were both calculated using the best nucleotide substitution model determined by software Modeltest version 3.7 (Posada and Crandall, 1998). BI was conducted using MRBayes 3.12 (Huelsenbeck and Ronquist, 2001) with MCMC estimation of posterior probability distributions. Four chains of the MCMC were run each for 1,000,000 generations and were sampled. For all analyses, the first 300 samples from each run were discarded as burn-in to ensure that the chains reached stationarity. Phylogenetic inferences were based on the trees sampled after 30,000 generations.

3. Results

3.1. Polyploidy and ITS coexistence following genomic merger

A primary objective of the present study was to assess the fate of divergent ITS sequences following their merger into a single nucleus at the time of polyploidization. Notably, among the six polyploids in the study, *O. punctata* was the only species in which two types of ITS sequences were detected representing the contributions of the two progenitor genomes. In all others, only a single basic sequence type was detected.

3.2. Polymorphism of ITS sequences

Estimates of nucleotide diversity (π) and haplotype numbers (H) are shown for each ITS genomic type in Table 1. Low levels of polymorphism, with π in the range of 0.0000 (in *O. officinalis*) to 0.0023 (in *O. punctata*) were found in diploid clones within species. Although high haplotype diversity was found within the BB diploid species *O. punctata*, (4 haplotypes in only 10 sequences), the interclonal differences were minor, involving only four point mutations. These point mutations, however, would not have been detected had we instead directly sequenced the PCR pool.

A variable level of polymorphism was observed among the polyploid species. The most polymorphic ITS sequences were found in *O. minuta* ($\pi = 0.0082$, $H = 19$), while the ITS fragments of each genomic type in *O. punctata* were the least polymorphic (B genome, $\pi = 0.0017$; C genome, $\pi = 0.0034$). Not surprisingly, ITS polymorphism levels were generally related to the number of sequenced clones for both diploid and polyploid species.

3.3. Phylogenetic analysis

A total of 152 ITS clones were condensed into the 63 haplotypes observed in this study (Table 1). A network of these haplotypes is shown in Fig. 1. Haplotypes from same species were generally grouped together. The most significant feature of the topology was the existence of three clades ("B", "C", and "D" of Fig. 1), with B separated from C and D by 12 and 18 mutations, respectively. Clade C includes haplotypes from the species with the C genome, i.e., the three CC-genome diploids *O. officinalis*, *O. rhizomatis* and *O. eichingeri*, African BBCC-genome polyploids (*O. punctata*), and American CCDD-genome polyploids (*O. alta* and *O. grandiglumis*). Clade B contained all species with the B genome, i.e., African BB diploids and BBCC polyploids (*O. punctata*) and Asian BBCC species (*O. malampuzhaensis* and *O. minuta*). The third clade (D) included sequences only from the remaining CCDD species (*O. latifolia*). Because the ITS haplotypes in this clade did not group with the C-genome clade, these haplotypes might be representative of those contributed by the "missing genome", i.e. the D genome.

Phylogenetic analysis based on 28 representative clones reinforces the results based on the haplotype network. The strict consensus of the eight most-parsimonious trees (length = 160, CI = 0.744, RI = 0.869) is shown in Fig. 2. As with the haplotype network, three monophyletic clades were also recovered from parsimony analysis. The most notable feature of the MP tree was that

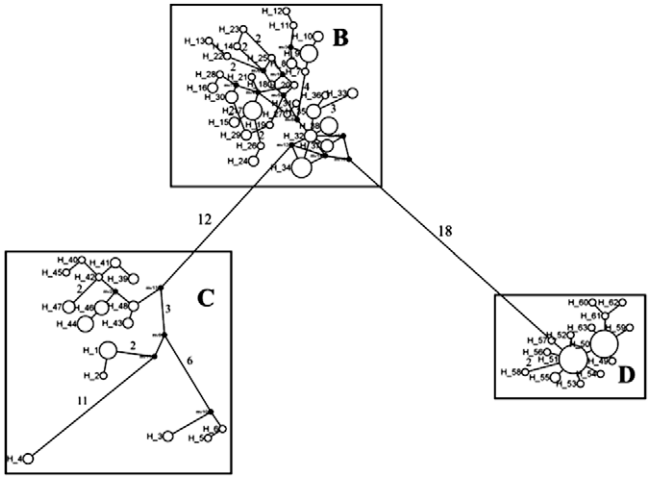

Fig. 1. Haplotype network based on the 63 haplotypes detected among the 152 ITS sequences. Open circles represent ITS haplotypes and the relative sizes of the circles represent the sample size of each haplotype in the study. Filled circles represent the median vectors. When more than one nucleotide difference exists between linked haplotypes, this is indicated by the numbers next to the lines. Boxes include corresponding genome types. Capital letters B, C and D denote specific *Orzya* genomes.

divergent ITS repeat types from polyploidization grouped with those from both putative parental species for one African BBCC species, whereas for other polyploids only one ITS repeat type was found, and as expected, in these cases the single ITS type grouped with sequences from one presumptive parental species (Asian BBCC or American CCDD species). Three CCDD species grouped with either the maternal (*O. alta* and *O. grandiglumis*) or paternal (*O. latifolia*) species. Similar topologies were also found using ML, with only a few differences in statistical support for some clades (Fig. 2). Bayesian inference based on the best evolutionary model of HKY+G (Hasegawa et al., 1985) with a transition–transversion ration of 1.921, and the sharp parameter of 0.014 generated only minor differences in topology from the former trees. The Bayesian inference weakly supported a clade including the Asian, CC diploids, especially *O. officianalis*, with two CCDD species (*O. alta* and *O. grandiglumis*), with 0.52 and 0.58 posterior probability (PP), and weakly supported the grouping C genome (CC diploids and African BBCC, two American CCDD polyploids) and B genome (BB diploids and three BBCC polyploids) clades with 0.56 PP (Fig. 2).

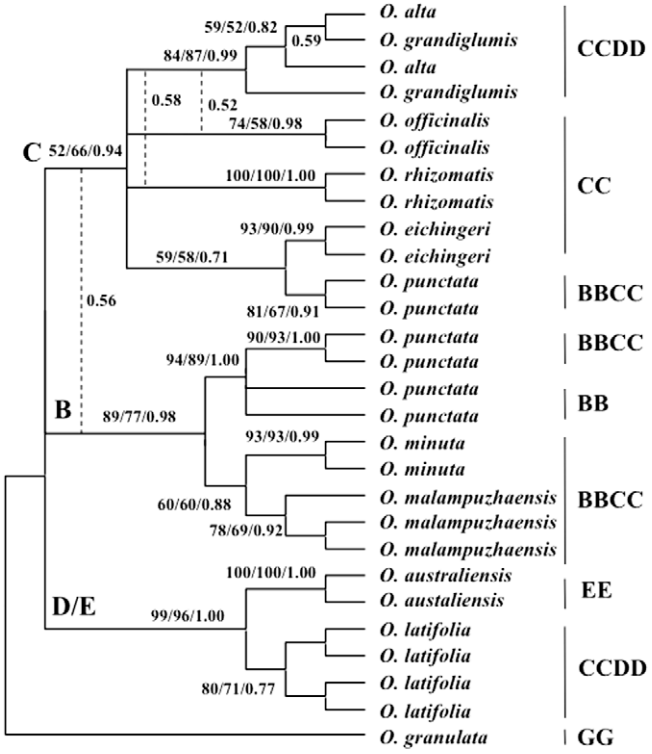

Fig. 2. Strict consensus tree of the eight most-parsimonious trees (160 steps, CI = 0.744, RI = 0.869) based on 28 representatives of ITS sequences. The topologies obtained by Bayesian inference ($-\ln L = 1791.07$) were the same except for the nodes indicated in the figures. The numbers near branches are bootstrap percentages of MP and ML trees, and are followed by Bayesian posterior probabilities. Dashed lines indicated the nodes supported by Bayesian inference. Capital letters above major clades and to the right of species names indicate corresponding genome types.

3.4. Copy number validation

As pointed out by Baldwin et al. (1995), PCR "selection" and "drift" can potentially lead to a pool of PCR products that includes a biased representation of the paralogs present in a genome. To mitigate against this possibility, we designed genome-specific primers based on different ITS types. For the BBCC genome species, the primer ITS4CF (5′TCGCGGTCCGAGCCGGTC3′) was designed that amplified the C type of ITS sequences in combination with ITS5. Although the fragments particular for the C genome were obtained from all CC-genome diploids and polyploids of *O. punctata*, they were not found in *O. malampuzhaensis* and *O. minuta* (Fig. 3).

For the CD genome species, we designed and used two genome-specific primers, ITS1CF (5′ACGGCGTCAAGGAACACATCGAC3′), specific to the C type of ITS sequence, and ITS1DF (5′CACCACGAGGTGGTTAGTAATTT3′) specific to the D type of ITS sequence. With the first primer, in combination with ITS4, we obtained PCR product from all CC-genome diploids and two CCDD species, i.e. *O. alta* and *O. grandiglumis*, but failed to obtain the corresponding product from *O. latifolia*. With the second primer and ITS4 combination, specific to the D genome, we obtained PCR product from *O. latifolia*, but failed to obtain the corresponding product from *O. alta* and *O. grandiglumis* (Fig. 4). To test whether the relative concentration of total DNA of C and D genomes would affect the PCR amplification of either type of ITS in the CCDD species, we mixed total DNA of *O. officinalis* (CC) and *O. latifolia* (CCDD) or in reverse with the following ratios: 1:1, 1:5 and 1:10 as the template for amplification. Combined with the two genome-specific primers, we evaluated the ability of PCR to amplify both ITS types depending on their relative ratios. The result was shown in Fig. 4, indicating PCR would work successfully for both types of ITS copies despite of very low proportion of one of genome DNA. These experiments confirmed that there was only one ITS copy in Asian polyploids possessing BBCC genome and American polyploids possessing CCDD genome.

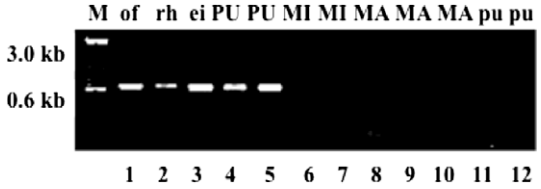

Fig. 3. PCR amplification products with primers ITS1 and ITS4CF. The latter primer is specific to C-genome repeat types in BBCC allopolyploids. M indicates a DNA size marker. Letters above the figure are species abbreviation, with capital letters denoting polyploid species. Notice the absence of amplification in *O. malampuzhaensis* and *O. minuta*, for which only B-type ITS sequences were recovered (cf. Figs. 1 and 2). Primers are explained in the text.

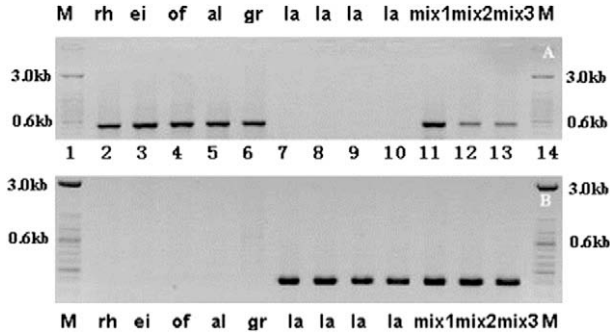

Fig. 4. PCR amplification using primers ITSCF/ITSDF and ITS4. (A) The profiles of PCR amplification products with primers ITSCF and ITS4. (B) The profiles of PCR amplification products using primers ITSDF and ITS4. M indicates a DNA size marker, while letters above the figure are species designations (cf. Table 1). Mixes 1–3 indicate products resulting from amplification following template mixing using the CC diploid *O. officinalis* and the CCDD polyploid *O. latifolia* in the proportions 1:1, 1:5 and 1:10, respectively. Primers are explained in the text.

4. Discussion

ITS sequences have been the most popular source of DNA data for plant phylogenetics for many years (Baldwin et al., 1995; Grebenstein et al., 1998; Jobst et al., 1998; Soltis and Soltis, 1998; Hershkovitz et al., 1999; Baum et al., 2001; Hörandl et al., 2005; Nieto Feliner and RosselLó, 2007). One important reason for their widespread exploitation is that ITS sequences are highly reiterated as components of rDNA repeats and hence often are subject to rapid homogenization through concerted evolution, which promotes intragenomic uniformity of repeats even between loci on non-homologous chromosomes (Baldwin et al., 1995). As a result, intra-individual polymorphism has generally been considered to be the exception rather than the rule for nrDNA (Hugall et al., 1999; Wissemann, 1999, 2003; Mayol and Rosselló, 2001; Comes and Abbott, 2001; Denk et al., 2002; Yonemori et al., 2002; Manen, 2004; Siripun and Schilling, 2006). Nevertheless, the process of ITS evolution is complicated, especially in allopolyploid species (Álvarez and Wendel, 2003; Denk et al., 2005; Volkov et al., 2007), which merge different ITS repeats contributed by parental donors into a single genome (e.g., Baldwin et al., 1995; Wendel et al., 1995; O'Kane et al., 1996; Buckler-IV et al., 1997; Mayol and Rosselló, 2001; Kovarik et al., 2005, 2008; Matyasek et al., 2007).

In principle, there are three evolutionary outcomes of this merger. First, two divergent rDNA copies may be retained and evolve independently without interaction, because concerted evolution fails to act across repeat units contributed by different parental species. Examples of this include *Arabidopsis* (O'Kane et al., 1996), *Silene* (Popp and Oxelman, 2001), and *Triticum* (Zhang et al., 2002). A second possibility is that one rDNA type would be lost either through loss of an entire duplicated array or via interlocus homogenization, the latter comprising directional concerted evolution. In allopolyploid cotton (Wendel et al., 1995), *Paeonia* (Sang et al., 1995), *Nicotiana* (Volkov et al., 1999), *Zingeria trichopoda* (Kotseruba et al., 2003), and *Persicaria* (Kim et al., 2008), for example, a single type of rDNA repeat from either the maternal or paternal parent is retained in different allopolyploids. The situation is more variable within the recently formed and model allopolyploids *Tragopogon mirus* and *T. miscellus*, where variable levels of repeat bias are observed among individuals and populations, with a trend toward elimination of repeats from one (*T. dubius*) of the two parental diploid genomes (Kovarik et al., 2005; Matyasek et al., 2007; Lim et al., 2008). A third evolutionary possibility is a mosaic of two different rDNA types, potentially homogenizing to a new rDNA type unlike that of either progenitor parent. In allopolyploid *Nicotiana* (Volkov et al., 2007), for example, 35S rDNA units have been replaced by novel variants. In the present study, the first two of these three classes of outcomes were observed in *Oryza* allopolyploids. It is difficult to entirely exclude the third class (fixation of a recombinant or mosaic repeat type) without extant diploids and a precise knowledge of the ancestral conditions, although inspection of our data yielded no evidence suggestive of this possibility.

Maintenance of both parental types of ITS repeats was observed in only one *Oryza* allopolyploid, the African, BBCC polyploid *O. punctata*. This was not the case, however, for the two Asian, BBCC polyploids, where directional concerted evolution has resulted in homogenization to only the B type of ITS repeat. Similarly, directional evolution appears to dominate ITS evolution within CCDD allopolyploids, where in each species only a single basic ITS type was recovered. Notably, this directional concerted evolution is bidirectional, with the American *O. alta* and *O. grandiglumis* exhibiting only a C-type ITS repeat, whereas the D-type repeat was observed in *O. latifolia*. Recently, a study based on FISH revealed that among the *Oryza* polyploids, only one allopolyploid species (*O. punctata*; BBCC genome) showed additivity of diploid parental rDNAs (Chung et al., 2008). These authors also found evidence of genome-diagnostic rDNA loci. One rDNA site at the end of the short arm of chromosome 4 appeared specific to species with the BB genome, as it was found in *O. punctata* (BB), *O. punctata* (BBCC) and *O. minuta* (BBCC), respectively. A second rDNA locus, in the proximal region of the short arm of chromosome 5, was specific to CC genome species, as it was observed in *O. officinalis* (CC), *O. punctata* (BBCC), and *O. grandiglumis* (CCDD). Interestingly, these putative CC genome repeat types were not detected in *O. minuta* (BBCC) and *O. latifolia* (CCDD), although both have CC genomes found. These data corroborate our results based on sequence analysis.

When considered together, a diversity of outcomes is revealed for ITS evolution following allopolyploidization in *Oryza*, which has several parallels in other angiosperm genera studied to date (Franzke and Mummenhoff, 1999; Kovarik et al., 2008). Perhaps the most relevant examples come from the analyses of *Nicotiana* allopolyploids (Kovarik et al., 2008), where, for example, the three natural allopolyploids *N. tabacum*, *N. rustica*, and *N. arentsii*, which vary in antiquity of origin, have experienced different degrees of parental gene replacement. The extent of rDNA sequence homogenization decreases in the order of *N. arentsii*, *N. tabacum*, and *N. rustica*. Kovarik et al. (2008) further suggested that expression patterns of ribosomal loci may be epigenetically established even in the initial hybrids, which then may influence subsequent evolutionary patterns of rDNA homogenization and retention; those that are epigenetically silenced were suggested to be less vulnerable to sequence homogenization but more subject to ultimate mutational obliteration. In this respect Kovarik et al. (2008) suggest that only about one million years would be required to for sequence deletion of epigenetically silenced arrays. Work in *Tragopogon* allopolyploids, many of which formed as recently as within the last century, illustrates that the temporal dependence of inter-array homogenization may operate on surprisingly brief timescales, in the initial generations following allopolyploid formation, where genomes experience a highly labile phase prior to stabilization (Kovarik et al., 2005; Matyasek et al., 2007; Lim et al., 2008). This is evidenced by highly variable levels of inter-array homogenization, loss of parental repeats, and a range of expression bias ranging from partial to complete nucleolar dominance. Similar examples of variable and rapid homogenization have been demonstrated in other systems, including in *Armeria*, where biased homogenization arises as early as the F_2 generation following interspecific hybridization (Fuertes Aguilar et al., 1999), and in *Cardamine*, where both biased homogenization and repeat maintenance was observed in newly derived allopolyploids (Franzke and Mummenhoff, 1999).

These interrelationships between age of formation, repeat loss vs. sequence homogenization, and epigenetic silencing of particular arrays represent promising directions for future studies of *Orzya* allopolyploids. Our previous work revealed that African and Asian polyploids with BBCC genomes originated at least three times (Bao et al., 2006). Given that these BBCC polyploids originated from different speciation events, and that the African polyploid exhibiting both B and C repeat types (*O. punctata*) might have originated more recently than their Asian counterparts that exhibited complete directional homogenization (the other BBCC allopolyploids), one possible explanation for incomplete homogenization in *O. punctata* is that there simply has not been sufficient time for interlocus homogenization to operate to completion.

One noteworthy result in the present study is the bi-directional interlocus concerted evolution pattern in polyploids having the genomic composition of CCDD. Bidirectional concerted evolution was first described in *Gossypium*, where Wendel et al. (1995) reported that among the five allopolyploid cottons with AADD genomes, four had ITS copies that had been homogenizing to an AA-genome repeat type, whereas one had ITS repeats that had concerted to a DD-genome repeat type. They suggested as possible mechanisms of interlocus homogenization mitotic or meiotic unequal crossing over between repeats located on different chromosomes. Based on rDNA-FISH, Chung et al. (2008) found that two rDNA loci on chromosome 5 and chromosome 10 inherited from the CC genome were lost in *O. latifolia* (CCDD) and the number and distribution of rDNA sites in *O. grandiglumis* were modified. They attributed this loss and modification to possible homoeologous pairing, unequal crossing over or amplification, or other cytogenetic anomalies. Alternatively, the possibility exists that functionally or selectively unequal repeated families could become united in common nucleolus as a consequence of polyploidization (or diploid hybridization) and thereby provide the opportunity for differential selection (Wendel et al., 1995). This has been noted more recently and discussed in light of the possible connections to expression dominance relationships, as proposed for *Tragopogon* (Matyasek et al., 2007) and *Nicotiana* (Kovarik et al., 2008) allopolyploids.

The present study adds to our growing understanding of the complexity of ribosomal DNA evolution in flowering plants, particularly when hybridization and/or allopolyploidization is involved. Here we have shown how a diverse array of ITS outcomes is possible even within a single genus, drawing attention both to this molecular evolutionary possibility as well as its implications for phylogenetics. Future work involving additional polyploids and the connections between gene expression and the fate of divergent ITS repeat types following genomic mergers will further illuminate these aspects of ITS evolution in plants.

Acknowledgments

The authors thank the International Rice Research Institute (Manila, Philippines) for providing plant materials for this study. This research was supported by the National Natural Science Foundation of China (30430030, 30770145); the Program for New Century Excellent Talents in University (NCET-06-0609); and the Doctoral Research Fund for Shandong Province (2007BS08020).

References

Aggarwal, R.K., Brar, D.S., Khush, G.S., 1997. Two new genomes in the *Oryza* complex identified on the basis of molecular divergence analysis using total genomic DNA hybridization. Mol. Gen. Genet. 254, 1–12.

Álvarez, I., 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., Hughes, C.E., 2003. Characterization of angiosperm nrDNA polymorphism, paralogy, and pseudogenes. Mol. Phylogenet. Evol. 29, 435–455.

Baldwin, B.G., Sanderson, M.J., Wojciechowski, M.F., Campbell, C.S., Donghue, M.J., 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. Ann. Mo. Bot. Gard. 82, 247–277.

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

Bao, Y., Ge, S., 2004. Origin and phylogeny of *Oryza* species with the CD genome based on multiple-gene sequence data. Plant Syst. Evol. 249, 55–66.

Bao, Y., Lu, B.-R., Ge, S., 2005. Identification of genomic constitutions of *Oryza* species with the B and C genomes by the PCR-RFLP method. Genet. Resour. Crop Evol. 52, 69–76.

Bao, Y., Zhou, H.-F., Hong, D.-Y., Ge, S., 2006. Genetic diversity and evolutionary relationships of *Oryza* species with the B- and C-genomes as revealed by SSR markers. J. Plant Biol. 49, 339–347.

Baum, B.R., Johnson, D.A., Bailey, L.G., 2001. Defining orthologous groups among multicopy genes prior to inferring phylogeny, with special emphasis on the Triticeae (Poaceae). Hereditas 135, 123–138.

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

Chung, M.C., Lee, Y.I., Cheng, Y.Y., Chou, Y.J., Lu, C.F., 2008. Chromosomal polymorphism of ribosomal genes in the genus *Oryza*. Theor. Appl. Genet. 116, 745–753.

Comes, H.P., Abbott, R.J., 2001. Molecular phylogeography, reticulation, and lineage sorting in Mediterranean Senecio sect. Senecio (Asteraceae). Evolution 55, 1943–1962.

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

Denk, T., Grimm, G., Stögerer, K., Langer, M., Hemleben, V., 2002. The evolutionary history of *Fagus* in western Eurasia: evidence from genes, morphology and the fossil record. Plant Syst. Evol. 232, 213–236.

Denk, T., Grimm, G.W., Hemleben, V., 2005. Patterns of molecular and morphological differentiation in Fagus: implications for phylogeny. Am. J. Bot. 92, 1006–1016.

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

Franzke, A., Mummenhoff, K., 1999. Recent hybrid speciation in Cardamine (Brassicaceae)—conversion of nuclear ribosomal ITS sequences in statu nascendi. Theor. Appl. Genet. 98, 831–834.

Fuertes Aguilar, J., Rossello, J.A., Nieto Feliner, G., 1999. Nuclear ribosomal DNA (nrDNA) concerted evolution in natural and artificial hybrids of *Armeria* (Plumbaginaceae). Mol. Ecol. 8, 1341–1346.

Ge, S., Sang, T., Lu, B.-R., Hong, D.-Y., 1999. Phylogeny of rice genomes with emphasis on origins of allopolyploid species. Proc. Natl. Acad. Sci. USA 96, 14400–14405.

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

Gopalakrishnan, R., Sharma, S.D., Shastry, S.V.S., 1965. Genome constitution of *Oryza malampuzhaensis* Krishn. et Chandra. Curr. Sci. 34, 128.

Grebenstein, B., Röser, M., Sauer, W., Hemleben, V., 1998. Molecular phylogenetic relationships in Aveneae (Poaceae) species and other grasses as inferred from ITS1 and ITS2 sequences. Plant Syst. Evol. 213, 233–250.

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

Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–174.

Hershkovitz, M.A., Zimmer, E.A., Hahn, W.J., 1999. Ribosomal DNA sequences and angiosperm systematics. In: Hollingworth, P.M., Bateman, R.M., Gornall, R.J. (Eds.), Molecular Systematics and Plant Evolution. Taylor & Francis, London, pp. 268–326.

Hörandl, E., Paun, O., Johansson, J.T., Lehnebach, C., Armstrong, T., Lockhart, P., 2005. Phylogenetic relationships and evolutionary traits in Ranunculus s.l. (Ranunculaceae) inferred from ITS sequence analysis. Mol. Phylogenet. Evol. 36, 305–327.

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

Hugall, A., Stanton, J., Moritz, C., 1999. Reticulate evolution and the origins of ribosomal internal transcribed spacer diversity in apomictic Meloidogyne. Mol. Biol. Evol. 16, 157–164.

Jobst, J., King, K., Hemleben, V., 1998. Molecular evolution of the internal transcribed spacers (ITS1 and ITS2) and phylogenetic relationships among species of Cucurbitaceae. Mol. Phylogenet. Evol. 9, 204–219.

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

Kotseruba, V., Gernand, D., Meister, A., Houben, A., 2003. Uniparental loss of ribosomal DNA in the allotetraploid grass *Zingeria trichopoda* ($2n = 8$). Genome 46, 156–163.

Kovarik, A., Pires, J.C., Leitch, A.R., Lim, K.Y., Sherwood, A.M., Matyasek, R., Rocca, J., Soltis, D.E., Soltis, P.S., 2005. Rapid concerted evolution of nuclear ribosomal DNA in two *Tragopogon* allopolyploids of recent and recurrent origin. Genetics 169, 931–944.

Kovarik, A., Dadejova, M., Lim, Y.K., Chase, M.W., Clarkson, J.J., Knapp, S., Leitch, A.R., 2008. Evolution of rDNA in Nicotiana allopolyploids: a potential link between rDNA homogenization and epigenetics. Ann. Biol. 101, 815–823.

Lim, K.Y., Soltis, D.E., Soltis, P.S., Tate, J., Matyasek, R., Srubarova, H., Kovarik, A., Pires, J.C., Xiong, Z., Leitch, A.R., 2008. Rapid chromosome evolution in recently formed polyploids in *Tragopogon* (Asteraceae). PLoS ONE 3, e3353.

Lu, B.R., Ge, S., 2005. Taxonomic treatment of *Porteresia coarctata* (Poaceae: Oryzeae). Nord. J. Bot. 23, 555–558.

Manen, J.-F., 2004. Are both sympatric species Ilex perado and Ilex canariensis secretely hybridizing? Indication from nuclear markers collected in Tenerife. BMC Evol. Biol. 4, 46.

Matyasek, R., Tate, J.A., Lim, Y.K., Srubarova, H., Koh, J., Leitch, A.R., Soltis, D.E., Soltis, P.S., Kovarik, A., 2007. Concerted evolution of rDNA in recently formed *Tragopogon* allopolyploids is typically associated with an inverse correlation between gene copy number and expression. Genetics 176, 2509–2519.

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

Nayar, N.M., 1973. Origin and cytogenetics of rice. In: Caspari, E.W. (Ed.), Advances in Genetics. Academic Press, New York, pp. 153–292.

Nei, M., Li, W.H., 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. USA 76, 5269–5273.

Nieto Feliner, G., Rosselló, 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.

O'Kane, S.L., Schaal, B.A., Al-Shehbaz, I.A., 1996. The origin of *Arabidopsis suecica* (Brassicaceae) as indicated by nuclear rDNA sequences. Syst. Bot. 21, 559–566.

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

Popp, M., Oxelman, B., 2001. Inferring the history of the polyploid *Silene aegaea* (Caryophyllaceae) using plastid and homoeologous nuclear DNA sequences. Mol. Phylogenet. Evol. 20, 474–481.

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

Rosselló, J.A., Cosín, R., Boscaiu, M., Vicente, O., Martínez, I., Soriano, P., 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., Cosín, R., Molins, A., 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.

Rozas, J., Rozas, R., 1999. DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics 15, 174–175.

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

Siripun, K.C., Schilling, E.E., 2006. Molecular confirmation of the hybrid origin of Eupatorium godfreyanum (Asteraceae). Am. J. Bot. 93, 319–325.

Soltis, D.E., Soltis, P.S., 1991. Multiple origins of the allopolyploid *Tragopogon mirus* (Compositae): rRNA evidence. Syst. Bot. 16, 407–413.

Soltis, D.E., Soltis, P.S., 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis, D.E., Soltis, P.S., Doyle, J.J. (Eds.), Molecular Systematics Plants II. DNA Sequencing. Kluwer Academy Publications, Boston, pp. 1–42.

Swofford, D.L., 2002. PAUP*: Phylogenetic Analysis using Parsimony (* and Other Methods), Vers. 4.0b10. Sinauer, Sunderland, MA, USA.

Thompson, J.D., Higgins, D.G., 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.

Vaughan, D.A., 1989. The Genus *Oryza* L. Current Status of Taxonomy. International Rice Research Institute, Manila, Philippines.

Vaughan, DA., 1994. The Wild Relatives of Rice. International Rice Research Institute, Manila, Philippines.

Volkov, R.A., Borisjuk, N.V., Panchuk, I.I., Schweizer, D., Hemleben, V., 1999. Elimination and rearrangement of parental rDNA in the allopolyploid *Nicotiana tabacum*. Mol. Biol. Evol. 16, 311–320.

Volkov, R.A., Komarova, N.Y., Hemleben, V., 2007. Ribosomal DNA in plant hybrids: inheritance, rearrangement. Expr. Syst. Biodiv. 5, 261–276.

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.

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

Wissemann, V., 1999. Genetic constitution of Rosa Sect. Caninae (*R. canina, R. jundzillii*) and Sect. Gallicanae (*R. gallica*). Angew. Bot. 73, 191–196.

Wissemann, V., 2003. Hybridization and the evolution of the nrITS spacer region. In: Sharma, A.K., Sharma, A. (Eds.), Plant Genome: Biodiversity and Evolution, Phanerogams, vol. I. Science Publishers, Enfield, pp. 57–71.

Yonemori, K., Honsho, C., Kanzaki, S., Eiadthong, W., Sugiura, A., 2002. Phylogenetic relationships of Mangifera species revealed by ITS sequences of nuclear ribosomal DNA and a possibility of their hybrid origin. Plant Syst. Evol. 231, 59–75.

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, W., Qu, L.J., Gu, H., Gao, W., Liu, M., Chen, J., Chen, Z., 2002. Studies on the origin and evolution of polyploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. Theor. Appl. Genet. 104, 1099–1106.

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.