SEQUENCE CORNER

Polyploidy origin of wheatgrass *Douglasdeweya wangii* (Triticeae, Poaceae): evidence from nuclear ribosomal DNA internal transcribed spacer and chloroplast *trnL–F* sequences

Quanlan Liu · Bao-Rong Lu · NingNing Zhang · Jie Liu · Ying Yu · Hongguang Wang · Xuebing Yan

Abstract To study hybrid speciation in wheatgrass *Douglasdeweya wangii* and to investigate the evolutionary pattern of nuclear ribosomal DNA (nrDNA) internal transcribed spacer sequences (ITSs) in allotetraploids, DNA sequence variation of ITSs and chloroplast *trnL–F* sequences from *D. wangii* and its putative donors were analyzed. The ITSs revealed that *D. wangii* had an StP genome composition. Most accessions of *D. wangii* had one parental ITS copy in their genome, one accession had two parental ITSs. The *trnL–F*

Communicated by K. Schneitz

Q. Liu (✉) · N. Zhang · J. Liu · Y. Yu
Department of Bioengineering and Biotechnology,
Qingdao University of Science and Technology,
Qingdao 266042, China
e-mail: liuquanlan@yahoo.com

Q. Liu
e-mail: liuquanlan@qust.edu.cn

B.-R. Lu
Key Laboratory of Education Ministry for Biodiversity Science
and Ecological Engineering, Institute of Biodiversity Science,
Fudan University,
200433 Shanghai, China

H. Wang
Department of Pharmacy,
Qingdao University of Science and Technology,
Qingdao 266042, China

X. Yan
College of Animal and Veterinary Science,
Henan Agricultural University,
450002 Zhengzhou, China

sequences revealed an especially close relationship of *Pseudoroegneria* to all *D. wangii* individuals included, and the two accessions of *Pseudoroegneria tauri* (PI401324 and PI401331) were maternal candidates of the studied *D. wangii* individuals. Both of ITS and *trnL–F* trees suggested multiple origins and recurrent hybridization of *D. wangii*. Thus, the results suggested that: (1) the St and P genome in allotetraploid *D. wangii* were donated by *Pseudoroegneria* and *Agropyron*, respectively;(2) *Pseudoroegneria* was the maternal donor of *D. wangii,* and *P. tauri* 26 (accession PI401324) and *P. tauri* 27 (accession PI401331) were most likely the potential candidates of maternal donors; (3) *D. wangii* individuals studied here showed multiple origins and experienced recurrent hybridization; and (4) bidirectional interlocus concerted evolution of ITSs had occurred in most *D. wangii* accessions, while in one accession concerted evolution among homeologous loci did not occur.

Keywords *Douglasdeweya wangii* · Internal transcribed spacer (ITS) · Chloroplast *trnL–F* · Phylogeny · Hybridization

Introduction

Hybridization and polyploidization are thought to play an important role in angiosperm evolution. Many studies on the evolution of polyploid complexes have shown that recurrent formation and multiple origins of polyploids are the rule rather than the exception (e.g., Soltis and Soltis 1999). The origin of polyploids and the mechanisms underlying the establishment of newly evolved taxa and populations are among the most challenging questions in plant sciences (e.g., Soltis and Soltis 1999). Molecular tools

have greatly improved our knowledge about hybrid speciation, as reviewed by Soltis and Soltis (2009). DNA analyses of the nuclear and plastid genomes have greatly increased the possibility of detecting and distinguishing evolution events of polyploids, such as parental donors, hybridization, and introgression, following paternal and maternal genome lineages. Through such analyses, several polyploid species within the Triticeae have been characterized in significant detail, including those of the genera *Elymus* (Mason-Gamer et al. 2002; Liu et al. 2006), *Douglasdeweya* (Baum and Johnson 2008), etc. In most of these studies, the chloroplast DNA (cpDNA) sequences, particularly the noncoding regions such as the intron of *trnL* (UAA) and the intergenic spacer of *trnL* (UAA)–*trnF* (GAA) served as valuable source of markers for identifying the maternal donors of polyploids; also nuclear ribosomal DNA (nrDNA) internal transcribed spacer region sequences (ITSs) had been used to investigate origins of polyploids. In these and others' studies, the phenomenon of concerted evolution of ITSs has been demonstrated (e.g., Dover 1982). Concerted evolution describes the molecular process of DNA sequence homogenization among different loci within the Triticeae (e.g., Liu et al. 2006).

Douglasdeweya wangii C. Yen, J. L. Yang &B.R. Baum is a newly established species of the newly erected genus *Douglasdeweya* C. Yen, J.L. Yang & B.R. Baum. It was originally included in *Pseudoroegneria* (Nevski) Á. Löve until recently when a new treatment was made by Yen et al. (2005). Cytological data and molecular analysis of 5S rDNA sequences suggested that *D. wangii* has the StP genome composition, suggesting that the St genome was donated by the genus *Pseudoroegneria* and the P genome by the genus *Agropyron* (Yen et al. 2005; Baum and Johnson 2008). However, hybrid speciation of *D. wangii* is still unknown. Knowledge of the molecular phylogeny of *D. wangii* and its putative diploid donors will provide a better understanding of its polyploid origin. To gain some insights into the ITSs evolution of natural polyploid individuals, we selected here the ITS region to study the change of parental ITS copies in *D. wangii*. In addition, we used the chloroplast *trnL–F* sequences to discern the possible maternal taxon of *D. wangii*. Thus, the specific objectives of this study were to: (1) to study hybrid speciation and the maternal donor of *D. wangii*; (2) to investigate the evolutionary pattern of the ITS region in *D. wangii*.

Materials and methods

Plant materials

Seed materials of *D. wangii* are rare to find, only six accessions were included in this study. They were analyzed together with 23 *Pseudoroegneria* taxa (St), two *Agropyron* taxa (P), and one accession of *Hordeum bogdanii*. *H. bogdanii* was used as an outgroup selected based on the previous phylogenetic studies of Triticeae (Hsiao et al. 1995). All seed materials were kindly provided by American National Plant Germplasm System (Pullman, Washington, USA). Information on the accessions used in this study is presented in Electronic supplementary Table S1 (online).

Internal transcribed spacer amplification, cloning, and sequencing

Total genomic DNA was extracted from fresh leaves, following the method as described by Liu et al. (2006). NrDNA ITS region was amplified by the primers of ITS4 and ITS5 (Hsiao et al. 1995). The polymerase chain reaction (PCR) amplification of ITS was carried out in a total reaction volume of 25 μL containing 1× reaction buffer, 1.5 mM $MgCl_2$, 0.5 μM of each primer, 200 μM of each dNTP (TaKaRa Inc., Dalian, Liaoning, China), 0.5 U of Ex Taq Polymerase (TaKaRa Inc., Dalian, Liaoning, China), with an addition of 8% dimethyl sulfoxide (DMSO) and water to the final volume. The thermocycling profile of an initial denaturation step at 94°C for 3 min, followed by 30 cycles of 30 s at 94°C, 40 s at 55°C, 1 min at 72°C and a final extension step of 10 min at 72°C. PCR reactions from the *D. wangii* and *Pseudoroegneria* accessions were run in triplicates in different thermocyclers and the PCR products were combined in an attempt to offset the potential effects of PCR drifts (Wagner et al. 1994). The PCR products were purified and linked into a pMD-T vector. Purified plasmid DNAs were digested with *Eco*RI and *Hin*dIII. For each accession used in this study, sic to eight cloned PCR products were sequenced to include all the possible ITS sequences, and run on an ABI 3730 sequencer.

CpDNA amplification and sequencing

The chloroplast tRNA genes *trnL–F* were amplified using the primers c and f of Taberlet et al. (1991). Amplification of the cpDNA was performed in a total reaction volume of 25 μl with the same components as described in the ITS amplification, except that DMSO was not added. The PCRs were performed as for the ITS, except that the annealing and extension times were 90 s each. The PCR products were cleaned as described in the previous section, the primers c and f of Taberlet et al. (1991) were used to sequence both strands of the PCR fragments to unambiguously identify all sites. Sequencing was run on an ABI 3730 sequencer.

Data analysis

The ITS and *trnL–F* sequences were aligned with CLUSTAL X version 1.81 and refined manually. In the case of

multiple identical sequences resulting from cloned PCR products from one accession, only one sequence was included in the data set. The boundaries of the ITSs (ITS1-5.8S-ITS2) and *trnL–F* (*trnL* intron-*trnL* 3′exon-intergenic spacer-*trnF*5′ exon) were determined according to Hsiao et al. (1995) and Ogihara et al. (2002), respectively. Gaps were coded as missing data. The basic sequence statistics, including nucleotide frequencies and variability in different regions of the sequences were computed by MEGA 3. The haplotype analysis of each data set was conducted by the software program DnaSP 4.0.

Phylogenetic analyses of the sequence data were performed using the maximum parsimony and neighbor-joining methods. Maximum parsimony analysis was conducted in PAUP version 4.0b10 (Swofford 1998). Heuristic research was performed with tree bisection-reconnection branch swapping, MULPARS option, ACCTRAN optimization, and 100 random addition replicates. Statistical support of the branches was tested with 500 repeats, using the same parameters as above. Neighbor-Joining analysis was performed using the neighbor-joining Program of the PHYLIP package, and carried out with 1,000 bootstrap replicates.

Results and discussion

Sequence variation

The length of ITSs including the 5.8S rDNA gene is highly conserved among the taxa studied (ITS1/5.8S rDNA/ITS2 (length in bp): *D. wangii*: 263–267:164:271–272; *Pseudoroegneria*, 266–268:164:271–273; *Agropyron*, 263:164:272). Aligned ITSs are 706 bp (ITS1, 268 bp; 5.8S rDNA, 164 bp; ITS2, 274 bp), 74 sites were constant, 35 were parsimony-informative, and 97 occurred only once. The number of gaps relative to the outgroup after alignment were eight (five in ITS1 and three in ITS2). The average of G+C content was 61.1% (ITS1, 60.7%; 5.8S rDNA, 60.2%; ITS2, 62.4%).

For clarity, different sequences from the same accessions was named as "accession name+A, or+B, or+C, etc." We analyzed all ITS types using the software program DnaSP 4.0. In total, we detected 49 different ITS types that were distributed among the species as follows: *D. wangii* with six, *Pseudoroegneria* taxa with 40, *Agropyron* taxa with two, and *H. bogdanii* with 1 ITS type (Electronic supplementary Fig. S1, online). It was notable that the ITS type of *D. wangii* 6 was the same as the ITS type of *Pyramidula strigosa* 19A. The distribution of the variable nucleotides among these taxa is shown in Electronic supplementary Fig. S1 (online). This analysis demonstrates that *D. wangii* resamples the variation of the parental taxa

with the exception of sites 24, 595, 600, and 640 (Electronic supplementary Fig. S1, online).

The length of the sequenced chloroplast *trnL–F* varied from 836 to 855 bp in all accessions (*D. wangii*, 843–847; *Pseudoroegneria*, 844–855; *Agropyron*, 836). Aligned *trnL–F* sequences were 866 bp, 824 sites were constant, 14 were parsimony-informative, and 28 occurred only once. The number of gaps relative to the outgroup after alignment were 65. The average of G+C content was 29.8%.

We analyzed all *trnL–F* types. In total, we detected 24 different *trnL–F* types, which were distributed among the several species as follows: *D. wangii* with five, *Pseudoroegneria* taxa with 17, *Agropyron* taxa with one, and *H. bogdanii* with one *trnL–F* type (Electronic supplementary Fig. S2, online). The results showed that *D. wangii* 6 had the same *trnL–F* type with that of *P. tauri* 26. In the *Pseudoroegneria* parental taxa, some accessions had a single *trnL–F* copy type; in the *Agropyron* parental taxa, two *Agropyron* species had the same *trnL–F* type. The distribution of the variable nucleotides among these taxa is shown in Electronic supplementary Fig. S2 (online).

Phylogenetic analysis

Maximum parsimony analysis of ITSs resulted in 32 equally most parsimonious trees. Each of the trees was 303 steps with a consistency index (CI) of 0.4719 and a retention index (RI) of 0.4754. In one of the most parsimonious trees (Fig. 1), two major clades were identified in the phylogenetic tree, which correspond to the two genomic types (St and P). The first clade (with 87% bootstrap support) consisted of *Agropyron* species and four accessions of *D. wangii* (*D. wangii* 1, 2, 3, and 4A), named P clade. The second clade (with 92% bootstrap support) included *Pseudoroegneria* species and three accessions of *D. wangii* (*D. wangii* 4B, 5, and 6), named as St clade. Neighbor-Joining analysis generated a similar topology with minor variation in bootstrap values.

Analysis of *trnL–F* sequences yielded 2 equally most parsimonious trees, with a tree length of 51, a CI of 0.9608, and a RI of 0.9524. One of the most parsimonious *trnL–F* trees was randomly chosen for the phylogenetic analysis (Fig. 2). As shown in the tree, all six accessions of *D. wangii* and *Pseudoroegneria* species formed a large and highly supported clade (92% bootstrap support), named as St clade. Only *Agropyron* diploid species formed P clade. Neighbor-joining analysis generated exactly the same topology with minor variation in bootstrap values. The *trnL–F* gene tree offers an opportunity to identify the maternal donor of *D. wangii*, because the chloroplast genome is maternally inherited in grasses (Mason-Gamer et al. 2002). The topology of *D. wangii* suggests that *Pseudoroegneria* species (St) served as the maternal donor
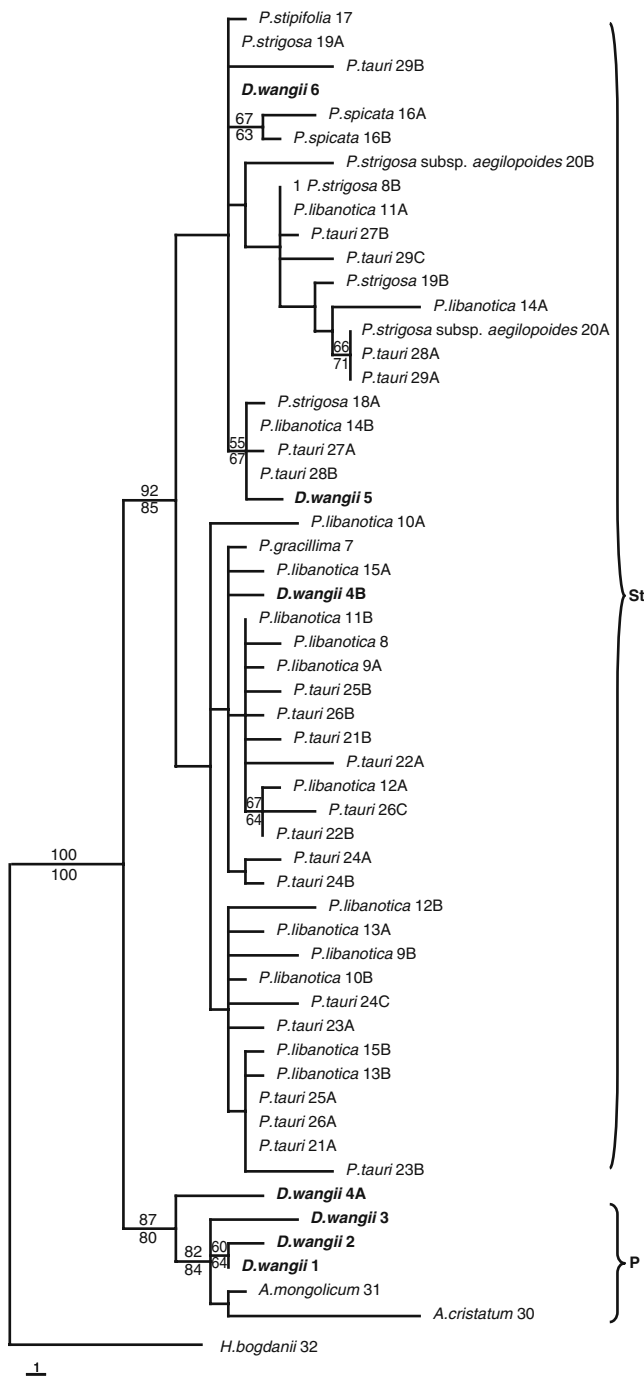
Fig. 1 One of 32 most parsimonious (MP) trees inferred from ITS sequences of *Douglasdeweya wangii* and its putative donors in this study (tree length=303, consistency index=0.4719, retention index= 0.4754). Numbers above and below the branches indicate bootstrap values >50% by MP and NJ analyses, respectively. Branch lengths are proportional to the number of nucleotide substitutions; the scale bar at the upper-left corner indicates one substitution. Numbers following species names correspond to the numbers in the first column in the Electronic supplementary Table S1. Names in bold font indicate individuals of *D. wangii*. The genome type (St or P) of a monophyletic group is given to the right

during speciation of *D. wangii*. It should be notified that, in the *trnL–F* tree, *P. tauri* 27 and *D. wangii* 4 grouped into one subclade; whereas *P. tauri* 26 and the five accessions of *D. wangii* (*D. wangii* 1, 2, 3, 5, and 6) formed another subclade. These two subclades demonstrated that *P. tauri* 26 (accession PI401324) and *P. tauri* 27 (accession PI401331) were the potential individuals as maternal donor of *D. wangii*.

Origins of wheatgrass *D. wangii* allotetraploids

Based on the genomic classification in the Triticeae, *D. wangii* was treated separately from *Pseudoroegneria* and grouped in the new genus *Douglasdeweya* (Yen et al. 2005). Only limited molecular studies addressing the genomic constitution of *D. wangii* are reported (Baum and Johnson 2008). Little is known about phylogeny of *D. wangii* and its related genera at the molecular level. The analyses of ITSs and *trnL–F* sequences will provide opportunities for understanding their ancestral donors and polyploidization events in the speciation processes.

In the ITS tree, two genomes presented in *D. wangii* formed two distinct clades. For example, *D. wangii* 4 was included in the St and P clades; *D. wangii* 5 and 6 were grouped within the St clade; *D. wangii* 1, 2 and 3 were included in the P clade. These topologies indicated that ITSs in different *D. wangii* accessions showed a clear linkage with their related diploid ancestors, which further support the allotetraploid origin of *D. wangii* (Yen et al. 2005; Baum and Johnson 2008). In addition, the ITS tree in this study also suggested a multiple origin of *D. wangii* resulting from recurrent hybridization, which can be shown by different accessions of the same species, will appear at different clades of a phylogenetic tree (Soltis and Soltis 1999). For example, *D. wangii* 4, 5, and 6 were grouped in different subclades of the St clade. Similarly, the *trnL–F* tree also indicated the multiple origin of *D. wangii*. For example, *D. wangii* 4 and *P. tauri* 27 were grouped together with 62% bootstrap support; while other accessions of *D. wangii* and *P. tauri* 26 were grouped together with 86% bootstrap support.

Internal transcribed spacer type evolution

Recent research suggested that hybridization and polyploidization can lead to genome complicated change within hybrids (e.g., Soltis and Soltis 1999; Soltis and Soltis 2009). On the other hand, empirical studies of the fate of nrDNA loci after polyploidy indicated the evolution of nrDNA loci exhibit a range of pattern, from the maintenance of both homeologous loci to the rapid loss of a locus or interlocus homogenization between homeologous loci (e.g., Rauscher et al. 2004). All these studies indicated the complexity of the nrDNA sequence evolution.

Fig. 2 One of 2 most parsimonious (MP) trees inferred from *trnL–F* sequences of *Douglasdeweya wangii* and its putative donors used in this study (tree length=51, consistency index=0.9608, retention index=0.9524). Numbers above and below the branches indicate bootstrap values >50% by MP and NJ analyses, respectively. Branch lengths are proportional to the number of nucleotide substitutions; the scale bar at the upper-left corner indicates one substitution. Numbers following species names correspond to the numbers in the first column in Electronic supplementary Table S1. Names in bold font indicate individuals of *D. wangii*. The genome type (St or P) of a monophyletic group is given to the right
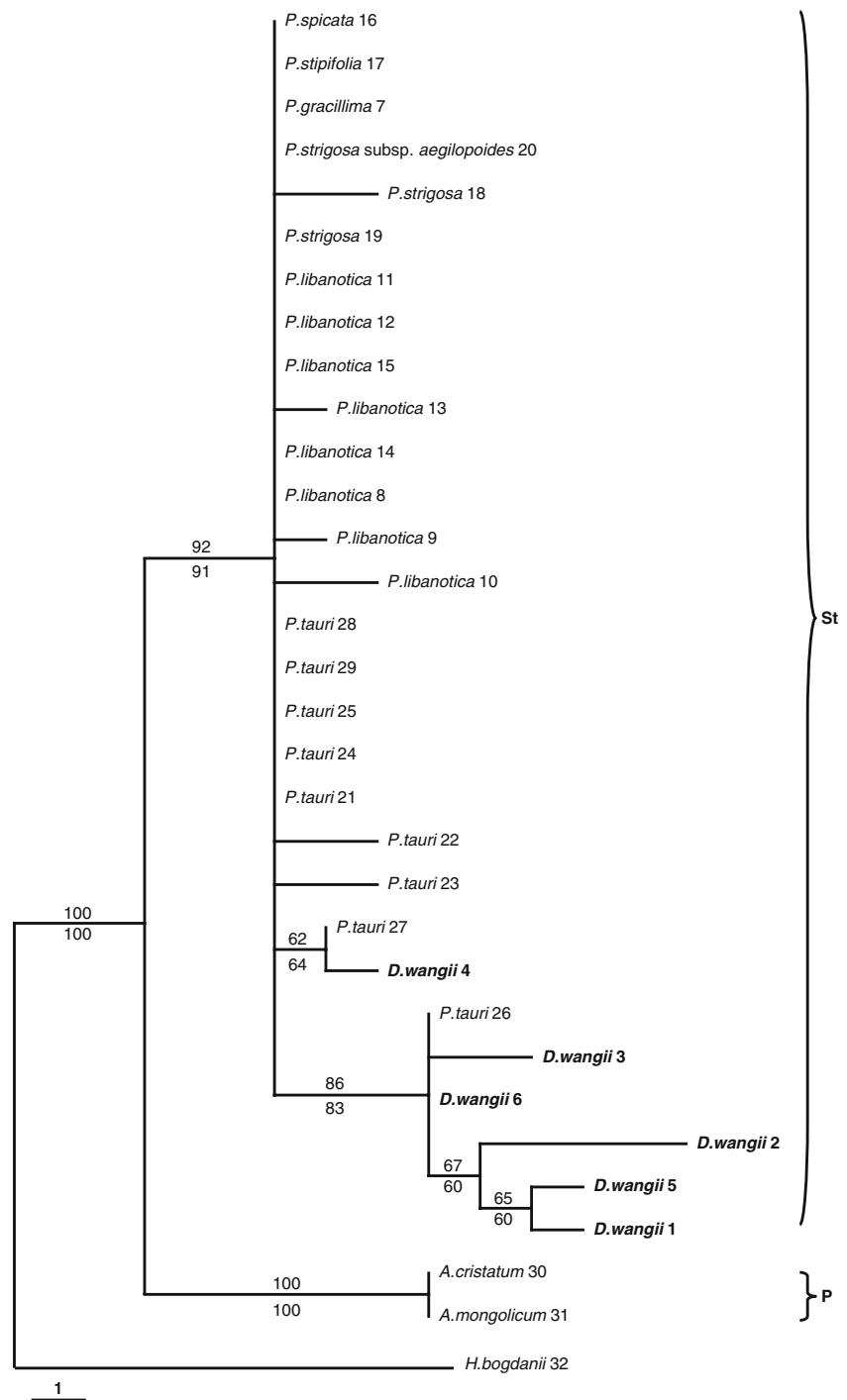
A previous study in which ITSs amplified from polyploid *Elymus* species and its related diploids were used to investigate putative diploid donors of polyploids (Liu et al., 2006). This study showed that the ITS primers and experiment procedure were effective to obtain the related diploid copies of polyploids. However, when using the same primers and experimental procedure in this study, the ITSs of *D. wangii* exhibited high heterogeneity. For example, three *D. wangii* individuals

(*D. wangii* 1, 2, and 3) only had the P genomic ITSs; *D. wangii* 5 and 6 only had the St genomic ITSs; and *D. wangii* 4 had two genomic (P and St) ITSs. There are three possible explanations for this phenomenon. Firstly, the bidirectional concerted evolution of ITSs had occurred between the St and P genomes in some allotetraploid *D. wangii* individuals, i.e. ITS of some *D. wangii* accessions has concerted to a St genome repeat type, whereas ITS from other accessions has become

homogenized to a P genome repeat type. Secondly, the concerted evolution of ITSs among homeologous loci did not occur in some accessions. For example, *D. wangii* 4 had both homeologous loci. These two explanations were supported by others' studies (e.g., Rauscher et al. 2004). In glycine, different homeologues of ITS sequences have been favored among different accessions of the same polyploid species (Rauscher et al., 2004). The third explanation is that PCR selection and PCR drift had occurred during the process of PCR amplification. PCR selection is unlikely since the same primer favored different repeats in different accessions. PCR drift is also unlikely to account for most of this variation. In addition, to avoid PCR drift, the PCR product was designed as a mixture of three PCR products amplified from three different thermocyclers. A same experimental approach to avoid PCR drift was successfully used by others (Mason-Gamer 2001; Liu et al. 2006). Therefore, we prefer the first and second explanations. However, the chloroplast experiment did not reveal a similar heterogeneity. Unlike the biparentaly inherited ITSs, the *trnL–F* sequence is maternally inherited (e.g., Mason-Gamer et al. 2002; Liu et al. 2006). In addition, the low differentiation of the chloroplast *trnL–F* sequences further supported the close genetic similarity of *Pseudoroegneria* species suggested by previous cytological investigation (Jensen et al. 1995).

References

Baum BR, Johnson DA (2008) Molecular confirmation of the genomic constitution of *Douglasdeweya* (Triticeae: Poaceae): demonstration of the utility of the 5S rDNA sequences as a tool for haplome identification. Mol Genet Genomics 279:621–628

Dover G (1982) Molecular drive: cohesive mode of species evolution. Nature 9:111–116

Hsiao C, Chatterton NJ, Asay KH, Jensen KB (1995) Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences. Genome 38:221–223

Jensen KB, Curto M, Asay K (1995) Cytogenetics of Eurasian bluebunch wheatgrass and their relationship to North American bluebunch and thickspike wheatgrasses. Crop Sci 35:1157–1162

Liu QL, Ge S, Tang HB, Zhang XL, Zhu GF, Lu BR (2006) Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast *trnL–F* sequences. New Phytol 170:411–420

Mason-Gamer RJ (2001) Origin of North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. Syst Bot 26:757–768

Mason-Gamer RJ, Orme NL, Anderson CM (2002) Phylogenetic analysis of North American *Elymus* and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. Genome 45:991–1002

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S et al (2002) Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. Mol Genet Genomics 266:740–746

Rauscher JT, Doyle JJ, Brown AHD (2004) Multiple origins and nrDNA internal transcribed spacer homeologue evolution in the *Glycine tomentella* (Leguminosae) allopolyploid complex. Genetics 166:987–998

Soltis DE, Soltis PS (1999) Polyploidy: recurrent formation and genome evolution. Trends Ecol Evol 14:348–352

Soltis PS, Soltis DE (2009) The role of hybridization in plant speciation. Ann Rev Plant Biol 60:561–588

Swofford DL (1998) PAUP. Phylogenetic analysis using parsimony (and other methods). Sinauer Associates, USA

Taberlet PL, Gielly GP, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol Biol 17:1105–1109

Wagner A, Blackstone N, Cartwright P, Dick M, Misof B, Snow P, Wanger GP, Bartels J, Murtha M, Pendleton J (1994) Surveys of gene families using polymerase chain reaction: PCR selection and PCR drift. Syst Biol 43:250–261

Yen C, Yang JL, Baum BR (2005) *Douglasdeweya*: a new genus, with a new species and a new combination (Triticeae: Poaceae). Can J Bot 83:413–419