Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast *trnL-F* sequences

Quanlan Liu[1], Song Ge[2], Haibao Tang[1], Xianglin Zhang[3], Genfeng Zhu[3] and Bao-Rong Lu[1]

[1]The Ministry of Education Key Laboratory for Biodiversity Science and Ecological Engineering, Institute of Biodiversity Science, Fudan University, Shanghai 200433, China;[2] Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China;[3] Chinese National Human Genome Center at Shanghai, Shanghai 201203, China

Author for correspondence:
Bao-Rong Lu
Tel./Fax: +86 21 65643668
Email: brlu@fudan.edu.cn

Received: *22 October 2005*
Accepted: *9 December 2005*

Summary

• To estimate the phylogenetic relationship of polyploid *Elymus* in Triticeae, nuclear ribosomal internal transcribed spacer (ITS) and chloroplast *trnL-F* sequences of 45 *Elymus* accessions containing various genomes were analysed with those of five *Pseudoroegneria* (St), two *Hordeum* (H), three *Agropyron* (P) and two *Australopyrum* (W) accessions.

• The ITS sequences revealed a close phylogenetic relationship between the polyploid *Elymus* and species from the other genera. The ITS and *trnL-F* trees indicated considerable differentiation of the StY genome species.

• The *trnL-F* sequences revealed an especially close relationship of *Pseudoroegneria* to all *Elymus* species included. Both the ITS and *trnL-F* trees suggested multiple origins and recurrent hybridization of *Elymus* species.

• The results suggested that: the St, H, P, and W genomes in polyploid *Elymus* were donated by *Pseudoroegneria*, *Hordeum*, *Agropyron* and *Australopyrum*, respectively, and the St and Y genomes may have originated from the same ancestor; *Pseudoroegneria* was the maternal donor of the polyploid *Elymus*; and some *Elymus* species showed multiple origin and experienced recurrent hybridization.

Key words: *Elymus*, nuclear ribosomal internal transcribed spacer (ITS), chloroplast *trnL-F*, phylogeny, genomic differentiation, hybridization, polyploidization.

New Phytologist (2006) **170**: 411–420

doi: 10.1111/j.1469-8137.2006.01665.x

Introduction

Polyploidy, resulting from either duplication of a single but complete genome (autopolyploidy) or from combination of two or more differentiated genomes (allopolyploidy), is a prominent mode of speciation (Stebbins, 1971; Masterson, 1994; Soltis & Soltis, 2000). About 70% of angiosperms are of polyploid origin (Masterson, 1994; Soltis & Soltis, 2000; Wendel, 2000), which has significantly enriched diversity of plants. A better understanding of the processes of polyploidization and rapid diversification of the descendants of a single polyploidization event is therefore of widespread evolutionary interest (Wendel, 2000; Soltis *et al.*, 2003). Recent studies using genetic markers in many genera suggest that recurrent origins for polyploid species are the rule rather than the exception (Soltis & Soltis, 2000), and that genetic diversity within recent polyploids is adequate to support rapid adaptive evolution (Doyle *et al.*, 2003; Soltis *et al.*, 2003).

The wheat tribe (Poaceae: Triticeae), an important gene pool for genetic improvement of cereal crops (Dewey, 1984; Lu, 1993, 1994), includes many autopolyploid and allopolyploid taxa. Data from extensive cytogenetic analyses have been used to illustrate systematic relationships of the tribe and to clarify the ancestry of many polyploid species. One complex group of polyploids within Triticeae is the genus *Elymus* that, following the taxonomic delimitation by Löve (1984)

based essentially on genomic constitutions, includes approx. 150 perennial species distributed in a wide range of ecological habitats over the temperate and subtropic regions. *Elymus* has its origin through a typical allopolyploidy process (Dewey, 1984; Löve, 1984). Cytological studies suggest that five basic genomes, namely, the St, Y, H, P and W in various combinations constitute *Elymus* species (Lu, 1994). The St genome is a fundamental genome that exists in all *Elymus* species and is donated by the genus *Pseudoroegneria* (Dewey, 1967). The H, P and W genomes are derived from the genera *Hordeum*, *Agropyron* and *Australopyrum* of Triticeae, respectively (Dewey, 1971; Jensen, 1990; Torabinejad & Mueller, 1993). However, the donor of the Y genome that is present in the majority of the Asiatic *Elymus* species has not yet been identified, although extensive investigations have been carried out (Lu, 1993, 1994).

One important group of *Elymus* includes tetraploids with the StY genomes. About 30 StY genome *Elymus* species are found restrictedly in the temperate Asia, where more than half of the known *Elymus* species originated (Salomon & Lu, 1992; Lu, 1994). A large data set from cytological analyses of artificial hybrids among the StY genome *Elymus* species clearly indicates that the degree of chromosome pairing in the hybrids gradually decreases with increase in geographical distance from the locality of their parental species (Lu & Salomon, 1992; Lu, 1993). This means that the StY genomes in tetraploid *Elymus* species have been modified to a large extent and have relatively high genetic diversity. This phenomenon has not been found in the StH genomes, which have relatively high homology among different tetraploid *Elymus* species distributed in Asia (Lu *et al.*, 1992). Knowledge of the molecular phylogeny of the StY genome *Elymus* will provide a better understanding of their genetic differentiation.

Molecular phylogenetic studies have successfully revealed the origins and evolutionary history of polyploids in plants, clarified the nature of different polyploids, and identified their parental lineages and the hybridization events involved in their formation (Soltis & Soltis, 1993; Wendel, 2000; Soltis *et al.*, 2003). Comparative phylogenies between nuclear and chloroplast/mitochondrial sequences have become a powerful tool to identify the mode of polyploidization in particular groups (Ge *et al.*, 1999; Mason-Gamer, 2001; Popp & Oxelman, 2001; Mason-Gamer, 2004; Rauscher *et al.*, 2004). Among the available nuclear sequences, internal transcribed spacer (ITS) sequences have been used successfully in studying phylogenetic and genomic relationships of plants at lower taxonomic levels (Baldwin *et al.*, 1995; Hsiao *et al.*, 1995; Wendel *et al.*, 1995; Zhang *et al.*, 2002; Hao *et al.*, 2004). 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) are also valuable source of markers for identifying the maternal donors of polyploids with additional capacity to reveal phylogenetic relationships of related species (Sang *et al.*, 1997; Mason-Gamer *et al.*, 2002; Xu & Ban, 2004).

In this study, we sequenced and analysed the nuclear ribosomal ITS and chloroplast *trnL-F* fragments for 30 *Elymus* polyploids and their putative diploid donors to explore the origin and relationships of the polyploid *Elymus* species. The objectives of this study were (1) to reveal relationships of the St, Y, P, W and H genomes of *Elymus* in relation to their putative diploid ancestors; (2) to elucidate the phylogenetic relationships of the StY genome Asiatic tetraploids; and (3) to determine the maternal genomic donor of *Elymus* polyploids.

Materials and Methods

Plant materials

A total of 57 Triticeae accessions were used in this study, including 45 *Elymus* accessions with different genomic combinations (i.e. the StY, StStY, StH, StPY, StWY and StHY genomes), five species of the related genus *Pseudoroegneria* (St and St_1St_2 genomes), three species of *Agropyron* (P genome), two species of *Hordeum* (H genome) and two species of *Australopyrum* (W genome). *Bromus catharticus* was used as the outgroup based on previous phylogenetic studies of Poaceae (Hsiao *et al.*, 1995; Gaut, 2002). All seed materials were collected from the field or provided by Dr B. Salomon of the Swedish University of Agricultural Sciences, Sweden, and Drs H. Y. Zhou and H. Q. Zhang of Sichuan Agriculture University, China. All the accessions sequenced in this study, with their scientific names, geographic origins and GenBank accession numbers are listed in Table 1. The voucher specimens of this study are deposited in the Swedish University of Agricultural Sciences, Sweden and Sichuan Agriculture University, China.

DNA extraction and purification

Seeds were germinated and grown in a growth chamber with 15 h in light and 9 h in dark at 25°C. Leaf samples collected from each accession at the seedling stage were ground in liquid nitrogen in a 1.5-ml microfuge tube, and DNA was extracted and purified using a slight modification of the cetyltrimethylammonium bromide (CTAB) procedure outlined in Doyle and Doyle (1990).

ITS amplification, cloning, and sequencing

The internal transcribed spacer (ITS) region of the nuclear ribosomal DNA was amplified by polymerase chain reaction (PCR) using the primers of ITS4 and ITS5 (Hsiao *et al.*, 1995). The PCR amplification of ITS DNA 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 units of Ex *Taq* Polymerase (TaKaRa Inc.), with an addition of 8% dimethyl sulfoxide (DMSO) and water to the final volume. The thermocycling profile consisted of an initial denaturation step at 94°C for 3 min, followed by 35 cycles of 0.5 min at 94°C, 1 min at

Table 1 Species of *Elymus* and other closely related genera used in this study

No.	Species	Accession No.	Genome	Origin	GenBank Accession No. ITS	GenBank Accession No. *trnL-F*
Pseudoroegneria (Nevski) A. Löve						
1	*P. spicata*	PI547161	St	Oregon, USA	AY740793	AF519159*
2	*P. libanotica*	PI228389	St	Iran	AY740794	AY730567
3	*P. strigosa*	PI499637	St	Urumqi, Xinjiang, China	AY740795	AF519155*
4	*P. alashanica*	Z2006	St_1St_2	Yinchuan, Ninxia, China	AY740796, AY740797	AY73069
5	*P. elytrigioides*	Z2005	St_1St_2	Changdu, Tibet, China	AY740798, AY740799	AY730568
Hordeum L.						
6	*H. bogdanii*	PI531761	H	China	AY740876	AY740789
7	*H. brevisubulatum*	Y1604	H	Fuyun, Xinjiang, China	AY740877	AY740790
Agropyron Gaertner						
8	*A. cristatum*	H10154	P	Altai, Xinjiang, China	AY740890	AY740791
9	*A. cristatum*	H10066	P	Altai, Xinjiang, China	AY740891	AY740792
10	*A. mongolicum*	–	P	–	L36482*	AF519117*
Australopyrum (Tsvelev) A. Löve						
11	*A. retrofractum*	Crane 86146	W	–	–	AF519118*
12	*A. pectinatum*	–	W	–	L36483*, L36484*	–
Elymus L.						
13	*E. antiquus*	H7087	StY	Lixian, Sichuan, China	AY740814, AY740815	–
14	*E. antiquus*	H3400	StY	Sichuan, China	AY740818, AY740819	AY730581
15	*E. anthosachnoides*	Y2236	StY	Yajiang, Sichuan, China	AY740820, AY740821	AY740770
16	*E. barbicallus*	H3267	StY	China, seeds from D. R. Dewey, 1988, D 2509	AY740824, AY740825	–
17	*E. barbicallus*	H3268	StY	China, seeds from D. R. Dewey, 1988, D 3512	AY740822, AY740823	AY730580
18	*E. brevipes*	Y2245	StY	Yajiang, Sichuan, China	AY740826, AY740827	AY740771
19	*E. burchan-buddae*	Y3049	StY	Hongyuan, Sichuan, China	AY740872, AY740873	–
20	*E. burchan-buddae*	Y2219	StY	Batang, Sichuan, China	AY740870, AY740871	AY740772
21	*E. burchan-buddae*	Y2207	StY	Zogang, Tibet, China	AY740874, AY740875	–
22	*E. ciliaris*	H7000	StY	Beijing, China	AY740830, AY740831	AY740773
23	*E. dolichatherus*	H8024	StY	Zhaojie, Sichuan, China	AY740834, AY740835	–
24	*E. dolichatherus*	Y1411	StY	Wenchuan, Sichuan, China	AY740836, AY740837	AY730574
25	*E. grandis*	H3879	StY	Lingtong, Shaanxi, China	AY740828, AY740829	AY730572
26	*E. nakaii*	H7386	StY	Maowen, Sichuan, China	AY740856, AY740857	–
27	*E. nakaii*	H7371	StY	Li xian, Sichuan, China	AY740854, AY740855	AY730585
28	*E. pendulinus*	H8986	StY	Changdu, Tibet, China	AY740846, AY740847	AY730582
29	*E. pendulinus*	Y1412	StY	Wenchuan, Sichuan, China	AY740848, AY740849	–
30	*E. shandongensis*	H3202	StY	Wuhan, Hubei, China	AY740816, AY740817	AY730583
31	*E. tibeticus*	H8927	StY	Luhuo, Sichuan, China	AY740862, AY740863	–
32	*E. tibeticus*	H8366	StY	Gongbogyamda, Tibet, China	AY740864, AY740865	–
33	*E. abolinii*	H3306	StY	Alma-Ata, Medeo, Kazakstan	AY740899, AY740900	–
34	*E. abolinii*	H8491	StY	Xinyuan, Xinjiang, China	AY740812, AY740813	AY730584
35	*E. canaliculatus*	H4116	StY	NWFP, Pakistan	AY740832, AY740833	AY740774
36	*E. fedtschenkoi*	H7510	StY	Habahe, Xinjiang, China	AY740838, AY740839	AY740775
37	*E. fedtschenkoi*	H4114a	StY	NWFP, Pakistan	AY740840, AY740841	–
38	*E. glaberrimus*	Y2042	StY	Jeminay, Xinjiang, China	AY740844, AY740845	AY740777
39	*E. gmelinii*	H1033	StY	Altai, Xinjiang, China	AY740842, AY740843	AY740776
40	*E. longearistatus*	H3261	StY	Northern Iran	AY740804, AY740805	AY740778
41	*E. longearistatus*	H4114b	StY	NWFP, Pakistan	AY740806, AY740807	AY740779
42	*E. macrochaetus*	H10303	StY	Gissar, Tadzhikistan	AY740850, AY740851	AY730573
43	*E. macrochaetus*	H10208	StY	Gissar, Tadzhikistan	AY740852, AY740853	–
44	*E. nevskii*	H3305	StY	Chatkal, Uzbekistan	AY740860, AY740861	AY740780
45	*E. nevskii*	H10213	StY	Gissar, Tadzhikistan	AY740858, AY740859	AY730571
46	*E. semicostatus*	H3288	StY	Mandi, Pakistan	AY740800, AY740801	AY730576
47	*E. semicostatus*	H4101	StY	Hazara, Pakistan	AY740802, AY740803	AY730575
48	*E. validus*	H4078	StY	Babusar, Gilgit, Pakistan	AY740866, AY740867	AY730578
49	*E. validus*	H4100	StY	Naran village, Hazara, Pakistan	AY740868, AY740869	AY730579
50	*E. caucasicus*	H3207	StY	Dilidjan, Armenia	AY740808, AY740809	AY730577
51	*E. panormitanus*	H4152	StY	Crimea, Ukraine	AY740810, AY740811	AY730570

Table 1 *continued*

No.	Species	Accession No.	Genome	Origin	GenBank Accession No. ITS	GenBank Accession No. *trnL-F*
52	*E. tschimganicus*	H3302	StStY	Xinjiang, China	AY740878, AY740879, AY740880	AY740783
53	*E. caninus*	PI564910	StH	Russian Federation	AY740897, AY740898	AY740781
54	*E. himalayanus*	H4134	StHY	Astor valley, Gilgit, Pakistan	AY740881, AY740882, AY740883	AY740782
55	*E. rectisetus*	H3152	StWY	Lake Lyndon, Australia	AY740893, AY740894 AY740896, AY740786	AY740895
56	*E. melantherus*	ZY3146	StPY	Sichuan, China	AY740884, AY740885, AY740886	AY740785
57	*E. rigidulus*	ZY3113	StPY	Xiahe, Gansu, China	AY740887, AY740888, AY740889	AY740784
Bromus L.						
58	*B. catharticus*	S20004	–	Kunming, Yunnan, China	AF521898*	AY829228

Numbers before species' names serve as identifiers of specific accessions that correspond to the numbers in Figs 1,2, and 3. GenBank accessions with an asterisk (*) represent previously published sequences from the GenBank (http://www.ncbi.nlm.nih.gov).

55°C, 1 min at 72°C and final extension step of 10 min at 72°C. PCR reactions from the polyploid *Elymus* 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 using a gel extraction kit (TaKaRa Biotechnology (Dalian) Co., Ltd, Dalian, China) and linked into a pMD-T vector according to the manufacturer's instruction (TaKaRa Biotechnology (Dalian) Co., Ltd). Transformation, plating and isolation of plasmids were performed as described in Sambrook *et al.* (1989). Purified plasmid DNAs were digested with *Eco*RI and *Hin*dIII. For each of the *Elymus* species, 10–15 cloned PCR products were sequenced to include all the possible ITS sequences from the donor species, using ABI BigDye terminators according to the product instructions, and run on an ABI 3730 sequencer.

CpDNA amplification and sequencing

The chloroplast tRNA genes *trnT*, *trnL-5′*, *trnL-3′* and *trnF*, along with their intervening noncoding regions, were amplified using the primers a and f, or 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 nuclear ribosomal DNA, except the annealing and extension times were 1.5 min each. The PCR products were cleaned as described in the previous section, the primers b, d, e and f of Taberlet *et al.* (1991) were used to sequence both strands of the PCR fragments to unambiguously identify all sites, and the sequencing reactions were as described earlier.

Phylogenetic analysis

The ITS and *trnL-F* sequences were aligned with CLUSTAL X (Thompson *et al.*, 1999) and refined manually. The boundaries of the ITS region (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 binary characters by their presence/absence, and were used for the phylogenetic analyses. The basic sequence statistics, including nucleotide frequencies, transition/transversion (ns : nv) ratio and variability in different regions of the sequences were computed by MEGA 3 (Kumar *et al.*, 2004).

The aligned sequences were used as the input data for PAUP 4.0 (Swofford, 1998), MEGA 3 and the PHYLIP software package (Felsenstein, 1995). Parsimony analyses were performed by heuristic search with tree bisection-reconnection (TBR) branch swapping, MULPARS option, ACCTRAN optimization, and 100 random addition replicates. Topological robustness was assessed by bootstrap analysis with 500 replicates using simple taxon addition. The aligned sequences were also analyzed with the Neighbor-Joining Program of the PHYLIP package, and carried out with 1000 bootstrap replicates.

Results

Variation in ITS and *trnL-F* sequences

The ITS sequences in this study included three regions: (1) 43 nucleotides of the 18S rRNA gene; (2) the complete sequences of ITS1, 5.8S rRNA gene and ITS2; and (3) 58 nucleotides of the 26S rRNA. Sequences of the 18S rRNA and the 26S

rRNA genes showed no variation for all accessions included in this study. The length of sequences ranged from 212 to 221 bp in the ITS1 region, and from 216 to 218 bp in the ITS2 region. The 5.8S rRNA gene was 164 bp long and completely identical for all the cloned sequences of the 58 accessions. The average of G + C content was 61.2%. Sequence alignment necessitated gaps of one to eight bases in length. Of 223 variable sites, 115 were parsimoniously informative.

The *trnL-F* fragment sequenced in this study included four regions: (1) the partial *trnL* intron; (2) the *trnL* 3′ exon; (3) the *trnL–trnF* intergenic spacer; and (4) the partial *trnF* exon with 40 bp. The length of the sequenced chloroplast *trnL-F* varied from 859 to 882 bp in all accessions. The average of G + C content was 30.2%. Of the 210 variable sites, 107 were parsimoniously informative, including the polymorphisms introduced by insertions/deletions.

Phylogenetic analysis of the nuclear ITS sequences

We first constructed the phylogeny of all the diploid species based on their ITS sequences in order to reveal relationships of the putative genomic donors of *Elymus*. Three equally most parsimonious trees were obtained with the tree length of 147 steps, a consistency index (CI) of 0.857 and a retention index (RI) of 0.764. As shown in the strict consensus tree (Fig. 1), species with the same genomes formed highly supported monophylies, indicating that the four genomes have been well differentiated and corresponded to different morphologically recognized genera. The H genome (*Hordeum*) was at the basal position, followed by the St genome (*Pseudoroegneria*) that was sister to the other two related P and W genomes (*Agropyron* and *Australopyrum*).

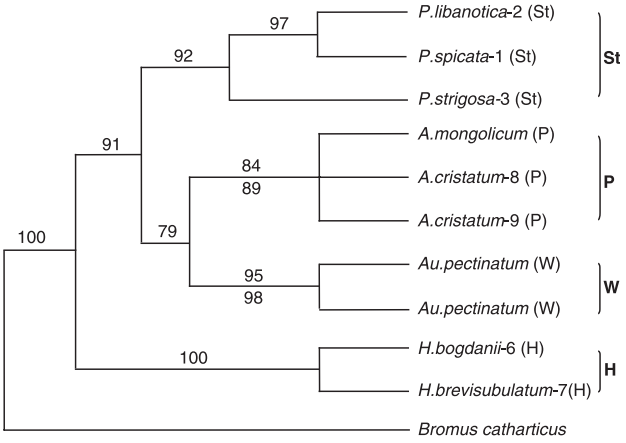

Fig. 1 The strict consensus tree of three most parsimonious (MP) trees inferred from the ITS sequences of the diploid species (Tree length = 147, consistency index (CI) = 0.857, retention index (RI) = 0.764). The topologies obtained by Neighbor-Joining method (NJ) are the same except for some nodes having different bootstrap values. Numbers above and below the branches indicate bootstrap values > 50% by MP and NJ analyses, respectively. Numbers after the species names refer to accession numbers as indicated in Table 1. Capital letters in parentheses following the species names indicate the genome type of the species. The genome type (St, P, W or H) of a monophyletic group is given to the right.

To further analyse genomic relationships and the origin of the polyploid *Elymus*, ITS sequences of all the polyploid species (*Elymus* polyploids with the StH, StY, StStY, StHY, StPY and StWY genomes and *Pseudoroegneria* tetraploids with the St_1St_2 genomes) were included in the phylogenetic analysis, together with those of diploids containing the St, H, P and W genomes. To ensure obtaining all possible types of ITS sequences of a polyploid, 10–15 clones from each of the selected *Elymus* species were sequenced. In the case of multiple identical sequences resulting from cloned PCR products of one accession, only one sequence was included in the data set. Consequently, 107 unique sequences were obtained and used for phylogenetic analyses.

Maximum parsimony analysis resulted in 440 equally most parsimonious trees. Each of the trees was 431 steps with a CI of 0.620 and a RI of 0.819. In one of the most parsimonious tree (Fig. 2), all the homeologous ITS sequences from polyploid accessions grouped with those of the diploid parental clades expected from cytological studies. Four major clades with high bootstrap support (83–100%) were found, which correspond to the four genomic types (H, P, W and St). The first clade consisted of the *Hordeum* and *Elymus* species with the StH and StHY genomes (100% bootstrap support). The second clade included the *Australopyrum* and StWY genome *Elymus* species (96% bootstrap support), while the third clade included the *Agropyron* and StPY genome *Elymus* species (95% bootstrap support). The fourth (largest) clade (83% bootstrap support) comprised the *Pseudoroegneria* species (St and St_1St_2 genomes) and all polyploid *Elymus* with the StStY, StY, StH, StHY, StPY, and StWY genomes. Three subclades (A, B, and C subclades) were recognized in this clade. It is worth mentioning that no obvious Y-genome specific clade was detected in the phylogenetic tree. Neighbor-Joining analysis generated a similar topology with minor variation in bootstrap values.

For clarity, we named the four clades using genomic symbols of the diploid species, i.e. the H, W, P, and St clades, respectively (Fig. 2). The phylogenetic tree showed that all *Elymus* polyploids, except for those with the St and Y genomes, had two distinct types of ITS sequences with one forming a clade with its respective putative diploid donors and the other grouped with the St genome clade. By contrast, all *Elymus* species with the Y and St genomes (StY and StStY) were retained in the St clade. In addition, a well-supported branch (97% bootstrap) was found within the St clade, including the St_1St_2 genome *Pseudoroegneria* and the StY genome *Elymus* tetraploids that shared an 8-bp deletion in their ITS sequences (Fig. 2).

Phylogenetic analysis of the chloroplast *trnL-F* sequences

The chloroplast *trnL-F* sequences of all *Elymus* polyploids and their putative diploid donor species were included for

phylogenetic analysis. Maximum parsimony analysis resulted in 316 equally most parsimonious trees with 307 steps, a consistency index of 0.785, and a retention index of 0.816. It was evident from the phylogenetic tree (Fig. 3) that the diploid species grouped into four distinct clades corresponding to the St, P, W and H genomes, respectively. The tetraploid *Pseudoroegneria* (St_1St_2) and all *Elymus* species formed a large but highly supported clade (100%) together with the diploid *Pseudoroegneria* species, suggesting a close relationship between *Pseudoroegneria* and *Elymus* species in terms of their chloroplast genomes. This clade was named as St clade because all diploid and polyploid accessions contained the St genome. The diploid *Hordeum* species (H clade) was the earliest divergent lineage, followed by the diploid *Agropyron* (P clade) and *Australopyrum* (W clade) species that were sisters to the St clade. Neighbor-Joining analysis produced a similar topology with only very minor differences in bootstrap supports (Fig. 3).

Discussion

Phylogenetic relationships of *Elymus* and its proposed diploid ancestors

The genus *Elymus* consists of polyploids that are widely distributed over different continents and includes a large number of endemic species. Only a few molecular studies addressing phylogenetic relationships of the StH and StHY genome *Elymus* species are reported (Mason-Gamer, 2001; Mason-Gamer *et al.*, 2002; McMillan & Sun, 2004; Xu & Ban, 2004). Little is known about phylogeny of the Asiatic StY genome *Elymus* species at molecular level. Analyses of ITS sequences collected from a wide range of polyploid *Elymus* species and their related genera will provide opportunities for understanding their phylogenetic relationships, ancestral donors and polyploidization events in the speciation processes.

In the diploid and polyploid ITS trees, four of the five genomes presented in *Elymus* formed distinct clades, the exception being the Y genome. Of the four clades, the H genome clade was basal, with the W, P, and St clades being successively more distantly related. There was no obvious Y genome clade, and all the StY species were placed in the St genome clade. These results indicate that ITS sequences of all the genomes derived from the diploid ancestors have remained clearly differentiated in the polyploid *Elymus*. This can be reflected by the fact that all the allopolyploid species (except for those with the Y genome) contained two distinct types of ITS sequences, with one type in the St clade and the others in different genomic clades (H, P or W clade), respectively. This strongly suggests that ITS sequences in different *Elymus* species showed a clear linkage with those in their diploid ancestors. This is illustrated by the fact that the StH, StHY, StPY and StWY genome *Elymus* species were simultaneously clustered in both of their ancestral groups, indicating that two distinct types of ITS sequences exist in these polyploid *Elymus*. Obviously, the homogenization of ITS sequences is not significant in the polyploid *Elymus*. This provides strong evidence that the polyploid *Elymus* species are derived from polyploidization through hybridization between different ancestral genera, as indicated by cytological analyses (Dewey, 1984; Lu, 1994). For example, the StH genome *Elymus* species are derived from the hybridization between *Pseudoroegneria* (St) and *Hordeum* (H). Although the rDNA sequences are considered to undergo a rapid homogenization through concerted evolution, growing evidence shows that ITS polymorphism or incomplete homogenization is the rule rather than the exception (Buckler *et al.*, 1997; Hershkovitz *et al.*, 1999). Directional and bidirectional interlocus concerted evolution following allopolyploid speciation have been documented (Wendel *et al.*, 1995; Fulnecek *et al.*, 2002), but allopolyploid species often maintain both parental sequences of the ITS region. The phenomenon of incomplete homogenization is useful for understanding the origin of hybrids and genomic constructions in polyploidy species (Ainouche & Bayer, 1997; Hershkovitz *et al.*, 1999; Popp & Oxelman, 2001; Yonemori *et al.*, 2002; Koch *et al.*, 2003). It is evident that the homogenization of ITS sequences in the allopolyploid *Elymus* is not completed, which probably suggests the recent origin of these polyploids because sufficient time is needed to allow ITS sequences to be homogenized (Ainouche & Bayer, 1997).

By contrast, all the tetraploid StY genome *Elymus* species had only one genomic type of ITS sequences since all these species were included in the St clade (Fig. 2). This was also true of the tetraploid *Pseudoroegneria* species. However, the hexaploids that combined the St and Y genomes with one of the three differentiated genomes (H, P or W genome) had two genomic types of ITS sequences. There are two possible explanations for this phenomenon. One of the explanations is that the St and Y genomes may have the same origin, because the St genome clade in this study did not show any St- and Y-genome-associated subclades. This explanation is supported by the

Fig. 2 One of the 440 most parsimonious (MP) trees inferred from the internal transcribed spacer (ITS) sequences of all the accessions used in this study (Tree length = 431, consistency index (CI) = 0.620, retention index (RI) = 0.819). The topologies obtained by Neighbor-Joining method (NJ) are the same except for some nodes having different bootstrap values. 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 after species names refer to the accession numbers shown in Table 1. Capital letters in parentheses indicate the genome type of the species. Names in boldface indicate species of *Pseudoroegneria* and other diploid genera. The arrow indicates the 8-bp deletion shared by species in the subclade. *, polyploid species with the 'H' genome; **, polyploid species with the 'W' genome; ***, polyploid species with the 'P' genome. The genome type (St, P, W or H) of a monophyletic group is given to the right.

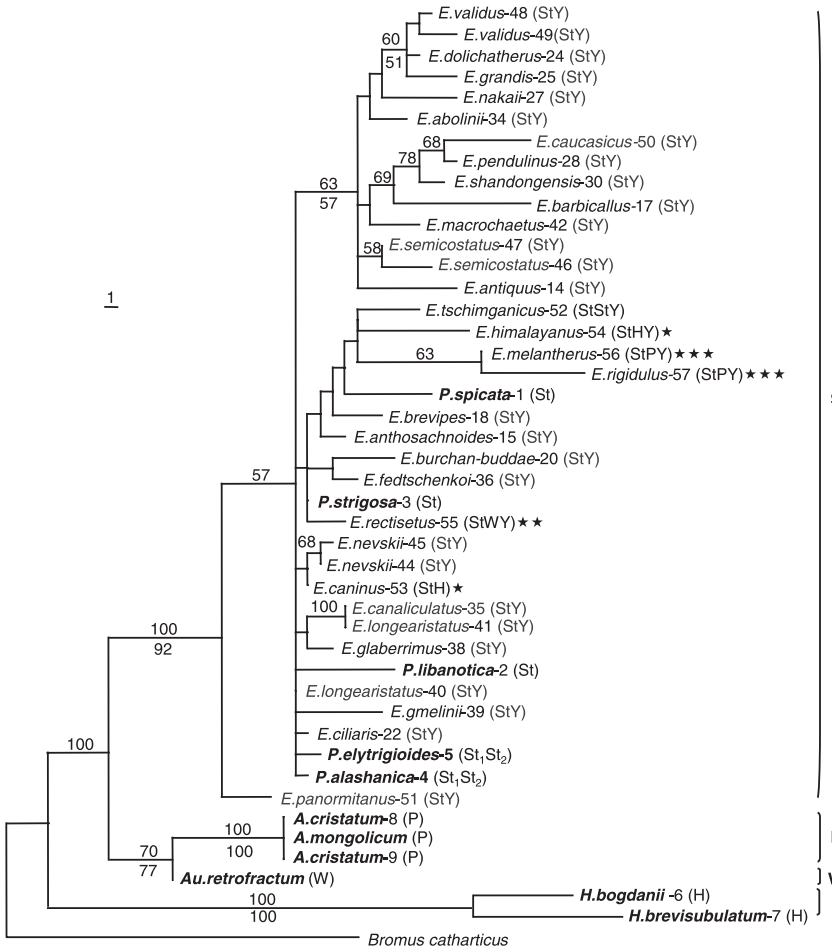

Fig. 3 One of 316 most parsimonious (MP) trees inferred from the *trnL-F* sequences of all the accessions used in this study (Tree length = 307, consistency index (CI) = 0.785, retention index (RI) = 0.816). The topologies obtained by Neighbor-Joining method (NJ) are the same except for some nodes having different bootstrap values. Numbers above and below the branches indicate bootstrap values greater than 50% by MP and NJ analyses, respectively. Branch lengths are proportional to the number of nucleotide substitutions, and the scale bar at the upper-left corner indicates one substitution. Capital letters in parentheses indicate the genome type of the species. Names in boldface indicate species of *Pseudoroegneria* and other diploid genera. *, polyploid species with the 'H' genome; **, polyploid species with the 'W' genome; ***, polyploid species with the 'P' genome. The genome type (St, P, W or H) of a monophyletic group is given to the right.

relatively close affinities between the St and Y genomes reported by Lu *et al*. (1992), and Lu & Bothmer (1991) based on cytological investigations, although a close affinity between the St and P genomes are also reported (Wang *et al*., 1985). In addition, no diploid Y genome species has been found so far despite great efforts world-wide, further supporting the this explanation. It is possible that homogenization of ITS sequences has occurred between the St and Y genomes in polyploid *Elymus* species. However, since there is no (or extremely low) chromosome pairing between the St and Y genomes under normal conditions (the situation is the same between the St and H, P or W genomes) the likelihood of homogenization of the ITS sequences between the St and Y genomes seems to be low in the polyploid *Elymus* species. Previous studies also indicate that homogenization of ITS sequences in allopolyploids rarely results in the formation of a single genomic type of sequences (Wendel *et al*., 1995; Rauscher *et al*., 2004). Therefore, we prefer the explanation that the St and Y genomes may have the same origin.

Differentiation of the StY genome *Elymus* tetraploids

Previous cytological investigations suggest that the StY genomes in *Elymus* tetraploids are considerably differentiated in relation to their geographical distribution (Lu & Salomon, 1992; Lu, 1993). The analysis of ITS sequences based on nearly all the StY genome *Elymus* species covering a wide range of distribution also demonstrated genomic differentiation of these species. It is obvious that a subclade including two *Pseudoroegneria* (St_1St_2) tetraploids and eight *Elymus* (StY) species are supported with high bootstrap support (97%) and are differentiated from other StY species (Fig. 2). This provides an evidence of genomic differentiation of the StY genome species from molecular data. Genomic differentiation among species through polyploidy or at the same ploidy level usually results in great diversity of polyploid species, as suggested by Soltis & Soltis (2000). However, because of the weak resolution of the current ITS data in the St clade, whether or not the differentiation of the StY genome *Elymus* species is associated with their geographic distribution, as suggested by previous cytological observations, needs to be verified further with additional evidence.

Multiple origin is a very important evolutionary process of polyploid species, which emphasizes 'each polyploid species forms over and over again from different parental genotypes generating a diverse array of polyploid genotypes' (Soltis & Soltis, 1999). The ITS tree in this study also suggested a multiple origin of some StY genome species resulting from

recurrent hybridization, which can be shown by different accessions of the same species, will appear at different clades of a phylogenetic tree. For example, in this study, different accessions of *Elymus antiquus* (−13 and −14) and *Elymus pendulinus* (−28 and −29) were grouped in different subclades of the St clade. This helps to explain the abundant genetic diversity within an *Elymus* species. The recurrent hybridization also promoted rapid adaptation of the *Elymus* species to different ecological habitats, resulting in the formation of many endemic genotypes and species.

It is worth pointing out that the two *Pseudoroegneria* and eight StY genome *Elymus* species from the central and eastern Asia consistently had an 8-bp deletion in their ITS sequences and clustered distinctly into one subclade. It is most likely that the 8-bp deletion had already occurred in the *Pseudoroegneria* species before being passed on to some *Elymus* species during the polyploidization process. If this assumption holds true, the St genome in *Elymus* should have been derived from more than one *Pseudoroegneria* species/populations through hybridization. In other words, *Elymus* species with the 8-bp deletion in the ITS sequences may have evolved from hybridization of the *Pseudoroegneria* ancestor with the 8-bp deletion, whereas other *Elymus* species without the 8-bp deletion might originate from hybridization of other *Pseudoroegneria* ancestors. This could also explain the rich diversity and wide adaptation of the StY genome *Elymus* species (Lu, 1994).

The maternal donor of *Elymus* species

The *trnL-F* gene tree represents a maternal genealogy of the *Elymus* species, because the chloroplast genome is maternally inherited in grasses (Ge *et al*., 1999; Mason-Gamer *et al*., 2002). This offers an opportunity to identify the maternal parents of the *Elymus* species. All polyploid species including the *Pseudoroegneria* and *Elymus* species formed a highly supported monophyletic group (100% and 92% bootstraps for MP and NJ trees, respectively) on the maternal *trnL-F* tree. This suggests that the *Pseudoroegneria* species (St genome) served as the maternal donor during the polyploid speciation of the *Elymus* species (tetraploids and hexaploids). This result, in conjunction with the biparentally inherited ITS tree (Fig. 2), implies that diploid *Hordeum* (H genome), *Agropyron* (P) and *Australopyrum* (W) species were the paternal parents for the *Elymus* species with the H, P and W genomes. This conclusion is in good agreement with previous studies by McMillan & Sun (2004) and Mason-Gamer *et al.* (2002) using restriction fragment length polymorphism (RFLP) analysis of cpDNA and chloroplast DNA sequence data, where *Pseudoroegneria* was suggested as the chloroplast genome donor of the northern American StH genome *Elymus* and two StY genome species. A recent study based on analysis of partial *trnL-F* sequences of a few *Elymus* species by Xu & Ban (2004) also presented similar results.

The phylogenetic tree based on *trnL-F* sequences of the *Elymus* species also indicated a multiple origin of polyploids in the evolutionary process of some *Elymus* species. For example, different accessions of *Elymus longearistatus* were clustered in the different subclades of the St clade. In addition, all *Elymus* species used in this study were scattered in different subclades in the *trnL-F* tree, which may also suggest different maternal lineages of the polyploid genus. Conversely, in the ITS tree, at least two hybridization events involving *Elymus caninus* (StH) and *Elymus himalayanus* (StHY) occurred to form these polyploids with the H genome species because they appeared in different subclades of the St clade (*E. caninus* in the A-subclade and *E. himalayanus* in the B-subclade). In addition, two hybridization events involving *Elymus melantherus* (StPY), *Elymus rigidulus* (StPY), and *Elymus rectisetus* (StWY) can be identified for generating the polyploid species with the P and W genomes. Therefore, nuclear DNA and cpDNA sequences showed that many *Elymus* species had a multiple origin and experienced recurrent hybridization between species with different genomes. This suggests that hybridization and polyploidization were the major driving force in the diversity and evolution of the genus *Elymus*.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (Grant no. 30270092) and the Program for Key International S & T Cooperation Project of China (Grant no. 2001CB711103). Prof. M. E. Barkworth of the Utah State University, USA, provided critical comments. Dr N. Clarke of the Norwegian Forest Research Institute, Norway, kindly edited English of the manuscript.

References

Ainouche ML, Bayer RJ. 1997. On the origins of the tetraploid *Bromus* species (section *Bromus*, Poaceae): insights from the internal transcribed spacer sequences of nuclear ribosomal DNA. *Genome* **40**: 730–743.

Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF, Campbell CS, Donoghue MJ. 1995. The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* **82**: 247–277.

Buckler ESIV, Ippolito A, Holtsford TP. 1997. The evolution of ribosomal DNA: Divergent paralogues and phylogenetic implications. *Genetics* **145**: 821–832.

Dewey DR. 1967. Synthetic hybrids of *Agropyron scribneri* × *Elymus juncea*. *Bulletin of the Torrey Botanical Club* **94**: 388–395.

Dewey DR. 1971. Synthetic hybrids of *Hordeum bogdanii* with *Elymus canadensis* and *Sitanion hystrix*. *American Journal of Botany* **58**: 902–908.

Dewey DR. 1984. The genomic system of classification as a guide to intergeneric hybridization with the perennial Triticeae. In: Gustafsen JP, ed. *Gene manipulation in plant improvement*. New York, NY, USA: Plenum Press, 209–279.

Doyle JJ, Doyle JL. 1990. Isolation of plant DNA from fresh tissue. *Focus* **12**: 13–15.

Doyle JJ, Doyle JL, Rauscher JT, Brown AHD. 2003. Diploid and polyploid reticulate evolution throughout the history of the perennial soybeans (*Glycine* subgenus *Glycine*). *New Phytologist* **161**: 121–132.

Felsenstein J. 1995. *Phylip: phylogeny inference package, ver 3.57c*. Seattle, WA, USA: University of Washington Press.

Fulnecek J, Lim KY, Leitch AR, Kovarik A, Matyasek R. 2002. Evolution and structure of 5S rDNA loci in allotetraploid *Nicotiana tabacum* and its putative parental species. *Heredity* **88**: 19–25.

Gaut BS. 2002. Evolutionary dynamics of grass genomes. *New Phytologist* **154**: 15–28.

Ge S, Sang T, Lu BR, Hong DY. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences, USA* **96**: 14400–14405.

Hao G, Yuan YM, Hu CM, Ge XJ, Zhao NX. 2004. Molecular phylogeny of. *Lysimachia* (Myrsinaceae) based on chloroplast *trnL-F* and nuclear ribosomal ITS sequences. *Molecular Phylogenetics and Evolution* **31**: 323–339.

Hershkovitz MA, Zimmer EA, Hahn WJ. 1999. Ribosomal DNA sequences and angiosperm systematics. In: Hollingsworth PM, Bateman RM, Gornall RJ, eds. *Molecular systematics and plant evolution*. London, UK: Taylor & Francis, 268–326.

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. 1990. Cytology and taxonomy of *Elymus kengii*, *E. grandiglumis*, *E. alatavicus* and *E. batalinii* (Triticeae: Poaceae). *Genome* **33**: 668–673.

Koch MA, Dobes C, Mitchell-Olds T. 2003. Multiple hybrid formation in natural populations: Concerted evolution of the internal transcribed spacer of nuclear ribosomal DNA (ITS) in North American *Arabis divaricarpa* (Brassicaceae). *Molecular Biology and Evolution* **20**: 338–350.

Kumar S, Tamura K, Nei M. 2004. MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics* **5**: 150–163.

Löve A. 1984. Conspectus of the Triticeae. *Feddes Repertorium* **95**: 425–521.

Lu BR. 1993. Genomic relationships within the *Elymus parviglumis* group (Triticeae: Poaceae). *Plant Systematics and Evolution* **187**: 191–211.

Lu BR. 1994. The genus *Elymus* L. in Asia. Taxonomy and biosystematics with special reference to genomic relationships. In: Wang RRC, Jensen KB, Jaussi C, eds. *Proceedings of the 2nd International Triticeae Symposium, Logan, UT, USA*. Logan, UT, USA: The Utah State University Press 219–233.

Lu BR, von Bothmer R. 1991. Cytogenetic studies of the intergeneric hybrids between *Secale cereale* and *Elymus caninus*, *Elymus brevipes*, and *E. tsukushiensis* (Triticeae: Poaceae). *Theoretical and Applied Genetics* **81**: 524–532.

Lu BR, Salomon B. 1992. Differentiation of the SY genomes in Asiatic. *Elymus. Hereditas* **116**: 121–126.

Lu BR, Salomon B, von Bothmer R. 1992. Cytogenetic studies of progeny from the intergeneric crosses *Elymus* × *Hordeum* and *Elymus* × *Secale*. *Genome* **33**: 425–432.

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

Mason-Gamer RJ. 2004. Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. *Systematic Biology* **53**: 25–37.

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.

Masterson J. 1994. Stomatal size in fossil plants: evidence for polyploidy in majority of angiosperms. *Science* **264**: 421–424.

McMillan E, Sun GL. 2004. Genetic relationships of tetraploid *Elymus* species and their genomic donor species inferred from polymerase chain reaction-restriction length polymorphism analysis of chloroplast gene regions. *Theoretical and Applied Genetics* **108**: 535–542.

Ogihara Y, Isono K, Kojima T, Endo A, Hanaoka M, Shiina T, Terachi T, Utsugi S, Murata M, Mori N, Takumi S, Ikeo K, Gojobori T, Murai R, Murai K, Matsuoka Y, Ohnishi Y, Tajiri H, Tsunewaki K. 2002. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. *Molecular Genetics and Genomics* **266**: 740–746.

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

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

Salomon B, Lu BR. 1992. Genomic groups, morphology, and sectional delimitation in Eurasian *Elymus* (Poaceae: Triticeae). *Plant Systematics and Evolution* **180**: 1–13.

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. Cold Spring. Harbor, NY, USA: Cold Spring Harbor Laboratory Press.

Sang T, Crawford DJ, Stuessy T. 1997. Chloroplast DNA phylogeny, reticulate evolution, and biogeography of *Paeonia* (Paeoniaceae). *American Journal of Botany* **84**: 1120–1136.

Soltis DE, Soltis PS. 1999. Polyploidy: recurrent formation and genome evolution. *Trends in Ecology and Evolution* **14**: 348–352.

Soltis DE, Soltis PS, Tate JA. 2003. Advances in the study of polyploidy since *Plant speciation. New Phytologist* **161**: 173–191.

Soltis PS, Soltis DE. 1993. Molecular data and the dynamic nature of polyploidy. *Critical Reviews in Plant Sciences* **12**: 243–273.

Soltis PS, Soltis DE. 2000. The role of genetics and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences, USA* **97**: 7051–7057.

Stebbins GL. 1971. *Chromosomal evolution in higher plants*. Edward Arnold, London.

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

Taberlet PL, Gielly GP, Bouvet J. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105–1109.

Thompson JD, Plewniak F, Poch O. 1999. A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Research* **27**: 2682–2690.

Torabinejad J, Mueller RJ. 1993. Genome constitution of the Australian hexaploid grass, *Elymus scabrus* (Poaceae: Triticeae). *Genome* **36**: 147–151.

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. *Systematic Biology* **43**: 250–261.

Wang RR-C, Dewey DR, Hisao C. 1985. Intergeneric hybrids of *Agropyron* and *Pseudoroegneria. Botanical Gazette* **146**: 268–274.

Wendel JF. 2000. Genome evolution in polyploids. *Plant Molecular Biology* **42**: 225–249.

Wendel JF, Schnabel A, Seelanan T. 1995. Bi-directional interlocus concerted evolution following allopolyploid speciation in *cotton* (Gossypium). *Proceedings of the National Academy of Sciences, USA* **92**: 280–284.

Xu DH, Ban T. 2004. Phylogenetic and evolutionary relationships between *Elymus humidus* and other *Elymus* species based on sequencing of non-coding regions of cpDNA and AFLP of nuclear DNA. *Theoretical and Applied Genetics* **108**: 1443–1448.

Yonemori K, Honsho C, Kanzakis 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 Systematics and Evolution* **231**: 59–75.

Zhang W, Qu LJ, Gu H, Gao W, Liu M, Chen J, Chen Z. 2002. Studies on the origin and evolution of tetraploid wheats based on the internal transcribed spacer (ITS) sequences of nuclear ribosomal DNA. *Theoretical and Applied Genetics* **104**: 1099–1106.

