Polyploidy, Introgression, and Complex Phylogenetic Patterns within *Elymus*

R. J. Mason-Gamer, M. M. Burns and M. Naum

Department of Biological Sciences, The University of Illinois at Chicago, 60607 Chicago, IL, U.S.A., e-mail: robie@uic.edu

Abstract: The reticulate phylogeny of the *Triticeae* is demonstrated by extensive conflict among diploid gene trees, and by the presence of numerous allopolyploid combinations. *Elymus* is a geographically widespread allopolyploid genus which includes multiple distinct genomic combinations. By definition, all *Elymus* species include at least one set of *Pseudoroegneria* genomes, but these may be combined with genomes from one or more of several other *Triticeae* genera. Because allopolyploids represent a confluence of divergent evolutionary lineages, molecular phylogenetic analyses of allopolyploid species often include nuclear gene sequences, which potentially represent all of the distinct lineages that coexist within an individual plant. Here, we present data from two nuclear genes, β-amylase and starch synthase, in an attempt to characterize differences between cytogenetically distinct **StStYY**, **StStHH**, and **StStStStHH** species, and between the North American and Eurasian tetraploid species. Patterns of diversity among species also shed light on possible introgression following polyploidization.

Keywords: *Triticeae*; *Elymus*; phylogeny; polyploid; β-amylase; starch synthase

In the genomic system of *Triticeae* classification (Dewey 1984; Löve 1984; Barkworth & Dewey 1985), *Elymus* includes about 150 allopolyploid species with at least one set of *Pseudoroegneria* (**St**) genomes (Löve 1984). Under this definition, *Elymus* is heterogeneous and polyphyletic, and exhibits three levels of genomic heterogeneity. First, each allopolyploid individual represents a confluence of two or more distinct genomes; cytogenetic studies have been used extensively to characterize such plants (e.g., Dewey 1968, 1969, 1970, 1974; Jensen 1989; Jensen & Hatch 1989; Jensen & Wang 1991; Lu 1993a; Lu & von Bothmer 1993; Salomon & Lu 1992, 1994a, b). Second, there are different combinations of genomes in *Elymus*; the **St** genome, present in all *Elymus* species by definition, can be found in combination with one or more of several other *Triticeae* genomes (e.g., Dewey 1968, 1974, 1984; Jensen & Hatch 1989; Lu *et al.* 1990; Lu & von Bothmer 1993; Salomon & Lu 1992, 1994a, b). Third, there is heterogeneity within genomic combinations (e.g., Dewey 1968, 1974; Jaaska 1998; Lu 1993a, b; Lu & von Bothmer 1993; Salomon & Lu 1992, 1994b; Svitashev *et al.* 1996), consistent with multiple origins, diversification following polyploidization, introgression into polyploid lineages, or a combination of these. We report on ongoing molecular phylogenetic analyses of *Elymus*, including North American and Eurasian **StStHH** tetraploids, **StStYY** tetraploids, and a putative **StStStStHH** hexaploid. Our goals are to illuminate relationships among these groups, and to assess the effects of introgression using comparisons among gene trees.

MATERIAL AND METHODS

The analysis of our current representatives of *Elymus* (Table 1) employed two phylogenetic markers. The first is a 1.5 kb portion of a β-amylase

gene, including two partial exons, two full exons, and three introns of approximate lengths of 150, 450, and 240 base pairs (bp) (Mason-Gamer 2005). The β-amylase genes form a small gene family in wheat, barley, and rye. The copy used here has been mapped to the *Triticeae* group 2 homoeologous chromosomes (Sharp *et al.* 1988), and is homologous to genes with "tissue-ubiquitous" expression, while other members of the family are expressed in endosperm (e.g. Shewry *et al.* 1988; Daussant & Lauriére 1990). The second marker is a 1.3 kb portion of the granule-bound starch synthase gene, including two partial exons, four full exons, and five ~100 bp introns (Mason-Gamer *et al.* 1998). This gene maps to the *Triticeae* group 7 homoeologous chromosomes or a portion of chromosome 4 translocated from, and thus homoeologous to, the group 7 chromosomes (Devos & Gale 1997; Korzun *et al.* 1997). Amplification, cloning, and sequencing methods follow Mason-Gamer (2005 – β-amylase) and Mason-Gamer (2001, 2004 – starch synthase).

Data analyses were carried out using PAUP* 4.0 (Swofford 2002). Tree searches were performed with maximum parsimony with characters equally weighted. Parsimony bootstrap support was estimated using 1000 fast stepwise-addition replicates. A broad sample of diploid Triticeae species was analyzed simultaneously with sequences from allopolyploid *Elymus*. A list of the diploid representatives included in each full analysis can be found in Mason-Gamer (2005; β-amylase), and Mason-Gamer (2004; starch synthase). Subtrees corresponding to the **St** genome (*Pseudoroegneria* + *Elymus*), and the **H** genome (*Hordeum* + *Elymus*) were examined and compared.

RESULTS AND DISCUSSION

From each of the full *Triticeae* gene trees (not shown), subtrees corresponding to the **St** genome (*Pseudoroegneria* and related sequences from *Elymus*) and the **H** genome (*Hordeum* and related sequences from *Elymus*) were examined (Figure 1). We ask whether: (1) different genome combinations originate from markedly different *Pseudoroegneria* ancestors; (2) differentiation within *Elymus* corresponds to geography and/or genome complement; and (3) introgression following polyploidization can be detected in comparisons among trees. These results are part of an in-progress analysis of *Elymus*; the complete results will be published at a later date. The present results must be viewed as preliminary: sampling, vouchering, and Genbank submission are incomplete, and we have employed simplistic phylogenetic analyses so far. In spite of these caveats, we have uncovered some intriguing phylogenetic patterns, which will be further investigated in future analyses.

In the **St** clade from the β-amylase gene tree (Figure 1A), all tetraploids are together in a monophyletic group with all of the *Pseudoroegneria* sequences. Within this clade, species are not distinguished by either geographic origin (North America vs. Eurasia) or by genome complement (**StStHH** vs. **StStYY**). Sequences from eight *E. repens* individuals form a monophyletic clade, entirely separate from the tetraploids.

In the β-amylase **H** clade (Figure 1B), as expected, the **StStYY** species are not represented (but they do yield sequences that form a distinct clade, presumed to represent the **Y** genome; not shown). The **StStHH** tetraploids form a monophyletic group along with one genome from the allotetraploid *H. jubatum*. Within this group, two of the three North American *Elymus* tetraploids form a clade, while the remaining North American and Eurasian species are only slightly differentiated from one another. As in the β-amylase **St** clade, the *E. repens* sequences are entirely separate from the tetraploids. In contrast to the **St** clade, however, the **H**-genome sequences from *E. repens* are polyphyletic, and furthermore, the maximum pairwise divergence within *E. repens* is twice the maximum seen among all eight **StStHH** tetraploid species (4.2% and 2.1%, respectively).

The **St** and **H** clades from the β-amylase tree both suggest that the tetraploid *Elymus* species, regardless of geographic origin, evolved from a fairly homogeneous gene pool, and that hexaploid *E. repens* is not an evolutionary derivative of an ancestor of any of these tetraploid species. The two clades differ, however, in the placement of *E. repens* sequences. The **St**-genome sequences are monophyletic and similar, while the **H**-genome sequences are polyphyletic and divergent. Thus, *E. repens* has drawn from a uniform pool of β-amylase **St**-genome sequences, and from a diverse and polyphyletic pool of **H**-genome sequences.

The **St** clade from the starch synthase gene tree has an unexpected phylogenetic pattern (Figure 1C). The *Pseudoroegneria* and related *Elymus* sequences fall into two distinct clades, but whether or not they are sister to one another remains unclear.

Table 1. *Elymus* species included in the analysis

Genome complement *Elymus* species	Accession Number	Collection Location	Sequences			
			β-amylase		GBSSI	
StStYY			**St**	**H**	**St**	**H**
E. abolinii	PI 531555	Northwest China	X	na	X	na
E. caucasicus	PI 531573	Estonia		na	X	na
E. ciliaris 1	PI 531575	Northeast China	X	na		na
E. ciliaris 2	PI 531577	Japan	X	na		na
E. ciliaris 5	PI 531576	Estonia	X	na		na
E. gmelinii	PI 499477	Northwest China	X	na	X	na
E. longearistatus	PI 401277	Iran		na	X	na
E. nevskii	PI 314620	Iraq	X	na	X	na
E. pendulinus	PI 499452	Northcentral China	X	na	X	na
StStHH – Eurasian						
E. brachyaristatus	PI 499411	Northwest China	X	X	X	X
E. caninus 1	PI 314205	Uzbekistan	X	X	X	X
E. caninus 2	PI 314612	Kazakhstan	X	X	X	X
E. caninus 4	PI 499413	Northwest China	X	X	X	X
E. caninus 5	PI 531571	Poland	X	X	X	X
E. dentatus 1	PI 628702	Southcentral Russia	X	X	X	
E. dentatus 2	PI 531599	Pakistan	X	X		X
E. mutabilis 1	PI 628704	Southcentral Russia	X	X	X	X
E. mutabilis 2	PI 499449	Northwest China	X	X	X	X
E. sibiricus 1	PI 628699	Southeast Russia		X	X	X
E. sibiricus 3	PI 499461	Northcentral China	X	X		X
StStHH – North America						
E. elymoides	PI 531606	Washington USA			X	X
E. glaucus 4	RJMG 130	Idaho USA			X	X
E. glaucus 6	W6 10215	Colorado USA			X	X
E. glaucus 7	PI 593652	Oregon USA				X
E. hystrix	MEB 97-87	Utah USA			X	
E. lanceolatus 1	W6 14220	Idaho USA	X	X	X	X
E. lanceolatus 2	W6 14218	Utah USA			X	X
E. riparius	RJMG 160	Connecticut USA				X
E. trachycaulus 1	PI 372500	Northwest Territory Canada	X	X	X	X
E. trachycaulus 3	PI 452446	Alberta Canada			X	X
E. virginicus 4	RJMG 161	Connecticut USA			X	X
E. virginicus 5	RJMG 162	Connecticut USA			X	
E. virginicus 9	RJMG 168	Maine USA			X	X
E. wawawaiensis 1	PI 285272	Washington USA	X	X	X	X
E. wawawaiensis 3	PI 598812	Oregon USA			X	X
StStStHH						
Elymus repens 1	RJMG 119	Idaho USA	X		X	X
Elymus repens 2	RJMG 123	Idaho USA	X	X	X	
Elymus repens 3	RJMG 131	Idaho USA	X	X	X	
Elymus repens 4	RJMG 159	Wisconsin USA	X	X		X
Elymus repens 5	RJMG 166	Maine USA	X	X		X
Elymus repens 6	RJMG 167	Maine USA			X	X
Elymus repens 8	PI 440065	Russian Federation	X			
Elymus repens 9	PI 317409	Afghanistan	X	X		
Elymus repens 10	PI 380623	Iran	X			

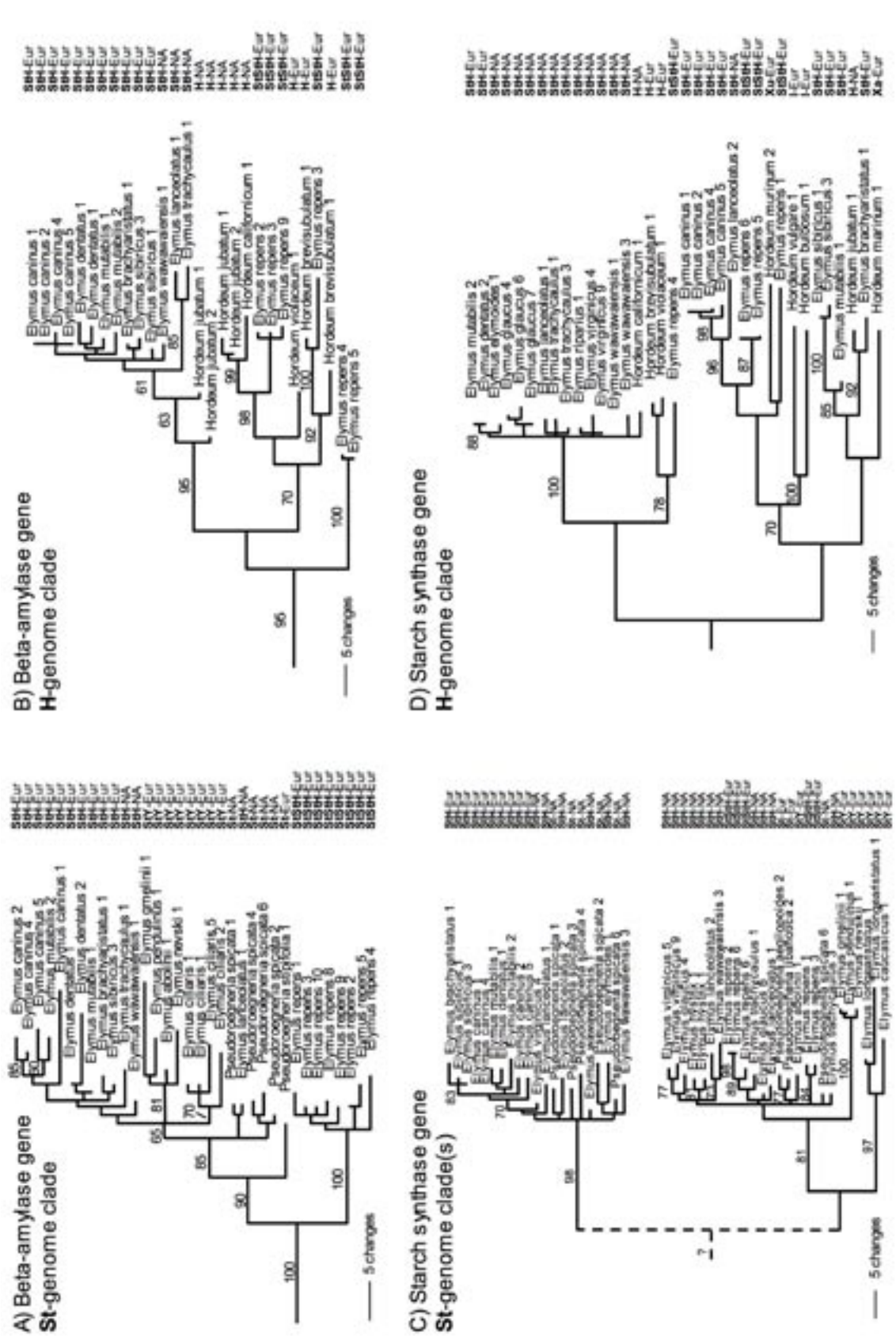

Figure 1. Subclades from an analysis of *Elymus* and a broad sample of the Triticeae. A and B – β-amylase subtrees. A – *Pseudoroegneria* and related *Elymus* sequences; B – *Hordeum* and related *Elymus* sequences. C and D – Starch synthase subtrees. C – *Pseudoroegneria* and related *Elymus* sequences; D – *Hordeum* and related *Elymus* sequences

In an analysis of the diploid Triticeae and North American *Elymus*, the two **St** clades were together on the maximum likelihood tree, but with very weak support, (Mason-Gamer 2001). In the present analysis of the entire tribe (not shown) the clades are separated, but their placement among the other diploids is very weakly supported. Several North American *Elymus* species, and some individuals, contain sequences from both clades. The split between the two clades appears to predate the origin of the North American species. This is consistent with several scenarios, including (1) a single origin of North American *Elymus* from a polymorphic *Pseudoroegneria* gene pool, and maintenance of the polymorphism within and among extant species; (2) multiple origins from a polymorphic *Pseudoroegneria* gene pool; or (3) introduction of an **St** variant into a North American *Pseudoroegneria*/*Elymus* gene pool that was at one time homogeneous. Both of the variants are found among the Eurasian species as well, but in a clearer pattern. All of the Eurasian **StStHH** species are in one clade, while the Asian **StStYY** sequences are in the other, suggesting that these two species groups acquired their starch synthase **St** sequences from distinct *Pseudoroegneria* gene pools. A survey of *Pseudoroegneria* from European and Asian *Pseudoroegneria* may help to clarify this pattern.

Within the starch synthase **H** clade (Figure 1D), as expected, there are no sequences from the **StStYY** taxa. As in the β-amylase tree, the **StStYY** tetraploids yield distinct, presumed **Y**-genome starch synthase sequences. The relationships among the **StStHH** species are not clear-cut, but the weak basal split within the clade is partially consistent with a North American - Eurasian distinction. Most of the North American individuals are extremely similar to one another and to the North American native *H. californicum*. Included within this group, however, are two Eurasian individuals. The remaining Eurasian species are distinct from this group and from one another. The polyphyly of the **H**-genome sequences from Eurasian *Elymus* is consistent with multiple origins of **StStHH** *Elymus* in Eurasia, involving different *Hordeum* genome donors. However, because the polyphyletic pattern sharply contrasts with the other clades (Figure 1A–C), on which the Eurasian **StStHH** tetraploids are very similar, it seems more likely that the group arose once (or multiple times from related donors), and subsequently acquired additional **H**-genome starch synthase sequences through introgression. The **H**-genome starch synthase sequences from the hexaploid species *E. repens* are polyphyletic (Figure 1D), and like the β-amylase **H**-genome sequences (Figure 1B), they indicate that *E. repens* has drawn from multiple *Hordeum* lineages. The polyphyly of the **H**-genome sequences is, again, consistent with multiple origins of *E. repens*, involving different *Hordeum* progenitors. However, the pattern is more likely indicative of introgression from *Hordeum*, because the unique and unexpected genomic complexity of *E. repens* (Mason-Gamer 2004) seems unlikely to have arisen multiple times.

When viewed together, the trees shed light on possible introgression into *Elymus*, particularly involving *Hordeum*. We are completing the sequencing of both genes for the taxa included here (Table 1), and expanding our sample of *Elymus*, *Pseudoroegneria*, and *Hordeum*. We hope to fill in some of the geographic and taxonomic details regarding the ongoing evolutionary interactions between *Elymus* and its diploid progenitors.

Acknowledgment: Sincere thanks go to the Organizing and Scientific Committees of the 5th International Triticeae Symposium. The research was supported by the National Science Foundation (DEB-9974181 and DEB-0426194).

REFERENCES

Barkworth M.E., Dewey D.R. (1985): Genomically based genera in the perennial Triticeae of North America: identification and membership. American Journal of Botany, **72**: 767–776.

Daussant J., Lauriére C. (1990): Detection and partial characterization of two antigenically distinct b-amylases in developing kernels of wheat. Planta, **181**: 505–511.

Devos K.M., Gale M.D. (1997): Comparative genetics in the grasses. Plant Molecular Biology, **35**: 3–15.

Dewey D.R. (1968): Synthetic *Agropyron-Elymus* hybrids: III. *Elymus canadensis* × *Agropyron caninum*, *A. trachycaulum*, and *A. striatum*. American Journal of Botany, **55**: 1133–1139.

Dewey D.R. (1969): Synthetic hybrids of *Agropyron caespitosum* × *Agropyron spicatum*, *Agropyron caninum*, and *Agropyron yezoense*. Botanical Gazette, **130**: 110–116.

Dewey D.R. (1970): Genome relations among *Elymus canadensis*, *Elymus triticoides*, *Elymus dasystachyus*, and *Agropyron smithii*. American Journal of Botany, **57**: 861–866.

Dewey D.R. (1974): Cytogenetics of *Elymus sibiricus* and its hybrids with *Agropyron tauri*, *Elymus canadensis*, and *Agropyron caninum*. Botanical Gazette, **135**: 80–87.

Dewey D.R. (1984): The genomic system of classification as a guide to intergeneric hybridization within the perennial Triticeae. In: Gustafson J.P. (ed.): Gene Manipulation in Plant Improvement. Proc. 6th Stadler Genetics Symp. Columbia University Press, New York: 209–279.

Jaaska V. (1998): Isoenzyme data on the origin of North American allotetraploid *Elymus* species. In: Jaradat A.A. (ed.): Triticeae III. Science Publishers, Enfield, New Hampshire: 209–216.

Jensen K.B. (1989): Cytology, fertility, and origin of *Elymus abolinii* (Drob.) Tzvelev and its hybrids with *Pseudoroegneria spicata*, *E. lanceolatus*, *E. dentatus* ssp. *ugamicus*, and *E. drobovii*. Genome, **32**: 468–474.

Jensen K.B., Hatch S.L. (1989): Genome analysis, morphology, and taxonomy of *Elymus gmelinii* and *E. strictus* (Poaceae: Triticeae). Botanical Gazette, **150**: 84–92.

Jensen K.B., Wang R.R.-C. (1991): Cytogenetics of *Elymus caucasicus* and *Elymus longearistatus* (Poaceae, Triticeae). Genome, **34**: 860–867.

Kleinhofs A. (1997): Integrating barley RFLP and classical marker maps. Barley Genetics Newsletter, **27**: 105–112.

Korzun V., Malyshev S., Voymokov A., Börner A. (1997): RFLP-mapping of three mutant loci in rye (*Secale cereale* L.) and their relation to homoeologous loci within the Gramineae. Theoretical and Applied Genetics, **95**: 468–473.

Löve Á. (1984): Conspectus of the Triticeae. Feddes Repertorium, **95**: 425–521.

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

Lu B.-R. (1993b): Meiotic studies of *Elymus nutans* and *E. jacquemontii* (Poaceae, Triticeae) and their hybrids with *Pseudoroegneria spicata* and seventeen *Elymus* species. Plant Systematics and Evolution, **186**: 193–212.

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

Lu B.-R., von Bothmer R. (1993): Meiotic analysis of *Elymus caucasicus*, *E. longearistatus*, and their interspecific hybrids with twenty-three *Elymus* species (Triticeae, Poaceae). Plant Systematics and Evolution, **185**: 35–53.

Lu B.-R., Yan J., Yang J.-L., Flink J. (1990): Biosystematic studies among *Roegneria pendulina* Nevski, *R. ciliaris* (Trin.) Nevski, and *R. kamoji* Owhi in the tribe Triticeae, Gramineae. Acta Botanica Yunnanica, **12**: 161–171.

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

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

Mason-Gamer R.J. (2005): The β-amylase genes of grasses and a phylogenetic analysis of the Triticeae (Poaceae). American Journal of Botany, **92**: 1045–1058.

Mason-Gamer R.J., Weil C.F., Kellogg E.A. (1998): Granule-bound starch synthase: structure, function, and phylogenetic utility. Molecular Biology and Evolution, **15**: 1658–1673.

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

Salomon B., Lu B.-R. (1994a): Genomic relationships between species of the *Elymus semicostatus* group and *Elymus* sensu lato (Poaceae). Plant Systematics and Evolution, **191**: 199–211.

Salomon B., Lu B.-R. (1994b): Interspecific hybridizations among species of the *Elymus semicostatus* and *Elymus tibeticus* groups (Poaceae). Plant Systematics and Evolution, **189**: 1–13.

Sharp P.J., Kreis M., Shewry P.R., Gale M.D. (1988): Location of b-amylase sequences in wheat and its relatives. Theoretical and Applied Genetics, **75**: 286–290.

Shewry P.R., Parnar S., Buxton B., Gale M.D., Liu C.J., Hejgaard J., Kreis M. (1988): Multiple molecular forms of b-amylase in seeds and vegetative tissues of barley. Planta, **176**: 127–134.

Svitashev S., Salomon B., Bryngelsson T., von Bothmer R. (1996): A study of 28 *Elymus* species using repetitive DNA sequences. Genome, **39**: 1093–1101.

Swofford D.L. (2002): PAUP* v. 4.0. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland, Massachusetts.