Syst. Biol. 53(1):25–37, 2004

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150490424402

Reticulate Evolution, Introgression, and Intertribal Gene Capture in an Allohexaploid Grass

ROBERTA J. MASON-GAMER

University of Illinois at Chicago, Department of Biological Sciences, MC 066, 845 W. Taylor Street, Chicago, Illinois 60607, USA; E-mail: robie@uic.edu

Abstract.— Recent molecular phylogenetic studies of polyploid plants have successfully clarified complex patterns of reticulate evolution. In this study of *Elymus repens*, an allohexaploid member of the wheat tribe Triticeae, chloroplast and nuclear DNA data reveal an extreme reticulate pattern, revealing at least five distinct gene lineages coexisting within the species, acquired through a possible combination of allohexaploidy and introgression from both within and beyond the Triticeae. Earlier cytogenetic studies of *E. repens* suggested that *Hordeum* (genome **H**) and *Pseudoroegneria* (**St**) were genome donors to *E. repens*. Chloroplast DNA data presented here (from the *rpoA* gene and from the region between *trnT* and *trnF*) identify three potential maternal genome donors (*Pseudoroegneria*, *Thinopyrum*, and *Dasypyrum*), and information from previous molecular work suggests that, of these, *Pseudoroegneria* is the most likely maternal donor. Nuclear starch synthase gene data indicate that both *Hordeum* and *Pseudoroegneria* have contributed to the nuclear genome of *E. repens*, in agreement with cytogenetic data. However, these data also show unexpected contributions from *Taeniatherum*, and from two additional donors of unknown identity. One of the sequences of unknown origin falls within the Triticeae, but is not closely associated with any of the sampled diploid genera. The second falls outside of the clade containing Triticeae and its outgroup *Bromus*, suggesting the acquisition of genetic material from a surprisingly divergent source. Bias toward the amplification of certain starch synthase variants has complicated attempts to thoroughly sample from within individuals, but the data clearly indicate a complex pattern of reticulate evolution, consistent not only with allohexaploidy, but also with introgression from unexpectedly divergent sources. [Chloroplast DNA; *Elymus repens*; GBSSI; phylogeny; Poaceae; polyploidy; reticulate evolution; Triticeae.]

Morphological, cytogenetic, and allozyme data have been widely used to clarify the origins and evolutionary history of polyploid plant taxa. Building upon these foundations, phylogenetic analyses of DNA sequence and restriction site data have led to an even more detailed understanding of these complex lineages. Molecular phylogenetic studies have successfully revealed the phylogenetic affinities of several polyploid species, with some providing support for existing hypotheses (e.g., Ge et al., 1999; Mason-Gamer, 2001), while others have uncovered unexpected relationships among polyploids and diploids (e.g., Ainouche et al., 1999; Ballard and Sytsma, 2000; Emshwiller and Doyle, 2002; Smedmark et al., 2003), including especially complex relationships involving multiple unexpected sources (e.g., Cronn et al., 2003). Still other studies have demonstrated fascinating phenomena such as multiple origins of allopolyploid genome combinations (Wendel et al., 1991; Soltis et al., 1995; Cook et al., 1998; Doyle et al., 1999, 2002), hybridization among polyploids (Doyle et al., 1999, 2002; Ferguson and Sang, 2001), chloroplast DNA (cpDNA) introgression into polyploid lineages (Wendel, 1989), and gene silencing following polyploidization (Ford and Gottlieb, 2002; Adams et al., 2003; Osborn et al., 2003). In this study, molecular phylogenetic analyses of chloroplast and nuclear DNA sequence data are used to investigate the origin and evolution of *Elymus repens* (L.) Gould, an allohexaploid member of the wheat tribe, Triticeae.

The wheat tribe is best known for the economic importance of its most familiar members, wheat, barley, and rye. The group is of additional interest from an evolutionary perspective, however, because of its complex reticulate history. Molecular phylogenetic studies of the tribe (Hsiao et al., 1995; Kellogg and Appels, 1995; Petersen and Seberg, 1997, 2002; Mason-Gamer and Kellogg, 2000; Mason-Gamer, 2001; Mason-Gamer et al., 2002) yield incongruent chloroplast DNA and nuclear gene trees, suggesting extensively reticulate relationships among the diploid genera (Kellogg et al., 1996; Mason-Gamer and Kellogg, 1996). Furthermore, auto- and allopolyploidy, involving many different genome combinations, account for about 75% of Triticeae species with known chromosome number (Löve, 1984).

One complex group of polyploids within the Triticeae is classified as *Elymus*, based on cytogenetic data. Under the genomic definition, all members of *Elymus* are allopolyploids that include at least one set of genomes from *Pseudoroegneria* (genome designation **St**). The **St** genome can be found in numerous tetraploid, hexaploid, and octoploid combinations with a variety of other genomes from within the tribe, including **Y** (from an unknown donor), **H** (*Hordeum*), **P** (*Agropyron*), and **W** (*Australopyrum*). While many of these cytogenetically defined entities have yet to be examined in a molecular phylogenetic context, recent studies of presumed **StStHH** allotetraploid species native to North America reveal *Pseudoroegneria* and *Hordeum* to be their closest diploid relatives (Mason-Gamer, 2001; Mason-Gamer et al., 2002), thus confirming hypotheses based on cytogenetic data.

Elymus repens, the focus of this article, is a morphologically variable, allohexaploid native of Europe and Asia. It was introduced to the United States at least a century ago and has since become a problematic weed, especially in the northern part of the country (Batcher, 2002). It is listed as a noxious weed in 11 states, and is a restricted or prohibited seed contaminant in 49 states and the District of Columbia (USDA and NRCS, 2002). Cytogenetic studies of *E. repens* have revealed it to be hexaploid ($2n = 42$), with two sets of similar genomes

and a third set distinct from the other two (Cauderon, 1958; Dewey, 1961). At meiosis, *E. repens* chromosomes normally form 21 bivalents, although quadrivalents or other irregularities are occasionally seen (Cauderon, 1958; Dewey, 1961). Dewey (1970, 1976) examined *E. repens* X *Pseudoroegneria spicata* and *E. repens* X *P. stipifolia* hybrids and concluded that *Pseudoroegneria* (genome **St**) is the likely donor of the two similar genomes of *E. repens*, but he remained uncertain about the origin of the third. Cauderon and Saigne (1961) examined patterns of genome pairing between *E. repens* and *Hordeum secalinum* and concluded that the one distinct chromosome set of *E. repens* was derived from *Hordeum* (genome **H**), but Dewey (1984) doubted this possibility based on morphological grounds. Assadi and Runemark (1995) examined hybrids between *E. repens* and the presumed **StStStStHH** allohexaploid *E. transhyrcanus*, and concluded that *E. repens* shared the **StStStStHH** genomic constitution. Most recently, genomic *in situ* hybridization (GISH) of *Hordeum californicum* probes to the *E. repens* genome confirmed the presence of the **H** genome (Ørgaard and Anamthawat-Jónsson, 2001), supporting Cauderon and Saigne (1961), but *Pseudoroegneria spicata* probes hybridized poorly, placing *Pseudoroegneria*'s role as a genome donor in doubt. While the existing studies are not in complete agreement, the genomic complement of *E. repens* is often assumed to be **StStStStHH**.

In this study, chloroplast and nuclear DNA sequence data were used to investigate the phylogenetic relationships between *E. repens* and its potential diploid progenitors in light of previous hypotheses based on cytogenetic and GISH data. The purposes of the study were as follows: (1) To identify the maternal parent to *E. repens* using cpDNA data. To achieve this, sequences of two cpDNA loci were obtained from *E. repens* and analyzed within a broad sample of diploid genera from throughout the tribe. The expectation, based on the cytogenetic studies discussed above, was that the chloroplast genome of *E. repens* would be most closely related to that from *Hordeum* or *Pseudoroegneria*. (2) To identify the combination of taxa that contributed to the nuclear genome of *E. repens*. For this goal, multiple cloned sequences of the single-copy granule-bound starch synthase gene (GBSSI or *waxy*) from *E. repens* were analyzed within a data set representing most of the diploid Triticeae genera. The preliminary expectation, based on the existing cytogenetic and GISH data, was that *E. repens* has an **StStStStHH** genomic complement. Thus, each individual was expected to have three homoeologous sets of GBSSI alleles, with two sets closely related to *Pseudoroegneria* (and possibly phylogenetically indistinguishable from one another), and the third closely related to *Hordeum*. Together, the cpDNA and GBSSI data sets have revealed the history of *E. repens* to be much more complex than had been suspected based on cytogenetic data, involving the combination of at least five distinct gene lineages. The results should prompt a reassessment of the ability of the members of this group, and possibly plants in general, to acquire genetic material from divergent sources.

MATERIALS AND METHODS

Sampling

Six *E. repens* individuals were sampled; all are introduced accessions from the northern United States (Table 1). *Elymus repens* sequences were analyzed along with samples representing nearly all of the diploid genera in the Triticeae. Sample information and Genbank accession data for most of the diploids have been presented elsewhere (GBSSI: Mason-Gamer, 2001; cpDNA: Mason-Gamer et al., 2002). New GBSSI sequences from six additional diploid individuals are from *Cutandia memphitica* clone a (J. I Davis, unpubl., Genbank accession AY362758), *Bromus tectorum* (plant accession Kellogg s.n., AY362757), *Pseudoroegneria strigosa* ssp. *aegilopoides* clone 2c (PI531755, AY360823), *P. libanotica* clone 2p (PI380644, AY360824), and *Taeniatherum caput-medusae* clone 2b (RJMG189, AY360847) and clone 2e (RJMG189, AY360848). An analysis of a divergent group of *E. repens* sequence variants was carried out

TABLE 1. *Elymus repens* individuals included in this study.

Individual	Location	rpoA	trnT–trnL	Starch synthase	
1-RJMG119	Bonner Co. ID	AY362780	AY362786	a1	AY360825
				c	AY360826
				e	AY360827
				f[a]	AY360849
				g[a]	AY360858
				q	AY360828
2-RJMG123	Boundary Co. ID	AY362781	AY362787	aa	AY360830
				bb	AY360831
				cc	AY360832
				g[a]	AY360850
				h[a]	AY360859
				hh	AY360829
3-RJMG131	Payette Co. ID	AY362782	AY362788	a	AY360833
				aa	AY360834
				dd	AY360835
				f[a]	AY360851
				g[a]	AY360860
				h[a]	AY360852
4-RJMG159	Dane Co. WI	AY362783	AY362789	a	AY360836
				aa	AY360838
				b[a]	AY360853
				dd	AY360839
				i[a]	AY360854
				t	AY360837
5-RJMG166	Washington Co. ME	AY362784	AY362790	b[a]	AY360861
				cc	AY360841
				dd	AY360842
				hh	AY360843
				j[a]	AY360855
				k	AY360840
6-RJMG167	Washington Co. ME	AY362785	AY362791	a[a]	AY360856
				a2	AY360844
				dd	AY360845
				g[a]	AY360857
				hh	AY360846

[a]Shorter fragments derived from the F-for/K-bac primer combination.

within a taxonomically broader sample of GBSSI sequences from within the Poaceae (selected from Mason-Gamer et al., 1998; GPWG, 2001; Mason-Gamer, 2001; Mathews et al., 2002; J. I Davis, unpubl.). The broader sample included representatives of subfamilies Pooideae (which includes the Triticeae), Bambusoideae, Danthonioideae, Chloridoideae, Panicoideae, Pharoideae, and Anomochlooideae. Alignments were deposited with EMBL under accession numbers ALIGN_000623 (chloroplast *rpoA* gene), ALIGN_000628 (chloroplast tRNA genes and spacers), ALIGN_000622 (Triticeae GBSSI including introns), and ALIGN_000619 (Poaceae GBSSI exons only).

Molecular Methods

Chloroplast DNA.—Two cpDNA loci were sequenced from *E. repens*: (1) the gene for the α-subunit of RNA polymerase (*rpoA*) and its short flanking spacers (Petersen and Seberg , 1997) and (2) the region between the tRNA genes *trnT* and *trnF*, which includes the *trnL* gene and intron and the *trnT–trnL* and *trnL–trnF* intergenic spacers (Taberlet et al., 1991). Amplification and sequencing of the chloroplast *rpoA* gene were carried out as in Mason-Gamer et al. (2002). Briefly, the *rpoA* gene and its flanking regions were amplified in 50-μl reactions with primers rpoA1, located at the 3′ end of the *rps11* gene, and rpoA2, located at the 5′ end of the *petD* gene (Petersen and Seberg, 1997; Table 2). The resulting 1.45 kb products were cleaned using 2.5 units of exonuclease I, 1 unit of shrimp alkaline phosphatase (SAP), in a 1X concentration of SAP reaction buffer. Cleaned products were directly sequenced using the rpoA1, rpoA2, rpoA4, rpoA5, rpoA8, and rpoA9 primers (Petersen and Seberg, 1997; Table 2). The tRNA genes and spacers were amplified as in Mason-Gamer et al. (2002) using primers a and f (Taberlet et al., 1991; Table 2). Following cleaning as above, the resulting 1.7 kb products were sequenced using the a, b, c, d, e, and f primers (Taberlet et al., 1991; Table 2). Both cpDNA products were sequenced using the ABI BigDye cycle sequencing kit, following enclosed instructions except that reagents were reduced in amount to one quarter or one eighth, in a total reaction volume of 10 μl. Template amounts were estimated by eye at 10–20 ng per reaction. Sequencing reactions were run on an ABI 377 sequencer or an MJ Research Base Station according to instructions. Sequences from each clone were edited and joined in Sequencher version 3.0 (Genecodes Corporation, Ann Arbor, MI).

TABLE 2. Sequences of primers mentioned in the text. Primer names follow the original publications.

Region	Primers	Sequence
GBSSI[a]	F-for	TGCGAGCTCGACAACATCATGCG
	H-for	AAGGCGCTGAACAAGGAGGCG
	J-bac	ACGTCGGGGCCCTTCTGCTC
	K-bac	GCAGGGCTCGAAGCGGCTGG
	L1-for	GCAAGACCGGGTTCCACATGG
	L2-bac	CGCTGAGGCGGCCCATGTGG
	M-bac	GGCGAGCGGCGCGATCCCTCGCC
*rpo*A[b]	rpoA1	ACACCTATGCCGCATAATGG
	rpoA2	GTTAGGTATTGGAGCAACAT
	rpoA4	AGATCTTCTTCAGTCTTACT
	rpoA5	GAAGTTATCCTATAGATGCT
	rpoA8	TCAAGTTAGCTATAGGTTGT
	rpoA9	GGACACTCTAGAGAAGCATC
tRNA genes/ spacers[c]	a	CATTACAAATGCGATGCTCT
	b	TCTACCGATTTCGCCATATC
	c	CGAAATCGGTAGACGCTACG
	d	GGGGATAGAGGGACTTGAAC
	e	GGTTCAAGTCCCTCTATCCC
	f	ATTTGAACTGGTGACACGAG

[a] Mason-Gamer et al., 1998.
[b] Petersen and Seberg, 1997.
[c] Taberlet et al., 1991.

Starch synthase gene.—Amplification reactions were carried out as in Mason-Gamer et al. (1998) in 10-μl reactions using primers F-for and M-bac (Table 2). These yield a product of about 1250 bp, including part of exon 9, introns 9–13, exons 10–13, and part of exon 14. The resulting products were checked on agarose minigels and cleaned with Geneclean (Bio101, La Jolla, CA) according to manufacturer's instructions. Cleaned PCR products were cloned into pGEM-T Easy vectors (Promega, Madison, WI). Ligation reactions, transformations, and plating were carried out according to the instructions included with the Promega pGEM-T Easy cloning kit, except that the final volumes of the ligation and transformation reactions were halved.

Because *E. repens* is an allohexaploid, each PCR reaction from a single individual amplifies a variable mixture of starch synthase gene sequences. Thus, preferential amplification of one sequence variant, whether by chance or because of differences in amplification or cloning efficiency, can lead to an over-representation of that variant in the final reaction mixture. These processes have been referred to as PCR drift and PCR selection, respectively (Wagner et al., 1994). In an attempt to offset the random effects of PCR drift, three PCR reactions from each *E. repens* individual were combined for use in ligation reactions. Because PCR with the F-for and M-bac primers (or the subsequent cloning of these products) resulted in a marked bias in the frequency of one of the starch synthase variants in all six individuals (see "Results" section), additional amplifications were done as above using a reverse primer specifically chosen to exclude the over-represented gene variant. The primer, K-bac, sits on a deletion that is shared only by the over-represented *E. repens* variant and its closest diploid relative. The F-for/K-bac PCR products, about 600 bp in length, include part of exon 9, exon 10, introns 9 and 10, and part of exon 11. The F-for/K-bac products were cleaned and cloned as above.

Cloned products were prepared for sequencing either by plasmid purification using either Qiagen or Promega Wizard miniprep kits, or by PCR amplification of inserts directly from colonies using the original amplification primers. Products of PCR amplification were cleaned with 2.5 units of exonuclease I and 1 unit of SAP, in 1X SAP buffer. From each individual, sequences were obtained from between 11 and 27 F-for/M-bac clones and between 7 and 11 F-for/K-bac clones. Sequencing

reactions and sequence editing were carried out as described for the cpDNA markers.

Phylogenetic Analyses

General approach.—Three data sets were analyzed: the combined cpDNA data set, the GBSSI intron and exon sequences from *E. repens* and the Triticeae, and the GBSSI exons only from a subsample of *E. repens* and the Triticeae within a broader sample of grasses. The same general approach, described here, was used for all three analyses; differences in details specific to each data set are presented in later paragraphs. Phylogenetic analyses were performed using PAUP* 4.0b10 (Swofford, 2002). All data sets were initially analyzed using cladistic maximum parsimony (MP) heuristic searches with all characters equally weighted. Prior to further analyses using maximum likelihood (ML), nucleotide frequencies at informative sites were examined for deviation from stationarity for all three data sets. Ten of the most parsimonious trees were chosen at random for the evaluation of 16 maximum-likelihood models of sequence evolution (e.g., Swofford et al., 1996; Frati et al., 1997; Sullivan et al., 1997). Substitution models included (1) Jukes–Cantor (JC; Jukes and Cantor, 1969), (2) Kimura two-parameter (K2P; Kimura, 1980), (3) Hasagawa–Kishino–Yano (HKY; Hasagawa et al., 1985), and (4) general time-reversible (GTR; Yang, 1994a). Under each substitution model, the following models of among-site rate heterogeneity were examined: no variation among sites; some proportion of sites assumed to be invariable (I; Hasagawa et al., 1985); rate variation among sites allowed to follow a gamma distribution (Γ; Yang, 1994b); and some invariable sites, with rate variation among the remaining sites following a gamma distribution (I+Γ; Gu et al., 1995; Waddell and Penny, 1996). For each data set, likelihood scores and model parameters were estimated on all ten saved MP trees, and the scores from the best tree were used to evaluate the models. The model with the best resulting score was compared to the two next-best models using a likelihood ratio test (Yang et al., 1995). For all three data sets, the score for the GTR+I+Γ model was significantly better ($P < 0.01$), following a Bonferroni correction for two non-independent comparisons, so the GTR+I+Γ model was used for subsequent ML analyses. Heuristic ML analyses were run under starting model parameters estimated on the best of the ten MP trees. Parameters were readjusted on the resulting ML tree, and used as input parameters in a new search. Searches were stopped when the ML tree score was identical to that from the previous search; this required between two (cpDNA and GBSSI exons) and three (GBSSI with introns) ML searches.

Because of the length of time required to estimate ML bootstrap support for these data under GTR+I+Γ with full heuristic searches, nodal support was instead estimated under a minimum-evolution (ME) criterion, with distances estimated using ML under the GTR+I+Γ model, using parameters of sequence evolution estimated from the final ML tree. Support values were based on full heuristic ME searches on 1,000 bootstrap replicates. Starting trees were obtained by stepwise addition with 10 trees held at each step.

Chloroplast DNA data.—Chloroplast DNA sequences from *E. repens* were manually added to an existing data set that includes a broad representation of the monogenomic Triticeae (Mason-Gamer et al., 2002); *E. repens* sequences are listed in Table 1. Gaps were treated as missing data. Nucleotide frequencies at informative sites did not significantly differ among taxa. Phylogenetic analyses were carried out as described above; two sequential ML searches were required to obtain the final ML tree ($-\ln L = 6744.869$). The optimized sequence parameters on the tree were used for ML estimation of ME distances in the bootstrap analysis: nucleotide frequencies A = 0.36445, C = 0.14414, G = 0.16535, and T = 0.32605; substitution frequencies AC = 0.90158, AG = 1.03432, AT = 0.60729, CG = 0.54794, CT = 1.58101, and GT = 1.00000; proportion of invariant sites = 0.709; and gamma distribution shape parameter = 0.924.

Starch synthase gene data from the Triticeae.—GBSSI sequences (with introns included) were manually added to an existing aligned data set representing most of the diploid genera in the Triticeae (Mason-Gamer, 2001); *E. repens* sequences are listed in Table 1. Fragments from the F-for/K-bac amplifications (superscript labels in Table 1) were included when they represented a variant that was not recovered from the same individual in the F-for/M-bac reactions. Within the introns, there were numerous insertions and deletions; these were treated as missing data. While alignment was generally straightforward in spite of length variation, there were some regions where patterns of length variation were complex enough to make homology assessment difficult. Regions judged to be ambiguous in terms of alignment were excluded from phylogenetic analyses, and include aligned positions 906–1016, 1133–1219, and 1369–1504. Nucleotide frequencies at the remaining informative sites did not differ significantly among the sequences.

Because of the potential for PCR-mediated recombination during amplification from a mixed population of sequences (Bradley and Hillis, 1997; Judo et al., 1998; Cronn et al., 2002; Doyle et al., 2002), several steps were taken in order to identify potential recombinant sequences, after which they were eliminated from the matrices used in the phylogenetic analyses. First, *E. repens* sequences with unresolved positions in the MP 50% majority rule bootstrap consensus trees were considered to be potential recombinants. Second, MP bootstrap analyses were run with the potential recombinants deleted singly, in pairs, and in groups in order to assess the affects of their removal on support for clades containing parental diploids and apparently non-recombinant *E. repens* sequences. When removal of a sequence resulted in a marked increase in bootstrap support for a clade, the sequence was hypothesized to be a recombinant. These possible recombinants were divided into halves and thirds, and each segment was run as a separate sequence in order to detect different phylogenetic histories in different segments of the same sequence. Finally, the putative recombinants were defined in terms of synapomorphies of the clades

of parental diploids, and inspected for combinations of synapomorphies of two or more clades. Eight sequences determined to be recombinants following these steps were removed from the analyses (results not shown).

Phylogenetic analyses were run as described earlier; the final tree ($-\ln L = 7537.304$) was obtained after three sequential ML searches. The optimized sequence parameters on this tree were used for ML estimation of ME distances in the bootstrap analysis: nucleotide frequencies A = 0.23452, C = 0.27664, G = 0.30388, and T = 0.18496; substitution frequencies AC = 1.15402, AG = 3.29127, AT = 0.73103, CG = 1.41517, CT = 4.68789, and GT = 1.00000; proportion of invariant sites = 0.331; and gamma distribution shape parameter = 0.727.

Starch synthase exon data from the Poaceae.—The Poaceae GBSSI exon sequences were analyzed to provide a broader phylogenetic context for a divergent F-for/K-bac sequence type found in four of the six *E. repens* individuals. The data set was analyzed as above; two ML searches were required to obtain the final tree ($-\ln L = 5586.823$). The optimized sequence parameters on the tree were used for ML estimation of ME distances in the bootstrap analysis: nucleotide frequencies A = 0.21861, C = 0.29454, G = 0.32070, and T = 0.16615; substitution frequencies AC = 1.06754, AG = 2.61243, AT = 0.63955, CG = 1.69737, CT = 4.01614, and GT = 1.00000; proportion of invariant sites = 0.338; and gamma distribution shape parameter = 0.848. In this data set, there was significant non-stationarity in nucleotide frequencies at informative sites, which could be attributed to the two basal-most members of the group, *Pharus* and *Anomochloa*. Upon removal of these sequences, nucleotide frequencies did not differ significantly, and ML trees generated as above and rooted with the Panicoids yielded an identical topology (with comparable support values) to the tree presented here, on which *Pharus* and *Anomochloa* are included.

RESULTS

Chloroplast DNA Data

The sequences of the *rpoA* gene and the tRNA genes and spacers from the diploid Triticeae, and the results of the corresponding phylogenetic analyses, were described in detail elsewhere (Mason-Gamer et al., 2002). The *E. repens rpoA* and tRNA gene/spacer products were 1354 and 1680 bp, respectively, and were easy to align with those from the diploids. There were very few differences among the *E. repens* cpDNA sequences; all were in the tRNA gene/spacer data set. Individual 3 had a single-base deletion in a poly-T region (T_{10} vs. T_{11}) and individual 4 had a single-base deletion in a poly-C region (C_7 vs. C_8). These differences do not affect the phylogenetic analysis (gaps were counted as missing data), and individuals 2, 3, 4, 5, and 6 were otherwise identical in both data sets. These five individuals were therefore represented by a single sequence in the phylogenetic analyses and on the resulting trees. Individual 1 differed from the others by two substitutions, and by a single-base deletion in a poly-A region (A_{11} vs. A_{12}). The ML tree (Fig. 1) places all of the *E. repens* individuals in a clade that includes *Pseudoroegneria*, *Thinopyrum*, and *Dasypyrum*, suggesting that the maternal genome donor to *E. repens* is from within one of these three genera.

Starch Synthase Data

Amplifications with the GBSSI F-for and M-bac primers yielded products of the expected length of about 1250 bp. Reactions produced a heterogeneous mix of GBSSI sequences from *E. repens*, as is expected in the case of an allopolyploid. However, the F-for and M-bac amplifications, or the subsequent cloning of the products, were sharply biased toward gene copies most similar to those from the diploid genus *Taeniatherum*. Out of 107 cloned F-for/M-bac products screened from *E. repens*, 93 were *Taeniatherum*-like (Table 3). Later amplifications using the F-for and K-bac primer combination nearly eliminated the *Taeniatherum*-like copies. The F-for/K-bac fragment mixture was enriched for the other variants of the gene, along with an additional variant not seen in the F-for/M-bac amplifications (labeled "Unknown 2" in Table 3). While the F-for/K-bac primer pair successfully excluded *Taeniatherum*-like sequences, the combination may itself give biased results. From within the four plants that have the "Unknown 2" sequence (individuals 1, 2, 3, and 5), 30 F-for/K-bac clones were of that type, while 31 clones comprised all of the other types combined (Table 3). The "Unknown 2" sequences were relatively highly diverged from the rest of the copies, and their intron sites were difficult to align with those from the rest of the Triticeae. The first starch synthase tree presented below (Fig. 2), based on exons and introns, excludes these sequences. The second tree (Fig. 3), based only on exons,

TABLE 3. Number of clones representing the different haplotypes found in *E. repens*. Clones from the F-for/M-bac primer combination are listed in plain text first; those from the F-for/K-bac combination are in boldface.

Plant	*Hordeum* FM—**FK**	*Pseudoroegneria* FM—**FK**	*Taeniatherum* FM—**FK**	Unknown 1 FM—**FK**	Unknown 2 FM—**FK**	Total of each clone FM—**FK**	Total per plant FM + **FK**
1	1—**3**	2—**3**	11—**0**	1—**0**	0—**5**	15—**11**	26
2	0—**0**	0—**1**	15—**1**	3—**0**	0—**8**	18—**10**	28
3	0—**0**	0—**1**	11—**0**	0—**1**	0—**9**	11—**11**	22
4	1—**1**	0—**0**	15—**0**	2—**6**	0—**0**	18—**7**	25
5	0—**1**	0—**0**	16—**0**	2—**2**	0—**8**	18—**11**	29
6	1—**1**	0—**5**	25—**0**	1—**5**	0—**0**	27—**11**	38
Total	3—**6**	2—**10**	93—**1**	9—**14**	0—**30**	107—**61**	168
Per clade	9	12	94	23	30	168	

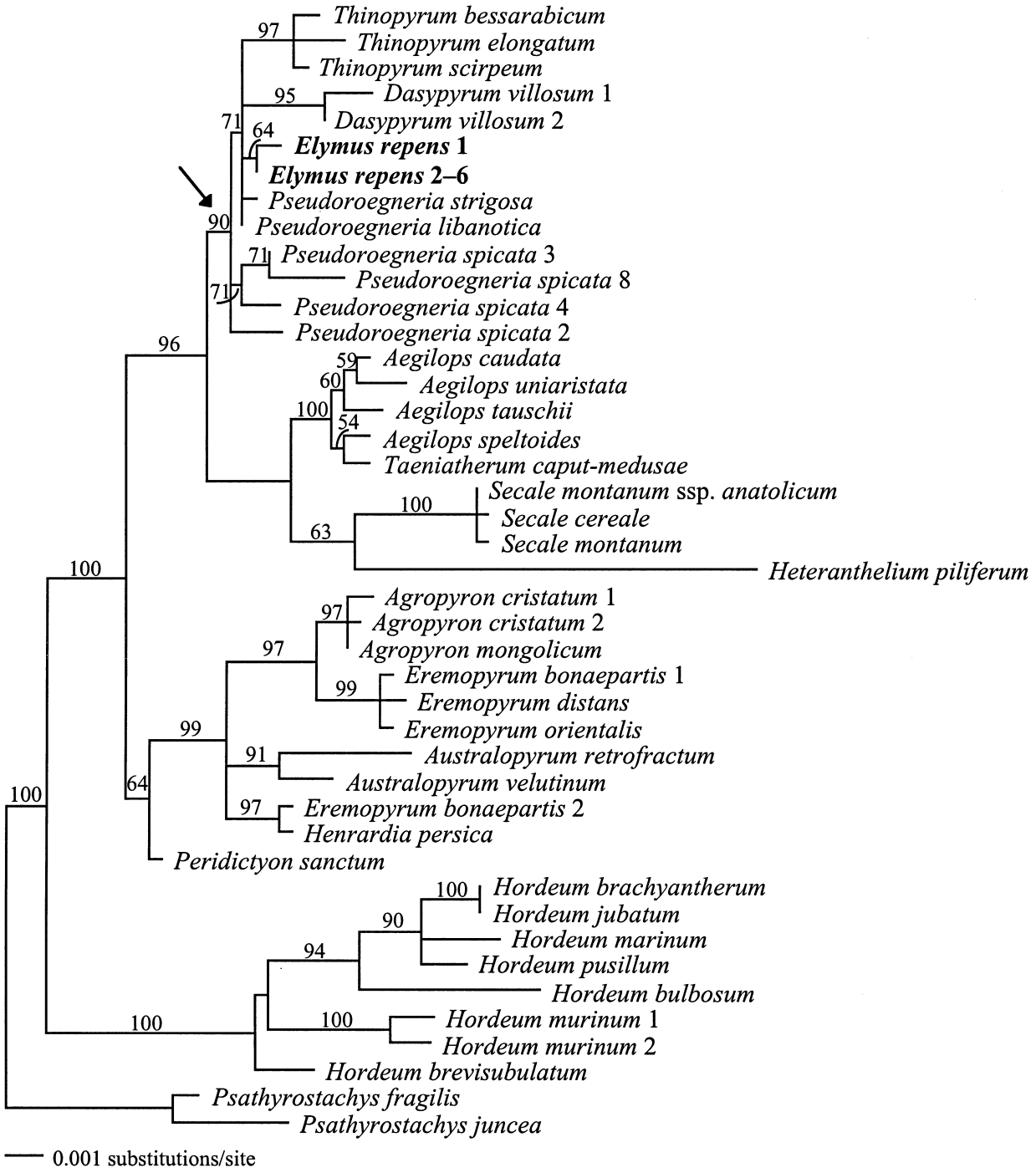

FIGURE 1. Chloroplast genome tree showing the placement of *Elymus repens* with *Dasypyrum*, *Thinopyrum*, and *Pseudoroegneria*. The tree was estimated from sequences of the *rpoA* gene, and the *trnT*, *trnL*, and *trnF* genes and their intervening spacers. Sequences were analyzed using a ML heuristic search under the GTR+I+Γ model of sequence evolution, and bootstrap support is based on heuristic ME searches of 1,000 bootstrap replicates, with ME distances estimated under the GTR+I+Γ model. *Elymus repens* individuals 2, 3, 4, 5, and 6 were treated as a single taxon. Other taxa are from Mason-Gamer et al. (2002).

includes the "Unknown 2" sequences within a taxonomically broader sample of grasses.

The *E. repens* starch synthase sequences form five clades, including four clades within the Triticeae (Fig. 2) and one outside of the tribe (Fig. 3), but not all of the individual plants are represented in every clade. The four clades within the Triticeae, starting from the top of Figure 2, are as follows: (1) *Hordeum* and *E. repens*: The *E. repens* sequences, representing four of the six individuals (#1, 4, 5, and 6), form a very weakly supported group

within the well-supported *Hordeum* + *E. repens* clade, but the current data do not clearly resolve their exact placement within the *Hordeum* group. (2) *Pseudoroegneria* and *E. repens*: Four of the six *E. repens* individuals (#1, 2, 3, and 6) are represented in this clade. The group is itself divided into two clades: all but one of the North American *P. spicata* sequences form one clade, while the two Eurasian species, *P. strigosa* and *P. libanotica*, form a second clade along with the *E. repens* sequences. (3) *Taeniatherum* and *E. repens*: This sequence type, recovered from each of the six *E. repens* individuals, is markedly over-represented in the F-for/M-bac PCR amplifications (Table 3). The variants within this clade, including those from *Taeniatherum* itself, are probably non-functional: all have a 10-bp deletion in exon 11, and all but the *Taeniatherum* 2e clone have a 13-bp deletion in exon 14. The *E. repens* 1aa clone has an additional 18-bp deletion in exon 13. (4) *Elymus repens* only ("Unknown 1"): This clade, representing an unknown diploid donor, includes sequences from all six *E. repens* individuals. This clade is itself divided into two clades; two of the individuals (#2 and #4) are represented in both. The sequence variants within the "Unknown 1" clade vary in length: the F-for/M-bac sequences 1q, 2cc, 5hh, and 5k are 1243 bp; 6hh and 2aa are 1260 bp; and 4a is 1231 bp.

The "Unknown 2" sequences in Table 3 were recovered from the F-for/K-bac amplification reactions, and are distinct from other Triticeae GBSSI sequences. This sequence variant, recovered from four of the six *E. repens* individuals, is shorter than the other *E. repens* sequences, and thus it can be identified among the other F-for/K-bac products on agarose gels. Based on visual inspection of gels, the shorter "Unknown 2" sequence did not amplify in individuals 4 and 6, so the lack of the variant in these individuals (Table 3) was not merely due to a sampling artifact from the cloning process.

A phylogenetic analysis of GBSSI exons from a broad sample of grasses (Fig. 3) suggested that the "Unknown 2" sequences were most similar not to a member of the Triticeae, but to *Cutandia memphitica*, a species within the same subfamily (Pooideae) but outside of the Triticeae. Sequences from all four *E. repens*-containing clades within the Triticeae (Fig. 2) are also included in Fig. 3: *E. repens* 1c, 3a, 2aa, and 6dd represent, respectively, the *Pseudoroegneria*, *Taeniatherum*, unknown donor, and *Hordeum* clades in Fig. 2. The *E. repens* sequences 2h, 5b, 1g, and 3g are "Unknown 2" sequences, and fall outside of the Triticeae and its sister tribe, Bromeae. Furthermore, the "Unknown 2" sequences share with *Cutandia* the loss of intron 10, which is present in all of the sampled Triticeae and in almost all grasses examined to date. (I have encountered one other loss of intron 10 previously, in *Arundinella nepalensis*, a member of the subfamily Panicoideae; Genbank accession AF079237.) The "Unknown 2" sequences of *E. repens* have three unique length differences in exon 11 relative to other grasses examined to date, suggesting that this variant is non-functional: a 4 bp insertion, a 1 bp insertion, and a 9 bp deletion.

DISCUSSION

Biased Recovery of Sequence Variants

It is not clear why the primer sets used here yield markedly biased samples of sequence variants from within *E. repens*. Two obvious factors to consider are differences in fragment length, under the assumption that shorter fragments amplify more efficiently than longer ones, and differences among the variants in terms of primer affinities. Because the F-for and M-bac primers are the most external primers used, differences among gene variants in terms of primer mismatch are not known. The observed biases are consistent with size differences among the amplified variants, but the differences are small. Of the four main variant types amplified by the F-for/M-bac primer pair (i.e., those corresponding to the four clades in Fig. 2), the over-represented *Taeniatherum*-like variant is the shortest at 1205 bp. This is 39 bp shorter than the *Pseudoroegneria*-like variant, 26–55 bp shorter than the "Unknown 1" variants, and only 15–16 bp shorter than the *Hordeum*-like variants. Clones obtained from the PCR reactions using the F-for and K-bac primers (with K-bac designed to exclude *Taeniatherum*) were again biased toward the shortest fragment, in this case the "Unknown 2" variant (560–564 bp vs. 634–644 bp for all other variants). It may be that length differences provide enough advantage to the shorter variants to allow them to dominate the PCR and/or ligation reactions. The bias toward the *Taeniatherum*-like copy in the F-for/M-bac reactions, however, seems extreme in light of the mere 15–16 bp difference in length between it and the *Hordeum*-like variants. In North American *Elymus* tetraploids (Mason-Gamer, 2001), which combine *Hordeum* and *Pseudoroegneria* genomes, little or no PCR bias was observed in F-for/M-bac reactions in spite of the fact that the *Hordeum*-like sequences are 18–38 bp shorter than the *Pseudoroegneria*-like variants.

While length differences can potentially lead to PCR bias, many other factors can lead to biased amplification. These include certain single-base substitutions (Liu et al., 1997; Barnard et al., 1998), effects of flanking regions (Hansen et al., 1998), and variation in G+C content (Reysenbach et al., 1992). Clearly, there are many reasons why PCR-based sampling of variation may yield a set of products whose frequencies do not correspond to those in the genomic DNA, and caution must be taken when interpreting PCR product mixtures as samples of intra-individual diversity. Furthermore, failure to obtain a particular gene copy from an individual may not indicate the lack or loss of that gene copy; in such cases primers specific to the "missing" gene copy may clarify the situation (e.g., Ge et al., 1999; Ferguson and Sang, 2001; Doyle et al., 2002). In *E. repens*, sets of primers specific to each of the five observed GBSSI clades may be needed to clarify the patterns of variation within each individual, and thus allow further study of the variation within the species as a whole.

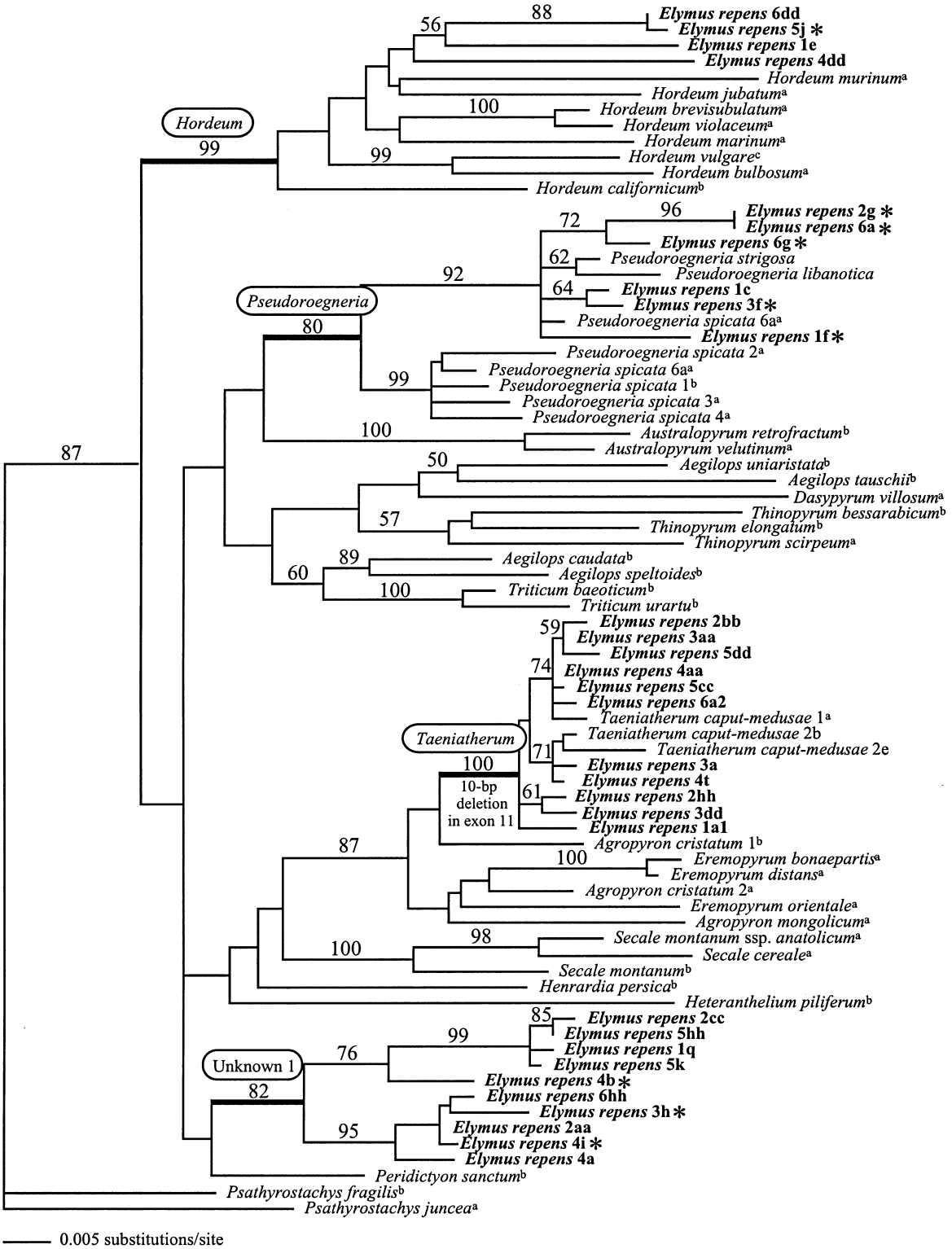

FIGURE 2. Starch synthase gene tree based on both introns and exons, showing the placement of *Elymus repens* sequences (boldface) into four clades within the Triticeae. The tree was estimated using an ML heuristic search under the GTR+I+Γ model of sequence evolution, and bootstrap support is based on heuristic ME searches of 1000 bootstrap replicates, with ME distances estimated under the GTR+I+Γ model. Each *E. repens* sequence is followed by a numerical identifier that corresponds to a particular individual. Letters following numerical identifiers are specific to cloned sequences from within that individual. Asterisks indicate fragments amplified with the F-for/K-bac primer pair, which are about half the length of the F-for/M-bac sequences. [a]Mason-Gamer, 2001; [b]Mason-Gamer et al., 1998; [c]Rohde et al., 1988.

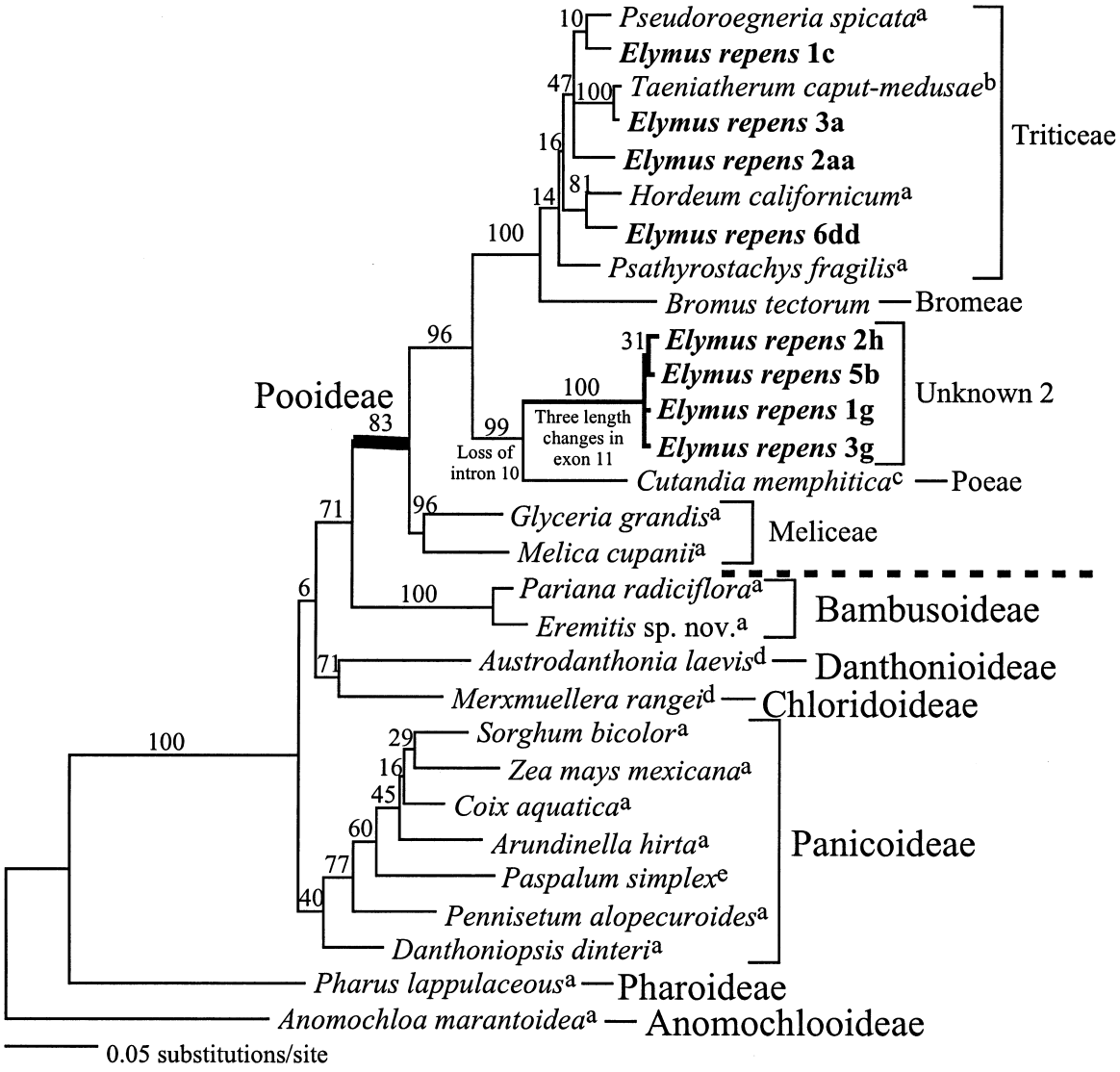

FIGURE 3. Starch synthase gene tree based on exons only, showing the placement of the divergent "Unknown 2" *E. repens* sequences within the subfamily Pooideae (labeled node). Labels on the right identify tribes within the Pooideae and the other sampled subfamilies. The four *E. repens* sequences within the Triticeae represent the four labeled clades in Fig. 2. The "Unknown 2" sequences were recovered from PCR reactions using the F-for/K-bac primer pair; all other sequences are from F-for/M-bac reactions. The tree was estimated using a ML heuristic search under the GTR+I+Γ model of sequence evolution, and bootstrap support is based on heuristic ME searches of 1,000 bootstrap replicates, with ME distances estimated under the GTR+I+Γ model. The intron loss and length changes marked on the tree were not used in the estimation of the tree. [a]Mason-Gamer et al., 1998; [b]Mason-Gamer, 2001; [c]J. I Davis, unpubl.; [d]GPWG, 2001; [e]Mathews et al., 2002.

Phylogenetic Affinities of Elymus repens

The cpDNA data are consistent with the cytogenetic results of Dewey (1970, 1976), from which he concluded that *Pseudoroegneria* was a genome donor to *E. repens*. The results, however, are more complicated in that they indicate three possible maternal genome donors: *Pseudoroegneria*, *Dasypyrum*, and *Thinopyrum*. There have been relatively few cytogenetic studies of hybrids involving *Dasypyrum*, and I am unaware of any that address its relationship to *E. repens*. The possible presence of a *Thinopyrum* genome within *E. repens* was suggested by the FISH data of Ørgaard and Anamthawat-Jónsson (2001). Cauderon and Saigne (1961) and Heneen (1963), on the other hand, examined meiotic cells from *E. repens* X *Thinopyrum junceiforme* hybrids and concluded that *E. repens* does not include a *Thinopyrum* nuclear genome, and Dewey (1980) later drew similar conclusions based on *E. repens* X *T. curvifolium* hybrids. The conclusion that *Pseudoroegneria* is the maternal donor is favored here not because of its consistency with chromosome pairing data, which form the basis for the original hypothesis, but because *Dasypyrum* and *Thinopyrum* do not form a clade with *Pseudoroegneria* on any other molecular phylogenetic trees to date (Hsiao et al., 1995; Kellogg and Appels, 1995; Petersen and Seberg, 1997, 2002; Mason-Gamer and Kellogg, 2000; Mason-Gamer, 2001) or on

the most comprehensive morphology-based tree (Seberg and Frederiksen, 2001). Therefore, the close relationships among *Pseudoroegneria*, *Dasypyrum*, and *Thinopyrum* on the cpDNA tree appear to reflect introgression of the *Pseudoroegneria* chloroplast genome into *Dasypyrum* and *Thinopyrum*, rather than an overall, genome-wide relatedness (Kellogg et al., 1996). *Elymus repens* itself, as discussed below, yields GBSSI sequences that group with *Pseudoroegneria*, but none that group with either *Dasypyrum* or *Thinopyrum*.

Results from GBSSI gene sequences are more complicated than those from the chloroplast genome. This is to be expected, since nuclear sequences can reflect multiple donors to a polyploid taxon; at least *Pseudoroegneria*- and *Hordeum*-like sequences are expected within *E. repens*, based on existing cytogenetic data. However, the situation in *E. repens* is more complex than can be explained by its hexaploid origin alone, in that five genome donors are implicated in its origin, including four from within the Triticeae and one from outside of the tribe.

Within the Triticeae, the *E. repens* sequences fall into four clades, labeled in Fig. 2. The first group of sequences, representing four of the six *E. repens* individuals, are closely related to those from *Hordeum*. The presence of the **H** genome is consistent with cytogenetic (Cauderon and Saigne, 1961) and GISH (Ørgaard and Anamthawat-Jónsson, 2001) data. Although this conflicts with Dewey's intuition (Dewey, 1984), he did not cite explicit data supporting the absence of the *Hordeum* genome. *Elymus repens* and *Hordeum* are morphologically dissimilar, but given the general morphological similarity between *E. repens* and some of the **StStHH** tetraploid species of *Elymus*, the presence of the **H** genome within *E. repens* is not a startling result. Individuals 2 and 3 might, in principle, lack a *Hordeum*-like sequence altogether, but it is as reasonable to assume that the sequence simply has not been sampled. One effect of the observed bias in frequencies of recovered sequence variants (Table 3) is that it is difficult to thoroughly sample all of the existing gene variants within each individual. Given the small numbers of *Hordeum*-like sequences recovered from the individuals in which it has been found (4 out of 26 clones; 2 out of 25; 1 out of 29; and 2 out of 38 in plants 1, 4, 5, and 6, respectively), sampling effects seem a reasonable explanation for the missing sequences in individuals 2 and 3.

A second GBSSI sequence type, resembling the *Pseudoroegneria* sequences, has been recovered from four of the six *E. repens* individuals. Thus, like the cpDNA data, the GBSSI data support the initial hypothesis, based on cytogenetic data (Dewey, 1970, 1976), that *Pseudoroegneria* was involved in the evolution of this polyploid species. As with the *Hordeum*-like gene copies, *Pseudoroegneria*-like copies were not recovered from all individuals; this may be another illustration of the sampling problem related to biased recovery of sequence variants. While this gene variant may be lacking from individuals 4 and 5, it seems as likely that it has simply not been sampled.

The apparent involvement of *Taeniatherum* (genome **Ta**) as a third participant in the evolution of *E. repens* comes as a surprise. Although *Taeniatherum*, like *E. repens*, has become a problematic introduced weed in North America (Maurer et al., 1988), they are very different in terms of their morphological characteristics and life history features. There is little cytogenetic basis for hypothesizing one way or the other about the presence of the **Ta** genome in *E. repens*. *Taeniatherum* is a small genus of diploids, with a single variable species (Frederiksen, 1986; Frederiksen and von Bothmer, 1986), and has been the subject of relatively few cytogenetic analyses. One study found little or no genome pairing between *Taeniatherum* and several other genera in the tribe (Frederiksen and von Bothmer, 1989), but I am unaware of any cytogenetic studies specifically addressing the presence of a **Ta** genome in *E. repens*. All of the sequences in the *Taeniatherum* + *E. repens* clade have between 1 and 3 deletions within exons, and appear to be non-functional. Because the gene is apparently nonfunctional in both *E. repens* and in *Taeniatherum*, the loss of function appears to predate the origin of *E. repens* (or, at least, its acquisition of the *Taeniatherum*-like copy); thus, they do not represent a case of gene silencing following polyploidization (Ford and Gottlieb, 2002).

Another unexpected result is the apparent genetic contribution from a fourth, unknown donor ("Unknown 1" in Fig. 2 and Table 3). At least one copy from this sequence group was recovered from each of the six *E. repens* individuals. Although the monogenomic genera on the tree represent a reasonably comprehensive sampling of the phylogenetic diversity throughout the tribe, it is certainly possible that there is a closely related extant donor of the "Unknown 1" sequences that was not included in the analysis. Alternatively, it is possible that this group of sequences was derived from a now-extinct ancestor.

Because of the consistent overrepresentation of *Taeniatherum*-like sequences, the *Hordeum*-like, *Pseudoroegneria*-like, and Unknown 1 sequences have not been exhaustively sampled from each individual. This makes it difficult not only to determine whether each is truly present within each individual, but also to assess levels of variation in the corresponding clades. In two of these three clades, potentially intriguing patterns of variation are already apparent. In the *Pseudoroegneria* clade, two of the four *E. repens* individuals (1 and 6) each exhibit two distinct variants. Sequences in the "Unknown 1" clade differ from one another by both length changes and single-base substitutions. Three of the six individual plants in this clade are variable, and three distinct variants have been recovered from individual 4. Further exploration of variation within each clade awaits an approach that will overcome the observed sequence-specific bias. A possible solution lies in the use of a different set of primers for each clade, based on character states unique to each.

The fifth and final group of *E. repens* sequences is placed not within the Triticeae but with *Cutandia*. The present analysis (Fig. 3), based on a small sample of grasses, does not suggest that *Cutandia* itself is the closest extant relative to the divergent *E. repens* sequences. In fact, as is clear in Fig. 3, the "Unknown 2" sequences are

far more divergent from *Cutandia* than the other *E. repens* sequences are from their putative donors within the Triticeae. With three deletions in exon 11, the "Unknown 2" sequences are apparently non-functional. Because their donor is unknown, it remains uncertain whether their loss of function occurred before or after their acquisition by *E. repens*. A more comprehensive sample of GBSSI sequences from within the subfamily Pooideae will be required in order to determine more precisely the origin of these sequences; only then will speculation about the processes leading to this broadly reticulate pattern be warranted. Regardless of the exact identity of the donor, it is clear that *E. repens* has managed to acquire genetic material from a surprisingly divergent source.

The presence of five distinct sequence types within *E. repens* is at odds with its hexaploid genomic configuration, from which at most three distinct categories of variants were expected. The observed pattern is not consistent with gene duplication. Duplication following polyploidization would give rise to additional sequences within existing clades, whereas the variants in *E. repens* form five distinct clades. Alternatively, a gene duplication event early in the evolution of the Triticeae would be evident in the diploids as well as in *E. repens*. One possibility is that the observed diversity represents allelic variation, with certain alleles at one or more of the three homoeologous loci having been acquired by introgression. While this is speculative, it is in keeping with the demonstrated reticulate history of the tribe (Kellogg et al., 1996; Mason-Gamer and Kellogg, 1996). In this scenario, the variation within *E. repens* represents both the confluence of entire genomes via hexaploidy, and the acquisition of parts of genomes through introgression. If some pairs of variants indeed represent heterogeneous alleles at a locus, then another explanation for the lack of some variants in some plants (in addition to sampling artifacts as suggested earlier) would be that the plant is homozygous at one or more of the three homoeologous GBSSI loci. The relative effects of allopolyploidy and introgression will be better characterized with data from additional nuclear markers representing other chromosomes.

With a sample of six individuals, it is difficult to predict whether the observed patterns are characteristic of the entire species. *Elymus repens* is not native to the United States, and it may or may not be significant that all of the sampled individuals in the present study are from introduced populations. It is tempting to speculate that the success and rapid spread of *E. repens* throughout much of North America might be due, in part, to its ability to acquire genetic material from divergent sources. Preliminary data from three Eurasian specimens (Mason-Gamer, unpubl.) reveal all four of the clades within the Triticeae, but not the "Unknown 2" sequence from outside the tribe, leading to a tentative working hypothesis that the "Unknown 2" sequence was acquired subsequent to the introduction of the species to North America. Ultimately, understanding the molecular variation within *E. repens* in terms of its present distribution will require data from more *E. repens* accessions representing widely dispersed native Eurasian sites and additional sites in North America. Because hybridization, including allopolyploidy, is postulated to play a role in the evolution of plant invasiveness (e.g., Ellstrand and Schierenbeck, 2000), it will be particularly interesting to see whether introgression and/or changes in genomic content have accompanied the introduction and rapid spread of *E. repens* within North America.

CONCLUSION

In their groundbreaking review of the mechanisms of polyploid formation, Harlan and deWet (1975:382) pointed out that opportunities for wide hybridization were probably greatest among plants at high polyploid levels: "High polyploids can withstand the shock of alien germplasm ... and the widest crosses are likely to be most successful at that level." Their comment was based on a broad review of crossing and cytogenetic data from many diverse plant taxa, but now the detailed molecular dissection of polyploid taxa has become a realistic, complementary approach to the study of polyploids. *Elymus repens* may serve as an illustration of Harlan and deWet's complex view of polyploids, showing that polyploid plants have the potential to acquire genetic diversity via processes beyond polyploidy alone, perhaps through a combination of allopolyploidy and introgression. *Elymus repens* is not the first polyploid in which molecular phylogenetic analyses have uncovered a history of both polyploidy and introgression (e.g., *Gossypium gossypioides*; Cronn et al., 2003). It may, however, be the most far-reaching examined so far in terms of its sources of genetic variation.

ACKNOWLEDGMENTS

Thanks to Mary Barkworth, Megan Helfgott, and Toby Kellogg for discussion on the Triticeae; to Steve Brunsfeld for discussion on polyploidy; and to Jerry Davis for insights into the Pooideae and for providing an unpublished *Cutandia* GBSSI sequence. J. Chris Pires and an anonymous reviewer provided thoughtful reviews of the manuscript, and Associate Editor Peter Linder and Editor Chris Simon provided valuable input. The research was funded by NSF DEB-9974181.

REFERENCES

Adams, K. L., R. Cronn, R. Percifield, and J. F. Wendel. 2003. Genes duplicated by polyploidy show unequal contributions to the transcriptome and organ-specific reciprocal silencing. Proc. Natl. Acad. Sci. USA 100:4649–4654.

Ainouche, M. L., R. J. Bayer, J. P. Gourret, A. Defontaine, and M. T. Misset. 1999. The allotetraploid invasive weed *Bromus hordeaceus* L. (Poaceae): Genetic diversity, origin, and molecular evolution. Folia Geobot. 34:405–419.

Assadi, M., and H. Runemark. 1995. Hybridisation, genomic constitution, and generic delimitation in *Elymus* s. l. (Poaceae: Triticeae). Pl. Syst. Evol. 194:189–205.

Ballard, H. E., and K. J. Sytsma. 2000. Evolution and biogeography of the woody Hawaiian violets (*Viola*, Violaceae): Arctic origins, herbaceous ancestry, and bird dispersal. Evolution 54:1521–1532.

Barnard, R., V. Futo, N. Pecheniuk, M. Slattery, and T. Walsh. 1998. PCR bias toward the wild-type K-*ras* and *p*53 sequences: Implications for PCR detection of mutations and cancer diagnosis. BioTechniques 25:684–691.

Batcher, M. S. 2002. Element stewardship abstract for *Elytrigia repens* (L.) Desv. ex B.D. Jackson. The Nature Conservancy, Arlington, VA.
Bradley, R. D., and D. M. Hillis. 1997. Recombinant DNA sequences generated by PCR amplification. Mol. Biol. Evol. 14:592–593.
Cauderon, Y. 1958. Étude cytogénétique des *Agropyrum* français et de leurs hybrides avec les blés. Ann. Amélior. Plantes 8:389–567.
Cauderon, Y., and B. Saigne. 1961. New interspecific and intergeneric hybrids involving *Agropyron*. Wheat Information Service 12:13–14.
Cook, L. M., P. S. Soltis, S. J. Brunsfeld, and D. E. Soltis. 1998. Multiple independent formations of *Tragopogon* tetraploids (Asteraceae): Evidence from RAPD markers. Mol. Ecol. 7:1293–1302.
Cronn, R., M. Cedroni, T. Haselkorn, C. Grover, and J. F. Wendel. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. Theor. Appl. Genet. 104:482–489.
Cronn, R., R. L. Small, T. Haselkorn, and J. F. Wendel. 2003. Cryptic repeated genomic recombination during speciation in *Gossypium gossypioides*. Evolution 57:2475–2489.
Dewey, D. R. 1961. Hybrids between *Agropyron repens* and *Agropyron desertorum*. J. Hered. 52:13–21.
Dewey, D. R. 1970. A cytogenetic study of *Agropyron stipifolium* and its hybrids with *Agropyron repens*. Bull. Torrey Bot. Club 97:315–320.
Dewey, D. R. 1976. Derivation of a new forage grass from *Agropyron repens* X *Agropyron spicatum* hybrids. Crop Sci. 16:175–180.
Dewey, D. R. 1980. Hybrids and induced amphiploids of *Agropyron repens* X *A. curvifolium*. Crop Sci. 20:473–478.
Dewey, D. R. 1984. The genomic system of classification as a guide to intergeneric hybridization within the perennial Triticeae. Pages 209–279 *in* Gene manipulation in plant improvement, Proc. 6th Stadler genetics symposium (J. P. Gustafson, ed.). Columbia University Press, New York.
Doyle, J. J., J. L. Doyle, and A. D. H. Brown. 1999. Origins, colonization, and lineage recombination in a widespread perennial soybean complex. Proc. Natl. Acad. Sci. USA 96:10741–10745.
Doyle, J. J., J. L. Doyle, A. D. H. Brown, and R. G. Palmer. 2002. Genomes, multiple origins, and lineage recombination in the *Glycine tomentella* (Leguminosae) polyploid complex: Histone H3-D gene sequences. Evolution 56:1388–1402.
Ellstrand, N. C., and K. A. Schierenbeck. 2000. Hybridization as a stimulus for the evolution of invasiveness in plants? Proc. Natl. Acad. Sci. USA 97:7043–7050.
Emshwiller, E., and J. J. Doyle. 2002. Origins of domestication and polyploidy in oca (*Oxalis tuberosa*: Oxalidaceae). 2. Chloroplast-expressed glutamine synthase data. Amer. J. Bot. 89:1042–1056.
Ferguson, D., and T. Sang. 2001. Speciation through homoploid hybridization between allotetraploids in peonies (*Paeonia*). Proc. Natl. Acad. Sci. USA 98:3915–3919.
Ford, V. S., and L. D. Gottlieb. 2002. Single mutations silence *PgiC2* genes in two very recent allotetraploid species of *Clarkia*. Evolution 56:699–707.
Frati, F., C. Simon, J. Sullivan, and D. L. Swofford. 1997. Evolution of the mitochondrial cytochrome oxidase II gene in Collembola. J. Mol. Evol. 44:145–158.
Frederiksen, S. 1986. Revision of *Taeniatherum*. Nord. J. Bot. 6:389–397.
Frederiksen, S., and R. von Bothmer. 1986. Relationships in *Taeniatherum* (Poaceae). Can. J. Bot. 64:2343–2347.
Frederiksen, S., and R. von Bothmer. 1989. Intergeneric hybridization between *Taeniatherum* and different genera of Triticeae, Poaceae. Nord. J. Bot. 9:229–240.
Ge, S., T. Sang, B.-R. Lu, and D.-Y. Hong. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. Proc. Natl. Acad. Sci. USA 96:14400–14405.
GPWG (Grass Phylogeny Working Group). 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). Ann. Missouri Bot. Gard. 88:373–457.
Gu, X., Y.-X. Fu, and W.-H. Li. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. Mol. Biol. Evol. 12:546–557.
Hansen, M. C., T. Tolker-Nielsen, M. Givskov, and S. Molin. 1998. Biased 16S rDNA PCR amplification caused by interference from DNA flanking the template region. FEMS Microbiol. Ecol. 26:141–149.
Harlan, J. R., and J. M. J. deWet. 1975. On Ö. Winge and a prayer: The origins of polyploidy. Bot. Rev. 41:361–390.
Hasagawa, M., H. Kishino, and T. Yano. 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22: 160–170.
Heneen, W. K. 1963. Meiosis in the interspecific hybrid *Elymus farctus* X *E. repens* (= *Agropyron junceum* X *A. repens*). Hereditas 49:107–118.
Hsiao, C., N. J. Chatterton, K. H. Asay, and K. B. Jensen. 1995. Phylogenetic relationships of the monogenomic species of the wheat tribe, Triticeae (Poaceae), inferred from nuclear rDNA (internal transcribed spacer) sequences. Genome 38:211–223.
Judo, M. S. B., A. B. Wedel, and C. Wilson. 1998. Stimulation and suppression of PCR-mediated recombination. Nucleic Acids Res. 26:1819–1825.
Jukes, T. H., and C. R. Cantor. 1969. Evolution of protein molecules. Pages 21–132 *in* Mammalian protein metabolism (N. H. Munro, ed.). Academic Press, New York.
Kellogg, E. A., and R. Appels. 1995. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. Genetics 140:325–343.
Kellogg, E. A., R. Appels, and R. J. Mason-Gamer. 1996. When gene trees tell different stories: The diploid genera of Triticeae (Gramineae). Syst. Bot. 21:312–347.
Kimura, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. 16:111–120.
Liu, Q., E. C. Thorland, and S. S. Sommer. 1997. Inhibition of PCR amplification by a point mutation downstream of a primer. BioTechniques 22:292–300.
Löve, Á. 1984. Conspectus of the Triticeae. Feddes Repert. 95:425–521.
Mason-Gamer, R. J. 2001. Origin of North American species of *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. Syst. Bot. 26:757–768.
Mason-Gamer, R. J., and E. A. Kellogg. 1996. Chloroplast DNA analysis of the monogenomic Triticeae: Phylogenetic implications and genome-specific markers. Pages 301–325 *in* Methods of genome analysis in plants (P. P. Jauhar, ed.). CRC Press, Inc., Boca Raton, Florida.
Mason-Gamer, R. J., and E. A. Kellogg. 2000. Phylogenetic analysis of the Triticeae using the starch synthase gene, and a preliminary analysis of some North American *Elymus* species. Pages 102–109 *in* Grasses: Systematics and evolution (S. W. L. Jacobs and J. Everett, eds.). CSIRO Publishing, Collingwood VIC.
Mason-Gamer, R. J., N. L. Orme, and C. M. Anderson. 2002. Phylogenetic analysis of North American *Elymus* and the monogenomic Triticeae (Poaceae) using three chloroplast DNA data sets. Genome 45:991–1002.
Mason-Gamer, R. J., C. F. Weil, and E. A. Kellogg. 1998. Granule-bound starch synthase: Structure, function, and phylogenetic utility. Mol. Biol. Evol. 15:1658–1673.
Mathews, S., R. E. Spangler, R. J. Mason-Gamer, and E. A. Kellogg. 2002. Phylogeny of Andropogoneae inferred from phytochrome B, GBSSI, and *ndhF*. Int. J. Plant Sci. 163:441–450.
Maurer, T., M., J. Russo, and A Godell. 1988. Element stewardship abstract for *Taeniatherum caput-medusae*. The Nature Conservancy, Arlington, VA.
Ørgaard, M., and K. Anamthawat-Jónsson. 2001. Genome discrimination by *in situ* hybridization in Icelandic species of *Elymus* and *Elytrigia* (Poaceae:Triticeae). Genome 44:275–283.
Osborn, T. C., J. C. Pires, J. A. Birchler, D. L. Auger, Z. J. Chen, H.-S. Lee, L. Comai, A. Madlung, R. W. Doerge, V. Colot, and R. A. Martienssen. 2003. Understanding the mechanisms of novel gene expression in polyploids. Trends Genet. 19:141–147.
Petersen, G., and O. Seberg. 1997. Phylogenetic analysis of the Triticeae (Poaceae) based on *rpo*A sequence data. Mol. Phyl. Evol. 7:217–230.
Petersen, G., and O. Seberg. 2002. Molecular evolution and phylogenetic application of *DMC1*. Mol. Phyl. Evol. 22:43–50.
Reysenbach, A.-L., L. J. Giver, G. S. Wickham, and N. R. Pace. 1992. Differential amplification of rRNA genes by polymerase chain reaction. Appl. Environ. Microbiol. 58:3417–3418.
Rohde, W., D. Becker, and F. Salamini. 1988. Structural analysis of the waxy locus from *Hordeum vulgare*. Nucleic Acids Res. 16:7185–7186.
Seberg, O., and S. Frederiksen. 2001. A phylogenetic analysis of the monogenomic Triticeae (Poaceae) based on morphology. Bot. J. Linn. Soc. 136:75–97.

Smedmark, J. E. E., T. Eriksson, R. C. Evans, and C. S. Campbell. 2003. Ancient allopolyploid speciation in Geinae (Rosaceae): Evidence from nuclear granule-bound starch synthase (GBSSI) gene sequences. Syst. Biol. 52:374–385.

Soltis, P. S., G. M. Plunkett, S. J. Novak, and D. Soltis. 1995. Genetic variation in *Tragopogon* species: Additional origins of the allotetraploids *T. mirus* and *T. miscellus* (Compositae). Amer. J. Bot. 82:1329–1341.

Sullivan, J., J. A. Markert, and C. W. Kilpatrick. 1997. Phylogeography and molecular systematics of the *Peromyscus aztecus* species group (Rodentia: Muridae) inferred using parsimony and likelihood. Syst. Biol. 46:426–440.

Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates.

Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis. 1996. Phylogenetic inference. Pages 407–514 *in* Molecular systematics, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17:1105–1109.

USDA and NRCS. 2002. The PLANTS database, version 3.5 *at* http://plants.usda.gov. National Plant Data Center, Baton Rouge.

Waddell, P. J., and D. Penny. 1996. Evolutionary trees of apes and humans from DNA sequences. Pages 53–73 *in* Handbook of symbolic evolution (A. J. Locke and C. R. Peters, eds.). Clarendon Press, Oxford.

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

Wendel, J. F. 1989. New World tetraploid cottons contain Old World cytoplasm. Proc. Natl. Acad. Sci USA 86:4132–4136.

Wendel, J. F., J. M. Stewart, and J. H. Rettig. 1991. Molecular evidence for homoploid reticulate evolution among Australian species of *Gossypium*. Evolution 45:694–711.

Yang, Z. 1994a. Estimating the pattern of nucleotide substitution. J. Mol. Evol. 39:105–111.

Yang, Z. 1994b. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. J. Mol. Evol. 39:306–314.

Yang, Z., N. Goldman, and A. Friday. 1995. Maximum likelihood trees from DNA sequences: A peculiar statistical estimation problem. Syst. Biol. 44:384–399.

First submitted 23 February 2003; reviews returned 9 June 2003; final acceptance 14 September 2003
Associate Editor: Peter Linder