

The Evolution of North American *Elymus* (Triticeae, Poaceae) Allotetraploids: Evidence from Phosphoenolpyruvate Carboxylase Gene Sequences

D. MEGAN HELFGOTT¹ and ROBERTA J. MASON-GAMER

University of Illinois at Chicago, Department of Biological Sciences, (MC 066), 845 West Taylor Street, Chicago, Illinois 60607

¹Author for Correspondence. Current address: Department of Biological Sciences, 2325 North Clifton; McGowan Rm. 118, DePaul University, Chicago, Illinois 60614 (dhelpgot@depaul.edu)

Communicating Editor: Gregory M. Plunkett

ABSTRACT. Cytogenetic studies of North American *Elymus* suggest that the genus is an allopolyploid derivative of *Pseudoroegneria* (**St**) and *Hordeum* (**H**). To test this, we conducted a phylogenetic analysis of North American *Elymus* species within a broad sample of diploid Triticeae taxa using cloned sequences from one member of the nuclear gene family encoding phosphoenolpyruvate carboxylase. The phylogeny supports the hypothesis that *Pseudoroegneria* and *Hordeum* are the diploid progenitors of the North American *Elymus* tetraploids. Each tetraploid *Elymus* individual has two distinct forms of the gene, and each form is in a strongly supported clade with sequences from either *Pseudoroegneria* or *Hordeum*, suggesting that these *Elymus* species have an **St + H** genomic content. This pattern is consistent with a single (or multiple very similar) polyploid ancestor(s) of the North American tetraploids, confirming earlier results based on granule-bound starch synthase I gene sequence data. We also examined the utility of the phosphoenolpyruvate carboxylase gene to reconstruct the evolutionary history of the Triticeae by comparing it to starch synthase gene sequence data. Both nuclear data sets are phylogenetically informative, but suggest somewhat different evolutionary histories among genera within the tribe.

Allopolyploid taxa can be difficult to interpret within phylogenies of their diploid relatives. The problems with placing hybrid taxa within the bifurcating trees that are obtained from the most common methods of phylogenetic analysis have been discussed (e.g., Hull 1979; Cronquist 1987) and examined experimentally (e.g., McDade 1992, 1997). Analyses of individual gene trees, however, allow these problems to be circumvented, as long as their tree-like history has not been disrupted by recombination. Thus, molecular phylogenetic data have recently revealed the reticulate histories of several polyploid species or groups, including, for example, *Gossypium* L. (e.g., Cronn et al. 1996, 2003; Seelanan et al. 1997; Small and Wendel 2000), *Geum* L. (Smedmark et al. 2003), *Glycine* Willd. (e.g., Doyle et al. 2002; Rauscher et al. 2002), *Oxalis* L. (Emshwiller and Doyle 1998, 2002), *Oryza* L. (Ge et al. 1999), and *Paeonia* L. (e.g., Sang and Zhang 1999).

In the Triticeae, data from both chloroplast (Mason-Gamer and Kellogg 1996a; Petersen and Seberg 1998; Mason-Gamer et al. 2002) and nuclear genes (Hsiao et al. 1995; Kellogg and Appels 1995; Petersen and Seberg 2002) have greatly increased overall understanding of the evolution of the diploid members of the tribe, revealing surprising patterns of reticulate evolution (e.g., Kellogg et al. 1996). The polyploids, however, have been the focus of far fewer molecular phylogenetic studies, yet they may hold the key to understanding the history of the tribe as a whole. Of the approximately 350 Triticeae species whose chromosome number is known, 75% are of polyploid origin (e.g., Löve 1984). The present study focuses on North American members of *Elymus* L., an entirely polyploid genus of

about 150 species, found in temperate regions through much of the world.

The circumscription of *Elymus* varies greatly among treatments. In North America, one widely used definition follows the *Manual of Grasses of the United States* (Hitchcock 1951), and is based on morphological characteristics, including multiple spikelets per node and a non-disarticulating rachis. The genomic definition of *Elymus* (e.g., Dewey 1984; Löve 1984; Barkworth and Dewey 1985), on the other hand, is based on presumed overall genome similarity, deduced from the degree of chromosome pairing in meiotic cells of interspecific hybrids. Under this system, the circumscription of North American *Elymus* is very different from the morphology-based circumscription (Barkworth and Dewey 1985; Fig. 1 of Mason-Gamer 2001). Within the genomic classification system, *Elymus* is defined as those allopolyploid species with at least one set of *Pseudoroegneria* (Nevski) Á.Löve (**St**) genomes. When considered worldwide, *Elymus* may be tetraploid, hexaploid, or octoploid, and may combine the **St** genome with the **H** genome from *Hordeum* L., **P** from *Agropyron* Gaertner, **W** from *Australopyrum* (Tzvelev) Á.Löve, and **Y** from an unknown donor (e.g., Dewey 1984). The North American native *Elymus* species, which are the focus of this study, are nearly all tetraploid with a presumed **StStHH** genomic complement. Data from allozymes (Jaaska 1995, 1998), and from the nuclear *waxy* gene encoding granule-bound starch synthase I (GBSSI), were in agreement that *Elymus* is derived from *Pseudoroegneria* and *Hordeum* (Mason-Gamer 2001), but the demonstrated gene tree conflict within the Triticeae (Kellogg et al. 1996; Mason-Gamer and

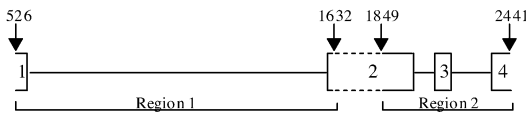


FIG. 1. A map of the *PepC* region (exons 1–4) used in this study.

Kellogg 1996b) precludes drawing strong phylogenetic conclusions from the analysis of a single gene.

Data from an additional nuclear marker are needed to confirm the genomic content of North American *Elymus*. We have turned to sequence data from the gene encoding phosphoenolpyruvate carboxylase (*PepC*; EC 4.1.1.31). The *PepC* genes form a small multigene family in the plants studied so far (e.g., Cushman and Bohnert 1989a, 1989b; Crétin et al. 1991; Kawamura et al. 1992; Lepiniec et al. 1992, 1993; Ernst and Westhoff 1997). *Sorghum vulgare* Pers. was the first species for which all three genes in the *PepC* gene family were characterized (Lepiniec et al. 1993). Members of the family are approximately 6 kb in length and include ten exons (Lepiniec et al. 1993). Each copy codes for a functionally distinct *PepC* isoform, including a widely-expressed housekeeping form, an inducible form expressed in the root, and a light-regulated form involved in C_4 photosynthesis (e.g., Lepiniec et al. 1994). Several *PepC* isoforms have been described in other grasses as well, including maize and wheat (Izui et al. 1986; Hudspeth and Grula 1989; Kawamura et al. 1992; Gonzalez et al. 1998).

Sequence data from the *PepC* gene have already been shown to be phylogenetically useful at a number of taxonomic levels. Gehrig et al. (2001) constructed phylogenies using *PepC* sequence data spanning the plant kingdom to gain insights into the evolution of the *PepC* gene family itself. The phylogenetic analyses of all available *PepC* sequences from seed plants by Svensson et al. (2003) revealed that grass sequences form three distinct clades relative to the other seed plant sequences, suggesting that the grass isoforms represent gene duplications that occurred prior to the origin of grasses. Besnard et al. (2002) used sequences from C_4 -specific isoforms of *PepC* to confirm the relationships among grass tribes that had been previously proposed based on other molecular markers. Malcomber (2002) employed *PepC* sequence data to examine evolutionary relationships within the genus *Gaertnera* Lam. (Rubiaceae).

The utility of single- and low-copy genes for phylogenetic analyses is increased if the genes are well characterized, and if there are enough available sequence data for designing amplification primers. The availability of sequences representing multiple *PepC* isoforms from several diverse grasses allowed us to develop primers specific to a single *PepC* gene copy, the apparent ortholog of the housekeeping form in *Sor-*

ghum (GenBank accession X59925; Lepiniec et al. 1991) and *Zea* (X61489; Kawamura et al. 1992). The present phylogenetic analysis, based on a portion of this gene, addresses two objectives: 1) to confirm the genomic constitution of North American *Elymus* species, and 2) to examine the phylogenetic utility of the housekeeping isoform of the *PepC* gene by comparing *PepC* and *waxy* gene sequence data from the Triticeae.

MATERIALS AND METHODS

Taxon Sampling. Sequences from the *PepC* gene were sampled from 27 species (Table 1). Ingroup members included seven species of *Elymus*, multiple accessions of the putative parent species in *Hordeum* and *Pseudoroegneria*, and a broad sample of other monogenic genera in the Triticeae. *Bromus tectorum* L. (Bromeae) was included as an outgroup (Kellogg 1992; Davis and Soreng 1993). Sequences were submitted to GenBank under the accession numbers AY553236–AY553269 and AY548399–AY548432, and the data matrices were deposited at TreeBASE (study accession number S1053, matrix accession numbers M1794–M1795).

Molecular Data. Using three grass sequences from GenBank, we designed primers to amplify and sequence a 1.7-kb portion of the housekeeping *PepC* gene (Table 2). The data set consists of two non-contiguous regions, corresponding to base pairs 526–1632 and 1849–2441 of *Triticum aestivum* (Genbank accession AJ007705) (Fig. 1). This copy of *PepC* is similar to sequences mapped to chromosome 9 of the rice genome (AP005802 and AP005781), and using Gale and Devos's (1998) map of twelve grass genomes, we postulate that it is on the long arm of the group 5 Triticeae chromosomes.

DNA Extraction, Amplification, and Sequencing. Total genomic DNA was isolated from individual plants following Doyle and Doyle (1987). Gene amplification was performed following Mason-Gamer et al. (1998) with the addition of 5% DMSO. In all cases, *PepC* regions 1 and 2 were amplified separately. We used *PepC* primers 219F or 467F(1) and 1672R(1) or 1672R(2) to amplify region 1, and *PepC* 1827F and 2443R for region 2 (Table 2). Because *Elymus* species are presumed allopolyploids, amplification reactions were run in triplicate and combined before cloning to counter the potential effects of PCR drift (Wagner et al. 1994; Mason-Gamer 2001).

The PCR products were purified using GeneClean (Bio101) according to the manufacturer's protocol, and the cloning reactions were carried out using Promega's pGem-T Easy cloning kit according to instructions, except that the final volumes of the ligation and transformation reactions were halved. Target colonies were isolated using QIAGEN or Promega miniprep kits. In a few cases, cloned inserts were amplified directly from colonies using the original PCR primers, and cleaned using exonuclease I (Exo) and shrimp alkaline phosphatase (SAP) as in Mason-Gamer (2004). For three diploid taxa (*Bromus tectorum*, *Secale cereale*, *S. montanum* subsp. *montanum*), *PepC* region 2 was sequenced directly from a pooled set of PCR products cleaned using the Exo-SAP protocol.

Individuals were sequenced using ABI Big Dye Terminators according to the manufacturer's instructions, except that sequencing reagent volumes were quartered, and the final reaction volume was 10 μ L. Both strands were sequenced for all taxa. The sequencing reactions were visualized using an ABI model Prism-377 or MJ Research Basestation automated sequencer; then assembled using Sequencher vers. 3.0 or 4.1 (Gene Codes Corp.).

In *Elymus*, we initially screened clones for sequence variation using a single sequencing primer (*PepC*-467F for region 1 or *PepC*-1827F for region 2), until both **St** and **H** sequence types were found, or until all clones generated for a specific individual were examined. Once the **St** and **H** clones were found within an individual, both forms were completely sequenced using both internal and external primers. Then all sequences were carefully inspected for mosaic sequence patterns combining the **St** and **H** genome in one recombinant molecule.

TABLE 1. List of taxa, collection information, and GenBank numbers. Voucher numbers containing names are collectors' accession numbers. Voucher H 5555 is from the Swedish-Danish Triticeae Consortium; sample provided by Dr. O. Seberg. The remaining "PI" voucher numbers are Plant Introduction accessions from the National Plant Germplasm System (U.S. Dept. Agriculture; <http://www.ars-grin.gov/npgs/>); samples were provided by the USDA-ARS. Data are presented in the following sequence: species name, *PepC* voucher, Clone # Region-I, Clone # Region-II, *Elymus* genome, GenBank No. Region-I, Region-II, Used in *PepC-waxy* comparison (yes/no), *waxy* voucher, *waxy* GenBank No.

Elymus. *E. californicus* (Bolander) Gould, *Barkworth s.n.* (UTC), 1645B, 637C, N.A., AY553241, AY548404, Yes, Same, AY011012. *E. canadensis* L., PI 531568 (ID), 973A, 534A, St, AY553248, AY548411, Yes, *Barkworth 97-86* (UTC), AY556481. *E. canadensis*, PI 531568 (ID), 973B, 534F, H, AY553242, AY548405, No, —, —. *E. canadensis*, PI 578675 (ID), 215C, 48C, H, AY553243, AY548406, No, —, —. *E. elymoides* (Rafin.) Swezey, PI 531606 (ID), 648C, 632C, St, AY553249, AY548412, Yes, Same, AY010992. *E. elymoides*, PI 531606 (ID), 957A, 632A, H, AY553244, AY548407, Yes, Same, AY010965. *E. glaucus* Buckley, W6 10215 (ID), 219A, 51A, St, AY553250, AY548413, Yes, *MasonGamer 130* (ID), AY010979. *E. hystrix* L., *Barkworth 97-87* (UTC), 981C, 622C, St, AY553251, AY548414, Yes, Same, AY010982. *E. hystrix*, *Barkworth 97-87* (UTC), 981B, 622A, H, AY553245, AY548408, No, —, —. *E. lanceolatus* (Scribn. & Smith) Gould, PI 531623 (ID), 1049B, 604B, St, AY553252, AY548415, Yes, W6 14220 (ID), AY010984. *E. lanceolatus*, PI 531623 (ID), 1049C, 604J, H, AY553246, AY548409, Yes, Same, AY010969. *E. wawawaiensis* J. Carlson ex. *Barkworth*, PI 598812 (ID), 229A, 74A, St, AY553253, AY548416, Yes, Same, AY010990. *E. wawawaiensis*, PI 598812 (ID), 229B, 74P, H, AY553247, AY548410, Yes, Same, AY010978

Triticaceae. *Aegilops comosa* Sm., G 602 (GH), 1305A, 1296A, N. A., AY553236, AY548399, Yes, Same, AF079263. *Agropyron cristatum* (L.) Gaertn., PI 279802 (GH), 1242B, 540A, N. A., AY553237, AY548400, Yes, Same, AF079271. *Australopyrum velutinum* (Nees) B. Simon, D 2873-2878 (GH), 1209B, 1128A, N. A., AY553238, AY548401, Yes, Same, AY011004. *Dasyphyrum villosum* (L.) Candargy, D 2990 (GH), 1641A, 1134A, N. A., AY553240, AY548403, Yes, PI 470279 (GH), AY556480. *Eremopyrum orientale* (L.) Jaub. & Spach, H 5555 (GH), 884D, 543E, N. A., AY553254, AY548417, Yes, Same, AY011007. *Heterantherium piliferum* (Banks & Sol.) Hochst., PI 402352 (GH), 1301B, 1292A, N. A., AY553255, AY548418, Yes, Same, AF079277. *Hordeum brachyantherum* subsp. *californicum* (Cov. & Steb.) Bothmer, MA-138-1-40 (GH), 339B, 312B, N. A., AY553256, AY548419, Yes, Same, AF079273. *H. jubatum* L., *MasonGamer 106* (ID), 940C, 316A, N. A., AY553257, AY548420, Yes, Same, AY010963. *H. marinum* Hudson, PI 304346 (ID), 1246C, 301A, N. A., AY553258, AY548421, Yes, Same, AY010959. *H. murinum* L., Ciho 15683 (ID), 989A, 297A, N. A., AY553259, AY548422, Yes, Same, AY010960. *H. vulgare* L., *MasonGamer 107* (ID), 1346BB, 1314A, N. A., AY553260, AY548423, Yes, Rohde et al. 1988, X07931. *Leymus racemosus* subsp. *sabulosus* (M.Bieb.) Tzvelev, R-20-21-25 (GH), 1592B, 1718A, N. A., AY553261, AY548424, No, —, —. *Peridictyon sanctum* (Janka) Seberg, Fred. & Baden, *Jensen 248* (GH), 1341A, 1157A, N. A., AY553262, AY548425, Yes, Same, AF079278. *Pseudoroegneria spicata* (Pursh) A. Löve, PI 610986 (ID), 352A, 325A, N. A., AY553263, AY548426, Yes, Same, AY010999. *P. spicata*, D 2844 (GH), 804B, 610A, N. A., AY553264, AY548427, Yes, Same, AY011000. *Secale cereale* L., *Kellogg s.n.* (GH), 2729B, 2766, N. A., AY553266, AY548429, Yes, Same, AY011009. *S. montanum* subsp. *montanum* Guss., T 36554 (GH), 2734C, 2770, N. A., AY553267, AY548430, No, —, —. *S. montanum* subsp. *anatolicum* (Boiss.) Tzelev, PI 206992 (GH), 2739A, 1145A, N. A., AY553265, AY548428, Yes, PI 206991 (GH), AY011008. *Taeniatherum caput-medusae* (L.) Nevski, *MasonGamer 189d* (ID), 1588A, 1141A, N. A., AY553268, AY548431, Yes, PI 208075 (ID), AY011010. *Thinopyrum elongatum* (Host) D.R.Dewey, *MasonGamer 113* (ID), 1460A, 1116F, N. A., AY553269, AY548432, Yes, PI 531719 (GH), AF079284. *Triticum aestivum* L., —, —, —, N. A., AJ007705, AJ007705, No, —, —.

Bromeae. *Bromus tectorum* L., *Kellogg s.n.* (GH), 1662C, 2774, N. A., AY553239, AY548402, Yes, Same, AY362757.

Sequence Analysis. Boundaries between the *PepC* introns and exons were determined by comparison with published *Triticum aestivum* (AJ007705) and *Sorghum vulgare* (X59925) sequences. Sequences were initially aligned using the default setting in Clustal W ver. 1.5 (Thompson et al. 1994), then adjusted manually using MacClade ver. 4 (Maddison and Maddison 2002). Three regions (35 characters; alignment positions 393–403, 996–972, and 1280–

1296) were judged to have an ambiguous alignment and were excluded prior to phylogenetic analyses. Percent of gap characters, missing characters, and nucleotide composition for each species, were determined using MacClade.

Phylogeny Reconstruction. Maximum likelihood (ML) methods were used to generate hypotheses about relationships among taxa using PAUP* ver. 4.0b10 (Swofford 2002). Initially, maximum

TABLE 2. List of primers.

Primer Name	Relationship to <i>Triticum aestivum</i> AJ007705	Primer (5'–3')
Region I (1–1562)	526–1632	
<i>PepC</i> -291F	219–240	ACTCCTGCCATCCGCCCTTCTAT
<i>PepC</i> -467F(1)	467–489	GCTGCTCGTCCCGCCAAGGTGT
<i>PepC</i> -796R	796–813	AAAAACACCAGATTACGCACAA
<i>PepC</i> -952R	952–973	TGACTGACACGATTTGAGATTC
<i>PepC</i> -1387R	1387–1408	GTCAAGGCATAGTCGTTTCAAG
<i>PepC</i> -1555R	1555–1575	AATCCATCAATCAACATAGAGA
<i>PepC</i> -1672R(1)	1672–1696	GCTTGTATCATCTTCCCGAGTTCA
<i>PepC</i> -1672R(2)	1672–1696	GCTTGTATCATCTTCCCGAGTTCA
Region II (1566–2185)	1849–2441	
<i>PepC</i> -1827F	1827–1848	ARAYTCRGCAATCACAGAACT
<i>PepC</i> -2443R	2443–2464	TRATCCCAATGTTCTTCAATGC

parsimony (MP) analyses were performed to obtain trees on which starting parameters for ML analyses would be estimated. Parsimony analyses assumed equal weights for characters and character state changes, and gaps were treated as missing data. An initial tree was obtained via random taxon stepwise addition, and a heuristic tree search was performed using MULTREES, ACCTRAN optimization, and tree-bisection-reconnection (TBR) branch swapping (Hendy and Penny 1982). The possibility of multiple tree islands was explored by running 10,000 random addition replicates, with twenty trees held at each step.

Parameters for 16 models of sequence evolution were estimated for all most-parsimonious trees (e.g., Swofford et al. 1996; Frati et al. 1997; Sullivan et al. 1997). Four models of nucleotide substitution—Jukes-Cantor (Jukes and Cantor 1969), Kimura two-parameter (Kimura 1980), Hasegawa-Kishino-Yano (Hasegawa et al. 1985), and general time reversible (GTR; Yang 1993)—were examined. Each substitution model was paired with each of four models of among-site rate variation: 1) no rate heterogeneity; 2) some sites invariable (I; Hasegawa et al. 1985) with equal rates of change among the remaining sites; 3) rate heterogeneity among sites following a gamma distribution (Γ ; Yang 1994); and 4) some sites invariable, with gamma-distributed variation among the remaining sites (I+ Γ ; Gu et al. 1995; Waddell and Penny 1996). Using the tree with the highest ML scores across all 16 models examined, we compared the models with the three highest scores employing a likelihood ratio test, and compared the results to a χ^2 distribution (Felsenstein 1981; Huelsenbeck and Crandall 1997; Huelsenbeck and Rannala 1997; Sanderson 1998; but see Goldman 1993).

Among the models of molecular evolution tested, the one incorporating the most parameters (GTR+I+ Γ) produced the highest log likelihood value (-7444.50). However, this value was not significantly different from the one obtained using GTR+ Γ alone (-7445.08), so the GTR+ Γ model was chosen for all subsequent ML analyses. Two sequential ML analyses were conducted. The following estimated model parameters were used as initial settings in the first ML search: [R (AC) = 1.08, R (AG) = 2.97, R (AT) = 0.42, R (CG) = 1.38, R (CT) = 3.62, R (GT) = 1.00, pi (A) = 0.25, pi (C) = 0.20, pi (G) = 0.21, pi (T) = 0.33, α = 0.49], using a heuristic tree searching strategy (the initial tree obtained via stepwise random addition with twenty trees held at each step, and TBR branch swapping). This search resulted in a single tree (-7444.62), with the following estimated ML parameters: [R (AC) = 1.08, R (AG) = 3.04, R (AT) = 0.44, R (CG) = 1.41, R (CT) = 3.67, R (GT) = 1.00, pi (A) = 0.25, pi (C) = 0.20, pi (G) = 0.21, pi (T) = 0.33, α = 0.48]. These estimates were used as the starting point for the second ML analysis, which produced a tree with the same score as the first.

Branch support for the parsimony tree was estimated using the non-parametric bootstrap (Felsenstein 1985; Sanderson and Donoghue 1989). Bootstrap values were calculated from 10,000 data sets using the "fast bootstrap" tree search strategy implemented in PAUP*. Posterior probability values for clades on the ML tree were inferred with Mr. Bayes (ver. 2.0; Huelsenbeck and Ronquist 2001) using the GTR + Γ (Yang 1993) model, a random starting tree, and a uniform prior probability. The analysis included three chains (two hot and one cold), which ran for a million generations, with trees saved every 1000 generations. We determined the plateau point of the chains at 4000 generations by graphing the weighted maximum likelihood scores obtained, and discarding the four trees saved before the plateau. A 95% majority-rule consensus tree was calculated from the set of remaining trees, and the group frequencies were used as posterior probability estimates.

Congruence Testing Between *PepC* and *waxy*. Previous studies comparing molecular data sets of the Triticeae uncovered incongruence among some of the gene trees, suggesting that portions of the nuclear genome have different histories (Kellogg et al. 1996; Mason-Gamer and Kellogg 1996b). We tested whether the *PepC* data were congruent with sequence data from a 1.3-kb portion of the single-copy nuclear *waxy* gene, previously obtained for the same members of the Triticeae (Mason-Gamer and Kellogg 2000; Mason-Gamer 2001). The *waxy* and *PepC* loci are believed to be on

different chromosomes. The *waxy* gene is found on the group 7 chromosomes in the Triticeae (Devos et al. 1995; Kleinohs 1997), or on a portion of chromosome 4 translocated from, and thus homoeologous to, the group 7 chromosomes (Devos et al. 1995; Korzun et al. 1997). The *PepC* gene is hypothesized here to be on the long arm of chromosome 5 (see Materials and Methods, "Molecular Data").

We tested congruence between the *PepC* and *waxy* data sets using the Shimodaira-Hasegawa (SH; Shimodaira and Hasegawa 1999) test with a specific model of evolution as implemented in PAUP* ver. 4.0b10 (Swofford 2002). Prior to analyses, the *waxy* and *PepC* data sets were reduced to 29 taxa, which matched exactly to the level of species (Table 1). Two tree topologies were evaluated (*PepC* and *waxy*), obtained from MP analyses of each data partition following the same protocol used to analyze the full *PepC* data set, except that only 100 random addition replicates were run. We determined the appropriate model of sequence evolution for both the *waxy* and *PepC* data partitions, and calculated which parsimony tree had the highest ML scores using the same procedure employed to analyze the complete *PepC* data set. The two highest-scoring trees (one from each data partition) were used as the two test trees.

Two separate SH tests were conducted. The ML scores for each test tree were estimated first under the model of sequence evolution best fitting the *PepC* data, and then under the model that best fit the *waxy* data. In both cases, the log-likelihood differences between the trees were determined. The first of these was compared to the RELL bootstrap distribution of log-likelihood differences generated using the *PepC* data and ML parameters, while the second was compared to the distribution generated using the *waxy* data and ML parameters. If the score difference between two trees fell within 95% of the appropriate RELL-generated scores in a one-tailed comparison, the estimated *p*-value was not considered significant at the 0.05 level. If both SH tests resulted in significant *p*-values, the trees were assumed to be incongruent.

Next, we used constraint analyses to examine the source of conflict between the *PepC* and *waxy* trees. Initially, branch support for the parsimony trees was estimated following the methods used in the analysis of the full *PepC* data set. Then, we constructed single-node constraint trees for each clade with bootstrap support above 50% in the reduced *PepC* and *waxy* trees. This yielded eight constraint trees based on the *PepC* tree and ten based on the *waxy* tree. Finally, we completed a series of constrained MP heuristic searches using stepwise addition, MULTREES, ACCTRAN optimization, and TBR branch swapping. We imposed each of the eight *PepC* constraint trees in MP searches of the *waxy* data, and each of the ten *waxy* constraint trees in MP searches of the *PepC* data. If the constrained trees were longer, it suggested that the corresponding nodes are sources of conflict between the data sets.

RESULTS

Sequence Length and Variation. Complete DNA sequences were generated for 27 species. The 1.7-kb portion of the *PepC* gene examined spans exons 1–4; however, the majority of the sequence data (80%) are derived from introns 1–3. Compared to the portion of the *waxy* gene used by Mason-Gamer (2001) for the same set of taxa (Table 3), the *PepC* gene is approximately 350 bp longer, has considerably more intron characters, and provides more parsimony-informative characters (although the percentage of parsimony-informative characters is somewhat lower). There is no evidence that either the **H** or the **St** copy of the *PepC* gene is non-functional; there are no length changes or stop codons in the exons examined.

All of the *Elymus* individuals examined have two distinct *PepC* gene variants. One (the **H** form) groups

TABLE 3. Data characteristics. ¹ Approximately half of region 2 is missing for *Secale montanum* subsp. *montanum*.

	<i>PepC</i>	<i>PepC</i> Reduced	<i>waxy</i>
Aligned Length	2185	2126	1530
Sequence Length—average (minimum/maximum)	1573 (1418–1693)	1587 (1429–1710)	1232 (1191–1299)
% Gap Characters—average (minimum/maximum)	26.7 (21.1–33.9)	25.4 (19.6–32.8)	19.5 (15.2–22.2)
% Missing Characters—average (minimum/maximum)	1.4 (0–33.1 ¹)	0.6 (0.0–4.1)	0.01 (0.0–0.1)
% Exon	19.5	20.1	49.9
GC Content	43.0 (42.1–44.4)	43.0 (42.1–43.9)	58.3 (56.8–59.5)
Parsimony Informative Characters—% (#)	9.7 (211)	9.0 (192)	12.2 (187)
Number of Parsimony Trees (steps)	3 (750)	6 (674)	30 (763)
Consistency Index/Retention Index	0.610/0.882	0.631/0.814	0.478/0.656

with members of *Hordeum* sect. *Critesion*, and the second (the **St** form) groups within *Pseudoroegneria*. However, for two *Elymus* individuals, we only have a partial sequence of one of the two forms: for *E. canadensis* we have **St** sequences from region 2 but not from region 1, and for *E. glaucus* we have an **H** copy from region 2 but not from region 1. Because approximately 60% of the nucleotides for these sequences are absent, they were not included in the phylogenetic analyses; only the **H** copy is included for *E. canadensis*, and only the **St** copy is included for *E. glaucus*.

Phylogenetic Analyses of *PepC* Data. Parsimony analysis of the *PepC* data yielded three most-parsimonious trees (750 steps; CI excluding uninformative characters = 0.610; RI = 0.822). The only topological difference between the three trees was the placement of two taxa within the *Elymus* + **H** clade (*E. lanceolatus* and *H. jubatum*). However, once nodes with zero-length branches were collapsed, there was only one tree, in which *E. lanceolatus* and *H. jubatum* were part of a polytomy within the *Elymus* + **H** clade. The final ML search resulted in a single tree (−7444.62; Fig. 2; see Materials and Methods, “Phylogeny Reconstruction” for estimated parameters).

The ML and parsimony trees have the same general topology (Fig. 2). In each, *E. californicus* is the first to diverge, followed by *Leymus racemosus*. Next, *Hordeum* forms a weakly supported clade sister to a second weakly supported clade containing the remaining members of the tribe. Many of the other nodes resolved in both the parsimony and ML trees are either incongruent or weakly supported, but every clade with high bootstrap (BS) or posterior probabilities (PP) was recovered by both methods of analysis.

The *PepC* data suggest that the *Elymus* tetraploids combine the **H** and **St** genomes. There is a strongly supported clade consisting of *Hordeum jubatum*, *H. brachyantherum*, and *Elymus* (BS = 98 and PP = 100). A second strongly supported group includes *Elymus* and *Pseudoroegneria spicata* (BS and PP = 100). *Elymus californicus* is the only *Elymus* species that does not group with *Hordeum* or *Pseudoroegneria*, but is found near the base of the tree. There are four other lineages with high bootstrap support and posterior probabili-

ties: 1) an *Agropyron cristatum* and *Eremopyrum orientale* clade (BS and PP = 100); 2) an *Aegilops comosa*, *Triticum aestivum*, and *Taeniatherum caput-medusae* clade (BS = 86 and PP = 100); 3) a *Secale* L. clade (BS = 97 and PP = 100) and 4) a *Hordeum murinum* and *H. vulgare* L. clade, sister to the *Elymus* + **H** lineage (BS = 83 and PP = 99).

Comparison of *PepC* and *waxy*. Sequence characteristics for the reduced *PepC* and *waxy* data sets are shown in Table 3. Parsimony analysis of the *PepC* data partition yielded six most-parsimonious trees (674 steps; CI excluding uninformative characters = 0.631, RI = 0.814), including two tree islands of three trees each. The strict consensus of all six trees is less resolved than the strict consensus from the entire *PepC* data set, but has the same general topology and well-supported nodes (Fig. 3). Among the six most-parsimonious trees, tree #6 had the highest log likelihood score and was chosen as the *PepC* tree for use in the SH test. Once again the log likelihood value obtained under the (GTR+I+ Γ) model was not significantly different than the one obtained using (GTR+ Γ) alone (−6887.34 vs. −6887.70), so we used the simpler model. Figure 3 depicts parsimony tree #6 under the ML parameters best fitting the *PepC* data: [R (AC) = 0.98, R (AG) = 2.83, R (AT) = 0.43, R (CG) = 1.33, R (CT) = 3.42, R (GT) = 1.00, pi (A) = 0.25, pi (C) = 0.20, pi (G) = 0.21, pi (T) = 0.33, α = 0.52].

Parsimony analysis of the reduced *waxy* data yielded 30 shortest trees (763 steps, CI excluding uninformative characters = 0.478, RI = 0.656). The strict consensus of these trees (Fig. 4) is similar to the strict consensus trees of tree islands 3 and 4 in Mason-Gamer's (2001) analysis of *waxy* data. All most-parsimonious *waxy* trees include a single well-supported *Hordeum* + *Elymus* clade (BS = 89). Within this group, the *Elymus* **H** sequences form a well-supported clade with *H. brachyantherum* (BS = 100), while *H. marinum*, *H. murinum*, *H. vulgare*, and *H. jubatum* form a weakly-supported clade. The *Elymus* and *Pseudoroegneria* sequences are grouped into two distinct, well-supported clades (**St1** and **St2**), but the **St** sequences do not form a monophyletic group. The **St1** clade (BS = 100) is sister to a weakly supported clade including *Aegilops* L.,

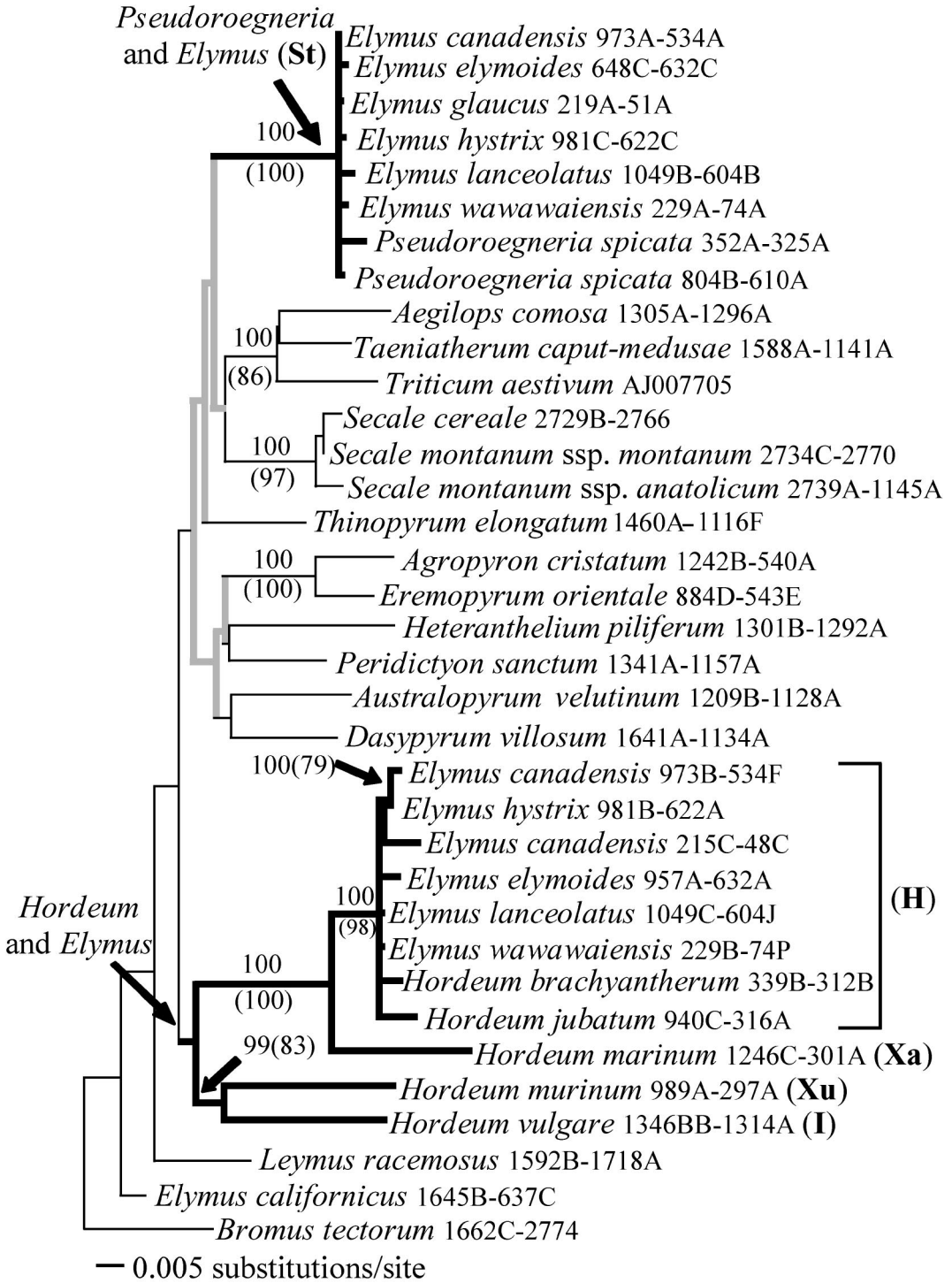


FIG. 2. The maximum likelihood tree based on *PepC* sequence data, generated under a GTR+ Γ model of evolution ($L = -7444.62$). Bold lines indicate the *Elymus* + *Pseudoroegneria* and *Elymus* + *Hordeum* clades. Genome designations are provided for *Elymus*, *Pseudoroegneria*, and *Hordeum* taxa. The posterior probabilities are above the branches and bootstrap values ($\geq 75\%$) are given in parentheses. Gray lines represent portions of the tree that are inconsistent with the parsimony results. *PepC* clone numbers are provided next to taxon names.

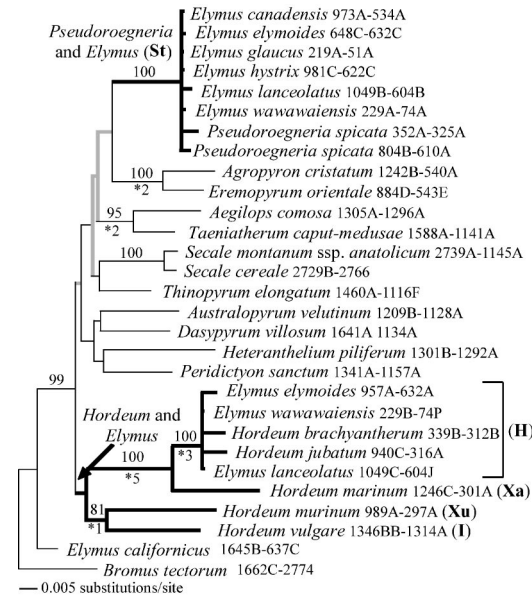


FIG. 3. The reduced *PepC* parsimony topology used as the test tree in the SH test, depicted using ML estimated parameters best fitting the *PepC* data under GTR+Γ. Bold lines indicate the *Elymus* + *Pseudoroegneria* and *Elymus* + *Hordeum* clades. Bootstrap values greater than 49% are above the nodes. Nodes that are sources of conflict in the constrained analyses are marked with asterisks and the resulting increase in tree length in each of the associated constrained analyses is given. Gray lines represent nodes that collapse in the strict consensus of all most parsimonious trees. Taxon labels as in Fig. 2.

Dasypyrum (Coss. & Durieu) P. Candargy, and *Thinopyrum* Å.Löve, and **St2** (BS = 98) is sister to *Australopyrum velutinum*. As in the *PepC* tree, *E. californicus* does not group with other *Elymus* species. The two *Secale* species form a strongly supported clade (BS = 100) sister to *Heteranthelium piliferum* (BS = 68). This group is in turn sister to a clade including *Agropyron cristatum*, *Eremopyrum orientale*, and *Taeniatherum caput-medusae* (BS = 68). Among the 30 most-parsimonious trees, tree #1 had the highest log likelihood score and was chosen as the *waxy* tree for use in the SH test. The most parameter rich model (GTR+I+Γ) had a significantly higher log likelihood value than the other models tested, and was used in the SH test. Figure 4 depicts parsimony tree #1 under the ML parameters best fitting the *waxy* data: [R (AC) = 1.50, R (AG) = 3.53, R (AT) = 1.36, R (CG) = 1.66, R (CT) = 4.87, R (GT) = 1.00, pi (A) = 0.22, pi (C) = 0.26, pi (G) = 0.30, pi (T) = 0.21, P-inv. = 0.35, α = 0.71].

Both SH tests yielded significant *p*-values. The SH test using the *PepC* data and associated ML parameters resulted in a difference in tree scores of 551.94, *P* < 0.001 (*PepC* score = -6887.70; *waxy* score = -7439.64). The SH test using the *waxy* data and associated ML parameters resulted in a difference in tree scores of

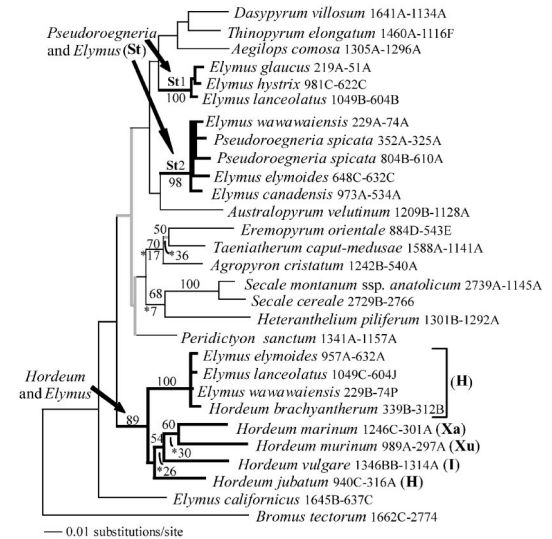


FIG. 4. The reduced *waxy* parsimony tree used as the test tree in the SH test, depicted using ML estimated parameters best fitting the *waxy* data under GTR+I+Γ. Bold lines indicate the *Elymus* + *Pseudoroegneria* and *Elymus* + *Hordeum* clades. Bootstrap values above 49% are above the nodes. Nodes that are sources of conflict in the constrained analyses are marked with asterisks and the resulting increase in tree length in each of the associated constrained analyses is given. Gray lines represent nodes that collapse in the strict consensus of all most parsimonious trees. Taxon labels as in Fig. 2.

397.26, *P* < 0.001 (*waxy* score = -5982.25; *PepC* score = -6379.50).

In constrained analyses of the *PepC* data, five of the ten nodes tested from the *waxy* tree (Fig. 4, asterisks) resulted in an increase in length relative to the unconstrained *PepC* trees (674 steps). In constrained analyses of the *waxy* data, five of the eight *PepC* tree nodes (Fig. 3, asterisks) resulted in increases in length over the unconstrained *waxy* trees (763 steps). The length increases associated with each node are shown in Figs. 3 and 4.

DISCUSSION

Evidence of St + H Genome Content. All tetraploid *Elymus* individuals examined here have two copies of the *PepC* isoform sequenced in the study. One copy forms a clade with *Hordeum* and the second with *Pseudoroegneria*, confirming that North American *Elymus* species do have an **St** + **H** genomic content, and therefore are allotetraploid derivatives of *Hordeum* and *Pseudoroegneria* (Fig. 2). The data thus confirm Dewey's (1984) and Löve's (1984) genomic concept of North American *Elymus*, in which all species (with the exception of the octoploid *E. californicus*) are allotetraploids containing the **St** and **H** genomes (Dewey 1984; Löve 1984). Some systematists have raised justifiable concerns regarding the use of genomic pairing data as the

primary criterion for both grouping and ranking of taxa in the Triticeae (e.g., Baum et al. 1987; Kellogg 1989; Seberg and Petersen 1998). However, the *PepC* phylogeny corroborates the cytogenetic findings, and is in agreement with previous studies based on both isozyme (Jaaska 1995, 1998) and *waxy* gene sequence data (Mason-Gamer 2001). A classification system based on chromosome pairing provides a reasonable starting point, in that it groups taxa based on a character that appears to reflect evolutionary relatedness. However, its utility is limited because it cannot provide insights into the hierarchical relationships among the taxa examined (e.g., Kellogg 1989).

Elymus californicus—is it a Member of *Elymus*?

The octoploid *Elymus californicus* is the only *Elymus* species in the present study in which the **St** and **H** genomes have not been detected. This is not surprising, because the classification of *E. californicus* has been unstable since it was first described in 1880, and its genomic makeup remains uncertain (Löve 1984; Mary Barkworth, pers. comm.). In an attempt to gain insights into the genomic complement of this enigmatic species, Jensen and Wang (1997) screened for the **St** and **Ns** genomes using genome-specific RAPD markers. Their study found no evidence of the **St** genome, but did detect a marker specific to the **Ns** genome, which is found in the diploid genus *Psathyrostachys* Nevski and in the allopolyploid genus *Leymus* Hochst. From this information Jensen and Wang (1997) suggested that *E. californicus* should be transferred to *Leymus*. The *PepC* data neither refute nor confirm these findings directly, placing *Elymus californicus* at the base of the tree. Nevertheless, the *PepC* phylogeny clearly demonstrates that *E. californicus* does not belong within *Elymus*. The *waxy* gene phylogeny by Mason-Gamer et al. (2001) was also consistent with Jensen and Wang's (1997) findings, revealing a weak association between *Psathyrostachys* and *E. californicus* in most of the most-parsimonious trees. Within the chloroplast DNA phylogeny by Mason-Gamer et al. (2002) the placement of *E. californicus* is equivocal; *E. californicus* is weakly grouped with *Hordeum* in some of the shortest trees. The *PepC* data add to the mounting evidence suggesting that *E. californicus* does not belong within *Elymus*, but additional sampling and phylogenetic analyses are needed to develop a firm understanding of how this species is related to other members of the Triticeae.

Elymus H and St Genomes. The *PepC* data indicate that all North American tetraploid species of *Elymus* have **St** + **H** genome content, but there is little resolution among *Elymus* individuals in either the **St** or the **H** clade. This pattern is consistent with the hypothesis that the *Elymus* tetraploids are formed from a single set of genome donors or multiple closely related sets of donors, but the specific progenitor species involved remain a mystery. The *Elymus H* genome sequences

form a strongly supported clade with the two North American *Hordeum* taxa included in the study (*H. brachyantherum* and *H. jubatum*), while the *Elymus St* sequences form an unresolved polytomy with *P. spicata*, the only North American *Pseudoroegneria* species (Fig. 2). The lack of resolution among the *Elymus St* copies is similar to that previously observed among chloroplast DNA sequences (Mason-Gamer et al. 2002). In a study based on the *waxy* gene (Mason-Gamer et al. 2001), **St** sequences formed two distinct clades, which did not correspond to geographical or morphological differentiation. It was unclear whether the pattern was the result of allelic variation or recent gene duplication prior to the divergence of the **St** group. The *PepC* and *waxy* studies have provided hints on which *Hordeum* species may be most closely related to *Elymus*, but additional sampling of both *Hordeum* and *Pseudoroegneria* is needed before we can develop a better picture of the origin of the **H** and **St** genomes.

Distinction Among Hordeum Genomes. A detailed discussion of *Hordeum* evolution in light of the *PepC* phylogeny is impossible given the limited number of taxa sampled. However, the data do resolve the four *Hordeum* genomes (**H**, **Xa**, **I**, and **Xu**), and provide some insights into how these groups are related (Fig. 2). Taxa with genomes **I** (*Hordeum vulgare*) or **Xu** (*Hordeum murinum*) form a strongly supported monophyletic group, which in turn is sister to a second well-supported clade that includes species with the **H** or **Xa** (*Hordeum maritimum*) genomes. This corroborates the study by Komatsuda et al. (1999) of *Hordeum* diploid taxa using *Vrs1* nuclear intron sequence data. The *PepC* results are also in accordance with Petersen and Seberg's (2003) analysis of diploid *Hordeum* taxa, in which they combined the chloroplast data of Doebley et al. (1992) and Komatsuda et al. (1999) with data from the nuclear marker *DMC1* (Petersen and Seberg 2002, 2003) and the chloroplast *rbcl* gene (Petersen and Seberg 1998). Nuclear and chloroplast DNA data were partly contradictory, but in the combined data set, the **Xu** and **I** genomes formed a sister group to the **H** and **Xa** taxa, although one **H** genome taxon (*H. brevisubulatum* Link.) was sister to *H. maritimum* (**Xa**). Data from the nuclear ribosomal internal transcribed spacers (Frank Blattner, pers. comm.) also recovered four major clades corresponding to the four *Hordeum* genomes, but the arrangement of the clades was somewhat different. Genomes **H** + **Xa** formed a clade sister to **Xu**, and this clade is sister to **I**, though the placement of the **Xu** clade is weakly supported. It is clear from these *Hordeum* studies that the genome groups are in fact phylogenetically informative.

Conflict Between PepC and waxy. Results from the SH test suggest that the *PepC* and *waxy* trees are significantly incongruent. Constrained analyses confirm these findings suggesting that there is disagreement

involving some of the clades in the *PepC* and *waxy* phylogenies (Figs. 3, 4). There are two cases within *Hordeum* where taxa have different positions in the two trees, but are members of clades with moderate to high bootstrap support. The first is *Hordeum marinum*: in the *PepC* tree, it is sister to the **H** clade, while in the *waxy* tree; it is sister to *H. murinum* (**Xu**). In the second case, the same accession of *H. jubatum* has a different position in the *PepC* and *waxy* trees. In the *PepC* phylogeny, it is nested within the well-supported **H** clade, while in the *waxy* phylogeny; it is found at the base of a grade including *H. marinum* (**Xa**), *H. murinum* (**Xu**), and *H. vulgare* (**I**). While the reason for the conflicting placement of *H. jubatum* is not known, it is an allotetraploid (e.g., Jørgensen 1986; Bothmer et al. 1987), and the two data sets may have sampled gene copies representing the two different, homoeologous genomes.

The placement of *Taeniatherum caput-medusae* is also at odds between the *PepC* and *waxy* phylogenies (Figs. 3, 4). In the *PepC* tree, *T. caput-medusae* forms a well-supported clade (BS = 95) with *Aegilops comosa*, while in the *waxy* tree it is sister to *Eremopyrum orientale* (BS = 50), which in turn is sister to *Agropyron cristatum* (BS = 70). The association seen in the *waxy* topology has not been seen in other gene trees, but a close relationship between *T. caput-medusae* and *Aegilops* was seen in the chloroplast DNA phylogeny (Mason-Gamer et al. 2002). The *PepC* data strongly conflict with the *waxy* data on the placement of *Taeniatherum*, with 36 additional steps required in the constrained analysis. The *waxy* data show far less conflict with *Taeniatherum*'s placement in the *PepC* tree, requiring only two additional steps in the constrained analysis. There may be some underlying signal within the *waxy* data set that supports a more congruent placement of *Taeniatherum*. In the *PepC* phylogeny, *E. orientale* forms a strongly supported clade with *A. cristatum* (BS = 100) rather than *Taeniatherum*, in agreement with the chloroplast DNA phylogeny (Mason-Gamer et al. 2002), and the *DMC1* phylogeny (Petersen and Seberg 2002), in which *Agropyron cristatum* was grouped with two other *Eremopyrum* species, *E. distans* (K.Koch) Nevski and *E. triticeum* (Gaertn.) Nevski.

Secale is well-supported on both trees (BS = 100), which is not surprising. A correlation between monogenic groups and clades on molecular phylogenetic trees has been previously reported (Hsiao et al. 1995; Kellogg and Appels 1995; Kellogg et al. 1996; Mason-Gamer and Kellogg 1996a, 1996b; Petersen and Seberg 1997). In the *waxy* tree, the *Secale* clade is sister to *Heteranthelium piliferum*, forming a weakly supported clade (BS = 68); this association was previously supported in Mason-Gamer's chloroplast DNA phylogeny.

Phylogenetic conflict was previously demonstrated among Triticeae gene trees, with the most extensive differences documented in comparisons between chlo-

roplast and nuclear data (Kellogg et al. 1996; Mason-Gamer and Kellogg 1996b). There were a few instances of statistically significant incongruence among the nuclear data sets, and the present comparison between *PepC* and *waxy* shows conflict involving several taxa. If the conflict is genuine, the nuclear gene trees may reflect differences among phylogenies of the chromosomes, or sections of chromosomes, on which the genes reside. The discordance between *PepC* and *waxy* gene tree topologies (as well as many others) highlights the need for continued research on chromosome evolution in the Triticeae in the hopes of better understanding mechanisms behind the complex historical patterns that are emerging from phylogenetic studies of this group.

Two different historical scenarios provide possible explanations for the reticulate history of the tribe (Kellogg et al. 1996). First, the pattern may result from periods of cladogenesis and dispersal, interspersed with hybridization and introgression between species. The level of hybridization slowed or stopped at some point in the past; most genera in the tribe are intersterile, and hybrids among them generally exhibit a lack of chromosome pairing (e.g., Dewey 1984 and references therein). However, some degree gene exchange among genera is possible, if a small percentage of viable hybrid seeds are produced through chance alone. It is also possible that present-day gene flow between diploid taxa is taking place through a polyploid intermediate. Gene flow between diploids via a tetraploid bridge requires the formation of partially fertile triploids, as has been reported for species of *Aegilops* by Vardi (1973), Feldman et al. (1979), and Kimber and Feldman (1987). There is also phylogenetic evidence of hybridization or introgression. Mason-Gamer and Kellogg (2000) reported that an accession of *Pseudoroegneria strigosa* (M. Bieb.) Á.Löve, which was expected to be a monogenic **St** species, has both **St** and **H** copies of the *waxy* gene and an **St** chloroplast. The result was consistent with both the allotetraploid-bridge and direct hybridization hypotheses.

A second possible explanation for the reticulate history of the Triticeae is that it reflects a burst of diversification with rapid lineage sorting, and fixation of different alleles into lineages, which quickly became intersterile (Kellogg et al. 1996). Given the need for extensive ancestral polymorphism in this hypothesis, it seems unlikely that rapid diversification and lineage sorting was the major driving force behind the reticulate history of the Triticeae. However, it is not clear how we could determine which scenario (lineage sorting and hybridization-introgression) has had the bigger impact on the history of the Triticeae. In fact, these hypotheses are not mutually exclusive, and it may be that both processes have played some role in shaping the history of the Triticeae.

In conclusion, this study verifies Dewey's (1982, 1983) and Löve's (1982, 1984) insightful chromosome-based revisions of the Triticeae. The *PepC* results clearly show that North American tetraploid members of *Elymus* have an **St + H** genomic content, and are therefore derived from *Hordeum* and *Pseudoroegneria*. However, the current data reveal little genetic differentiation within North American *Elymus*. Consequently, the relationships among *Elymus* species, and their relationships to specific progenitor species, await a better-resolved phylogeny. Additional studies, which integrate morphological, cytological, and phylogenetic information, are necessary to understand better the historical relationships within *Elymus*. The *PepC*-based analysis of North American *Elymus* is a small step forward in this process.

ACKNOWLEDGEMENTS. The authors wish to thank Mary Barkworth for sharing plant material, the USDA's Germplasm Resources Information Network for seeds, Gregory Plunkett and two anonymous reviewers for helpful comments on the manuscript. The National Science Foundation (DEB-9974181) funded this work.

LITERATURE CITED

- BARKWORTH, M. E. and D. R. DEWEY. 1985. Genomically based genera in the perennial Triticeae of North America: identification and membership. *American Journal of Botany* 72: 767–776.
- BAUM, B. R., J. R. ESTES, and P. K. GUPTA. 1987. Assessment of the genomic system of classification in the Triticeae. *American Journal of Botany* 74: 1388–1395.
- BESNARD, G., B. OFFMANN, C. ROBERT, and C. ROUCH. 2002. Assessment of the C_4 phosphoenolpyruvate carboxylase gene diversity in grasses (Poaceae). *Theoretical and Applied Genetics* 105: 404–412.
- BOTHMER, R. VON, J. FLINK, and T. LANDSTRÖM. 1987. Meiosis in *Hordeum* interspecific hybrids. II. Triploid hybrids. *Evolutionary Trends in Plants* 1: 41–50.
- CRÉTIN, C., S. SANTI, E. KERYER, L. LEPINIEC, D. TAGU, J. VIDAL, and P. GADAL. 1991. The phosphoenolpyruvate carboxylase gene family in *Sorghum*: promoter structure, amino acid sequences and expression of genes. *Gene* 99: 87–94.
- CRONN, R. C., R. L. SMALL, T. HASELKORN, and J. F. WENDEL. 2003. Cryptic repeated genomic recombination during speciation in *Gossypium gossypoides*. *Evolution* 57: 2475–2489.
- , X. ZHAO, A. H. PATERSON, and J. F. WENDEL. 1996. Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. *Journal of Molecular Evolution* 42: 685–705.
- CRONQUIST, A. 1987. A botanical critique of cladism. *Botanical Review* 53: 1–52.
- CUSHMAN, J. C. and H. J. BOHNERT. 1989a. Nucleotide sequence of the Ppc2 gene encoding a housekeeping isoform of PEPC from *Mesembryanthemum crystallinum*. *Nucleic Acids Research* 17: 6743–6744.
- and ———. 1989b. Nucleotide sequence of the gene encoding a CAM specific isoform of PEPC from *Mesembryanthemum crystallinum*. *Nucleic Acids Research* 17: 6745–6746.
- DAVIS, J. I. and R. J. SORENG. 1993. Phylogenetic structure in the grass family (Poaceae), as determined from chloroplast DNA restriction site variation. *American Journal of Botany* 80: 1444–1454.
- DEVOS, K. M., J. DUBCOVSKY, J. DVORAK, C. N. CHINOY, and M. D. GALE. 1995. Structural evolution of wheat chromosomes 4A, 5A, and 7B and its impact on recombination. *Theoretical and Applied Genetics* 91: 282–288.
- DEWEY, D. R. 1982. Genomic and phylogenetic relationships among North American perennial Triticeae. Pp. 51–88 in *Grasses and grasslands*, eds. J. R. Estes, R. J. Tyrl, and J. N. Brunken. Norman: Oklahoma University Press.
- . 1983. Historical and current taxonomic perspectives of *Agropyron*, *Elymus* and related genera. *Crop Science* 23: 637–642.
- . 1984. The genomic system of classification as a guide to intergeneric hybridization within the perennial Triticeae. Pp. 209–279 in *Gene manipulation in plant improvement, Proceedings of the 6th Stadler genetics symposium*, ed. J. P. Gustafson. New York: Columbia University Press.
- DOEBLEY, J., R. VON BOTHMER, and S. LARSON. 1992. Chloroplast DNA variation and the phylogeny of *Hordeum* (Poaceae). *American Journal of Botany* 79: 576–584.
- DOYLE, J. J. and J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.
- , ———, A. H. D. BROWN, and R. G. PALMER. 2002. Genomes, multiple origins, and lineage recombination in the *Glycine tomentella* complex: histone H3-D gene sequences. *Evolution* 56: 1388–1402.
- EMSHWILLER, E. and J. J. DOYLE. 1998. Origins of domestication and polyploidy in oca (*Oxalis tuberosa*: Oxalidaceae): nrDNA ITS data. *American Journal of Botany* 85: 975–985.
- and ———. 2002. Origins of domestication and polyploidy in oca (*Oxalis tuberosa*: Oxalidaceae). 2. Chloroplast-expressed glutamine synthetase data. *American Journal of Botany* 89: 1042–1056.
- ERNST K. and P. WESTHOFF. 1997. The phosphoenolpyruvate carboxylase (ppc) gene family of *Flaveria trinervia* (C_4) and *F. pringlei* (C_3): molecular characterization and expression analysis of the ppcB and ppcC genes. *Plant Molecular Biology* 34: 427–443.
- FELDMAN, M., I. STRAUSS, and A. VARDI. 1979. Chromosome pairing and fertility of F_1 hybrids of *Ae. longissima* and *Ae. searsii*. *Canadian Journal of Genetics and Cytology* 21: 261–272.
- FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.
- . 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783–791.
- FRATI F., C. SIMON, J. SULLIVAN, and D. L. SWOFFORD. 1997. Evolution of the mitochondrial cytochrome oxidase II gene in *Collembola*. *Journal of Molecular Evolution* 44: 145–158.
- GALE, M. D. and K. M. DEVOS. 1998. Plant comparative genetics after ten years. *Genome* 282: 656–659.
- GE, S., T. SANG, B.-R. LU, and D.-Y. HONG. 1999. Phylogeny of rice genomes with emphasis on origins of allotetraploid species. *Proceedings of the National Academy of Sciences, USA* 96: 14400–14405.
- GEHRIG, H. H., V. HEUTE, and M. KLUGE. 2001. New partial sequences of phosphoenolpyruvate carboxylase as molecular phylogenetic markers. *Molecular Phylogenetics and Evolution* 20: 262–274.
- GOLDMAN, N. 1993. Statistical tests of models of DNA substitution. *Journal of Molecular Evolution* 36: 182–198.
- GONZALEZ, M. C., L. OSUNA, C. ECHEVARRIA, J. VIDAL, and F. J. CEJUDO. 1998. Expression and localization of phosphoenolpyruvate carboxylase in developing and germinating wheat grains. *Plant Physiology* 116: 1249–1258.
- GU, X., Y.-X. FU, and W.-H. LI. 1995. Maximum likelihood estimation of the heterogeneity of substitution rate among nucleotide sites. *Molecular Biology and Evolution* 12: 546–557.
- HASEGAWA, M., H. KISHINO, and T. YANO. 1985. Dating of the human-ape split by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution* 22: 160–174.

- HENDY, M. D. and D. PENNY. 1982. Branch and bound algorithms to determine minimal evolutionary trees. *Mathematical Biosciences* 59: 277–290.
- HITCHCOCK, A. S. 1951. *Manual of the grasses of the United States*, ed. 2, revised by A. Chase. USDA Miscellaneous Publications 200. Washington, D. C.: U. S. Government Printing Office.
- HSHAO, 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.
- HUDSPETH, R. L. and J. W. GRULA. 1989. Structure and expression of the maize gene encoding the phosphoenolpyruvate carboxylase isozyme involved in C₄ photosynthesis. *Plant Molecular Biology* 12: 579–589.
- HUELSENBECK, P. H. and K. A. CRANDALL. 1997. Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecology and Systematics* 28: 437–466.
- and B. RANNALA. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* 276: 227–232.
- and F. RONQUIST. 2001. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754–755.
- HULL, D. L. 1979. The limits of cladism. *Systematic Zoology* 28: 416–440.
- IZUI, K., S. ISHIJIMA, Y. YAMAGUCHI, F. KATAGIRI, T. MURATA, K. SHIGESADA, T. SUGIYAMA, and H. KATUKI. 1986. Cloning and sequence analysis of cDNA encoding active PEPC of the C₄ photosynthesis in maize. *Nucleic Acids Research* 14: 1615–1628.
- JAASKA, V. 1995. Isozyme data on the diploid progenitors of allotetraploid *Elymus* species. Pp. 165–168 in *Proceedings of the 2nd International Triticeae Symposium*, eds. R. R.-C. Wang, K. B. Jensen, and C. Jauss. Logan, Utah: USDA Forage and Range Laboratory.
- . 1998. Isoenzyme data on the origin of North American allotetraploid *Elymus* species. Pp. 209–216 in *Triticeae III*, ed. A. A. Jaradat. Enfield, New Hampshire: Science Publishers Inc.
- JENSEN, R. and R.-C. WANG. 1997. Cytological and molecular evidence for transferring *Elymus coreanus* from the genus *Elymus* to *Leymus* and molecular evidence for *Elymus californicus* (Poaceae: Triticeae). *International Journal of Plant Sciences* 158: 872–877.
- JØRGENSEN, R. B. 1986. Relationships in the barley genus (*Hordeum*): an electrophoretic examination of proteins. *Hereditas* 104: 273–291.
- JUKES, T. H. and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 in *Mammalian protein metabolism*, ed. H. N. Munro. New York: Academic Press.
- KAWAMURA, T., K. SHIGESADA, H. TOH, S. OKUMURA, S. YANAGISAWA, and K. IZUI. 1992. Molecular evolution of phosphoenolpyruvate carboxylase for C₄ photosynthesis in maize: comparison of its cDNA sequence with a newly isolated cDNA encoding an isozyme involved in the anaerobic function. *Journal of Biochemistry* 112: 147–154.
- KELLOGG, E. A. 1989. Comments on generic genera in the Triticeae (Poaceae). *American Journal of Botany* 76: 796–805.
- . 1992. Tools for studying the chloroplast genome in the Triticeae (Gramineae): an EcoRI map, a diagnostic deletion, and support for *Bromus* as an outgroup. *American Journal of Botany* 79: 186–197.
- and R. APPELS. 1995. Intraspecific and interspecific variation in 5S RNA genes are decoupled in diploid wheat relatives. *Genetics* 140: 325–343.
- , ———, and R. J. MASON-GAMER. 1996. When gene trees tell different stories: the diploid genera of Triticeae (Gramineae). *Systematic Botany* 21: 321–347.
- KIMBER, G. and M. FELDMAN. 1987. *Wild wheat: an introduction*. Columbia: College of Agriculture, University of Missouri.
- KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* 16: 111–120.
- KLEINHOF, A. 1997. Integrating barley RFLP and classical marker maps. *Barley Genetics Newsletter* 27: 105–112.
- KOMATSUDA T., K.-I. TANNO, B. SALOMON, T. BRYNGELSSON, and R. VON BOTHMER. 1999. Phylogeny in the genus *Hordeum* based on nucleotide sequences closely linked to the *ovs1* locus (row number of spikelets). *Genome* 42: 973–981.
- KORZUN, V., S. MALYSHEV, A. VOYLOKOV, and A. BÖRNER. 1997. RFLP-based 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.
- LEPINIEC, L., E. KERYER, H. PHILIPPE, P. GADAL, and C. CRÉTIN. 1993. *Sorghum* phosphoenolpyruvate carboxylase gene family: structure, function and molecular evolution. *Plant Molecular Biology* 21: 487–502.
- , ———, D. TAGU, P. GADAL, and C. CRÉTIN. 1992. Complete nucleotide sequence of a *Sorghum* gene coding for the phosphoenolpyruvate carboxylase involved in C₄ photosynthesis. *Plant Molecular Biology* 19: 339–342.
- , S. SANTI, E. KERYER, V. AMIET, J. VIDAL, P. GADAL, and C. CRÉTIN. 1991. Complete nucleotide sequence of one member of the *Sorghum* phosphoenolpyruvate carboxylase gene family. *Plant Molecular Biology* 17: 1077–1079.
- , J. VIDAL, R. CHOLLET, P. GADAL, and C. CRÉTIN. 1994. Phosphoenolpyruvate carboxylase: structure, regulation and evolution. *Plant Science* 99: 111–124.
- LOVE, Á. 1982. Generic evolution of the wheatgrasses. *Biologisches Zentralblatt* 101: 199–212.
- . 1984. Conspectus of the Triticeae. *Feddes Repertorium* 95: 425–521.
- MADDISON, W. P. and D. R. MADDISON. 2002. MacClade, ver. 4.0: Analysis of phylogeny and character evolution. Sunderland: Sinauer Associates.
- MALCOMBER, S. T. 2002. Phylogeny of *Gaertnera* Lam. (Rubiaceae) based on multiple DNA markers: evidence of a rapid radiation in a widespread, morphologically diverse genus. *Evolution* 56: 42–57.
- MASON-GAMER, R. J. 2001. Origin of North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. *Systematic Botany* 26: 757–768.
- . 2004. Reticulate evolution, introgression, and intertribal gene capture in an allohexaploid grass. *Systematic Biology* 53: 25–37.
- and E. A. KELLOGG. 1996a. Chloroplast DNA analysis of the monogenomic Triticeae: phylogenetic implications and genome-specific markers. Pp. 301–325 in *Methods of genome analysis in plants*, ed. P. P. Jauhar. Boca Raton: CRC Press.
- and ———. 1996b. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Systematic Biology* 45: 524–545.
- and ———. 2000. Phylogenetic analysis of the Triticeae using the starch synthase gene, and a preliminary analysis of some North American *Elymus* species. Pp. 102–109 in *Grass systematics and evolution*, eds. S. W. L. Jacobs and J. Everett. Melbourne: CSIRO Publishing.
- , 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.
- , C. F. WEIL, and E. A. KELLOGG. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molecular Biology and Evolution* 15: 1658–1673.
- MCDADE, L. A. 1992. Hybrids and phylogenetic systematics II. The impact of hybrids on cladistic analysis. *Evolution* 46: 1329–1346.

- . 1997. Hybrids and phylogenetic systematics III. Comparison with distance methods. *Systematic Botany* 22: 669–683.
- PETERSEN G. and O. SEBERG. 1997. Phylogenetic analysis of the Triticeae (Poaceae) based on *rpoA* sequence data. *Molecular Phylogenetics and Evolution* 7: 217–230.
- and ———. 1998. Molecular studies on the phylogeny of the genus barley (*Hordeum*; Poaceae). Pp. 437–440 in *Molecular tools for screening biodiversity: plants and animals*, eds. A. Karp, P. G. Isaac and D. S. Ingram. New York: Chapman and Hall.
- and ———. 2002. Molecular evolution and phylogenetic application of *DMC1*. *Molecular Phylogenetics and Evolution* 22: 43–50.
- and ———. 2003. Phylogenetic analyses of the diploid species of *Hordeum* (Poaceae) and a revised classification of the genus. *Systematic Botany* 28: 293–306.
- RAUSCHER, J. T., J. J. DOYLE, and A. H. D. BROWN. 2002. Internal transcribed spacer repeat-specific primers and the analysis of hybridization in the *Glycine tomentella* (Leguminosae) polyploid complex. *Molecular Ecology* 11: 2691–2702.
- ROHDE, W., D. BECKER, and F. SALAMINI. 1988. Structural analysis of the *waxy* locus from *Hordeum vulgare*. *Nucleic Acids Research* 16: 7185–7186.
- SANDERSON, M. J. 1998. Estimating rate and time in molecular phylogenies; beyond the molecular clock. Pp. 242–264 in *Molecular systematics of plants II. DNA sequencing*, eds. D. E. Soltis, P. S. Soltis, and J. J. Doyle. Boston: Kluwer Academic Publishers.
- and M. J. DONOGHUE. 1989. Patterns of variation in levels of homoplasmy. *Evolution* 43: 1781–1795.
- SANG, T. and D. ZHANG. 1999. Reconstructing hybrid speciation using sequences of low copy nuclear genes: hybrid origins of five *Paeonia* species based on *Adh* gene phylogenies. *Systematic Botany* 24: 148–163.
- SEBERG, O. and G. PETERSEN. 1998. A critical review of concepts and methods used in classical genome analysis. *The Botanical Review* 64: 372–417.
- SEELANAN, T., A. SCHNABEL, and J. F. WENDEL. 1997. Congruence and consensus in the cotton tribe (Malvaceae). *Systematic Botany* 22: 259–290.
- SHIMODAIRA, H. and M. HASEGAWA. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* 16: 1114–1116.
- SMALL, R. L. and J. F. WENDEL. 2000. Phylogeny, duplication, and intraspecific variation of *Adh* sequences in New World diploid cottons. *Molecular Phylogenetics and Evolution* 16: 73–84.
- 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 synthase (GBSSI) gene sequences. *Systematic Biology* 52: 374–385.
- SULLIVAN, 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. *Systematic Biology* 46: 426–440.
- SVENSSON, P., O. E. BLÄSING, and P. WESTHOFF. 2003. Evolution of C_4 phosphoenolpyruvate carboxylase. *Archives of Biochemistry and Biophysics* 414: 180–188.
- SWOFFORD, D. L. 2002. PAUP* Phylogenetic analysis using parsimony (*and other methods). version 4b10. Sunderland: Sinauer Associates.
- , G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phylogenetic inference. Pp. 407–514 in *Molecular systematics*, 2nd ed., eds. D. M. Hillis, C. Moritz, and B. K. Mable. Sunderland: Sinauer Associates.
- THOMPSON, J. D., D. G. HIGGINS, and T. J. GIBSON. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673–4680.
- VARDI, A. 1973. Introgression between different ploidy levels in the wheat group. Pp. 131–141 in *Proceedings of the Fourth International Wheat Genetics Symposium*, eds. E. R. Sears, and L. M. S. Sears. Columbia: University of Missouri Press.
- WADDELL, P. J. and D. PENNY. 1996. Evolutionary trees of apes and humans from DNA sequences. Pp. 53–73 in *Handbook of symbolic evolution*, eds. A. J. Locke and C. R. Peters. Oxford: Clarendon Press.
- 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. *Systematic Biology* 43: 250–261.
- YANG, Z. 1993. Maximum likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites: approximate methods. *Molecular Biology and Evolution* 10: 1396–1401.
- . 1994. Estimating the pattern of nucleotide substitution. *Journal of Molecular Evolution* 39: 105–111.