Systematic Botany (2001), 26(4): pp. 757–768

Origin of North American *Elymus* (Poaceae: Triticeae) Allotetraploids Based on Granule-bound Starch Synthase Gene Sequences

ROBERTA J. MASON-GAMER

University of Idaho, Department of Biological Sciences, Moscow, Idaho 83844-3051;
Present Address: Department of Biological Sciences, University of Illinois at Chicago, M/C 066,
845 W. Taylor St., Chicago, Illinois 60607 (robie@uic.edu)

Communicating Editor: John V. Freudenstein

ABSTRACT. The current circumscription of *Elymus* based on cytogenetic analyses includes all allopolyploid Triticeae species containing the **St** (*Pseudoroegneria*) genome. In North American *Elymus*, the **St** genome is combined with **H** (from *Hordeum*) in an allotetraploid (**StStHH**) configuration. The goal of this study is to determine whether molecular phylogenetic analyses support existing cytogenetic data with regard to the evolutionary origin of North American *Elymus*. Analyses were performed using sequences from the nuclear starch synthase gene, and include multiple species of *Elymus*, *Pseudoroegneria*, and *Hordeum*, along with representatives of most of the other monogenomic genera in the Triticeae. The results support the hypothesis that *Pseudoroegneria* and *Hordeum* are the genome donors to the North American *Elymus* tetraploids. One species currently placed in *Elymus* (an octoploid, *Elymus californicus*) appears to be unrelated to the rest. The close relationships among *Elymus* **H**-genome sequences, with one exception, are consistent with a single origin of the group. The **St**-genome group consists of two well-defined clades, but support for the monophyly of the entire **St** group is weak. There are shortcomings associated with the current dependence on genome pairing data for grouping and ranking in the Triticeae, but molecular phylogenetic data suggest that, in many cases, groups delimited by cytogenetic data do in fact correspond to evolutionary lineages.

The origins of allopolyploid taxa and their relationships to their diploid relatives have been extensively addressed using cytogenetic and morphological characteristics. Molecular phylogenetic techniques are becoming well-established means for confirming, clarifying, or discovering the polyploid nature of plant species, and for testing hypotheses about their evolution and origins. Phylogenetic analyses of nuclear DNA markers are proving particularly valuable for examining the origins of allopolyploid species (e.g., Doyle et al. 1984, 2000; Soltis and Soltis 1991; Small et al. 1998; Ge et al. 1999).

The wheat tribe Triticeae (Poaceae) includes many different auto- and allopolyploid taxa. Data from cytogenetic analyses have been used extensively to study the evolutionary history of the tribe. In these studies, relationships among individuals are inferred from patterns of chromosome pairing during meiosis in the pollen mother cells of interspecific hybrids. A high degree of pairing is assumed to indicate high overall similarity between sets of chromosomes, and in turn, a close relationship between the corresponding taxa. A lesser ability of chromosomes to pair is assumed to reflect lower overall sequence similarity, and a more distant relationship. Cytogenetic techniques have also been widely used to clarify the ancestry of the many polyploid taxa in the group. When a polyploid of unknown genomic constitution is hybridized to a series of diploids and/or polyploids with known genomes, the patterns of meiotic pairing in the hybrids are assumed to reveal the genomic constitution of the unknown polyploid.

In the Triticeae, genome analysis has strongly influenced the taxonomy of the group (Dewey 1982, 1984; Löve 1982, 1984). Under the genomic system of classification, species are grouped into genera based on the ability of their chromosomes to pair at meiosis in interspecific hybrids. Genera are thus circumscribed as units between which hybrids exhibit incomplete or no chromosome pairing.

The genomic system of classification, however, has been opposed on many grounds. From the perspective of a field biologist, the system can be troublesome to apply. It is difficult to distinguish genomic genera using morphological characteristics, whereas the traditional genera were defined by a few convenient, obvious features. Because of these difficulties, some North American agrostologists largely disregard genomic classification, and instead follow Hitchcock's (1951) classification, which closely follows that of Bentham (Bentham 1881; Bentham and Hooker 1883; see Barkworth 2000 for a summary of the history of Triticeae classification). While genera in the genomic system are not distinguished by obvious characters, the cytogenetic data are, in fact, corroborated by a variety of other morphological and physiological characters (Barkworth 2000 and references therein).

For phylogeneticists, most of whom believe that classification should reflect phylogenetic relationships among taxa, the genomic method of classification raises additional questions, recently reviewed in detail by Seberg and Petersen (1998). Some concerns are based on practical questions about methodology, scoring and analysis of data, and interpretation of results. Other questions relate to the biology of chromosome pairing, and raise doubts about the correlation between degree

of chromosome pairing and overall sequence similarity (and, in turn, relatedness). Finally, there are philosophical problems regarding the use of a single continuous character for both grouping and ranking of taxa. The question underlying many of these concerns is whether patterns of chromosome pairing can be assumed to reflect phylogenetic relationships. Most proponents of genomic classification agree that classification should be based on phylogeny (e.g., Dewey 1982, 1984; Löve 1982), and assume that genomic pairing data do reflect phylogenetic relationships, but this assumption has been questioned (e.g., Seberg and Petersen 1998).

Because genome-pairing data form the basis of the generic classification in the Triticeae, it is important to address whether genome groups based on chromosome pairing data reflect the phylogeny as estimated from other sources of data. Molecular phylogenetic studies focusing on diploid, monogenomic members of the tribe (Kellogg and Appels 1995; Hsiao et al. 1995; Mason-Gamer and Kellogg 1996a, 1996b; Kellogg et al. 1996; Petersen and Seberg 1997) do, in many cases, support the monophyly of the genomic genera.

In this paper, I examine a complex of North American allopolyploid species in the genus *Elymus* L. within a phylogenetic context. Because the name *Elymus* has been applied in different ways, the traditional use of the name (i.e., Hitchcock 1951) will be compared briefly to its use within a genomic system of classification (Barkworth and Dewey 1985). This summary includes only North American species except where indicated.

In the traditional classification, most of the North American perennial members of the Triticeae other than *Hordeum* L. (wild barley) are separated into two genera based on the number of spikelets at each inflorescence node: *Agropyron* Gaertn. (one spikelet per node) and *Elymus* (two to six spikelets per node). Cytogenetic analyses (e.g., Dewey 1982, 1983a, 1983b, 1984) of North American species later suggested that:

(1) Most *Hordeum* species share a genome class, designated **H**, but a few species have divergent genomes, designated **I**, **Xa**, and **Xu**;

(2) one *Agropyron* species, *A. spicatum* (Pursh) Scribn. and J.G.Smith, is monogenomic, with genome **St**;

(3) most of the other *Agropyron* species and about half of the *Elymus* species are allotetraploid, combining the **St** and **H** genomes;

(4) the remaining *Elymus* species are also allotetraploid, but with a different genome complement (**NsNsXX**, derived from *Psathyrostachys* Nevski and an unknown donor); and

(5) a few *Agropyron* species, introduced to North America, are monogenomic for the **P** genome. (Other introduced species, with various genomic complements, were also traditionally placed in *Agropyron*, but their taxonomic fates in the genomic classification system are outside the scope of this summary.)

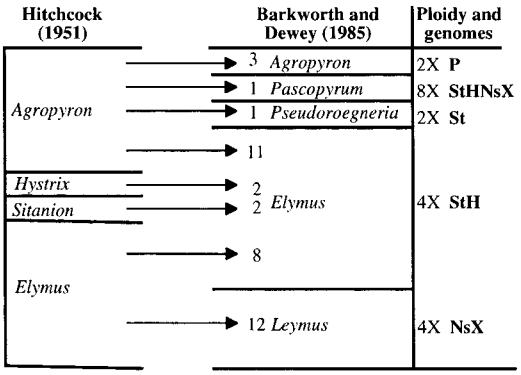

FIG. 1. Summary of major taxonomic differences between the morphological classification used in Hitchcock (1951) and the genomic genera of Barkworth and Dewey (1985). Numbers show approximate numbers of species.

Clearly, the cytogenetic results were inconsistent with the traditional classification. The resulting nomenclatural changes (Barkworth and Dewey 1985) are summarized here and in Fig. 1:

(1) the definition of *Hordeum* has remained largely unchanged;

(2) *A. spicatum* has been placed in *Pseudoroegneria* (Nevski) Á.Löve (= *P. spicata* (Pursh) Á.Löve);

(3) Allotetraploids with an **StStHH** genome combination are placed in *Elymus*, which therefore includes most of the former *Agropyron* and about half of the former *Elymus*;

(4) the remaining (**NsNsXX**) tetraploids formerly in *Elymus* have been placed in *Leymus* Hochst.; and

(5) *Agropyron* is restricted to include only monogenomic **P**-genome species, none of which are native to North America.

In this paper, the name *Elymus* is used in the sense of the genomic classification. Under this definition, all native North American species of *Elymus* are **StStHH** allotetraploids. The present work focuses on the phylogenetic relationships among *Hordeum* (monogenomic **H**, **I**, **Xu**, and **Xa**), *Pseudoroegneria* (monogenomic **St**), and North American *Elymus* (allotetraploid **StStHH**) within the context of the entire tribe. I address whether phylogenetic analyses of molecular data support studies of chromosome pairing with regard to the evolutionary origin of *Elymus*.

MATERIALS AND METHODS

Taxa. Starch synthase sequences were sampled from eight North American and Old World species of *Hordeum*, from the sole North American *Pseudoroegneria* species, and from nine North American species of *Elymus*. Most of the other monogenomic gen-

	No rate variation among sites	A proportion of sites (I) invariable	Rate variation among sites follows a gamma (Γ) distribution	A proportion of sites (I) invariable; rate variation among others follows a gamma (Γ) distribution
All substitutions equally probable; nucleotides occur in equal frequencies (JC)	JC	JC + I	JC + Γ	JC + I + Γ
Transistions, transversions differ; nucleotides occur in equal frequencies (K2P)	K2P	K2P + I	K2P + Γ	K2P + I + Γ
Transistions, transversions differ; nucleotides differ in frequency (HKY)	HKY	HKY + I	HKY + Γ	HKY + I + Γ
Substitution types (A/C, A/G, A/T, C/G, C/T, G/T) differ; nucleotides differ in frequency (GTR)	GTR	GTR + I	GTR + Γ	GTR + I + Γ

FIG. 2. Models of sequence evolution examined prior to ML searches. Substitution models include Jukes-Cantor (JC; Jukes and Cantor 1969), Kimura two-parameter (K2P; Kimura 1980), Hasagawa-Kishino-Yano (HKY; Hasagawa et al. 1985), and general time reversible (GTR; Yang 1994a). Models of among-site rate variation incorporate a proportion of invariable sites (I) with no additional rate variation (Hasagawa et al. 1985), rate variation among sites following a gamma distribution (Γ; Yang 1994b), and a combination that includes some proportion of invariable sites, with variation at the remaining sites assumed to follow a gamma distribution (I + Γ; Gu et al. 1995; Waddell and Penny 1996).

era in the Triticeae are represented to provide a broader phylogenetic context to the relationships between *Elymus*, *Hordeum*, and *Pseudoroegneria* (Table 1).

DNA Sequencing. Analyses were carried out using a 1.3 kb portion of the granule-bound starch synthase gene. The gene appears to be single-copy in grasses (Shure et al. 1983; Klosgen et al. 1986; Rohde et al. 1988; Clark et al. 1991; Wang et al. 1995), and has been used for phylogenetic studies of the grass family (Mason-Gamer et al. 1998) and of the Triticeae (Mason-Gamer and Kellogg 2000).

DNA samples were prepared from fresh, frozen, or dried leaf tissue using the methods of Doyle and Doyle (1987). Amplification reactions (10 μl) were carried out as in Mason-Gamer et al. (1998) using primers F-for (TGCGAGCTCGACAACATCATGCG) and M-bac (GGCGAGCGGCGCGATCCCTCGCC). PCR products were checked on agarose minigels and cleaned with Geneclean (Bio101) according to manufacturer's instructions. Cleaned PCR products were cloned into pGEM-T Easy vectors (Promega). Because *Elymus* species are presumed allopolyploids, PCR reactions from *Elymus* were run in triplicate and combined in an attempt to offset the potential effects of PCR drift, in which random events during early PCR cycles lead to an over-representation of one gene copy (Wagner et al. 1994). The three reactions were checked on a gel, and then mixed and cleaned for use in ligation reactions. Ligation reactions, transformations, and plating were carried out according to the instructions for the Promega pGEM-T Easy cloning kit, except that the final volumes of the ligation and transformation reactions were halved. Plasmids were isolated using Qiagen or Promega miniprep kits, digested with *Eco*RI, and checked on agarose gels.

Sequencing reactions were carried out using ABI BigDye terminators, and were run on an ABI 377 sequencer according to manufacturer's instructions except that sequencing reaction volumes were halved or quartered. In *Elymus*, four cloned PCR products were screened with the M-bac primer for sequences representing both genomes. If only one genome was represented, four additional clones were screened. Based on the results of screening, cloned sequences were fully sequenced with the F-for and M-bac primers along with H2-for (GAGGCCAAGGCGCTGAACAAGG), J-bac (ACGTCGGGGCCCTTCTGCTC), L1-for (GCAAGACCGGGTTCCACATGG), and L2-bac (CGCTGAGGCGGCCCATGTGG) (Mason-Gamer et al. 1998). Sequences from each individual were edited and joined in Sequencher version 3.0 (Genecodes Corporation). Sequences were aligned in Clustal V (Higgins et al. 1992) with gap open penalty = 10, gap extension penalty = 5, and equal weighting of transition and transversions. The resulting alignment was further adjusted by eye. Regions judged to be ambiguous in terms of alignment were excluded prior to phylogenetic analyses.

Phylogenetic Analyses. All phylogenetic analyses were performed using PAUP* 4.0b4a (Swofford 2000). Cladistic parsimony analyses were carried out with equal weighting of all characters, except those in ambiguously aligned regions. The possibility of multiple tree islands was explored using 1000 random addition replicates. For each replicate, 10 trees were held at each step of the stepwise addition to obtain the starting tree.

Ten trees were randomly chosen from each of four resulting islands for maximum likelihood (ML) estimation of parameters of sequence evolution. Several parameters, including nucleotide frequency, probabilities of different substitution types, and rate variation among sites were estimated for 16 models of sequence evolution on each of the trees using ML (Fig. 2; e.g., Swofford et al. 1996; Frati et al. 1997; Sullivan et al. 1997). Before a model was chosen for a ML search using the parameters estimated above, the appropriateness of the gamma distribution as a model of among-site rate variation was tested. A histogram of number of characters vs. number of steps was generated with MacClade 3.08 (Maddison and Maddison 1992), using the parsimony tree that gave the best likelihood score. A gamma distribution was fitted to the histogram using the program GAMMA (Sullivan et al. 1995), and compared to the histogram to determine whether the observed distribution of number of characters vs. number of steps differs significantly from a gamma distribution. The appropriateness of the incorporation of estimated base frequencies in a ML search was assessed by testing for stationarity of nucleotide frequencies among taxa in PAUP* 4.0 (Swofford 2000). Following these tests, estimated model parameters were used as settings in subsequent ML searches under the Kimura two-parameter (K2P; Kimura 1980) model with among-site rate variation assumed to follow a gamma distribution (Yang 1994b).

RESULTS

Starch Synthase Sequences. The cloned 1.3 kb portion of the starch synthase gene includes two partial exons, four full exons, and five introns (Mason-Gamer et al. 1998). Length changes are common in the introns, such that portions of several introns are difficult to align across the tribe; characters in ambiguous regions were excluded from phylogenetic analyses. There are

TABLE 1. List of taxa. Numbers following *Elymus* names are identifiers of specific individuals, and correspond to the numbers in Figs. 5 and 6. Vouchers stored at GH or ID. [1]Mason-Gamer et al. (1998); [2]Rohde et al. (1988).

Sample	No.	Collection	Genbank
Hordeum, Elymus, Pseudoroegneria			
Hordeum brevisubulatum (Trin.) Link	—	PI401387	AY010961
Hordeum bulbosum L.	—	PI440417	AY010962
Hordeum californicum Covas & Stebbins[1]	—	MA-138-1-40	AF079273
Hordeum jubatum L.	—	RJMG106	AY010963
Hordeum marinum Hudson	—	PI304346	AY010959
Hordeum murinum L.	—	CIho15683	AY010960
Hordeum violaceum Boiss. & Hohen	—	PI401390	AY010964
Hordeum vulgare L.[2]	—		X07932
Elymus californicus (Bolander) Gould	1a	MEBarkworth s.n.	AY011012
Elymus elymoides (Refin.) Swezey	1d2	PI531606	AY010965
	1a2		AY010992
Elymus glaucus Buckley	4b	RJMG130	AY010966
	4a		AY010979
Elymus glaucus	6a	W6-10215	AY010967
	6b		AY010980
Elymus glaucus	7b		AY010968
	7a	PI593652	AY010981
Elymus hystrix L.	1a		AY010982
	1c	MEBarkworth 97-87	AY010983
Elymus lanceolatus (Scribn. & Smith) Gould	1aa	W614220	AY010969
	1d		AY010984
	1a		AY010993
Elymus lanceolatus	2a	W614218	AY010970
	2c		AY010985
	2d		AY010994
Elymus riparius Wiegand	1a	RJMG160	AY010971
Elymus trachycaulus (Link) Gould ex Shinners	1a1	PI372500	AY010972
	1b1		AY010986
Elymus trachycaulus	3a	PI452446	AY010973
	3b		AY010974
	3d		AY010987
Elymus virginicus L.	4a	RJMG161	AY010975
	4d		AY010995
Elymus virginicus	5a	RJMG162	AF331953
	5b		AY010988
Elymus virginicus	9b	RJMG168	AY010976
	9a		AY010989
Elymus wawawaiensis J.Carlson ex Barkworth	1a	PI285272	AY010977
	1c		AY010996
Elymus wawawaiensis	3d	PI598812	AY010978
	3a		AY010990
	3b		AY010997
Pseudoroegneria spicata (Pursh) Á.Löve subsp. *spicata*[1]	1	PI232117	AF079281
Pseudoroegneria spicata subsp. *inermis* (Scribn. and J.G.Smith) Á.Löve	2	PI236681	AY010998
Pseudoroegneria spicata subsp. *spicata*	3	PI610986	AY010999
Pseudoroegneria spicata subsp. *spicata*	4	D2844	AY011000
Pseudoroegneria spicata subsp. *spicata*	6c	RJMG112	AY010991
	6a		AY011001
Other Representatives of Triticeae			
Aegilops caudata L.[1]	—	G758	AF079262
Aegilops speltoides Tausch[1]	—	Morrison s.n.	AF079267
Aegilops tauschii Coss.[1]	—	Morrison s.n.	AF079268
Aegilops uniaristata Vis.[1]	—	G1297	AF079270
Agropyron cristatum (L.) Gaertn.[1]	1	C-3-6-10	AF079271
Agropyron cristatum	2	PI281862	AY011002
Agropyron mongolicum Keng	—	D2774	AY011003
Australopyrum retrofractum (Vickery) Á.Löve[1]	—	Crane 86146	AF079272
Australopyrum velutinum (Nees) B.K.Simon	—	D2873-2878	AY011004
Dasypyrum villosum (L.) Candargy	—	PI251478	AF079274

TABLE 1. Continued.

Sample	No.	Collection	Genbank
Eremopyrum bonaepartis (Spreng.) Nevski	—	H5554	AY011005
Eremopyrum distans (C.Koch) Nevski	—	H5552	AY011006
Eremopyrum orientale (L.) Jaub. & Spach	—	H5555	AY011007
Henrardia persica (Boiss.) C.E.Hubb.[1]	—	H5556	AF079276
Heteranthelium piliferum (Banks & Sol.) Hochst.[1]	—	PI402352	AF079277
Peridictyon sanctum (Janka) Seberg, Fred. & Baden[1]	—	KJ248	AF079278
Psathyrostachys fragilis (Boiss.) Nevski[1]	—	C-46-6-15	AF09279
Psathyrostachys juncea (Fisch.) Nevski	—	PI206684	AF079280
Secale cereale L.	—	Kellogg s.n.	AY011009
Secale montanum Guss.[1]	—	PI440654	AF079282
Secale montanum subsp. *anatolicum* (Boiss.) Tzelev	—	PI206991	AY011008
Taeniatherum caput-medusae (L.) Nevski	—	MB-106-41-79	AY011010
Thinopyrum bessarabicum (Savul. & Rayss) Á.Löve[1]	—	PI531711	AF079283
Thinopyrum elongatum (Host) D.R.Dewey[1]	—	PI531719	AF079284
Thinopyrum scirpeum (C.Presl) D.R.Dewey	—	C-15-21-25	AY011011
Triticum baeoticum Boiss.[1]	—	Morrison s.n.	AF079285
Triticum urartu Tumanian[1]	—	Morrison s.n.	AF079287

five small deletions in exons, all of which are autapomorphic.

Observation of the aligned sequences reveals expected differences in rates of change among nucleotide sites, i.e., differences between introns and exons, and among first, second, and third positions in codons. The distribution of the number of characters vs. number of changes per character closely fits a gamma distribution ($P < 0.025$) with shape parameter $\alpha = 0.435$ (Fig. 3), and a test for stationarity of nucleotide frequencies among taxa does not reject stationarity ($P > 0.9999$).

Phylogenetic Analyses. Estimation of likelihood

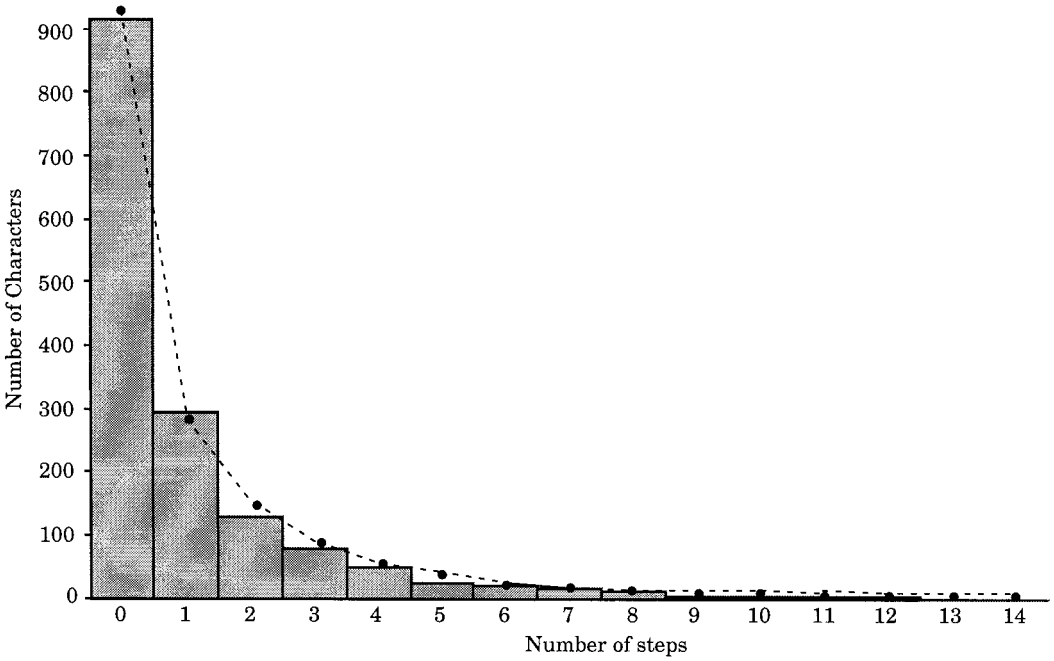

FIG. 3. Histogram of number of characters and number of steps per character estimated on one of the most parsimonious trees. Dots above bars represent points on a gamma-distribution ($\alpha = 0.435$) fitted to the data used to generate the histogram.

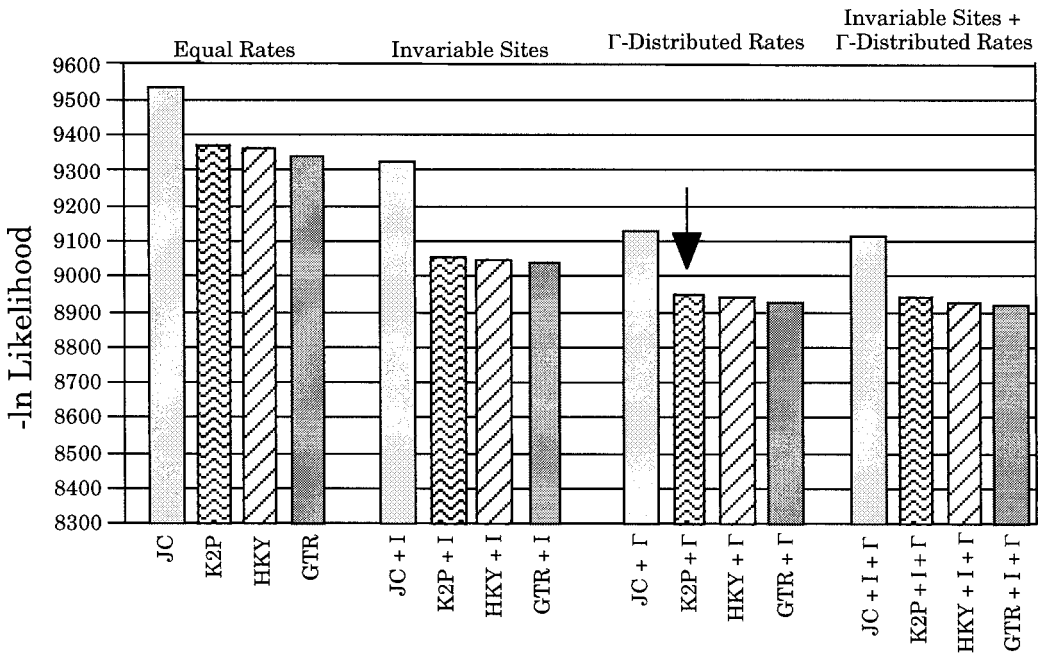

FIG. 4. Negative log-likelihood values estimated on one of the most-parsimonious trees. Lower values indicate a better fit of the model to the data. The arrow indicates the K2P + Γ model, used to generate the ML tree in Fig. 5.

scores under 16 models of sequence evolution (Fig. 2) shows, as expected, that the model incorporating the most parameters, GTR + I + Γ, best fits the data (Fig. 4). After consideration of the tradeoff between the better fit vs. the greater computational intensity of the most parameter-rich models, the simpler K2P + Γ model was chosen for a subsequent ML search. Although some of the other models result in better likelihood scores for the input trees, the two parameters accounting for nearly all of the improvement in score are differences in rates of transitions vs. transversions, and Γ-distribution of among-site rate variation (Fig. 4). These are the two parameters defining the K2P + Γ model (Fig. 2). This time-based tradeoff seems reasonable, because models that incorporate rate variation appear to be robust to uncertainties regarding the model itself or the parameter estimates (Sullivan and Swofford, in press). The parameter values for the search, estimated on the input trees using maximum likelihood, were fixed at a transition/transversion ratio of 1.409 with a gamma shape parameter of 0.391.

The resulting tree (Fig. 5) includes a monophyletic *Hordeum* + *Elymus* clade. Most of the remaining *Elymus* sequences group with *Pseudoroegneria* in a weakly-supported **St**-genome clade. This clade includes two relatively well-supported clades (''**St**1'' and ''**St**2'' in Fig. 5). *Elymus californicus* does not group with other *Elymus* species.

The cladistic parsimony analysis of 1000 random addition sequences yields 3172 most-parsimonious trees (MPTs) representing four tree islands with 2208, 828, 102, and 34 trees. Trees from islands 1 and 2 (not shown) are very similar to the ML tree in Fig. 5, with the main difference being the placement of *E. lanceolatus* 2a with *H. murinum* 2a, rather than with *H. jubatum* 1a. Based on parsimony bootstrap results, this difference is not well-supported.

Like the ML tree (Fig. 5), all MPTs include a single well-supported *Hordeum* + *Elymus* clade. The *Elymus* and *Pseudoroegneria* sequences are, as in the ML tree, grouped into two distinct clades (**St**1 and **St**2), but trees from different islands disagree as to whether these two clades are together monophyletic. On the trees from islands 1 and 2, **St**1 and **St**2 form a very weakly supported (42% bootstrap) **St** clade, but they are separated on the trees from islands 3 and 4 (Fig 6). Islands 1 and 2 also differ from islands 3 and 4 with regard to the monophyly of *Psathyrostachys*, here designated as an outgroup. (Differences between islands 1 and 2, and between islands 3 and 4, are trivial with regard to the questions addressed here: in islands 2 and 4, *Eremopyrum orientale* and *Agropyron mongolicum* group with other members of their respective genera, while in islands 1 and 3, they group with one another.) In all MPTs, as in the ML tree, *Elymus californicus* is the only *Elymus* species that does not group with *Pseudoroegneria* or *Hordeum*. Its exact placement within the Triticeae, however, is not resolved by these data.

DISCUSSION

Based on cytogenetic data, North American species of *Elymus* (other than *E. californicus*) are always tetra-

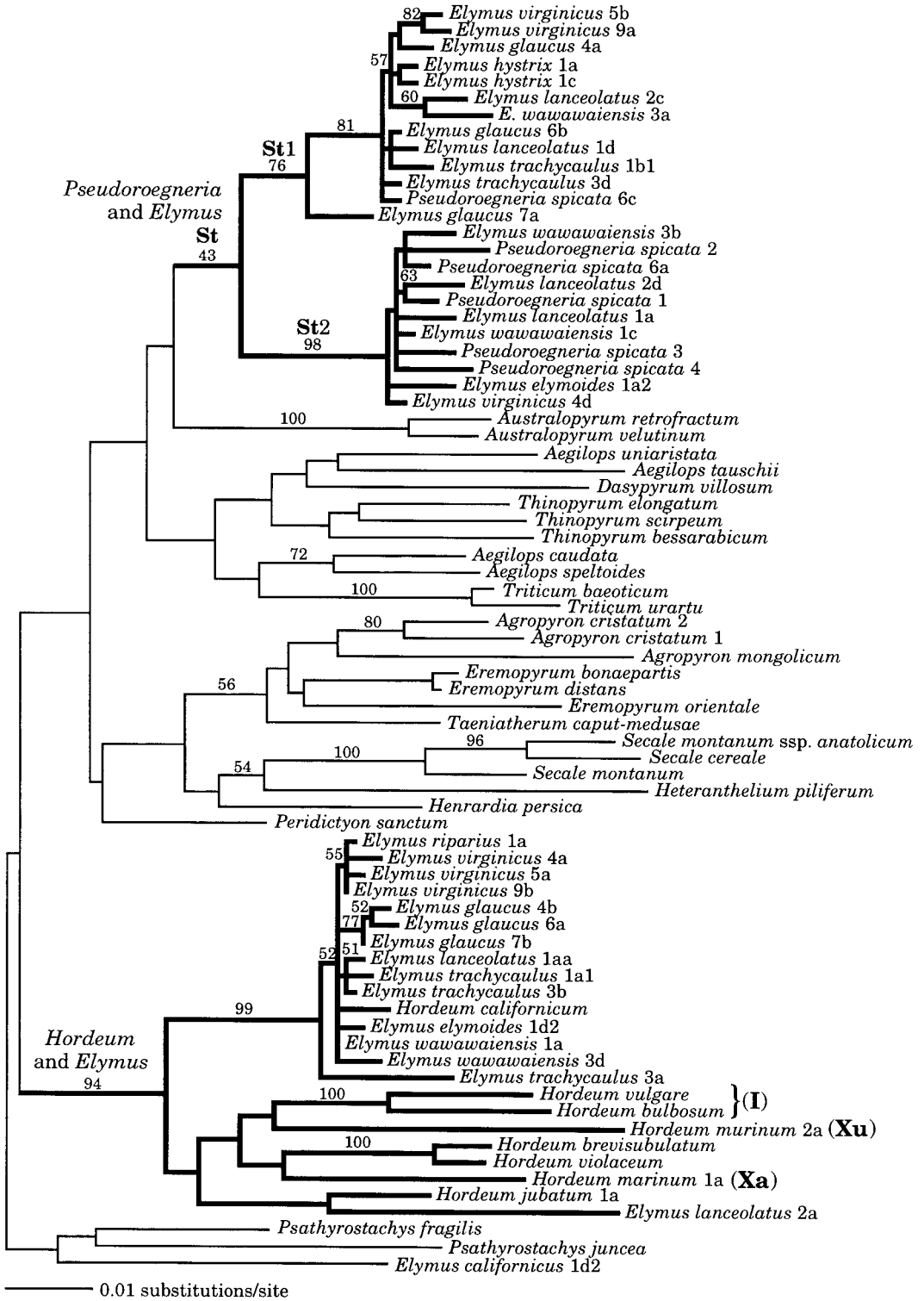

FIG. 5. Maximum likelihood tree estimated under the K2P + Γ model. Parameters estimated using ML: transition/transversion ratio = 1.409 and gamma shape parameter α = 0.391. Tree score: -ln likelihood = 8965.3094. Bold lines indicate the *Elymus*/*Pseudoroegneria* and *Elymus*/*Hordeum* clades. Genome designations are provided for those *Hordeum* species with genomes other than **H.** Parsimony bootstrap values greater than 50% are shown above nodes. Numbers following taxon labels are identifiers of specific plants, and correspond to the numbers in Table 1. Letters following plant identifier numbers indicate specific PCR-product clones from within individuals.

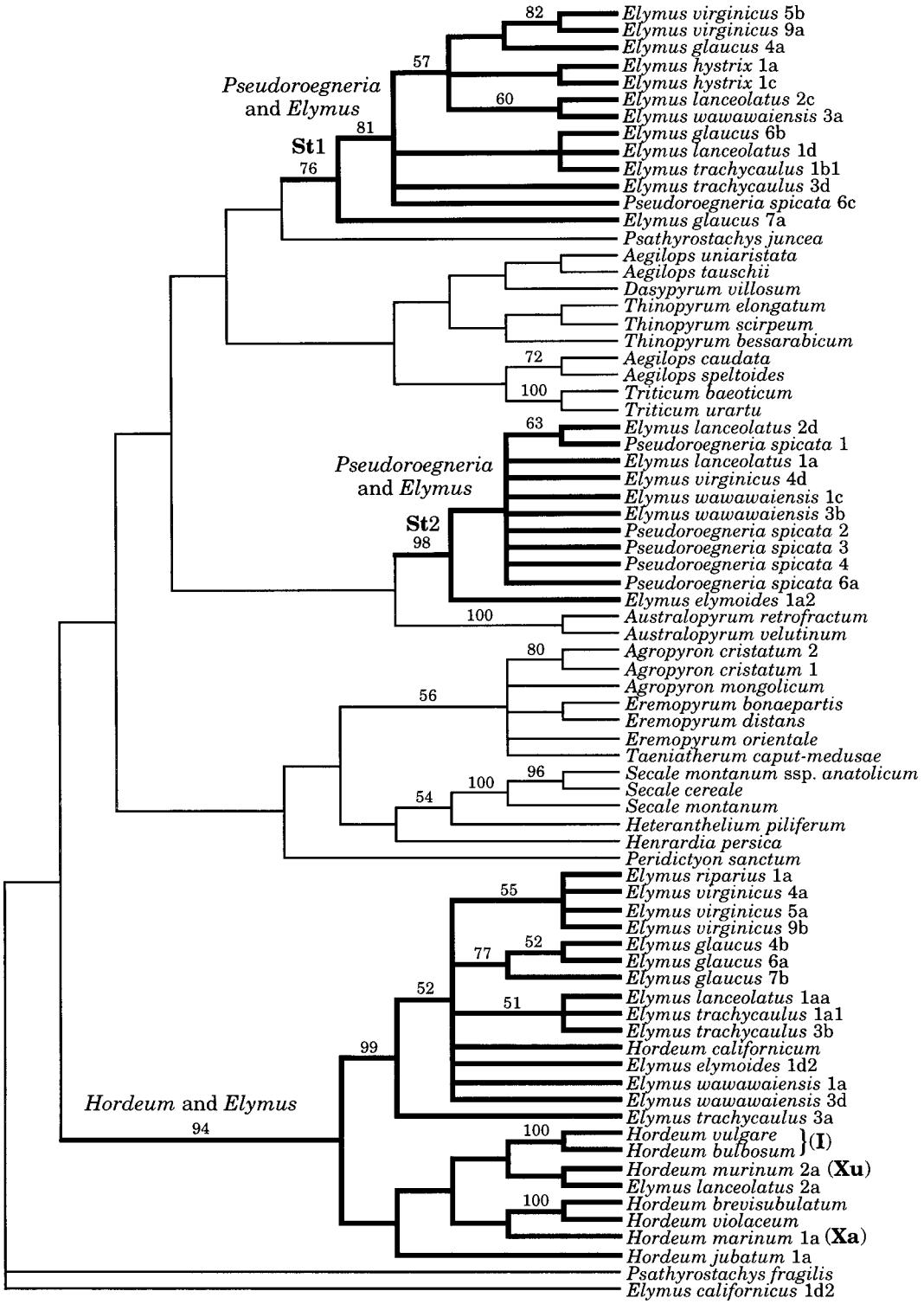

FIG. 6. Strict consensus of trees from islands 3 and 4 from the parsimony analysis, on which the **St** sequences do not form a monophyletic group. Bootstrap values above 50% are shown above nodes. Bold lines indicate the *Elymus*/*Pseudoroegneria* and *Elymus*/*Hordeum* clades. Genome designations are provided for those *Hordeum* species with genomes other than **H**. Parsimony bootstrap values greater than 50% are shown above nodes. Taxon labels as in Fig. 5.

ploid, and are presumed to include the **H** genome from *Hordeum* and the **St** genome from *Pseudoroegneria*. Concerns over the correlation between chromosome pairing with either sequence similarity or historical relatedness call into question the ability of cytogenetic data to clarify phylogenetic relationships and evolutionary processes (reviewed in Seberg and Petersen 1998). Most (but not all) of the *Elymus* individuals examined here have copies of the starch synthase gene that form clades with members of both *Hordeum* and with *Pseudoroegneria*. Within the larger phylogenetic context of the Triticeae, the data suggest that, in North America, *Elymus* has an **St** + **H** genomic content, and is therefore an allotetraploid derivative of *Hordeum* and *Pseudoroegneria* genomes. The molecular phylogenetic data therefore corroborate the cytogenetic data.

Based on the portion of the gene examined, there is no evidence for widespread loss of gene function in either genome, in the form of insertions/deletions in the exons, or acquisition of stop codons. Furthermore, the presence of both **St** and **H** sequences in most *Elymus* individuals provides evidence against homogenization via gene conversion. There are, however, two cases where a copy representing only one of the two genomes has been detected so far (*E. hystrix* with **St** only, and *E. riparius* with **H** only). It is possible that gene copies representing one of the two genomes were lost or have changed to the point where they cannot be amplified, or that one of the two copies was simply missed by chance in the sample of eight cloned PCR products screened. More extensive sampling is required to distinguish between these possibilities.

Elymus californicus is unusual within North American *Elymus* in that it is an octoploid (Löve 1984). It is also the only *Elymus* species in the present study in which neither the **H** nor the **St** genome is detected. This is in agreement with Jensen and Wang (1997), who screened *Elymus californicus* for an anonymous RAPD marker thought to be specific to the **St** genome and found it lacking (they did not address the presence of the **H** genome). They did detect a marker apparently specific to (and suggesting the presence of) the *Psathyrostachys* (**Ns**) genome. Their conclusions are consistent with the very weakly-supported association between *E. californicus* and *Psathyrostachys* on the starch synthase gene tree (Fig 5). Although its phylogenetic affinities are poorly resolved here, it is clear that *E. californicus* is not phylogenetically associated with the rest of *Elymus*. Furthermore, the starch synthase data do not support the traditionally recognized *Hystrix* Moench, which included *E. californicus* (= *H. californica* (Bolander) Kuntze) and *E. hystrix* (= *H. patula* Moench) along with several Asian species (Baden et al. 1998). Addressing the taxonomic implications for *H. californicus* will require a thorough analysis of its relationships to other members of the tribe.

Nearly all of the *Elymus* **H**-genome sequences are very closely related to one another, consistent either with a single **H**-genome donor, or multiple closely-related donors. The **H** sequence of *E. lanceolatus* 2 falls outside of the remaining *Elymus* **H**-genome sequences, suggesting either that this individual represents an independent origin of the **StStHH** tetraploid combination, or that the genome was later introduced via hybridization with another species of *Hordeum*. The starch synthase data strongly place most of the *Elymus* **H**-genome sequences with that of *H. californicum* rather than with *H. jubatum* (the only other North American *Hordeum* species included) or with the other, Old World, representatives. This may indicate *H. californicum*, a narrowly distributed western North American species, as a potential genome donor to *Elymus*, but the sampling scheme within *Hordeum* is currently limited. In light of the present results, it may be especially informative to focus on additional western North American species or populations in the future. Currently, however, it is impossible to draw detailed conclusions regarding species-level relationships between *Elymus* and its wild barley progenitors.

Relationships among *Hordeum* species cannot be addressed in detail here, given the limited sampling and lack of strong support within the genus. *Hordeum* is itself evolutionarily complex, with auto- and allopolyploidy and hybridization known to occur within the genus. However, the strong support of *H. vulgare* + *H. bulbosum* does corroborate results based on morphology (von Bothmer and Jacobsen 1985), chromosome pairing (von Bothmer et al. 1986, 1987), allozymes (Jørgensen 1986; Jaaska 1992), chloroplast DNA restriction sites (Doebley et al. 1992), C-banded karyotype analysis (Linde-Laursen et al. 1992) and *vrs1* intron sequences (Komatsuda et al. 1999). The weakly supported clade containing *H. vulgare* + *H. bulbosum* + *H. murinum* (= sect. *Hordeum*; von Bothmer and Jacobsen 1985) is also in agreement with several of these earlier studies.

The division of the **St**-genome starch synthase sequences into two distinct clades, as seen here, is not reflected in other data, including chloroplast DNA sequences (Mason-Gamer, unpubl. data) and cytogenetic data (Svitashev et al. 1996). Several individuals, representing both *Elymus* and *Pseudoroegneria*, have alleles from both of the **St** clades, so the pattern does not correspond to geographical or morphological differentiation. The two clades might represent either 1) allelic variation at a locus, or 2) a recent gene duplication prior to the divergence of the **St** group. If the gene has been duplicated, it is surprising that relatively few individuals have copies representing both clades, and I therefore favor the hypothesis of allelic variation. However, if duplication has occurred, followed by gene silencing or gene conversion, then gene copies from only one clade would be found within many of the individuals. Allelic variation vs. gene silencing cannot currently be distinguished, but a planned examination of progeny from plants exhibiting the polymorphism may provide useful information.

It remains unclear whether **St**1 and **St**2 form a single

St-genome clade. The **St** group is monophyletic on the ML tree, even when the search is initiated with starting trees from islands 3 or 4, on which the group is not monophyletic. However, parsimony bootstrap support for the **St** clade is low. Furthermore, only some of the MPTs from the cladistic parsimony analysis include a monophyletic **St**-genome clade, while others (Fig. 6) do not.

In general, the starch synthase results are straightforward regarding the origin of the North American species of *Elymus*. However, the relationships among *Elymus* species, and their relationships to specific progenitor species, await a better-resolved phylogeny. The current data suggest that there is little genetic differentiation among species, even though they encompass considerable morphological diversity.

Ultimately, phylogenetic and taxonomic analyses of the North American **StStHH** tetraploids represent the beginning of a complicated story. A clear understanding of *Elymus*, and of the highly reticulate Triticeae in general, will require a focus not just on relationships among organisms or among taxa, but on relationships within and among genome groups. The **St** genome provides the best example of this. It is found throughout the world in monogenomic diploids and autopolyploids (*Pseudoroegneria*), and in numerous allopolyploid combinations. Within the taxonomic entity *Elymus*, the **St** genome is found in combination with genomes **H** (from *Hordeum*), **P** (from *Agropyron*), **W** (from *Australopyrum* (Tzvelev) Á.Löve), and **Y** (from an unknown donor), and at various ploidy levels including **StStHH**, **StStYY**, **StStStStYY**, **StStStStHH**, **StStHHHH**, **StStHHYY**, **StStYYWW**, and **StStYYPP** (Dewey 1984; Jensen 1990a, 1990b; Torabinejad and Mueller 1993; Jensen et al. 1994; Jensen and Salomon 1995). The **St** genome is also a constituent of the North American allo-octoploid *Pascopyrum* Á.Löve (**StStHHNsNsXX**, where **Ns** is derived from *Psathyrostachys* and **X** from an unknown donor; Dewey 1984). Clearly, the **St** genome group as an evolutionary lineage cuts across taxonomic, phylogenetic, geographic, diploid/polyploid, and reproductive boundaries in the Triticeae.

In conclusion, sequence data from the starch synthase gene largely support the hypothesis, based on cytogenetic data, that the North American tetraploid species of *Elymus* are derived from *Hordeum* and *Pseudoroegneria*. A correlation between genome groups and clades on molecular phylogenetic trees had earlier been seen for monogenomic members of the tribe (Hsiao et al. 1995; Kellogg and Appels 1995; Kellogg et al. 1996; Mason-Gamer and Kellogg 1996a, 1996b; Petersen and Seberg 1997). Seberg and Petersen (1998) raise many well-founded concerns regarding the use of genomic pairing data as the dominant criterion for both grouping and ranking of taxa in the Triticeae. However, as Barkworth (2000) states, ''. . . genomic constitution provides limited information about evolution, but limited information is better than no information.'' It seems clear that cytogenetic data commonly delimit natural, evolutionary groups. Thus, they provide a solid basis for framing questions and designing sampling strategies for further analyses of phylogeny and evolution in this complex tribe of grasses.

ACKNOWLEDGEMENTS. I wish to thank Toby Kellogg for valuable insights into *Elymus*, the Triticeae, and reticulate evolution; Mary Barkworth for sharing plant material and her considerable knowledge of *Elymus* and the Triticeae; Mary Barkworth, John Freudenstein, and an anonymous reviewer for thoughtful comments on the manuscript; and the USDA's Germplasm Resources Information Network for seed. The work was funded by the Stillinger Foundation of the University of Idaho, and the National Science Foundation (DEB-9974181).

LITERATURE CITED

BADEN, C., S. FREDERIKSEN, and O. SEBERG. 1997. A taxonomic revision of the genus *Hystrix* (Triticeae, Poaceae). Nordic Journal of Botany 17: 449–467.

BARKWORTH, M. E. 2000. Changing perceptions of the Triticeae. Pp. 110–120 in *Grasses. Systematics and evolution,* eds. S. W. L. Jacobs and J. Everett. Collingwood, Victoria: CSIRO Publishing.

——— 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.

BENTHAM, G. 1881. Notes on Gramineae. Journal of the Linnean Society. Botany. London. 19: 14–134.

——— and J. D. HOOKER. 1883. Gramineae. Pp. 1074–1215 in *Genera Plantarum*, London: L. Reeve and Co.

BOTHMER, R. VON and N. JACOBSEN. 1985. Origin, taxonomy, and related species. Pp. 19–56 in *Barley*, ed. D. C. Rasmussin. Madison, WI: Agronomy Society of America.

———, J. FLINK, and T. LANDSTRÖM. 1986. Meiosis in interspecific *Hordeum* hybrids. I. Diploid combinations. Canadian Journal of Genetics and Cytology 28: 525–535.

———, ———, and ———. 1987. Meiosis in interspecific *Hordeum* hybrids. II. Triploid hybrids. Evolutionary Trends in Plants 1: 41–50.

CLARK, J. R., M. ROBERTSON, and C. C. AINSWORTH. 1991. Nucleotide sequence of a wheat (*Triticum aestivum* L.) cDNA clone encoding the waxy protein. Plant Molecular Biology 16: 1099–1101.

DEWEY, D. R. 1982. Genomic and phylogenetic relationships among North American perennial Triticeae. Pp. 51—81 in *Grasses and grasslands,* ed. J. R. Estes. Norman: Oklahoma University Press.

———. 1983a. New nomenclatural combinations in the North American perennial Triticeae. Brittonia 35: 30–33.

———. 1983b. 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: 576–684.

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.

———, R. N. BEACHY, and W. H. LEWIS. 1984. Evolution of rDNA in *Claytonia* polyploid complexes. Pp. 321—341 in *Plant biosystematics*, ed. W. F. Grant. Ottawa: Academic Press.

———, J. L. DOYLE, A. H. D. BROWN, and B. E. PFEIL. 2000. Con-

firmation of shared and divergent genomes in the *Glycine tabacina* polyploid complex (Leguminosae) using histone H3-D sequences. Systematic Botany 25: 437–448.

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.

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 Science U. S. A. 96: 14400–14405.

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.

HASAGAWA, M., H. KISHINO, and T. YANO. 1985. Dating the human-ape split by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22: 160–174.

HIGGINS, D. G., A. J. BLEASBY, and R. FUCHS. 1992. Clustal V: improved software for multiple sequence alignment. Computer Applications in the Biosciences 8: 189–191.

HITCHCOCK, A. S. 1951. *Manual of grasses of the United States, Second edition,* revised by A. Chase. Washington, D. C.: U. S. Government Printing Office.

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.

JAASKA, V. 1992. Isozyme variation in barley (*Hordeum* L.). 2. Aspartate aminotransferase and 6-phosphoglyconate dehydrogenase. Hereditas 116: 29–35.

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

———. 1990b. Cytology, fertility, and morphology of *Elymus kengii* (Keng) Tzvelev and *E. grandiglumis* (Keng) A. Löve (Poaceae: Triticeae). Genome 33: 408–421.

——— and B. SALOMON. 1995. Cytogenetics and morphology of *Elymus panormitanus* var. *heterophyllus* and its relationship to *Elymus panormitanus* (Poaceae: Triticeae). International Journal of Plant Science 156: 731–739.

——— and R. 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.

———, Z. W. LIU, B. R. LU, and B. SALOMON. 1994. Biosystematic study of hexaploids *Elymus tschimganicus* and *E. glaucissimus*. I. Morphology and genomic constitution. Chromosome Research 2: 209–215.

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. N. H. Munro. New York: Academic Press

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.

———, ———, and R. J. MASON-GAMER. 1996. When gene trees tell different stories: the diploid genera of Triticeae (Gramineae). Systematic Botany 21: 321–347.

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.

KLOSGEN, R. B., A. GIERL, Z. S. SCHWARZ-SOMMER, and H. SAEDLER. 1986. Molecular analysis of the waxy locus of *Zea mays*. Molecular and General Genetics 203: 237–244.

KOMATSUDA, T., K. TANNO, B. SALOMON, T. BRYNGELSSON, and R. VON BOTHMER. 1999. Phylogeny in the genus *Hordeum* based on nucleotide sequences closely linked to the *vrs1* locus (row number of spikelets). Genome 42: 973–981.

LINDE-LAURSEN, I., R. VON BOTHMER, and N. JACOBSEN. 1992. Relationships in the genus *Hordeum*: giemsa C-banded karyotypes. Hereditas 116: 111–116.

LÖVE, 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. 1992. MacClade version 3.0: Analysis of phylogeny and character evolution. Sunderland, Massachusetts: Sinauer Associates.

MASON-GAMER, R. J. 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, Florida: 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 *Grasses. Systematics and evolution*, eds. S. W. L. Jacobs and J. Everett. Collingwood, Victoria: CSIRO Publishing.

———, C. F. WEIL, and E. A. KELLOGG. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. Molecular Biology and Evolution 15: 1658–1673.

PETERSEN, G. and O. SEBERG. 1997. Phylogenetic analysis of the Triticeae (Poaceae) based on *rpo*A sequence data. Molecular Phylogenetics and Evolution 7: 217–230.

ROHDE, W., D. BECKER, and F. SALAMINI. Structural analysis of the waxy locus from *Hordeum vulgare*. Nucleic Acids Research 16: 7185–7186.

SEBERG, O. and G. PETERSEN. 1998. A critical review of concepts and methods used in classical genome analysis. The Botanical Review 64: 372–417.

SHURE, M., S. WESSLER, and N. FEDEROFF. 1983. Molecular identification and isolation of the *waxy* locus in maize. Cell 35: 225–233.

SMALL, R. L., J. A. RYBURN, R. C. CRONN, T. SEELANAN, and J. F. WENDEL. 1998. The tortoise and the hare: choosing between noncoding plastome and nuclear *Adh* sequences for phylogeny reconstruction in a recently diverged plant group. American Journal of Botany 85: 1301–1315.

SOLTIS, P. S. and D. E. SOLTIS. 1991. Multiple origins of the allotetraploid *Tragopogon mirus* (Compositae): rDNA evidence. Systematic Botany 16: 407–413.

SULLIVAN, J. and D. L. SWOFFORD. In press. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? Systematic Biology.

———, K. E. HOLSINGER, and C. SIMON. 1995 Among-site rate variation and phylogenetic analysis of 12S rRNA in Sigmodontine rodents. Molecular Biology and Evolution 12: 988–1001.

———, 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.

SVITASHEV, S., B. SALOMON, T. BRYNGELSSON, and R. VON BOTHMER. 1996. A study of 28 *Elymus* species using repetitive DNA sequences. Genome 39: 1093–1101.

SWOFFORD, D. L. 2000. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sunderland, Massachusetts: Sinauer Associates.

———, G. J. OLSEN, P. J. WADDELL, and D. M. HILLIS. 1996. Phy-

logenetic inference. Pp. 407–514 in *Molecular systematics*, Second edition, eds. D. M. Hillis, C. Moritz, and B. K. Mable. Sunderland, Massachusetts: Sinauer Associates.

TORABINEJAD, J. and R. J. MUELLER. 1993. Genome constitution of the Australian hexaploid grass, *Elymus scabrus* (Poaceae: Triticeae). Genome 36: 147–151.

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.

WANG, Z. Y., F. Q. ZHENG, J. P. GAO, D. P. SNUSTAD, M. G. LI, J. L. ZHANG, and M. N. HONG. 1995. The amylose content in rice endosperm is related to the post-transcriptional regulation of the *waxy* gene. Plant Journal 7: 613–622.

YANG, Z. 1994a. Estimating the pattern of nucleotide substitution. Journal of Molecular Evolution 39: 105–111.

———. 1994b. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods. Journal of Molecular Evolution 39: 306–314.