*Aliso* 23, pp. 335–348


# A PRELIMINARY PHYLOGENETIC ANALYSIS OF THE GRASS SUBFAMILY POOIDEAE (POACEAE), WITH ATTENTION TO STRUCTURAL FEATURES OF THE PLASTID AND NUCLEAR GENOMES, INCLUDING AN INTRON LOSS IN GBSSI

JERROLD I DAVIS[1,3] AND ROBERT J. SORENG[2]

[1]*L. H. Bailey Hortorium and Department of Plant Biology, Cornell University, Ithaca, New York 14853, USA;*
[2]*Department of Botany and U. S. National Herbarium, National Museum of Natural History, Smithsonian Institution, Washington, D.C. 20013-7012, USA, (sorengr@si.edu)*
[3]*Corresponding author (jid1@cornell.edu)*

ABSTRACT

Phylogenetic relationships in the grass family (Poaceae), with specific attention to the internal structure of subfamily Pooideae, are analyzed on the basis of nucleotide sequence variation in plastid-encoded genes (*mat*K, *ndh*F, *ndh*H, and *rbc*L). The resulting phylogenetic hypothesis was examined with attention to the taxonomic distributions of two inversions and an insertion/deletion within *ndh*F, the absence of intron 10 of the nuclear gene GBSSI (*waxy*), and positions of the boundaries between the Short Single Copy (SSC) region and the neighboring Inverted Repeat (IR) regions of the plastid genome, relative to the endpoints of *ndh*F and *ndh*H, which span these boundaries in some taxa. The PACCAD clade is resolved, and extension of the 3′-end of *ndh*F from the SSC region into the IR region is interpreted as a synapomorphy of this clade. The BEP clade also is resolved, with Ehrhartoideae placed as the sister of a clade in which Bambusoideae and Pooideae are sister groups. The loss of GBSSI intron 10 is interpreted as a synapomorphy of Poeae s.l., which includes the traditionally defined tribes Poeae, Aveneae, and Hainardieae, and the results support a novel set of relationships among the tribes of Pooideae, including the placement of Brachypodieae, Bromeae, Triticeae, and Poeae s.l. within a clade for which a three-nucleotide inversion in *ndh*F is interpreted as a synapomorphy, while a six-nucleotide inversion in *ndh*F marks a clade that includes all sampled members of subtribe Aveninae within Poeae s.l.

Key words: GBSSI, intron, *mat*K, *ndh*F, *ndh*H, phylogenetics, Poaceae, Pooideae, *rbc*L, systematics.

## INTRODUCTION

### *Phylogenetic Relationships*

The grass family (Poaceae) has been a constant subject of attention by systematists, who have evaluated relationships in this group for hundreds of years, on the basis of morphological, anatomical, and other features (see reviews by Stebbins and Crampton 1961; Soreng and Davis 1998; Grass Phylogeny Working Group [GPWG] 2001; and citations therein). In recent years, with the development of formal methods of phylogenetic analysis, several studies have focused on higher-level relationships in the grasses, using a variety of structural and molecular characters (e.g., Kellogg and Campbell 1987; Davis and Soreng 1993; Clark et al. 1995; Soreng and Davis 1998; Hilu et al. 1999; Hsiao et al. 1999; Mathews et al. 2000; Zhang 2000; GPWG 2001; Duvall et al. 2007). Most of these analyses provide similar or identical results with respect to many of the major groupings. For example, one group that now appears to be well established is the "PACCAD clade" (consisting of Panicoideae, Arundinoideae, Chloridoideae, Centothecoideae, Aristidoideae, and Danthonioideae; originally the PACC clade; Davis and Soreng 1993), which includes more than half of all species of the family (Clayton and Renvoize 1986). Another group that is usually resolved is subfamily Pooideae, which as circumscribed by the GPWG (2001) includes about one-third of all grass species (Clayton and Renvoize 1986). However, important differences exist among the results of these analyses, including the disparate relationships resolved among the PACCAD clade, Pooideae (sensu GPWG 2001), and two other subfamilies, Ehrhartoideae (formerly Oryzoideae) and Bambusoideae (sensu GPWG 2001), by Clark et al. (1995), Soreng and Davis (1998), Hilu et al. (1999), and the GPWG (2001). A "BEP clade" (consisting of Bambusoideae, Ehrhartoideae, and Pooideae; originally the BOP clade; Clark et al. 1995) frequently has been resolved, but alternative structures also have been observed.

Although monophyly of subfamily Pooideae (sensu GPWG 2001) has been supported by most of the analyses cited above, the relationships resolved among its constituent groups have varied widely. As the present contribution is a preliminary report of relationships resolved by DNA nucleotide sequence data, we do not provide a detailed review of the various relationships that have been supported by previous studies, but refer the reader to two analyses that have sampled broadly within this large subfamily, using restriction site data and morphology, and that discuss the problems attendant to the phylogeny of this group (Soreng and Davis 1998, 2000). One of the principal findings of the latter paper was that neither Poeae nor Aveneae are monophyletic as traditionally circumscribed (cf. Clayton and Renvoize 1986). Rather, elements of these two groups are intermixed within a clade that also includes the small tribe Hainardieae. We refer to this assemblage as Poeae s.l., and in the present work we examine relationships within this clade, within Pooideae, and within Poaceae as a whole, with particular attention to a series of structural characters of the plastid and nuclear genomes.

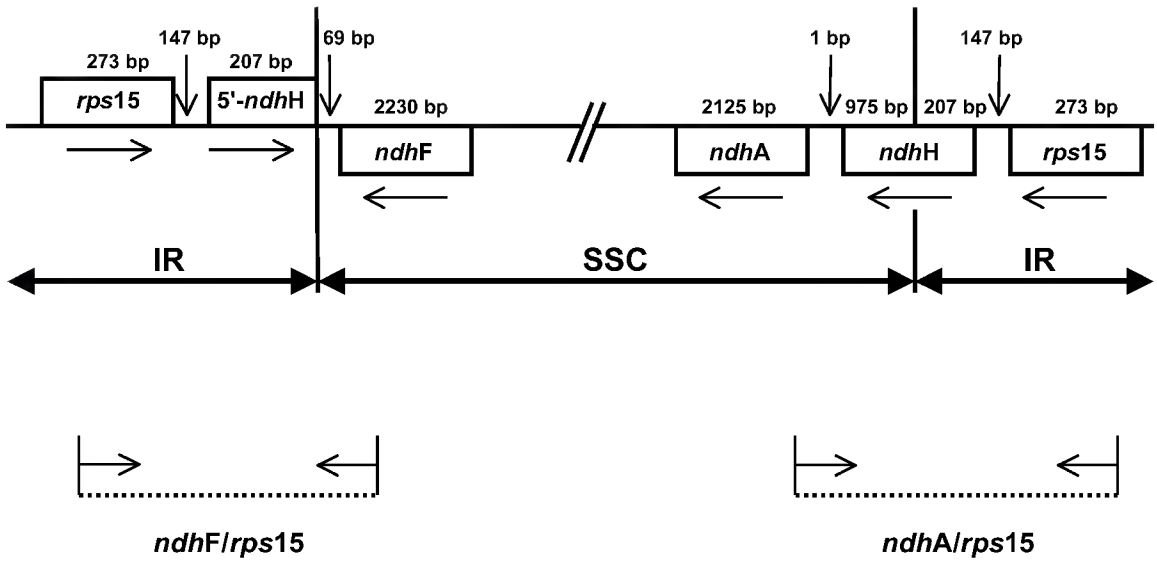


Fig. 1.—Map, not to scale, of a portion of the plastid genome of *Triticum aestivum* L., adapted from GenBank accessions AB042240 and NC002762. Upper portion of figure depicts genes and intergenic spacer regions near endpoints of the Short Single Copy (SSC) region and adjacent portions of flanking Inverted Repeat (IR) regions, with the central portion of the SSC region excised, and with boundaries between the regions indicated by vertical lines. Genes depicted above the horizontal line, with arrows pointing to the right, are encoded from left to right (i.e., their 5′-ends are on the left), and genes depicted below the horizontal line, with arrows pointing to the left, are encoded from right to left. Lengths of genes and intergenic spacers are indicated by numbers of base pairs (bp) above the horizontal line. Lower portion of figure depicts positions of two fragments that were amplified and sequenced to determine the locations of the 5′-endpoint of *ndh*H and the 3′-endpoint of *ndh*F relative to the SSC/IR boundaries (see text).

*Inversions and Gene Positions in the Plastid Genome*

The plastid genomes of grasses exhibit a set of three inversions that differentiate them from those of most other plants, and analyses of the distributions of these inversions have contributed to an understanding of relationships among grasses and other families. When the complete nucleotide sequence of the plastid genome of *Oryza* L. was published (Hiratsuka et al. 1989), it was observed to differ from that of *Nicotiana* L. by three inversions in the Large Single Copy (LSC) region. Subsequent analyses (Doyle et al. 1992; Katayama and Ogihara 1996; Michelangeli et al. 2003) have indicated that the three inversions demarcate three clades that appear to be perfectly internested. Until every species has been sampled, the precise status of every taxon with respect to these inversions cannot be known, but a fairly well-defined picture has emerged, as follows. The most inclusive of the three clades, marked by the largest of the inversions (28 kilobases [kb] in length), includes all species of Poaceae, Joinvilleaceae, and Ecdeiocoleaceae, and some or all elements of Restionaceae (cf. Katayama and Ogihara 1996; Michelangeli et al. 2003). This group includes four of the seven families of Poales sensu Dahlgren et al. (1985), but the absence of this inversion from the remaining families (e.g., Flagellariaceae), and possibly from some elements of Restionaceae, does not imply that Poales as circumscribed by Dahlgren et al. (1985) are non-monophyletic, for these taxa still may be the closest relatives of the clade that is marked by the inversion. However, the potential absence of this inversion from some species of Restionaceae does suggest that this family may not be monophyletic. The next most inclusive clade, marked by the second largest of the inversions (6 kb in length), includes all sampled species of Poaceae, Ecdeiocoleaceae, and Joinvilleaceae, but none of Restionaceae or any other family. The third most inclusive clade, marked by the smallest of the inversions (a few hundred base pairs [bp] in length, encompassing *trn*T), includes all sampled species of Poaceae and no others. The distributions of these inversions, in concert with morphological and nucleotide sequence data (Bremer 2002; Michelangeli et al. 2003), have provided compelling evidence that Joinvilleaceae and Ecdeiocoleaceae are the closest living relatives of the grasses. Consequently, elements of these two small families now are widely used as outgroups for phylogenetic studies of the grasses, and they are employed in that manner here.

Another example of structural variation among the plastid genomes of grasses involves the locations of two genes, *ndh*F and *ndh*H, relative to the boundaries of the major regions of this genome (Ogihara et al. 2002). These two genes are positioned near opposite ends of the Short Single Copy (SSC) region, close to the borders with the two Inverted Repeat (IR) regions that flank it (Fig. 1). The IR regions are identical copies of the same DNA sequence in reverse orientation (see Palmer 1983; Plunkett and Downie 2000; and citations within), so all genes and intergenic regions that lie within them are present as two copies in the plastid genome, oriented in opposite directions relative to the overall structure of the plastid genome (e.g., *rps*15 in Fig. 1). Examination of the first three plastid genome sequences of grasses to be published (*Oryza, Triticum* L. and *Zea* L.) indicates that a substantial portion of the 5′-end of *ndh*H extends into the IR region in *Oryza* and *Triticum,* and thus is duplicated in these taxa (Fig. 1; also see Ogihara et al. 2002: Fig. 5). In *Zea,* a portion of the 3′-end of *ndh*F extends into the IR

region, and thus is duplicated in a similar manner, but one nucleotide of *ndh*H in *Zea* also extends into the IR. Because different ends of the two genes lie near or within the IR region in all of these taxa (the 5′-end of *ndh*H and the 3′-end of *ndh*F), any portion of *ndh*H that extends into the IR region is read in one direction (from the 5′-end within the IR region towards the SSC region), while any portion of *ndh*F that extends into the IR region is read in the other direction (from the 5′-end in the SSC region towards the IR region). Thus, the solitary nucleotide of *ndh*H that lies within the IR region in *Zea* simultaneously encodes the first base pair of *ndh*H and an internal site within *ndh*F (i.e., the first site of *ndh*F that lies within the IR), as read in opposite directions. In summary, *ndh*H extends into the IR in all three of these grasses, though to varying extents (207 bp in *Triticum,* 163 bp in *Oryza,* and 1 bp in *Zea*), while *ndh*F extends into the IR only in *Zea.* Ogihara et al. (2002) concluded from this similarity between *Oryza* and *Triticum* that these two taxa are more closely related to each other than either is to *Zea,* but in the absence of evidence from additional taxa this conclusion is unsupported, for the similarity between *Oryza* and *Triticum* could be plesiomorphic within the smallest clade that includes all three of these genera. The taxonomic distribution of the positions of these two genes, relative to the SSC/IR boundaries, is examined in the present study, with attention to their phylogenetic implications.

### *GBSSI Intron 10*

There are many cases in which intron losses in the three principal genomes of plants have provided evidence of phylogenetic affinities (e.g., Wallace and Cota 1996; Frugoli et al. 1998; Itchoda et al. 2002). The phylogenetic results presented here are derived from structural and sequence variation in the plastid genome, but we have recently initiated a study of sequence variation within Poeae s.l. in the nuclear-encoded gene granule-bound starch synthase (GBSSI, or *waxy*). Variation patterns in GBSSI have proven useful in analyses at a range of taxonomic levels within the grasses, from the overall phylogenetic structure of the family to that among closely related species, including the details of polyploidization events (e.g., Mason-Gamer et al. 1998; GPWG 2001; Mason-Gamer 2001; Mathews et al. 2002; Ingram and Doyle 2003). In the present work we report a phylogenetically informative loss of GBSSI intron 10 within Poeae s.l.

## MATERIALS AND METHODS

### *Molecular Data*

Nucleotide sequence variation was examined in *matK, ndh*F, *ndh*H, and *rbc*L from 106 representative species of Poaceae and one species each from two related families, Joinvilleaceae and Ecdeiocoleaceae (Table 1). Of the 424 gene/taxon combinations, all but four sequences are in the working data set. In some cases (e.g., *Streptochaeta* Schrad. ex Nees; Table 1) the sequences for a given terminal in the analysis represent different species. This use of "conglomerate taxa" represents a compromise between the goal of consistency in sampling and that of including as many sequences as possible in the analysis.

Two of the four plastid-encoded genes lie within the LSC region (*mat*K and *rbc*L), while the other two (*ndh*F and *ndh*H) are principally in the SSC region, though each extends into the IR region in some taxa, as noted above. In the course of sequencing *ndh*F and *ndh*H, the regions adjacent to the 3′-end of the former and the 5′-end of the latter, collectively referred to here as the SSC/IR regions, also were sequenced (Fig. 1). The first of the two SSC/IR regions, the *ndh*F/*rps*15 region, is amplified with a primer situated at ca. nucleotide 80 of *rps*15, and another at ca. nucleotide 1810 of *ndh*F. The second SSC/IR region, the *ndh*A/*rps*15 region, is amplified with the same primer within *rps*15 that is used for the first SSC/IR region, in combination with a primer at ca. nucleotide 60 of *ndh*A. The positions of the boundaries between the two ends of the SSC region and the adjoining IR regions were determined, along with the positions of the 3′-terminus of *ndh*F and the 5′-terminus of *ndh*H relative to these boundaries, by comparing the sequences of these two regions, starting from the end of each that includes a portion of *rps*15, and reading toward the SSC/IR boundary. For the phylogenetic analysis, the presence vs. absence of any portion of *ndh*F within the IR region was coded as one binary character, and the presence vs. absence of any portion of *ndh*H within the IR region as another.

Most sequences used in the analysis were generated from total genomic DNA isolations of vouchered collections, following standard PCR and automated cycle-sequencing protocols, though some were obtained from GenBank (Table 1). Primers used for amplification and sequencing of *rbc*L and most of *ndh*F have been published previously (*rbc*L: Chase et al. 1993; Asmussen and Chase 2001; *ndh*F: Olmstead and Sweere 1994). Additional primers used for *mat*K and the SSC/IR regions (including *ndh*H and the 5′-end of *ndh*F) were developed for this project, and will be described in a forthcoming contribution. Sequences were aligned manually, and regions in which alignment was considered ambiguous were excluded from analyses.

In addition to the nucleotide sequence data and the positions of the endpoints of *ndh*F and *ndh*H relative to the SSC/IR boundaries, several additional structural features were scored as characters for the analysis. One of these is a 15-nucleotide insertion/deletion (indel) (Clark et al. 1995, 2000; also discussed as character 53 by GPWG 2001) corresponding to sites 1704–1718 in the *ndh*F sequence of *Triticum* (which is undeleted for these sites) in GenBank reference accession NC002762, and situated approximately at site 1704 in the *ndh*F sequence of *Oryza* (which is deleted for these sites; precise position is uncertain due to ambiguity in alignments) in GenBank reference accession NC001320.

Two additional structural variants in *ndh*F are interpreted as inversions of three and six nucleotides in length, respectively. The three-nucleotide inversion corresponds to sites 1918–1920 in the reference sequence of *Triticum.* For most taxa in the sample the sequence of these three nucleotides is either TAC (as in *Triticum*) or its reverse complement, GTA. Taxa with sequences other than these two combinations usually differ from one or the other of these three-nucleotide sequences at no more than one of the three sites, as would be consistent with the occurrence of point mutations having caused differentiation from one or the other of the two principal sequences. The six-nucleotide inversion

Table 1. Taxa sampled for plastid genome sequence variation, voucher collection information (herbarium acronyms as in Holmgren et al. 1990), and GenBank nucleotide sequence accession numbers. Asterisks identify taxa for which all sequences were obtained from GenBank; these names are accompanied in the left column by accession numbers of those sequences. Non-asterisked taxon names are those for which the authors generated one or more sequences, though additional sequences may have been obtained for these taxa from GenBank, as indicated by accession numbers with these names in the left column. GenBank accession numbers in the right column identify sequences generated by the authors from the specified plant accessions. GenBank numbers are absent for sequences generated by the authors but not yet released.

| Taxon | Voucher information for sequences generated by authors |
|---|---|
| *Achnatherum occidentalis* (Thurb. ex S. Watson) Barkworth subsp. *pubescens* (Vasey) Barkworth | *Soreng 7418* (US) |
| *Agrostis tenerrima* Trin. | *Soreng 3734* (BH) |
| *Aira caryophyllea* L. | *Soreng 5953b* (US) |
| *Alopecurus magellanicus* Lam. | *Soreng 3514* (BH) |
| *Ampelodesmos mauritanica* (Poir.) T. Durand & Schinz | *Soreng & Soreng 4029* (BH) |
| *Amphibromus scabrivalvis* (Trin.) Swallen | *Soreng 7013* (US) |
| *Amphipogon strictus* R. Br. [*rbc*L U88403.1]; **A. caricinus* F. Muell. [*mat*K AF31274.1] | *Linder 5634* (BOL) |
| *Anisopogon avenaceus* R. Br. | *Linder 5590* (BOL) |
| *Anomochloa marantoidea* Brongn. [*mat*K AF164381.1; *rbc*L AF021875.1] | *Davis 753* (BH) |
| *Anthoxanthum odoratum* L. | *Soreng 4292* (BH) |
| *Arctagrostis latifolia* (R. Br.) Griseb. | *Soreng 6016* (US) |
| *Arundo donax* L. [*mat*K AF164408.1; *rbc*L U31360.1] | *Crisp 278* (CANB) |
| *Avena sativa* L. 'ASTRO' | *Davis 759* (BH) |
| *Avenella flexuosa* (L.) Drejer. | *Soreng 7305b* (US) |
| *Bambusa multiplex* (Lour.) Raeusch. ex Schult. & Schult. f. | *Davis 770* (BH) |
| *Beckmannia syzigachne* (Steud.) Fernald | *Soreng 3513* (BH) |
| *Bellardiochloa variegata* (Lam.) Kerguélen | Grown from USDA Plant Intro. Sta. 253455 (no voucher) |
| *Brachyelytrum erectum* (Schreb.) P. Beauv. | *Soreng 3427a* (BH) |
| *Brachypodium pinnatum* (L.) P. Beauv. | *Davis 760* (BH); grown from USDA Plant Intro. Sta. 440170 [*ndh*F AY622312.1; *rbc*L AY632361.1] |
| *B. sylvaticum* (Huds.) P. Beauv. | *Soreng 5923b* (US) |
| *Briza minor* L. | *Davis 761* (BH); grown from USDA Plant Intro. Sta. 378653 |
| *Bromus inermis* Leyss. [*mat*K AF164398.1; *rbc*L Z49836.1] | *Davis 762* (BH); grown from USDA Plant Intro. Sta. 314071 |
| *B. suksodorfii* Vasey | *Soreng 7412* (US) |
| *Calamagrostis canadensis* Michx. | *Soreng 7414* (US) |
| *Calotheca brizoides* (Lam.) Desv. | *Soreng 7014* (US) |
| *Catabrosa aquatica* (L.) P. Beauv. | *Soreng 3861* (BH) |
| *Chascolytrum subaristatum* (Lam.) Desv. | *Soreng 7020* (US) |
| *Chasmanthium nitidum* (Baldwin) H. O. Yates; **C. laxum* (L.) H. O. Yates [*mat*K AF164414.1] | *Wipff & Jones 2075* (TAES) |
| *Chusquea* aff. *subulata* L. G. Clark | *Peterson & Judziewicz 9499* (US) |
| *Cutandia memphitica* (Spreng.) K. Richt. | *Boulos & Cope 17676* (E) |
| *Cynosurus cristatus* L. | Grown from RBG, Kew, seed bank 39006 (K) |
| *Dacytlis glomerata* subsp. *hackelii* (Asch. & Graebn.) Cif. & Giacom. | *Soreng 3692* (BH) |
| *Danthonia californica* Bol. | *Davis 763* (BH); grown from USDA Plant Intro. Sta. 232247 |
| *Deschampsia cespitosa* (L.) P. Beauv. subsp. *cespitosa* | *Soreng 7417* (US) |
| *Desmazeria sicula* (Jacq.) Dumort. | Grown from RBG, Kew, seed bank 17332 (K) |
| *Diarrhena obovata* (Gleason) Brandenberg | *Davis 756* (BH) |
| *Dielsiochloa floribunda* (Pilg.) Pilg. | *Peterson et al. 14566* (US) |
| *Dupontia fisheri* R. Br. | *Cayouette & Dalpé C6779-4* (DAO) |
| *Duthiea brachypodium* (P. Candargy) Keng & Keng f. | *Soreng 5358* (US) |
| *Ecdeiocolea monostachya* F. Muell. | *Conran et al. 938* (PERTH, ADU) [*ndh*F AY622313.1; *rbc*L AY123235.1] |
| *Echinopogon caespitosis* C. E. Hubb. | *Soreng 5900* (US) |
| *Ehrharta calycina* Sm. | Grown from USDA Plant Intro. Sta. 208983 |
| *Elymus trachycaulus* (Link) Gould ex Shinners | *Soreng 4291b* (BH) |
| *Eremitis* Döll sp. | US National Herbarium Greenhouse 153, *Soderstrom 2182* (US); or US Natinoal Herbarium Greenhouse 286 (no voucher) |
| *Festuca rubra* L. | *Soreng 7424* (US) |
| *Gastridium ventricosum* (Gouan) Schinz & Thell. | Grown from RBG, Kew, seed bank 5430 (K) |
| *Gaudinia fragilis* (L.) P. Beauv. | *Davis 764* (BH); grown from USDA Plant Intro. Sta. 442496 |

Table 1. Continued.

| Taxon | Voucher information for sequences generated by authors |
|---|---|
| *Glyceria grandis* S. Watson | No voucher [*ndh*F AY622314.1; *rbc*L AY632364.1] |
| *Guadua angustifolia* Kunth | *Peterson & Judziewicz 9527* (US) |
| **Guaduella marantifolia* Franch. [*ndh*F AF164777.1; *rbc*L AF164778.1] | All sequences obtained from GenBank |
| *Gynerium sagittatum* (Aubl.) P. Beauv. | *Clark & Asimbaya 1472* (ISC) |
| *Helictotrichon convolutum* (C. Presl) Henrard | *Soreng 3803* (BH) |
| *Hesperostipa comata* (Trin. & Rupr.) Barkworth | *Soreng 7431* (US) |
| *Holcus annuus* Salzm. ex C. A. Mey. | *Soreng 3642* (BH) |
| **Hordeum vulgare* L. [*mat*K AB078138.1; *ndh*F U22003.1; *ndh*H AJ011848.1; *rbc*L AY137456.1] | All sequences obtained from GenBank |
| *Joinvillea gaudichaudiana* Brongn. & Gris [*mat*K AF164380.1, published as *J. ascendens* Gaudich. ex Brongn. & Gris] | *Davis 751* (BH) |
| *Leucopoa kingii* (S. Watson) W. A. Weber | *Soreng 3515* (BH) |
| *Lithachne pauciflora* (Sw.) P. Beauv. [*mat*K AF164385.1] | *Clark 1297* (ISC) |
| *Littledalea tibetica* Hemsl. | *Soreng 5487, 5490, 5494* (US) |
| *Lolium perenne* L. | *Davis 765* (BH); grown from USDA Plant Intro. Sta. 418710 |
| *Lygeum spartum* L. | *Soreng 3698* (BH) |
| *Melica cupanii* Guss. | *Davis 766* (BH); grown from USDA Plant Intro. Sta. 383702 [*ndh*F AY622315.1; *rbc*L AY632365.1] |
| *Merxmuellera macowanii* (Stapf) Conert [*rbc*L U31438.1] | *Barker 1008* (BOL) |
| *M. rangei* (Pilg.) Conert [*rbc*L AY640153.1] | *Barker 960* (GRA) |
| *Mibora minima* (L.) Desv. | *Devesa 3885* (BH) |
| *Milium vernale* M. Bieb. | *Soreng 3748* (BH) |
| *Molineriella laevis* (Brot.) Rouy | *Soreng 3613* (BH) |
| *Molinia caerulea* (L.) Moench [*mat*K AF164411.1] | No voucher [*rbc*L AY632367.1] |
| *Nardus strictus* L. | *Royl & Schiers s. n.* (1988, B) |
| *Nassella pulchra* (Hitchc.) Barkworth | *Soreng 7407* (US) |
| *N. viridula* (Trin.) Barkworth | No voucher; grown from USDA Plant Intro. Sta. 387938 |
| *Olyra latifolia* L. [*mat*K AF164386.1] | *Peterson & Annable 7311* (US) |
| **Oryza nivara* Sharma & Shastry [all five genes NC005973.1] | All sequences obtained from GenBank |
| **O. sativa* L. [all five genes NC001320.1] | All sequences obtained from GenBank |
| *Oryzopsis asperifolia* Michx. | *Soreng 5989* (US) |
| *Parapholis incurva* (L.) C. E. Hubb. | Grown from RBG, Kew, seed bank 24867 (K) |
| *Pariana radiciflora* Sagot ex Döll [*mat*K AF164387.1] | *Clark & Zhang 1344* (ISC) [*rbc*L AY632369.1] |
| *Phaenosperma globosa* Munro ex Benth. | *Clark 1292* (ISC) [*rbc*L AY632370.1] |
| *Pharus latifolius* L. [*mat*K AF164388.1; *rbc*L AY357724.1] | No voucher |
| *Phleum pratense* L. | *Soreng 4293* (BH) |
| *Piptatherum miliaceum* (L.) Coss. | *Davis 767* (BH); grown from USDA Plant Intro. Sta. 284115 [*ndh*F AY622317.1] |
| *Pleuropogon refractus* (A. Gray) Benth. | *Soreng 3381* (BH) |
| *Poa alpina* L. | *Soreng 6115-1* (US) |
| *Polypogon monspeliensis* (L.) Desf. | No voucher |
| *Pseudosasa japonica* (Siebold & Zucc. ex Steud.) Makino ex Nakai | *Davis 771* (BH) |
| *Puccinellia distans* (Jacq.) Parl. | *Davis 755* (BH) |
| **Puelia ciliata* Franch. [*rbc*L AF164780.1]; **P. olyriformis* (Franch.) Clayton [*ndh*F AF182345.1] | All sequences obtained from GenBank |
| *Rostraria pubescens* (Desf.) Tzvelev | *Soreng 3793* (BH) |
| **Saccharum officinarum* L. [all five genes NC006084.1] | All sequences obtained from GenBank |
| *Schizachne purpurascens* (Torr.) Swallen | *Soreng 3348* (BH) |
| *Sclerochloa dura* (L.) P. Beauv. | *Soreng 3862* (BH) |
| *Sesleria caerulea* (L.) Ard. | *Davis 768* (BH); original collection *Scholz s. n.* (B) |
| *Sinochasea trigyna* Keng | *Soreng 5644* (US) |
| *Sphenopus divaricatus* (Gouan) Rchb. | *Soreng 3700* (BH) |
| *Sporobolus giganteus* Nash | *Peterson et al. 10008* (US) |
| *Stipa barbata* Desf. | *Davis 768* (BH); grown from USDA Plant Intro. Sta. 229468 |

Table 1. Continued.

| Taxon | Voucher information for sequences generated by authors |
|---|---|
| *Stipagrostis zeyheri* (Nees) DeWinter [*rbc*L U31378.1] | *Barker 1133* (BOL) |
| *Streptochaeta sodiroana* Hack.; **S. angustifolia* Soderstr. [*mat*K AF164382.1] | *Peterson & Judziewicz 9525* (US) [*ndh*F AY622318.1; *rbc*L AY632372.1] |
| *Timouria saposhnikovii* Roshev. | *Soreng 5448* (US) |
| *Torreyochloa pauciflora* (J. Presl) G. L. Church | *Davis 533* (BH) |
| *Trikeraia pappiformis* (Keng) P. C. Juo & S. L. Lu | *Soreng 5653* (US) |
| *Triplachne nitens* (Guss.) Link | *Soreng 3701* (BH) |
| *Trisetum canescens* Buckley | *Soreng 3383a* (BH) |
| **Triticum aestivum* L. [all five genes NC002762.1] | All sequences obtained from GenBank |
| *Uniola paniculata* L. [*mat*K AF144607.1] | No voucher [*rbc*L AY632373.1] |
| *Vahlodea atropurpurea* (Wahlenb.) Fr. ex Hartm. | *Soreng 6316* (S) |
| *Vulpia microstachys* (Nutt.) Munro | *Soreng 7406* (US) |
| **Zea mays* L. [all fives genes NC001666.2] | All sequences obtained from GenBank |

corresponds to sites 1932–1937 in the reference sequence of *Triticum*. The sequence of these six nucleotides is GAAAAA or its reverse complement, TTTTTC, in most taxa in the sample. As with the three-nucleotide inversion, most taxa with sequences other than these two combinations differ from one or the other of these sequences at no more than one of the sites (e.g., the sequence in this region is TAAAAA in *Triticum*). These two inversions were scored as binary characters for the analysis, as was the 15-nucleotide indel described above, and eight additional indels observed among the four genes. In the scoring of the two inversion characters, taxa with sequences corresponding exactly to the most commonly observed sequences described above, which are reverse complements (TAC or GTA for the three-nucleotide inversion, and GAAAAA or TTTTTC for the six-nucleotide inversion), and taxa differing from these four standard sequences at no more than one site, were scored as having one state or the other. Taxa with sequences other than these were scored as unknown.

One additional structural character, presence/absence of GBSSI intron 10, also was encoded as a binary character. Observations relating to this character were obtained from sequences in GenBank and from new sequences generated in the course of the present work, which otherwise are not included in the present analysis. Like the other structural characters, presence/absence of GBSSI intron 10 was encoded as a binary character and included in the analysis.

### *Data Analysis*

Nucleotide sequences of the four genes and scores for the structural characters described above were combined into one matrix that was subjected to cladistic analysis, with all characters weighted equally and treated as nonadditive (i.e., the states unordered) during tree searches, and with *Joinvillea* Gaudich. ex Brongn. & Gris as the outgroup for the purpose of rooting. Parsimony searches and a jackknife analysis were conducted following the removal of cladistically uninformative characters from the data set. Parsimony analyses were conducted with the multi-thread version of NONA vers. 1.6 (i.e., "PARANONA," compiled 26 Feb 1998; Goloboff 1993), using the default polytomy settings, which allow polytomies to occur (*poly* =), and which resolve a clade, rather than a polytomy, only when support for the resolution is unambiguous (*amb-*), with support for a group regarded as unambiguous only when the length of the group's subtending branch is greater than zero under all possible character optimizations. The search strategy involved 1000 individual search initiations, using random taxon entry sequences, with each initiation followed by tree-bisection-reconnection (TBR) swapping with up to 20 shortest trees retained and subjected to additional branch swapping, using the command *mult**, preceded by *rs 0* and *hold/20*. Six unique most-parsimonious trees were generated by these 1000 search initiations, and because all of these trees already had been subjected to complete rounds of branch swapping without generating additional trees, no further branch swapping was conducted.

Support for clades resolved by the cladistic analysis was assessed by jackknife analysis (Farris et al. 1996), using WinClada vers. 1.00.08 (Nixon 2002) running NONA as a daughter process for tree searches, and employing the same character and polytomy settings that were used in the basic analyses of relationships. The jackknife analysis consisted of 1000 replicates, each replicate consisting of four search initiations, with up to 20 trees retained during TBR swapping after each search initiation (*hold/20; mult*4*), followed by additional TBR swapping of all shortest trees, including those generated during this phase of swapping, with up to 100 trees retained (*hold 100; max**).

## RESULTS

### *Data Characteristics*

After the removal of ambiguously aligned regions, the total number of aligned nucleotide sites in each gene, along with the number of cladistically informative sites, and the percentage of aligned sites that are informative, are as follows: *ndh*F (2129, 594, 28%); *ndh*H (1206, 275, 23%); *mat*K (1590, 544, 34%); *rbc*L (1344, 237, 18%). Also, there are 13 cladistically informative structural characters. Thus, a total of 1663 informative characters were included in the analysis, with ca. 36% of these representing nucleotide sequence variation in *ndh*F, 33% in *mat*K, 17% in *ndh*H, 14% in *rbc*L, and less than 1% representing genomic structural characters. Cladistic analysis of the combined matrix of 1663 informative characters resolved six most-parsimonious trees of

length 6946, with a consistency index (CI) of 0.36 and a retention index (RI) of 0.73. Lengths, CIs, and RIs on these trees for the informative characters of the five data partitions are as follows: *ndh*F (2626–2627, 0.36, 0.73); *ndh*H (1261–1266, 0.31, 0.70); *mat*K (1987–1992, 0.42, 0.77); *rbc*L (1033–1034, 0.31, 0.71); structural characters (33, 0.51, 0.86). Because sequences of all four genes were not available for all 108 taxa, these numbers partially reflect the availability of more sequences for some genes than for others, but an analysis of the subset of 104 taxa that are complete for all four genes yielded numbers that differ only negligibly from those that are reported.

*Phylogenetic Relationships*

The consensus of the six most-parsimonious trees is depicted in Fig. 2 and 3. In these trees the grass family is resolved as monophyletic. Relationships among groups within Poaceae are described here in terms of the GPWG (2001) classification, to the extent that the present taxon sampling duplicates that of the GPWG. Sampling of Pooideae in the present analysis is substantially more extensive than in the GPWG analysis, and several taxa placed by the present analysis within a clade that is conventionally recognizable as Pooideae have seldom or never been included in Pooideae in previous treatments. Thus, we provisionally designate the clade that consists of *Brachyelytrum* P. Beauv. and its sister group (Fig. 2, 3) as Pooideae, and to minimize repetition a detailed description of that group is presented only in the Discussion.

Of the 12 subfamilies delimited by the GPWG (2001), all except Pharoideae, Centothecoideae, and Aristidoideae are represented in this analysis by more than one exemplar each, and thus subject to testing for monophyly. With the matter of Pooideae temporarily set aside, all but one of the remaining eight subfamilies are resolved as monophyletic (Fig. 2, 3). The exception is Anomochlooideae. The two elements of this subfamily that are sampled in the present study, representing the only two genera in the subfamily, are united as a clade in two of the six trees. In the other four, *Streptochaeta* is placed as the sister of a clade that includes all other grasses, with *Anomochloa* Brongn. as the next line to diverge from the clade that includes all remaining grasses. The consensus of these structures (Fig. 2) is a trichotomy at the base of the family, with these two genera and a clade consisting of all other grasses emerging from one point.

Apart from those genera placed by the present analysis with other elements of Pooideae (i.e., the clade consisting of *Brachyelytrum* and its sister group), this analysis includes only three genera that were not sampled by the GPWG (2001). The placement of these three genera by the present analysis is consistent with their conventional taxonomic placements (*Saccharum* L. in Panicoideae, and *Bambusa* Schreb. and *Guadua* Kunth in Bambusoideae).

Following the divergence of elements of Anomochlooideae from the clade that includes all other grasses, Pharoideae and Puelioideae are the next two groups to diverge, in succession, from the clade that includes all remaining grasses (Fig. 2). Of the remaining grasses, the PACCAD clade is resolved as monophyletic, as is the BEP clade, and these two clades are resolved as sister groups. The present sampling within the PACCAD clade is minimal, and we do not discuss relationships within this group further, except to note that a clade consisting of Centothecoideae and Panicoideae (the latter including *Gynerium* Willd. ex P. Beauv.; see below) is sister to one that includes the other four subfamilies, as in the results of the GPWG (2001). Within the BEP clade, Ehrhartoideae are placed as the sister of a clade in which Bambusoideae and Pooideae are sister taxa.

*Inversions and Gene Positions in the Plastid Genome*

The sequence in the three-nucleotide inversion region in *ndh*F is either TAC, or a sequence differing from TAC at one site, or "unknown" in all taxa within the clade that consists of *Brachypodium* P. Beauv. and its sister group (Fig. 3), with the one exception of *Hordeum* L. In particular, the sequence in both species of *Brachypodium* is TAC. The principal alternative sequence (GTA), or a sequence differing from GTA at one site, or "unknown" occurs in all other taxa in the sample that have been scored, including *Hordeum* and all four representatives of Meliceae (the sequence is GTA in all five of these taxa). The transformation from GTA to TAC therefore is interpreted as a synapomorphy of the clade that is depicted in Fig. 3 (internode C), and *Hordeum* is interpreted as having experienced a reversion to the plesiomorphic state.

The sequence of the six-nucleotide inversion region in *ndh*F occurs as state TTTTTC in *Avena* L., *Helictotrichon* Besser ex Schult. & Schult. f., *Trisetum* Pers., *Rostraria* Trin., and *Gaudinia* P. Beauv. In all other taxa in the sample the sequence in this region is GAAAAA, or a sequence differing from this sequence at one nucleotide, or a sequence scored "unknown." The transformation from GAAAAA to TTTTTC therefore is interpreted as a nonhomoplasious synapomorphy of the specified five-taxon clade (Fig. 3, internode E).

The deleted form of the 15-nucleotide indel in *ndh*F occurs in *Joinvillea, Ecdeiocolea* F. Muell., *Anomochloa, Streptochaeta, Pharus* P. Browne, both species of *Oryza,* and *Nardus* L. The undeleted form occurs in all other taxa in the sample. As previously reported by Clark et al. (1995, 2000) and the GPWG (2001), this character is interpreted most parsimoniously as an insertion in the lineage that includes all grasses except those of Anomochlooideae and Pharoideae, followed by a deletion within Ehrhartoideae, along the line leading to *Oryza* in the present analysis. The present data also imply a second deletion of this region, in *Nardus* (Fig. 2, internode A and two internodes A′). The GPWG (2001) analysis included an *ndh*F sequence from *Nardus,* but it was not definitive in the region of this indel, so this is the first report of a deletion event in this region in *Nardus.*

Among taxa that have been examined for the inclusion of portions of *ndh*F and *ndh*H within the IR region, *ndh*H extends into the IR region in nearly all, including *Joinvillea, Ecdeiocolea, Anomochloa, Streptochaeta,* and *Pharus.* In five taxa (*Amphipogon* R. Br., *Gynerium, Molinia* Schrank, *Sporobolus* R. Br., and *Pleuropogon* R. Br., the first four of which are elements of the PACCAD clade), the 5′-terminus of *ndh*H does not extend into the IR region. In light of the distribution of states of this character, the occurrence of a portion of *ndh*H within the IR region is interpreted as a

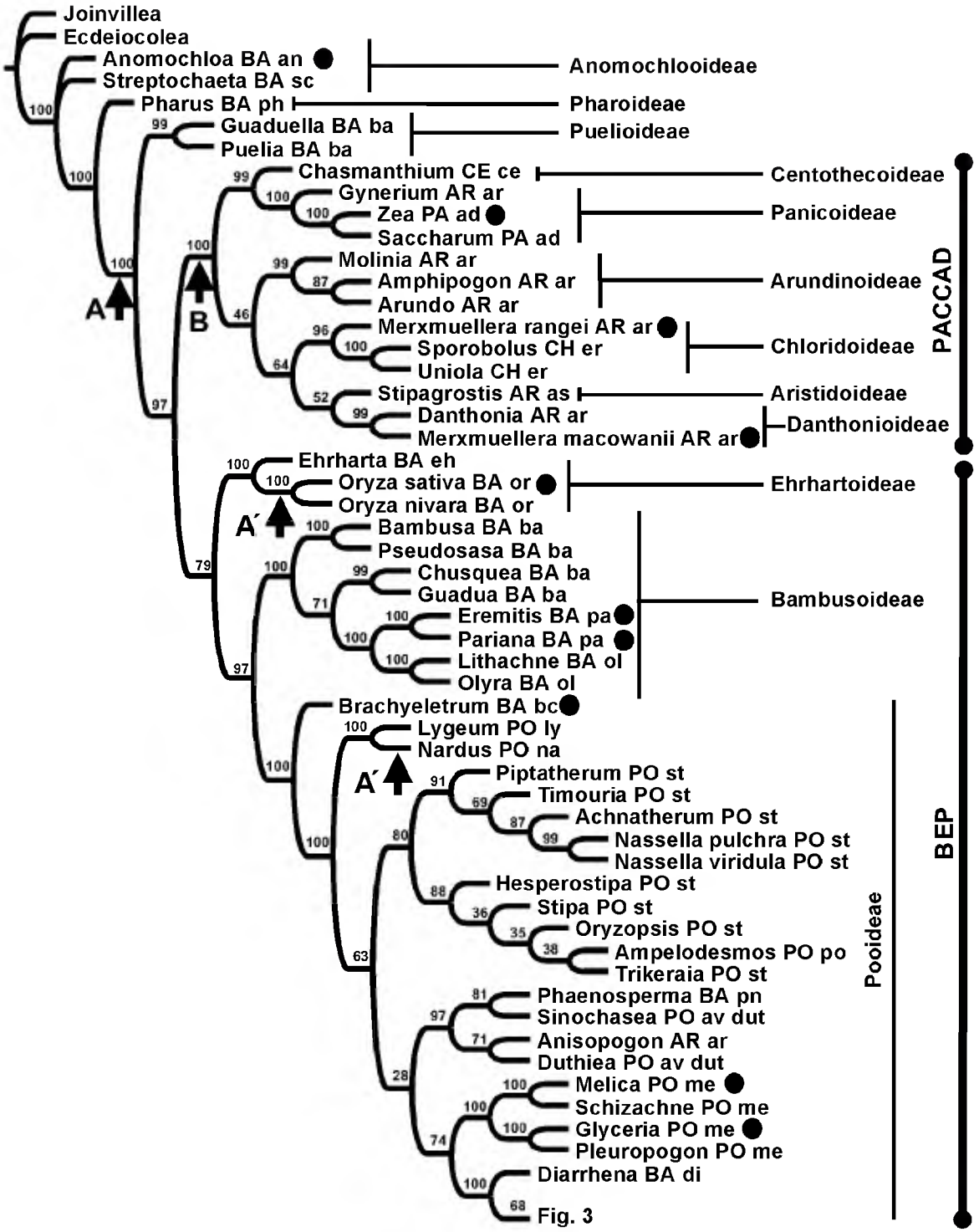


Fig. 2.—Consensus of six most-parsimonious trees resolved by cladistic analysis of nucleotide sequence variation in *ndh*F, *ndh*H, *mat*K, *rbc*L, and structural characters of the plastid and nuclear genomes from 108 taxa (Table 1). For representatives of Poaceae (i.e., all taxa except *Joinvillea* and *Ecdeiocolea*), taxonomy according to Clayton and Renvoize (1986; with some changes in spelling) is indicated by two-letter codes in upper case for subfamily (AR = Arundinoideae; BA = Bambusoideae; CE = Centothecoideae; CH = Chloridoideae; PA = Panicoideae; PO = Pooideae), two-letter codes in lower case for tribe (ad = Andropogoneae; an = Anomochloeae; ar = Arundineae; as = Aristideae; av = Aveneae; ba = Bambuseae; bc = Brachyelytreae; br = Bromeae; ce = Centotheceae; di = Diarrheneae; eh = Ehrharteae; er = Eragrostideae; ha = Hainardieae; ly = Lygeeae; me = Meliceae; na = Nardeae; ol = Olyreae; or = Oryzeae; pa = Parianeae; ph = Phareae; pn = Phaenospermateae; po = Poeae; sc = Streptochaeteae; st = Stipeae; tr = Triticeae), and three-letter codes in lower case for subtribes of Aveneae (alo = Alopecurinae; ave = Aveninae; dut = Duthieinae; pha = Phalaridinae). Taxonomy adopted by the GPWG (2001) is indicated by labels at right, with some modifications (see text). Numbers above internodes are jackknife percentages. Circles to the right of 24 names depict presence (closed) or absence (open) of GBSSI intron 10. Character transformations of selected characters are indicated by arrows as follow: A = 15 base pair (bp) insertion in *ndh*F; A′ = 15 bp deletion in *ndh*F; B = extension of a portion of *ndh*F from the Short Single Copy region into the Inverted Repeat region (three other occurrences are not mapped—see text).

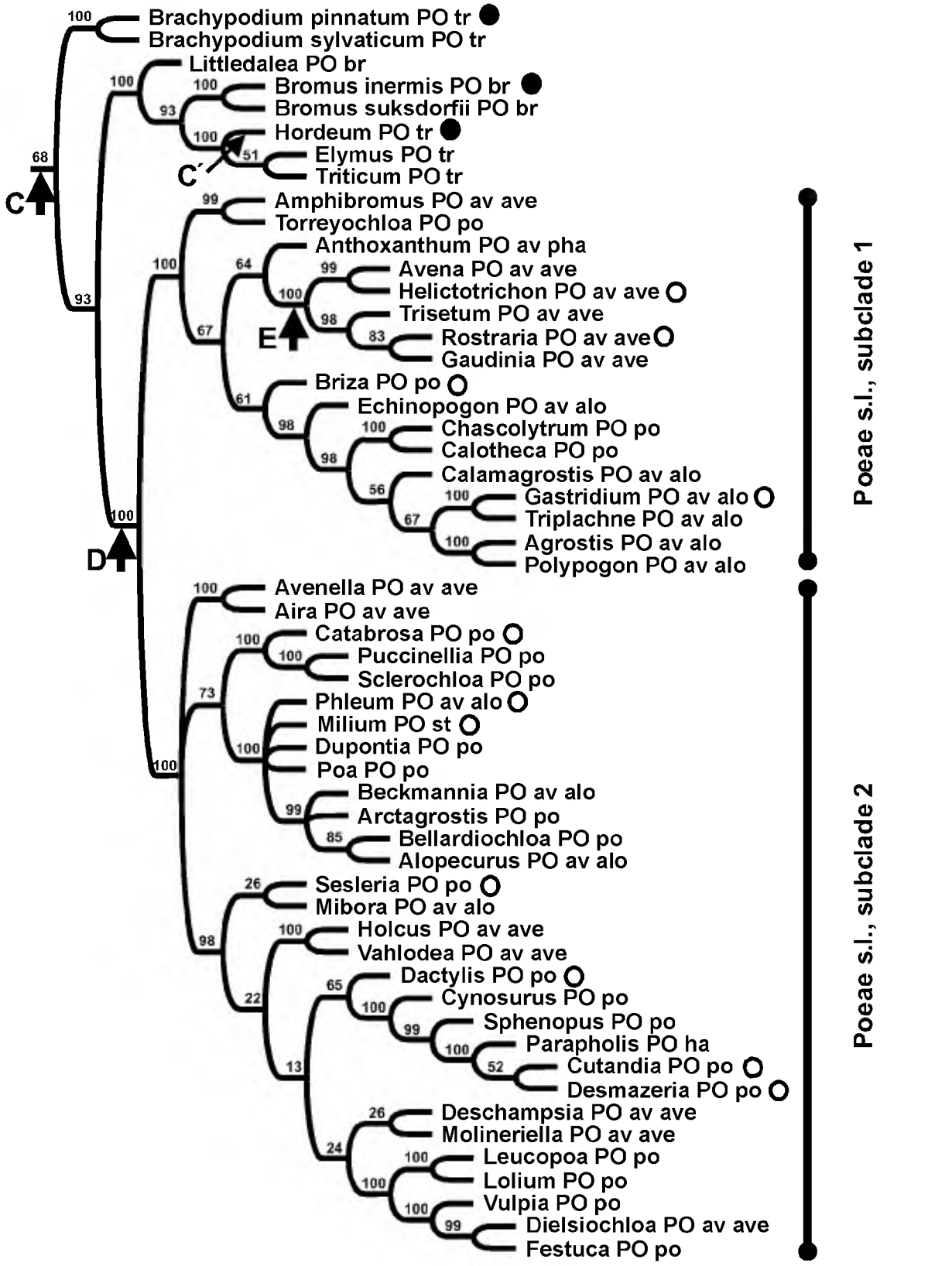


Fig. 3.—Continuation of consensus tree depicted in Fig. 2. Codes and labels are as in Fig. 2. Transformations of selected characters are indicated by arrows as follow: C = three-nucleotide inversion in *ndh*F; C′ = reversal of three-nucleotide inversion in *ndh*F; D = loss of GBSSI intron 10; E = six-nucleotide inversion in *ndh*F.

plesiomorphy for the grasses, shared at least with Joinvilleaceae and Ecdeiocoleaceae, and the 5′-end of *ndh*H is interpreted as having migrated out of the IR region once within Pooideae (in *Pleuropogon*) and at least three times within the PACCAD clade, where one of the transformations may be synapomorphic for *Amphipogon* and *Molinia.*

The 3′-end of *ndh*F extends into the IR region in all taxa of the PACCAD clade that have been examined, as well as in *Ehrharta* Thunb., *Olyra* L., and *Brachypodium pinnatum.* This gene is interpreted as having terminated within the SSC region in the earliest grasses, with its 3′-terminus having migrated into the IR region four times, with one of these occurrences being a synapomorphy of the PACCAD clade (Fig. 2, internode B; other three occurrences not depicted).

### *GBSSI Intron 10*

The presence or absence of intron 10 of GBSSI has been determined for 24 taxa in the sample (Fig. 2, 3). Of these 24 taxa, 13 have the intron and 11 lack it. Loss of the intron is interpreted as a nonhomoplasious synapomorphy of a clade within Pooideae (Fig. 3, internode D).

## DISCUSSION

Two of the four genes used in the analysis, *ndh*F and *mat*K, together contribute more than two-thirds of the characters in the analysis. One of these two, *ndh*F, is substantially longer than each of the other three, while *mat*K is the second longest. Also, *mat*K is more variable than any of the other three, as measured by the percentage of aligned sites that are cladistically informative. The other two genes, *ndh*H and *rbc*L, are similar to each other in length and in the percentages of characters that are cladistically informative, and together they contribute about a third of the total number of informative characters in the data set. Homoplasy levels are similar among three of the genes, while *mat*K, the outlier, has the highest consistency and retention indices of the four. Thus, *mat*K is both more variable and less homoplasious than the other three genes.

Relationships resolved by the present analysis should be regarded as tentative, and for this reason we do not discuss all noteworthy aspects of the phylogenetic results, particularly outside subfamily Pooideae. However, some aspects of these results deserve comment. The present analysis is consistent with the GPWG (2001) analysis in placing representatives of Anomochlooideae, Pharoideae, and Puelioideae as the first lineages to diverge within the grass family from a large clade that includes the remaining nine subfamilies (Fig. 2), and that is marked by the 15-nucleotide insertion in *ndh*F. Within this clade, the length of *ndhF* in this region is restored to its original (i.e., shorter) state in two of the sampled taxa, *Ehrharta* and *Nardus,* by what appear to be separate 15-nucleotide deletion events (Fig. 2, character transformation A′. Despite the general similarity of the present results to those of the GPWG analysis, with respect to the early diverging grass lineages, the absence in some trees of a clade consisting of *Anomochloa* and *Streptochaeta* suggests that these two anomalous grasses may not be closest relatives. The placement of these two genera as separately diverging lineages is consistent with the results of Hilu et al. (1999), whose analysis placed *Streptochaeta* as the sister of all other grasses, with *Anomochloa* placed as the next lineage to diverge from the group that includes all other grasses, as in some but not all of the most-parsimonious trees resolved by the present analysis. The vegetative and reproductive structures of *Anomochloa* and *Streptochaeta* are quite different in form (Judziewicz and Soderstrom 1989), and these taxa warrant further attention.

The PACCAD clade, a group consisting of six of the remaining nine subfamilies, as resolved by the GPWG (2001) and other analyses, also is resolved by the present analysis (Fig. 2). The extension of the 3′-end of *ndh*F into the IR region appears to be a synapomorphy of this clade (Fig. 2, character transformation B). *Gynerium,* a taxon of uncertain relationship, was resolved by the GPWG analysis as the sister of Panicoideae, and was treated there as incertae sedis. It has since been assigned formally to that subfamily (Sánchez-Ken and Clark 2001), within the tribe Gynerieae, and is treated here as an element of Panicoideae (Fig. 2). The placement of *Gynerium* by the present analysis, as sister of all other Panicoideae, is consistent with the results of the GPWG analysis, as is the placement of Centothecoideae (represented here by *Chasmanthium* Link) as sister of Panicoideae, with this grouping of two subfamilies placed as the sister of a clade that includes all remaining elements of the PACCAD clade. An alternative taxonomic treatment is provided by Zuloaga et al. (2003), who recognize an expanded Panicoideae that includes both Centotheceae (i.e., Centothecoideae) and Gynerieae.

With respect to placement of the endpoints of *ndh*F and *ndh*H, relative to the boundaries between the SSC and IR regions of the plastid genome, the present analysis supports a different interpretation than that of Ogihara et al. (2002). Those authors regarded the pattern that they observed as evidence of a closer relationship between *Oryza* and *Triticum* than between either of these taxa and *Zea* (i.e., the presence of a substantial portion of the 5′-end of *ndh*H in the IR region in *Oryza* and *Triticum,* as opposed to the presence of only one nucleotide of *ndh*H in the IR region in *Zea*; and the presence of the 3′-end of *ndh*F in the IR region in *Zea,* as opposed to termination of *ndh*F within the SSC region in *Oryza* and *Triticum*). The present analysis does support a closer phylogenetic relationship between *Oryza* and *Triticum* than between either of these taxa and *Zea,* as have most previous phylogenetic analyses that have addressed this matter (e.g., GPWG 2001 and several papers cited therein). Indeed, any analysis that resolves a BEP clade, with *Oryza* in Ehrhartoideae and *Triticum* in Pooideae, supports a closer relationship between these two taxa than between either of them and *Zea.* However, placement of the endpoints of *ndh*F and *ndh*H, relative to the boundaries between the major plastid genomic regions, does not in itself provide evidence in support of this relationship. As demonstrated here, the states of these characters, as observed in *Oryza* and *Triticum,* also occur in the early diverging lineages within the grasses, as well as outside the family, and thus are best interpreted as plesiomorphies of the grasses. What these features do provide, in terms of phylogenetic evidence, is support for monophyly of the PACCAD clade, as already indicated. Extension of the 3′-end of *ndh*F into the IR region is observed in all taxa of the PACCAD clade that have been examined, as well as in three other isolated taxa. These occurrences, in the

context of the relationships resolved by this analysis, are best explained as having arisen by four parallel migrations of the 3′-end of *ndh*F into the IR region. One of these occurrences marks the PACCAD clade, and the other three occurrences, once within each subfamily of the BEP clade, are interpreted as autapomorphies.

Like the analysis by the GPWG (2001), and various earlier analyses, the present study resolves the BEP clade. Within this clade, the present analysis resolves all three subfamilies as monophyletic (Fig. 2), as have several earlier analyses. However, the present analysis places Bambusoideae and Pooideae as sister taxa, with 97% jackknife support, with Ehrhartoideae placed as the sister of the clade that includes these two. In contrast, the GPWG (2001) analysis, like that of Clark et al. (1995) and others, placed Ehrhartoideae and Bambusoideae as sisters, with Pooideae the sister of the clade that consisted of those two subfamilies. Alternative phylogenetic structures for these three subfamilies have been resolved by other analyses, including those of various subsets of the GPWG data matrix, and in some analyses there is no BEP clade (e.g., Soreng and Davis 1998; Hilu et al. 1999). Even when the BEP clade is absent, the PACCAD clade generally is resolved, often with Pooideae placed as its sister, with Bambusoideae and Ehrhartoideae then situated either as sister taxa or as successively diverging lineages nearby. A complete review of the various arrangements that have been resolved among these taxa by other data sets is beyond the scope of this paper, and we simply note that relationships among Bambusoideae, Ehrhartoideae, Pooideae, and the PACCAD clade are not yet firmly established, and refer readers to the trees depicted in Appendix III of the GPWG (2001) analysis, and to the various contributions cited in that paper.

The most comprehensive previous analysis of phylogenetic relationships within Pooideae was that of Soreng and Davis (2000), which was based on morphological characters and restriction sites from the plastid genome. There the following set of relationship among conventionally delimited tribes and isolated genera was resolved within the subfamily: (Brachyelytreae ((Lygeeae Nardeae) ((*Anisopogon* R. Br. (Stipeae including *Ampelodesmos* Link)) (Diarrheneae ((Brachypodieae Meliceae) ((Bromeae Triticeae) Poeae s.l.)))))). Although *Brachyelytrum* was used as the outgroup for the published tree in that analysis, preliminary analyses utilizing taxa outside Pooideae established the sister group relationship between this genus and other members of the subfamily. The present analysis includes 77 representatives of Pooideae, 24 fewer taxa overall than the 101 representatives sampled by Soreng and Davis (2000), but with multiple species sampled within only three genera in the present study, as opposed to 18 genera sampled by two or three species each in the earlier analysis. Each analysis includes some taxa that the other does not, so all relationships resolved by one cannot be compared to those in the other. However, there is sufficient overlap in taxon sampling to facilitate comparison of the results. In each case, taxa of the traditionally recognized tribes Aveneae, Hainardieae, and Poeae (i.e., Poeae s.s.), plus *Milium* L., which was included in Stipeae by Clayton and Renvoize (1986), are placed within a clade that we continue to designate as the tribe Poeae s.l. In the present analysis we identify the apparently nonhomoplasious loss of GBSSI intron 10 as a synapomorphy of this clade (Fig. 3, character transformation D). This clade has high jackknife support (100%) with the intron character either included or excluded from the analysis. *Milium* is one of the 11 taxa for which the absence of this intron has been observed, and this congruence between a marker in the nuclear genome and the plastid-encoded genes that determine the structure of the tree suggests that this genus does belong within Poeae s.l.

The present analysis also agrees with that of Soreng and Davis (2000) in placing a clade consisting of Bromeae and Triticeae as sister of Poeae s.l. (Fig. 3), thus forming a group that we designate the BT/P clade (93% jackknife support). It should be noted, however, that Bromeae are not monophyletic in the present analysis, because of the placement of *Littledalea* Hemsl. as sister of *Bromus* L. plus Triticeae. A similar result is reported by Saarela et al. (2007), whose sampling within Bromeae is broader than that of the present analysis; *Littledalea* was not sampled by Soreng and Davis (2000). A critical difference between the results of the present analysis and those of Soreng and Davis (2000) lies in the placement of Brachypodieae (represented by *Brachypodium*), rather than a clade consisting of Brachypodieae plus Meliceae, as sister of the BT/P clade. The present analysis identifies a three-nucleotide inversion in *ndh*F as a synapomorphy of the clade that consists of Brachypodieae and its sister group, the BT/P clade (Fig. 3, character transformation C). This inversion has not been observed in any other species in the sample, including the various representatives of Meliceae. This character is homoplasious, since it is reversed to the uninverted state in *Hordeum* (Fig. 3, transformation C′), but the placement of Meliceae as sister of Brachypodieae would require an additional homoplasious transformation. Brachypodieae also were placed as the sister of Bromeae, Triticeae, and Poeae s.l., with Meliceae placed outside this group, by Hilu et al. (1999).

Several other differences within Pooideae, with respect to relationships among tribes and isolated genera, also exist between the trees resolved by the present analysis and those of Soreng and Davis (2000). The two analyses agree in placing Brachyelytreae as the sister of all other elements of Pooideae, and in placing Lygeeae and Nardeae as sisters of each other, with this small clade the sister of one that consists of all remaining elements of the subfamily (Fig. 2). The two analyses also agree in placing *Ampelodesmos* among genera conventionally assigned to Stipeae (Fig. 2). However, the analysis of Soreng and Davis (2000) places *Anisopogon* as the sister of this overall group, which itself is the sister of a clade that includes all representatives of Pooideae except Brachyelytreae, Lygeeae, and Nardeae. The present analysis differs in placing *Anisopogon* with three additional disparate genera, *Duthiea* Hack., *Sinochasea* Keng, and *Phaenosperma* Munro ex Benth. (Fig. 3). This group of four genera diverges just after Stipeae (including *Ampelodesmos*) from the clade that includes the BT/P clade. We tentatively refer to this group of four genera (plus other elements of subtribe Duthieinae sensu Clayton and Renvoize 1986) as Phaenospermateae. Of the taxa in this group, only *Anisopogon* was sampled by Soreng and Davis (2000), and the present analysis is the first report of a phylogenetic placement of *Duthiea* and *Sinochasea* on the basis of DNA sequence variation. The GPWG (2001) resolved a clade consisting of *Phaenosperma*

and *Anisopogon,* but did not sample the other two genera that are placed in the group by the present analysis. This set of genera and their presumed relatives have been difficult to classify on the basis of traditional characters. *Anisopogon* was placed in Arundinoideae by Clayton and Renvoize (1986) and Watson and Dallwitz (1992). *Duthiea* and *Sinochasea* (the latter as a synonym of *Pseudodanthonia* Bor & C. E. Hubb.) were placed, with *Metcalfia* Conert and *Stephanachne* Keng, in Pooideae tribe Aveneae subtribe Duthieinae by Clayton and Renvoize, while Watson and Dallwitz expressed their doubts by placing these either in Arundinoideae tribe Danthonieae or Pooideae tribe Aveneae, in both cases annotated with question marks. *Phaenosperma* was placed in Bambusoideae, in its own tribe, Phaenospermateae, by Clayton and Renvoize, and in the supertribe Oryzodae by Watson and Dallwitz.

The two remaining tribes of Pooideae, Diarrheneae and Meliceae, also are placed differently by the present analysis and that of Soreng and Davis (2000). In the present analysis, Meliceae and Diarrheneae diverge in succession from the large group that includes Brachypodieae and the BT/P clade (Fig. 2, 3), but the analysis of Soreng and Davis (2000) placed Diarrheneae as the first of these elements to diverge after Stipeae plus *Anisopogon,* with Meliceae plus Brachypodieae the next lineage to diverge.

In light of these various similarities and differences, the picture that emerges is one in which Brachyelytreae and a clade consisting of Lygeeae plus Nardeae appear to have diverged in succession from the lineage that includes all other members of Pooideae. Within the latter group are five smaller groups, these being Diarrheneae, Meliceae, Phaenospermateae (should it continue to be resolved), Stipeae (including *Amplelodesmos*), and a clade consisting of Brachypodieae and the BT/P clade, which represents the major radiation within the subfamily. Relationships among these five clades still are unclear.

We turn now to Poeae s.l., which corresponds closely to Poeae s.l. as resolved by Soreng and Davis (2000), though the taxon sampling varies between that analysis and the present one. Within Poeae s.l., the present analysis resolves two principal clades, labeled as Poeae s.l. subclades 1 and 2 in Fig. 3. To avoid excessive repetition, we refer to these clades through the remainder of the discussion as subclades 1 and 2. Subclade 1 consists mostly of elements of Clayton and Renvoize's (1986) tribe Aveneae (minus subtribe Duthieinae), but it also includes four taxa that those authors assigned to Poeae s.s. (*Torreyochloa* G. L. Church, *Briza* L. s.s., *Calotheca* Desv., and *Chascolytrum* Desv., the latter two of which they included in *Briza* s.l.). Subclade 2 includes most of the sampled taxa from Clayton and Renvoize's tribe Poeae, plus tribe Hainardieae, *Milium* (which they assigned to Stipeae), and several groups of genera that they assigned to Aveneae.

A reconsideration of relationships in Pooideae, on the basis of morphologically defined groupings and ongoing phylogenetic studies, is reflected in the revised classification of the subfamily proposed by Soreng et al. (2003), which is updated at: *http://nobot.mobot.org/W3T/Search/nwgclass.html.* Suprageneric taxa in the following discussion reflect groups that are recognized within this developing system. Within subclade 1, the present analysis and that of Soreng and Davis (2000) resolve two major groups of taxa traditionally assigned to Aveneae, which can be recognized as subtribes (though not all of the following genera were sampled in both studies): Aveninae (*Avena, Gaudinia, Helictotrichon, Rostraria, Trisetum*); and Agrostidinae (*Agrostis* L., *Ammophila* Host, *Calamagrostis* Adans., *Gastridium* P. Beauv., *Polypogon* Desf., *Triplachne* Link). The Aveninae clade is marked by the nonhomoplasious six-nucleotide inversion in *ndh*F described above (Fig. 3, character transformation E). Other isolated elements are placed among these clades, with positions of some of the genera differing between the studies. Subtribe Phalaridinae, represented by *Anthoxanthum* L. in the present study, was resolved by Soreng and Davis (2000) as sister of a clade that included elements of *Briza* s.l. plus Agrostidinae, but it is placed as sister of Aveninae in the present analysis. The six-nucleotide inversion is absent in *Anthoxanthum,* so this character does not favor either of these placements over the other. It is of interest that *Briza* s.l. is non-monophyletic in both studies, and that in the present study *Echinopogon* P. Beauv. (endemic to Australia) is placed, within a pectinate structure, between *Briza* s.s. from Europe and other elements of *Briza* s.l. from South America (including *Calotheca* and *Chascolytrum,* which were also associated with *Poidium* Nees in the earlier analysis). Also, *Amphibromus* Nees was the sister of all other elements of subclade 1 in the analysis of Soreng and Davis (2000), but in the present analysis it is joined in this position by *Torreyochloa* (i.e., these two taxa constitute a clade that is placed as sister of the rest of subclade 1). This clade of two genera also was resolved by Soreng and Davis (2000) in analyses of only the restriction site data, and these wetland genera are provisionally recognizable as constituting the small subtribe Torreyochloinae.

Within subclade 2, the present study and that of Soreng and Davis (2000) resolve some of the same groups of taxa: (A) intermixed elements of subtribes Alopecurinae s.s. (*Alopecurus* L., *Beckmannia* Host, *Phleum* L.), Poineae (*Arctagrostis* Griseb., *Bellardiochloa* Chiov., *Dupontia* R. Br., *Poa* L.), and Miliinae (*Milium*); (B) subtribes Dactylidinae (*Dactylis* L., which is placed here with restriction site data by Soreng and Davis 2000, but not when combined with morphology), Cynosurinae (*Cynosurus* L.), and Parapholiinae (*Parapholis* C. E. Hubb.), with *Sphenopus* Trin., *Cutandia* Willk., and *Desmazeria* Dumort.; (C) Loliinae (*Festuca* L., *Leucopoa* Griseb., *Lolium* L., *Vulpia* C. C. Gmel.) with *Dielsiochloa* Pilg.; (D) Puccinelliinae (*Catabrosa* P. Beauv., *Puccinellia* Parl., *Sclerochloa* P. Beauv.); and (E) Sesleriinae (incl. Miborinae; *Sesleria* Scop., *Mibora* Adans.). A fairly close relationship was detected between clades B and C in both analyses, but the sister of clade D was clade C in the analysis of Soreng and Davis (2000), while the sister of clade D is clade A in the present analysis, with 73% jackknife support (Fig. 3). *Sesleria* was resolved as the sister of clade A in the previous analysis, but in the present analysis, as part of group E (i.e., along with *Mibora*), it is placed with 98% jackknife support as sister to a clade that includes clades B and C, plus a clade consisting of *Holcus* L. (Holcinae) and *Vahlodea* Fr., and another, consisting of *Deschampsia* P. Beauv. and *Molineriella* Rouy (Fig. 3).

One of the most complex situations within Poeae s.l. involves Airinae (*Aira* L., *Deschampsia* s.l. [including *Ave-*

*nella* (Bluff & Fingerh.) Drejer and *Vahlodea*], Holcinae [*Holcus*], and *Molineriella*). The analysis of Soreng and Davis (2000) also placed *Avenula* (Dumort.) Dumort. in subclade 2, among genera of Airinae and Holcinae. In light of the discrepancy in the placement of *Deschampsia cespitosa* between the two studies (it was placed in subclade I by the analysis of Soreng and Davis 2000), we re-examined the seed sample sent as that species by the USDA Plant Introduction Station and used by Soreng and Davis (2000), and have determined that accession as a species of *Agrostis* s.s. (plants grown from this seed sample have not flowered). Thus, all verified accessions of *Deschampsia* and its segregates are placed by molecular characters within subclade 2. With this matter set aside, Airinae still appear to be nonmonophyletic on the basis of variation patterns in their chloroplast DNA. The sampling of elements that have been recognized within Aveneae subtribe Airinae is more complete in the present analysis than in that of Soreng and Davis (2000), and these taxa are scattered among three lineages within subclade 2. *Deschampsia* s.l. (including *Deschampsia* s.s., *Avenella,* and *Vahlodea*) is not monophyletic, and each of three representatives of this genus is placed in a small clade with another genus (*Molineriella, Aira,* and *Holcus,* respectively) in disparate locations within subclade 2, where they are removed from each other and from the other elements of Aveneae s.l. in subclade 1 (Fig. 3).

The present analysis confirms previous observations that the traditionally recognized tribes Poeae s.s. and Aveneae are phylogenetically intermixed, at least as determined by chloroplast DNA characters. Some of the cases described above may be attributable to error in the structures of these cladograms, relative to the actual history of diversification of the plastid genomes, while others may signify homoplastic changes in the morphological characters that underlie traditional classifications. Reticulation may be a third cause of these unexpected results.

## ACKNOWLEDGMENTS

We thank several colleagues for providing plant materials and DNA isolations used in this work, and Pilar Catalán for kindly inviting us to present this work at the symposium on Systematics of Pooideae at *Monocots III and Grasses IV.* We also thank the Smithsonian Institution (SI) Biodiversity Surveys and Inventories Programs for support to RJS for plant collecting in China, the SI Research Opportunities Fund for support to RJS for collecting in Australia, and the National Science Foundation for support of this project (grant DEB-0318686).

## LITERATURE CITED

ASMUSSEN, C. B., AND M. W. CHASE. 2001. Coding and noncoding plastid DNA in palm systematics. *Amer. J. Bot.* **88**: 1103–1117.

BREMER, K. 2002. Gondwanan evolution of the grass alliance of families. *Evolution* **56**: 1374–1387.

CHASE, M. W., D. E. SOLTIS, R. G. OLMSTEAD, D. MORGAN, D. H. LES, B. D. MISHLER, M. R. DUVALL, R. A. PRICE, H. G. HILLS, Y.-L. QIU, K. A. KRON, J. H. RETTIG, E. CONTI, J. D. PALMER, J. R. MANHART, K. J. SYTSMA, H. J. MICHAELS, W. J. KRESS, K. J. KAROL, W. D. CLARK, M. HEDRÉN, B. S. GAUT, R. K. JANSEN, K.-J. KIM, C. F. WIMPEE, J. F. SMITH, G. R. FURNIER, S. H. STRAUSS, Q.-Y. XIANG, G. M. PLUNKETT, P. S. SOLTIS, S. M. SWENSEN, S. E. WILLIAMS, P. A. GADEK, C. J. QUINN, L. E. EGUIARTE, E. GOLENBERG, G. H. LEARN, JR., S. C. H. BARRETT, S. DAYANANDAN, AND V. A. ALBERT. 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbc*L. *Ann. Missouri Bot. Gard.* **40**: 528–580.

CLARK, L. G., M. KOBAYASHI, S. MATHEWS, R. E. SPANGLER, AND E. A. KELLOGG. 2000. The Puelioideae, a new subfamily of Poaceae. *Syst. Bot.* **25**: 181–187.

———, W. ZHANG, AND J. F. WENDEL. 1995. A phylogeny of the grass family (Poaceae) based on *ndh*F sequence data. *Syst. Bot.* **20**: 436–460.

CLAYTON, W. D., AND S. A. RENVOIZE. 1986. Genera graminum: grasses of the world. *Kew Bull., Addit. Ser.* **13**: 1–389.

DAHLGREN, R. M. T., H. T. CLIFFORD, AND P. F. YEO. 1985. The families of monocotyledons. Springer-Verlag, Berlin, Germany. 520 p.

DAVIS, J. I., AND R. J. SORENG. 1993. Phylogenetic structure in the grass family (Poaceae) as inferred from chloroplast DNA restriction site variation. *Amer. J. Bot.* **80**: 1444–1454.

DOYLE, J. J., J. I. DAVIS, R. J. SORENG, D. GARVIN, AND M. J. ANDERSON. 1992. Chloroplast DNA inversions and the origin of the grass family (Poaceae). *Proc. Natl. Acad. Sci. U.S.A.* **89**: 7722–7726.

DUVALL, M. R., J. I DAVIS, L. G. CLARK, J. D. NOLL, D. H. GOLDMAN, AND J. G. SÁNCHEZ-KEN. 2007. Phylogeny of the grasses (Poaceae) revisited, pp. 237–247. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson [eds.], Monocots: comparative biology and evolution—Poales. Rancho Santa Ana Botanic Garden, Claremont, California, USA.

FARRIS, J. S., V. A. ALBERT, M. KÄLLERSJÖ, D. LIPSCOMB, AND A. G. KLUGE. 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12**: 99–124.

FRUGOLI, J. A., M. A. MCPEEK, T. L. THOMAS, AND C. R. MCCLUNG. 1998. Intron loss and gain during evolution of the catalase gene family in angiosperms. *Genetics* **149**: 355–365.

GOLOBOFF, P. A. 1993. NONA vers. 1.6. Tucumán, Argentina. Published by the author. Available at: *http://www.cladistics.org/education.html* (Jun 2005).

GRASS PHYLOGENY WORKING GROUP [GPWG]. 2001. Phylogeny and subfamilial classification of the grasses (Poaceae). *Ann. Missouri Bot. Gard.* **88**: 373–457.

HILU, K. W., L. A. ALICE, AND H. LIANG. 1999. Phylogeny of Poaceae inferred from *mat*K sequences. *Ann. Missouri Bot. Gard.* **86**: 835–851.

HIRATSUKA, J., H. SHIMADA, R. WHITTIER, T. ISHIBASHI, M. SAKAMOTO, M. MORI, C. KONDO, Y. HONJI, C. R. SUN, B. Y. MENG, Y. Q. LI, A. KANNO, Y. NISHIZAWA, A. IHRAI, K. SHINOZAKI, AND M. SUGIURA. 1989. The complete sequence of the rice (*Oryza sativa*) chloroplast genome: intermolecular recombination between distinct tRNA genes accounts for a major plastid DNA inversion during the evolution of the cereals. *Molec. Gen. Genet.* **217**: 185–194.

HOLMGREN, P. K., N. H. HOLMGREN, AND L. C. BARNETT. 1990. Index herbariorum. Part I: the herbaria of the world. Ed. 8. New York Botanical Garden, New York, USA. 693 p. Available at: *http://sciweb.nybg.org/science2/IndexHerbariorum.asp* (Jun 2005).

HSIAO, C., S. W. L. JACOBS, N. J. CHATTERTON, AND K. H. ASAY. 1999. A molecular phylogeny of the grass family (Poaceae) based on the sequences of nuclear ribosomal DNA (ITS). *Austral. Syst. Bot.* **11**: 667–688.

INGRAM, A. L., AND J. J. DOYLE. 2003. The origin and evolution of *Eragrostis tef* (Poaceae) and related polyploids: evidence from nuclear *waxy* and plastid *rps*16. *Amer. J. Bot.* **90**: 116–122.

ITCHODA, N., S. NISHIZAWA, H. NAGANO, T. KUBO, AND T. MIKAMI. 2002. The sugar beet mitochondrial *nad*4 gene: an intron loss and its phylogenetic implication in the Caryophyllales. *Theor. Appl. Genet.* **104**: 209–213.

JUDZIEWICZ, E. J., AND T. R. SODERSTROM. 1989. Morphological, anatomical, and taxonomic studies in *Anomochloa* and *Streptochaeta* (Poaceae: Bambusoideae). *Smithsonian Contr. Bot.* **68**: 1–52.

KATAYAMA, H., AND Y. OGIHARA. 1996. Phylogenetic affinities of the grasses to other monocots as revealed by molecular analysis of chloroplast DNA. *Curr. Genet.* **29**: 572–581.

KELLOGG, E. A., AND C. S. CAMPBELL. 1987. Phylogenetic analyses of the Gramineae, pp. 310–322. *In* T. R. Soderstrom, K. W. Hilu, C. S. Campbell, and M. E. Barkworth [eds.], Grass systematics and evolution. Smithsonian Institution Press, Washington, D.C., USA.

MASON-GAMER, R. J. 2001. Origin of North American *Elymus* (Poaceae: Triticeae) allotetraploids based on granule-bound starch synthase gene sequences. *Syst. Bot.* **26**: 757–768.

———, C. F. WEIL, AND E. A. KELLOGG. 1998. Granule-bound starch synthase: structure, function, and phylogenetic utility. *Molec. Biol. Evol.* **15**: 1658–1673.

MATHEWS, S., R. E. SPANGLER, R. J. MASON-GAMER, AND E. A. KELLOGG. 2002. Phylogeny of Andropogoneae inferred from phytochrome B, GBSSI, and *ndh*F. *Int. J. Pl. Sci.* **163**: 441–450.

———, R. C. TSAI, AND E. A. KELLOGG. 2000. Phylogenetic structure in the grass family (Poaceae): evidence from the nuclear gene phytochrome B. *Amer. J. Bot.* **87**: 96–107.

MICHELANGELI, F. A., J. I DAVIS, AND D. W. STEVENSON. 2003. Phylogenetic relationships among Poaceae and related families as inferred from morphology, inversions in the plastid genome, and sequence data from the mitochondrial and plastid genomes. *Amer. J. Bot.* **90**: 93–106.

NIXON, K. C. 2002. WinClada vers. 1.00.08. Ithaca, New York, USA. Distributed by the author. Available at: *http://www.cladistics.org/education.html* (Jun 2005).

OGIHARA, Y., K. ISONO, T. KOJIMA, A. ENDO, M. HANAOKA, T. SHIINA, T. TERACHI, S. UTSUGI, M. MURATA, N. MORI, S. TAKUMI, K. IKEO, T. GOJOBORI, R. MURAI, K. MURAI, Y. MATSUOKA, Y. OHNISHI, H. TAJIRI, AND K. TSUNEWAKI. 2002. Structural features of a wheat plastome as revealed by complete sequencing of chloroplast DNA. *Molec. Genet. Genomics* **266**: 740–746.

OLMSTEAD, R. G., AND J. A. SWEERE. 1994. Combining data in phylogenetic systematics: an empirical approach using three molecular data sets in the Solanaceae. *Syst. Biol.* **43**: 467–481.

PALMER, J. D. 1983. Chloroplast DNA exists in two orientations. *Nature* **301**: 92–93.

PLUNKETT, G. M., AND S. R. DOWNIE. 2000. Expansion and contraction of the chloroplast inverted repeat in Apiaceae subfamily Apioideae. *Syst. Bot.* **25**: 648–667.

SAARELA, J. M., P. M. PETERSON, R. M. KEANE, J. CAYOUETTE, AND S. W. GRAHAM. 2007. Molecular phylogenetics of *Bromus* (Poaceae: Pooideae) based on chloroplast and nuclear DNA sequence data, pp. 450–467. *In* J. T. Columbus, E. A. Friar, J. M. Porter, L. M. Prince, and M. G. Simpson [eds.], Monocots: comparative biology and evolution—Poales. Rancho Santa Ana Botanic Garden, Claremont, California, USA.

SÁNCHEZ-KEN, J. G., AND L. G. CLARK. 2001. Gynerieae, a new Neotropical tribe of grasses (Poaceae). *Novon* **11**: 350–352.

SORENG, R. J., AND J. I. DAVIS. 1998. Phylogenetics and character evolution in the grass family (Poaceae): simultaneous analysis of morphological and chloroplast DNA restriction site character sets. *Bot. Rev. (Lancaster)* **64**: 1–85.

———, AND ———. 2000. Phylogenetic structure in Poaceae subfamily Pooideae as inferred from molecular and morphological characters: misclassification vs. reticulation, pp. 61–74. *In* S. W. L. Jacobs and J. E. Everett [eds.], Grasses: systematics and evolution. CSIRO Publishing, Collingwood, Victoria, Australia.

———, P. M. PETERSON, G. DAVIDSE, E. J. JUDZIEWICZ, F. O. ZULOAGA, T. S. FILGUEIRAS, AND O. MORRONE (editors). 2003. Catalogue of New World grasses (Poaceae): IV. Subfamily Pooideae. *Contr. U.S. Natl. Herb.* **48**: 1–730.

STEBBINS, G. L., AND B. CRAMPTON. 1961. A suggested revision of the grass genera of temperate North America. *Recent Advances in Botany* **1**: 133–145.

WALLACE, R. S., AND J. H. COTA. 1996. An intron loss in the chloroplast gene *rpo*C1 supports a monophyletic origin for the subfamily Cactoideae of the Cactaceae. *Curr. Genet.* **29**: 275–281.

WATSON, L., AND M. J. DALLWITZ. 1992 onwards. The families of flowering plants: descriptions, illustrations, identification, and information retrieval, Vers. 14 Dec 2000. *http://biodiversity.uno.edu/delta/angio* [Now at: *http://delta-intkey.com* Mar 2006.]

ZHANG, W. 2000. Phylogeny of the grass family (Poaceae) from *rpl*16 intron sequence data. *Molec. Phylogen. Evol.* **15**: 135–146.

ZULOAGA, F. O., O. MORRONE, G. DAVIDSE, T. S. FILGUEIRAS, P. M. PETERSON, R. J. SORENG, AND E. J. JUDZIEWICZ (editors). 2003. Catalogue of New World grasses (Poaceae): III. Subfamilies Panicoideae, Aristidoideae, Arundinoideae, and Danthonioideae. *Contr. U.S. Natl. Herb.* **46**: 1–662.