ELSEVIER

A 313 plastome phylogenomic analysis of Pooideae: Exploring relationships among the largest subfamily of grasses

Lauren M. Orton [a,*], Patricia Barberá [b], Matthew P. Nissenbaum [a], Paul M. Peterson [c], Alejandro Quintanar [d], Robert J. Soreng [c], Melvin R. Duvall [a,e]

[a] *Plant Molecular and Bioinformatics Center, Biological Sciences, Northern Illinois University, 1425 W. Lincoln Hwy, DeKalb, IL 60115-2861, USA*
[b] *Department of Africa and Madagascar, Missouri Botanical Garden, St. Louis, MO 63110, USA*
[c] *Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington DC 20013-7012, USA*
[d] *Herbario MA, Unidad de Herbarios, Real Jardín Botánico de Madrid CSIC, 28014 Madrid, Spain*
[e] *Institute for the Study of the Environment, Sustainability and Energy, Northern Illinois University, 1425 W. Lincoln Hwy, DeKalb, IL 60115-2861, USA*

ABSTRACT

In this study, we analyzed 313 plastid genomes (plastomes) of Poaceae with a focus on expanding our current knowledge of relationships among the subfamily Pooideae, which represented over half the dataset (164 representatives). In total, 47 plastomes were sequenced and assembled for this study. This is the largest study of its kind to include plastome-level data, to not only increase sampling at both the taxonomic and molecular levels with the aim of resolving complex and reticulate relationships, but also to analyze the effects of alignment gaps in large-scale analyses, as well as explore divergences in the subfamily with an expanded set of 14 accepted grass fossils for more accurate calibrations and dating. Incorporating broad systematic assessments of Pooideae taxa conducted by authors within the last five years, we produced a robust phylogenomic reconstruction for the subfamily, which included all but two supergeneric taxa (Calothecinae and Duthieeae).

We further explored how including alignment gaps in plastome analyses oftentimes can produce incorrect or misinterpretations of complex or reticulate relationships among taxa of Pooideae. This presented itself as consistently changing relationships at specific nodes for different stripping thresholds (percentage-based removal of gaps per alignment column). Our summary recommendation for large-scale genomic plastome datasets is to strip alignment columns of all gaps to increase pairwise identity and reduce errant signal from poly A/T bias. To do this we used the "mask alignment" tool in Geneious software. Finally, we determined an overall divergence age for Pooideae of roughly 84.8 Mya, which is in line with, but slightly older than most recent estimates.

1. Introduction

From the late Cretaceous origins of Poaceae Barnhart, to modern day agriculture, the grass family is well known for the impacts it has had in both conquering diverse landscapes and its economic significance as biofuel or food sources for wide varieties of animal species, including humans (Bouchenak-Khelladi et al., 2010; Orton et al., 2017, 2019). The Poaceae comprise 12 subfamilies: Anomochlooideae Pilg. ex Potztal, Aristidoideae Caro, Arundinoideae Kunth ex Beilschm., Bambusoideae Luerss., Chloridoideae Kunth ex Beilschm., Danthonioideae H.P.Linder & N.P.Barker, Micrairoideae Pilg., Oryzoideae Kunth ex Beilschm., Panicoideae A.Braun, Pharoideae L.G.Clark & Judz., Pooideae Benth., and Puelioideae L.G.Clark & al. Among these, two large clades are

recognized (Kellogg, 2015; Soreng et al., 2015, 2017; Saarela et al., 2017): the BOP (Bambusoideae, Oryzoideae, and Pooideae) clade, and the PACMAD (Panicoideae, Aristidoideae, Chloridoideae, Micrairoideae, Arundinoideae, Danthonioideae) clade. The BOP is comprised of C_3 photosynthetic cool-season grasses that have wide distributions across temperate regions in primarily open to closed habitats. The PACMAD is comprised of both C_3 and C_4 photosynthetic species with cosmopolitan distributions and greatly varying habitats (Cotton et al., 2015; Gallaher et al., 2019). While relationships of these clades have been extensively studied, many of these studies utilized single or multi-loci data. More recently, complete genomic datasets (in particular, complete plastomes) focused on the complex subfamilial relationships in the BOP and PACMAD clades have been increasingly prevalent in

phylogenomic research (Leseberg and Duvall, 2009; Cahoon et al., 2010; Morris and Duvall, 2010; Burke et al., 2012, 2016a,b; Hand et al., 2013; Jones et al., 2014; Cotton et al., 2015; Saarela et al., 2015, 2018; Wysocki et al., 2015; Attigala et al., 2016; Duvall et al., 2016; Orton et al., 2017, 2019; Gallaher et al., 2019). Previous studies have also incorporated morphological data to provide an additional avenue of exploration with regard to relationships among and within these clades (Kellogg, 2015; Saarela et al., 2015; Soreng et al., 2015, 2017).

These complicated phylogenomic histories are oftentimes resolved with the addition of data to fill absences in previously underrepresented groups. In the case of the Poaceae, a vast number of species have no complete plastid genome (plastome) available in data repositories. Poaceae are comprised of roughly 12,000 species and 780 genera (Kellogg, 2015; Christenhusz & Byng, 2016; Soreng et al., 2017). The representation of grass plastomes on NCBI's GenBank (https://www.ncbi.nlm.nih.gov) is roughly 1,450 of the nearly 12,000 known species.

The subfamily Pooideae is the largest of the grass subfamilies, a member of the BOP clade, and is a group with a number of economically significant crop, pasture, and lawn grasses. Pooideae arguably has the widest impact, due to its size, diversity, and systematic/phylogenetic complexity, among grass subfamilies (Kellogg, 2015). Even with approximately 4,000 species and more than 200 genera (Saarela et al., 2015; Soreng et al., 2015, 2017; Orton et al., 2019), Pooideae complete plastomes are represented on NCBI's GenBank by only about 300 species, many of which are multiple sequences of cultivars or varieties belonging to a single species. This underrepresentation across the grass family and its subfamilies has presented challenges for reconstructing plastome phylogenomic histories.

Additionally, fossil specimens used for divergence date estimation calibration points among Poaceae, and especially from Pooideae, are limited (Prasad et al., 2005; Stromberg and McInerney, 2011; Iles et al., 2015; Poinar et al., 2015). The unfortunate side-effect is that the ages returned in these analyses are often unavoidably skewed (Burke et al., 2016a,b; Minaya et al., 2017; Orton et al., 2017; Schubert et al., 2019).

There are notable conflicts between plastome and nuclear phylogenies of Pooideae (Quintanar et al., 2007; Schneider et al., 2009; Saarela et al., 2010, 2017; Hochbach et al., 2015). Plastomes, being maternally inherited and highly conserved, provide a more stable phylogeny as a result of their uniparental genetic history. While nuclear phylogenies have been sought due to their wealth of genetic information, oftentimes, nuclear genomes have intricacies that make accurate assemblies difficult. Most nuclear genomes are assembled into multitudes of scaffolds making interpretations and phylogenomic analyses difficult to produce. For this reason, nuclear analyses have historically utilized single (i.e. nrDNA) or a few loci, limiting the information otherwise available in a complete genomic dataset (Quintanar et al., 2007; Schneider et al., 2009; Saarela et al., 2010, 2017). Additionally, many nuclear loci are subject to positive selection (Zhao et al., 2013), and although this is not exclusive to nuclear loci (plastid loci are also positively selected), generally the use of predominately positively selected sites greatly impacts phylogenetic signal and produces interpretation biases (Piot et al., 2017; Saarela et al., 2018). In contrast, plastomes contain the complete genetic information available in an organellar genome, and analyses of Poaceae are able to accommodate not only coding regions, but also noncoding sequences that are presumably less subject to selective effects to produce robust phylogenetic histories (Saarela et al., 2018). It is also important to note, that in many instances there is evidence of reticulation and single copy nuclear genes provide necessary phylogenetic signal to resolve these conflicts that arise between nuclear and chloroplast phylogenies (Kellogg, 2015). One relevant example is the relationship between Poeae R.Br. group 1 and 2 chloroplast type species which is strongly marked in plastome phylogenies (Orton et al., 2019), but not seen in nuclear phylogenies (Quintanar et al., 2007; Saarela et al., 2010, 2017). At the same time, steady advancements in computing technology have accommodated the increase in large-scale genetic data. These larger and more complete data sets, encompassing complete genomes, are becoming more prevalent in research; and analytical tools are more widely available to cater to this wealth of large-scale, genome level data.

Here we present the largest analysis of grass plastomes to date, focusing on the Poaceae with a targeted interest in resolving the complex and often reticulate relationships of the Pooideae subfamily, as well as interpreting relationships across the BOP clade. The goal of our work is to explore how plastome level data in large-scale studies will provide added resolution to reticulate phylogenomic relationships in groups plagued by complex phylogenomic histories, determine how alignment gaps impact phylogenetic reconstruction, and in particular, how they impact topology. We also compare our results with previous divergence estimation analyses for Pooideae, in which we expand upon both the molecular sampling (by including both coding and noncoding regions from complete plastomes) and fossil calibrations (by including additional accepted grass fossils).

2. Materials and methods

2.1. Sampling and DNA extraction

A total of 47 plastomes were targeted for DNA sequencing for this study. These species were chosen to bolster underrepresented tribes/subtribes and to seek additional resolution of phylogenomic relationships. Silica dried leaf tissues were collected or received from collaborators for these species (Supplemental Table 1). DNA was manually extracted through homogenization of tissue in liquid nitrogen and application of the DNeasy Plant Mini Kit protocol (Qiagen, Valencia, CA, USA). Total genomic DNAs in the extracts were quantified using the Qubit assay (Invitrogen, ThermoFisher Scientific, Wilmington, DE, USA), and diluted to 2.5 ng/µl in 20 µl sterile water.

Illumina sequencing libraries were prepared using the Nextera Prep Kit (Illumina Inc., San Diego, CA, USA; Caruccio, 2011), and of the 47 species 26 were sequenced from single end libraries, 16 were sequenced from paired end libraries, and 5 were previously sequenced by collaborators using MiSeq for mate-pair libraries. The DNA Clean and Concentrator kit (Zymo Research, Irvine, CA, USA) was used to purify DNA. Standard manufacturer protocols for each respective sample preparation kit were used for library preparation. Samples were then sent to the Iowa State University Core DNA Facility, Ames, IA, USA, for sequencing on the HiSeq 2500.

2.2. Plastome assembly and annotation

Once sequence read libraries were received, plastomes were assembled using *de novo* methods based on Wysocki et al. (2014). A proprietary Python language (van Rossum, 1995) based pipeline was used to streamline workflow. Programs included in the pipeline were: DynamicTrim v2.1 of the SolexaQA software suite (Cox et al., 2010), which performed initial quality trimming on reads using default settings. CutAdapt removed any remaining Illumina adapter sequences (Martin, 2011). LengthSort v2.1 (Cox et al., 2010) removed any sequences shorter than 25 bp in length. CDHit-EST of the CDHit package (Fu et al., 2012) identified and removed redundant sequences (parameter settings: sequence identity threshold (–c 1)). Assembly of contigs was done using SPAdes v. 3.8.1 software suite (Bankevich et al., 2012) (kmer parameter settings: -k 19,25,31,37,43,49,55,61,67,73) for all species, except *Echinaria capitata,* which was assembled under an older protocol using Velvet *de novo* assembler (Zerbino & Birney, 2008) following the methods as outlined in Wysocki et al. (2014). Remaining contigs were scaffolded using the anchored conserved region extension method (Wysocki et al., 2014) which uses highly conserved regions in grass plastomes to orient and scaffold contigs. Any remaining gaps between contigs were manually resolved through *in silico* genome walking, a method that locates regions of overlap in the quality processed reads to build across gapped regions until a complete plastome is produced.

Assembly verification was done by mapping quality-processed reads to the assembled plastome via Geneious Pro v. 11.0.5 software (Biomatters Ltd, Auckland, NZ; Kearse et al., 2012), and any assembly errors were identified and manually resolved. Following the methods of Burke et al. (2012), inverted repeat boundaries were located using NCBI's BLASTn utility by performing a local alignment of the plastome against itself, locating regions of plus/minus transitions, and annotated using Geneious software. Remaining annotations were identified and completed by aligning the newly completed plastome to a closely related Pooideae species in Geneious software and transferring annotations to the newly completed plastome (Wysocki et al., 2014). Coding sequences were adjusted when necessary to preserve stop codons and correctly position coding sequence boundaries to preserve reading frames.

2.3. Phylogenomic analyses

Phylogenomic reconstruction was completed on a stripped alignment matrix of the 313 taxa that was 78,002 bp in length (see below and Methods 2.6 for additional descriptions of how gaps in the alignment matrix were stripped from each alignment column). The 313 taxa included in this analysis were sampled to proportionally represent grass subfamilies based on availability of species both in GenBank and those for which complete plastomes were sequenced here (Supplemental Table 2). The alignment matrix was created using the MAFFT plugin available in Geneious v11.0.5 software (Katoh and Standley, 2013) following removal of Inverted Repeat A (IRa), to avoid redundant representation of inverted repeat sites. The alignment was then stripped in Geneious v11.0.5 by using the "mask alignment" tool and sequentially removing an increasing threshold percentage of gaps per alignment column ("stripping threshold", Duvall et al., 2020).

The Saarela et al. (2018) study conducted genome partition analyses of grass plastomes on a dataset with 250 taxa across Poaceae. In that study, 14 partitions ([*rbcL, ndhF, matK*], plastome coding sequences, plastome non-coding sequences, and complete plastome data) with gapped sites only, positively selected sites only, neither gapped sites nor positively selected sites, or both gapped and positively selected sites were analyzed in matrices that ranged from 3,476 bp to 197,529 bp. The results indicated that best practices in grass phylogenomics include utilizing complete plastome data while excluding ambiguously aligned regions by stripping alignment matrices.

The GTR + I + G model was chosen based on best-fit results of a jModelTest (Posada, 2008; Darriba et al., 2012) using the Akaike information criterion (Akaike, 1974). A maximum likelihood (ML) analysis was conducted using the RAxML – HPC2 v8.1.11 program (Stamatakis, 2014) available through XSEDE on the CIPRES science gateway (Miller et al., 2010). The GTRGAMMA model was selected and a ML best tree was produced and bootstrap analysis conducted. A consensus tree was generated from 1,000 bootstrap replicates using the Consense tool from the Phylip v3.66 software suite (Felsenstein, 2005), available on CIPRES. Trees were visualized using TreeGraph2 (Stöver and Müller, 2010) and FigTree v1.4.2 (Rambaut, 2014).

2.4. Bayesian evolutionary analysis sampling trees (BEAST2) divergence date analysis

A divergence date analysis was conducted using BEAST v2.4.8 (Bouckaert et al., 2014) available on CIPRES. Parameters and priors were set using the complementary BEAUTI program provided with the BEAST2 software package. Fossil dates were used as calibration points with the upper bound set to 110 Ma, which was the estimated age of the spikelet fossil identified in Poinar et al. (2015), to constrain the node age to the oldest identified and dated grass fossil, and prevent an unconstrained upper bound skewing the ages substantially older. The lower bound was set to the fossil age. These methods of calibration, which use fossil dates only to set minimum ages on the stem nodes of clades, make the fewest assumptions about the largely unknown relationships

between the ages of fossils and the ages of the groups to which they belong (Christin et al., 2014). See Table 1 for priors and fossil calibrations related to specific nodes. An uncorrelated relaxed log normal clock model and uniform distribution across priors (Christin et al., 2014) were chosen along with the GTR site model, while also including a gamma category count of six, 0.743488 estimated shape parameter, and 0.500577 estimated proportion of invariant sites. The estimated shape parameter and proportion of invariant sites were determined from the ML analysis and values input as BEAST2 program parameters. This clock model and prior distribution were selected based on Christin et al. (2014) where it was elaborated that an uncorrelated relaxed lognormal clock reduces the effects of substantial variability in evolutionary rates while a uniform distribution allows for fossil ranges in calibrations by providing upper and lower boundaries for the fossil age. The Yule tree prior was utilized as it more readily assesses relationships between single representatives across different species (Beast2 Tree Priors and Dating: https://www.beast2.org/tree-priors-and-dating/).

Additional parameter settings included: four MCMC chains of 500,000,000 generations, and trees logged every 100,000 generations. Trees were summarized using TreeAnnotator (Rambaut, 2014) and burn-in was set at 25%. A chronogram was created using R packages: Strap (Bell and Lloyd, 2015) and IPS (Heibl et al., 2019). The R package RWTY (Warren et al., 2017) was used to assess convergence.

The resulting tree and its associated posterior probability (PP)

Table 1
Fossil calibration points and associated information.

Fig. 3 Label	Fossil	Calibration Point	Lower bound (age, Ma)	Citation
1	Spikelet in amber	Spikelet clade (all grasses minus Anomochlooideae Pilg. ex Potztal)	97	Poinar et al. (2015)
2	Oryzeae phytoliths	Oryzeae Dumort.	66	Prasad et al. (2011)
3	*Stipa florissanti* (Knowlt.) MacGinitie	Stipeae Dumort.	34	MacGinitie (1953)
4	*Paleoeriocoma hitchcockii* M.K. Elias	*Oryzopsis* Michx.	8	Thomasson (1985)
5	*Berriochloa* spp.	*Hesperostipa* (M.K. Elias) Barkworth	26	Thomasson (1985)
6	*Festuca* cf. *amethystina* L.	*Festuca* L.	16	Juchniewicz (1975)
7	*Lygeum* sp.	*Lygeum* Loefl.	16	Jiménez-Moreno, Fauquette, & Suc (2010)
8	*Nassella* spp.	*Nassella* (Trin.) E. Desv.	8	Thomasson (1980)
9	*Chusquea* GSSCs	*Chusquea* Kunth	35	Strömberg (2005)
10	*Leersia seifhennersdorfensis* H.Walther	*Leersia* Sw.	30	Walther (1974)
11	*Distichlis* sp.	*Distichlis* Raf.	14	Dugas and Retallack (1993)
12	*Leptaspis* cf. *zeylanica* Nees	*Leptaspis* R.Br.	12	Jacobs and Kabuye (1987)
13	*Dichanthelium* sp.	*Dichanthelium* (Hitchc. & Chase) Gould + *Thyridolepis* S.T.Blake	8	Thomasson (1978)
14	*Setaria* sp.	*Cenchrus* L. + *Setaria* P.Beauv.	7	Elias (1942)

T1. Fossil calibration points, fossil and source information for all calibrations used. Citations listed in separate reference list. GSSCs = grass silica short cells.

values, served as the Bayesian Inference (BI) tree for this study. The ML and BI trees are available as Supplemental Figs. 1 and 2, respectively. Support values were noted in both ML (MLBV support less than 100% shown) and BI tree includes all PP values.

2.5. Neighbor-Net analysis

A neighbor-net analysis was conducted using the Splits-Tree program (Huson, 1998). Additionally, a phi test (Bruen et al., 2006) of recombination was done through the Splits-Tree program and a delta score with Q-Residual score was also calculated for the 313 taxa matrix to explore the level of reticulation across the dataset. All default parameters were used.

2.6. Alignment gaps analysis

Analyses of the effect of alignment gaps on resolution and topology were conducted by creating stripped nucleotide matrices from the "mask alignment" tool available in Geneious v.11.0.5. The "mask alignment" tool removes columns from the alignment that contain at least a specified percent of gaps. This tool only allows integer values (i.e. 5%, but not 5.5%). Masking thresholds were created for the first 15% of gapped positions stripped (1-15%), as well as for 0% (no gaps remaining), 25%, 50%, 75% and for 100% (all gaps remaining in the matrix), following the methods of (Duvall et al., 2020). As with phylogeny reconstruction, IRa was removed to avoid redundancy in inverted repeat site representation. Relative percentages of AT richness, matrix mean pairwise identity, and total number of gaps per stripping threshold alignment were catalogued in 10,000 bp segments using Geneious v11.0.5 statistics, and noted. Data were visualized using Circa (http://omgenomics.com/circa). Maximum likelihood analyses were done identically to the previously described methods (section 2.3), using RAxML-HPC2 on XSEDE for each alignment (20 total alignment matrices with varying stripping thresholds applied). Relationships among Pooideae species were manually noted and any differences in topologies across stripping thresholds were described in detail. Graphical representation of substantial topological inconsistencies was completed in a manner similar to the tree displayed in Carlsen et al. (2018). Support values were also examined to determine any relationships with low support that coincided with topological inconsistencies across stripping thresholds. Further comparisons to the Neighbor-Net tree were conducted to explore evidence of reticulation, similar to methods used in Mitchell et al. (2017) to explore hybridization as a potential hypothesis for differing tree topologies. Trees were visualized using FigTree v1.4.2 (Rambaut, 2014). These 20 alignment matrices are available via Dryad (https://doi.org/10.5061/dryad. p8cz8w9p0).

3. Results

3.1. Plastome phylogenomic analysis

This study utilized plastome level data for 313 taxa with representatives comprising: 164 Pooideae, 40 Bambusoideae, 15 Oryzoideae, 44 Panicoideae, 20 Chloridoideae, six Danthonioideae, six Arundinoideae, five Micrairoideae, six Aristidoideae, one Puelioideae, four Pharoideae, and two Anomochlooideae. Of the 164 Pooideae plastomes, 47 were sequenced and assembled for this study; and of those, 42 are species without a previously published plastome (Supplemental Tables 1 and 2). Both a maximum likelihood (ML) best tree and bootstrap tree were produced. Topological inconsistencies between the ML best tree and the bootstrap tree are noted in Supplemental Fig. 1 by the designation "NR" (not resolved). Support values are noted only in instances where a relationship has less than 100% support. Taxon names in bold correspond to species sequenced for this study. A BI tree with PP values is included in Supplemental Fig. 2.

Pooideae is monophyletic, sister to the Bambusoideae, and in turn,

that pair is sister to Oryzoideae. Likewise, the PACMAD clade is monophyletic as are all early diverging grass subfamilies. Within Pooideae, all major supertribes are monophyletic. However, non-monophyletic relationships become evident at the tribe and subtribe levels.

3.2. Bayesian evolutionary analysis sampling trees (BEAST2) divergence date analysis

This analysis utilized fossil calibration points from 14 recognized grass fossils (Table 1). Priors were calibrated to specific fossil dates. Pooideae had six fossil calibration points while the remaining eight fossil calibrations were distributed across the following subfamilies/clades: Bambusoideae, Oryzoideae, Panicoideae, Chloridoideae, Pharoideae, and the spikelet clade. See Table 2 for mean clade ages and 95% HPD intervals of BOP, PACMAD, and Pooideae supertribes and tribes. The results of this analysis indicate an origin for Pooideae around 84.8 Ma (76.3–93.3 Ma; 95% HPD interval) and all major supertribes having diverged by 30 Ma, except Triticodae which diverged at roughly 25 Ma (1.8–49.3 Ma; 95% HPD interval).

3.3. Neighbor-Net analysis

The neighbor-net analysis produced a separation network from the 313 plastome alignment. Contradictory patterns emerge among species within the early diverging tribes: Brachyelytreae Ohwi, Phaenospermateae Renvoize & Clayton, and Diarrheneae C.S.Campb. Similarly, these tribes also returned low support for their positions in the separation network bootstrap analysis (Brachyelytreae: 26.3; Phaenospermateae: 27.2; Diarrheneae: 27.1) (Supplemental Fig. 3). The Phi test did not find statistically significant evidence of recombination (p = 1.0). The delta score for the dataset was: 0.07428, and Q-residual score: 4.524E-4. These values indicate that there is minimal reticulation, and the data behave in a "tree-like" fashion.

3.4. Alignment gaps analysis

Twenty matrices and their resulting ML tree topologies were exhaustively catalogued through manual comparisons across all 20 matrices and trees to determine any topological inconsistencies based on different stripping thresholds. Matrices ranged in size from 78,002 sites (no gaps included) to 211,879 sites (all gaps included). Percent AT richness among the matrices consistently increased from 59.9% (no

Table 2

Results of BEAST divergence estimation analysis.

Crown Node ♠ = Clade ♦ = Supertribe ♠ = Tribe	95% HPD lower, Ma	95% HPD upper, Ma	Mean age, Ma
PACMAD♠	32.2	95.1	63.7
BOP♠	89.8	100.1	95.0
Brachyelytreae Ohwi♠	76.3	93.2	84.8
Nardodae Soreng♦	16.0	68.8	42.4
Phaenospermateae Renvoize & Clayton ♠	38.3	81.7	60.0
Melicodae Soreng♦	3.5	65.8	34.7
Stipodae L.Liu ♦	34.0	53.0	43.5
Diarrheneae C.S.Campb. ♠	7.0	66.4	36.7
Brachypodieae Harz♠	0.3	44.1	22.2
Triticodae T.D.Macfarl. & L. Watson ♦	1.8	49.3	25.6
Poodae L.Liu♦	47.8	62.4	55.1

T2. Results of BEAST divergence estimation analysis. Pooideae taxonomic classifications are predominantly sensu Soreng et al. (2017), refer to the Discussion for in-depth analyses, classifications, and taxonomic authorities of Pooideae taxa included in this study. 95% HPD lower and upper ranges are listed as well as mean age. All times in Ma (mega-annum).

gaps) to 62.5% (all gaps). Mean pairwise identity across the matrices consistently decreased from 97.0% (no gaps) to 88.7% (all gaps) (Fig. 2; Table 3). Additionally, visualization of these data in Fig. 2 clearly show the level of variability across alignments that contain gaps versus an alignment that contains no gaps. Coupled with tree topology comparisons, it becomes clear that interpretations of relationships can be clouded, particularly among taxa with potential hybridization or complex evolutionary histories, when alignment gaps are present in varying amounts across an analysis. All representative supertribes, tribes, super subtribes, and subtribes were exhaustively analyzed and differing topologies were noted (Fig. 1, Supplemental Table 3).

4. Discussion

4.1. PACMAD

4.1.1. Phylogeny overview

This study reiterates the complex phylogenomic interpretation of the Aristidoideae/Panicoideae relationship and its relation to the remaining CMAD subfamilies as elaborated in Duvall et al. (2020). All individual subfamilies are recognized as clades. However, sampling of the PAC-MAD clade here may have influenced the sister group topology of Aristidoideae/Panicoideae and the remaining subfamilies (CMAD). Here, the best tree returned the "Panicoideae sister" topology, while the bootstrap tree returned the "Aristidoideae sister" topology (Burke et al., 2016b; Saarela et al., 2018).

4.2. BOP

4.2.1. Phylogeny overview

The BOP clade, comprised of Bambusoideae, Oryzoideae, and Pooideae, is also returned as monophyletic. In turn, each of the subfamilies is monophyletic with (Oryzoideae, (Bambusoideae, Pooideae)). This larger clade includes a number of economically and ecologically significant species (Wysocki et al., 2015; Saarela et al., 2018; Orton et al., 2019).

4.3. Pooideae

4.3.1. Overview

The subfamily Pooideae is comprised of 15 tribes and 26 subtribes (Soreng et al., 2017), of which all but one subtribe and one tribe are represented here: Calothecinae Soreng and Duthieeae Röser & Jul. Schneider. Kellogg (2015) recognizes 10 tribes and 15 subtribes. Here we follow the more recent classifications by Soreng et al. (2017) and Tkach et al. (2020); the latter work added the subtribes: Hypseochloinae Röser & Tkach, Antinoriinae Röser & Tkach, Helictochloinae Röser & Tkach, Avenulinae Röser & Tkach, Brizochloinae Röser & Tkach. While *Helictochloa* is included as a representative of Helictochloinae, the remaining subtribe additions from Tkach et al. (2020) comprise single genera with one or two species and are not included in this study. However, we also accept many supergeneric taxa listed in Kellogg (2015). It should be noted, that the Calothecinae Soreng has only one recognized genus (*Chascolytrum* Desv.), which accounts for nine synonymous genera (Nogueira da Silva et al., 2020), and Duthieeae (despite having eight recognized genera), is a relatively new tribe in Pooideae, initially proposed by Pilger (1954) as a subtribe and more recently elevated to a tribe in 2011 (Schneider et al., 2011). Therefore, collection of specimens and complete plastome sequencing has lagged behind longer-established Pooideae for both Calothecinae and Duthieeae.
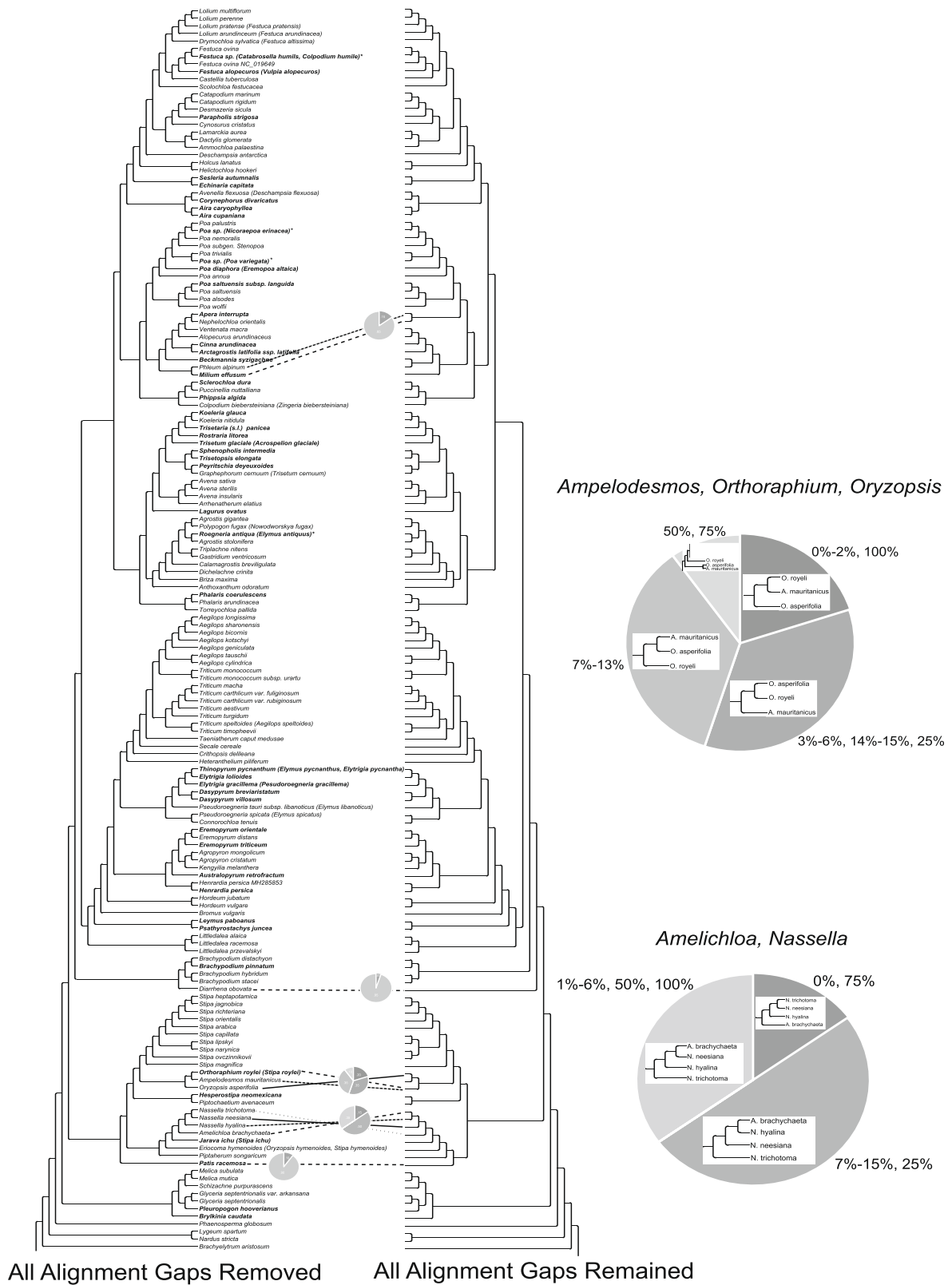
There are a number of well documented complex and reticulate relationships across the tribes of Pooideae, which have been repeatedly revised (Kellogg, 2015; Saarela et al., 2015; Soreng et al., 2015, 2017; Orton et al., 2019) as phylogenetic data of more species becomes readily available. Here we have increased sampling among Pooideae by 47 plastomes, and explored the effects of alignment gaps on topology in large-scale data of 313 taxa. By utilizing plastome data, expanding sampling, and exploring effects of alignment gaps, we have better clarified or confirmed previously unresolved or poorly supported relationships across Pooideae.

4.4. Neighbor-Net

The SplitsTree network, created through a neighbor-net analysis, depicts the relationships between taxa and determines conflicting patterns that emerge. As this study is one of the largest plastome separation networks, it is no surprise that there are a number of conflicts seen among taxa. The results of the delta and Q-residuals indicate the data cluster in a tree-like fashion and there is a high level of resolution among the included taxa, and no detectable reticulation despite evidence to suggest otherwise from previous studies (Mason-Gamer & Kellogg, 1996; Petersen et al., 2006; Mason-Gamer et al., 2010a,b; Fan et al., 2013). However, we have noted specific instances of known or suspected reticulate relationships or hybridizing complexes where there are identifiable points of conflict among taxa or clades (see below for descriptions, and Supplemental Fig. 3) despite the analysis returning no substantial evidence of reticulation per the neighbor-net analysis. These conflicts overwhelmingly appear among taxa with previously noted reticulate relationships or known hybridizing complexes, and were isolated to more narrow taxonomic classifications (tribe/subtribe levels such as Poeae, Diarrheneae, Triticeae, and within Stipeae/Ampelodesmeae). These results may also be a product of the uniparentally inherited, more stable plastome data as opposed to highly variable nuclear loci that often portray clear evidence of reticulation.

4.5. Alignment gaps

The twenty matrices and associated ML trees were extensively catalogued. This analysis expanded upon the levels of stripping thresholds seen in Duvall et al. (2020), particularly in the 1-15% stripping thresholds. As noted by Duvall et al. (2020), there are significant topology shifts within the top 11% of stripping thresholds, and the general trend in AT richness increases as more gapped positions are included in

Table 3
Descriptive information for stripping thresholds.

Stripping threshold (%)	Matrix length (bp)	AT (%)	Mean Pairwise Identity (%)
0 (no gaps)	78,002	59.9	97.0
1	89,911	60.7	96.5
2	94,194	61.0	96.2
3	97,388	61.2	95.9
4	99,720	61.4	95.6
5	101,398	61.5	95.4
6	102,645	61.6	95.3
7	103,868	61.7	95.1
8	104,874	61.8	94.9
9	105,897	61.8	94.7
10	106,822	61.9	94.6
11	107,370	61.9	94.5
12	107,776	61.9	94.4
13	108,247	62.0	94.3
14	108,632	62.0	94.2
15	109,003	62.0	94.1
25	111,634	62.1	93.3
50	115,178	62.3	92.0
75	119,086	62.4	90.5
100 (all gaps included)	211,879	62.5	88.7

T3. Descriptive statistics for each stripping threshold. Matrix length (in base pairs), percent AT richness, and mean pairwise identities are given.

Fig. 1. Mirrored tree depicting topological inconsistencies between alignments with all gaps removed (left side) and all gaps retained (right side). Positional lines are drawn across trees to show differences in the positions of taxa. Pie charts are situated on the positional lines to indicate percentage representation of topologies. Charts with only two topologies indicated are situated closer to the majority represented topology (i.e. if the majority represented topology was the tree derived from the alignment with all gaps remaining, the chart is placed closest to that tree). Pie charts with more than two topological inconsistencies represented are enlarged to the right of the mirror tree to indicate topology representation and stripping threshold (percentages labeled outside of each pie chart slice). Misidentified specimens are marked with an (*). Additional information on species misidentification (seed accessions, etc.) can be found in Supplemental Tables 1 and 2.

the analysis. Here, we have explored both AT richness and mean pairwise identity, as well as topological inconsistencies across stripping thresholds. These results are consistent with Duvall et al. (2020) with regard to AT richness, as the overall trend shows decreasing AT richness as the number of alignment gaps included in the analysis decreases. There are two stripping thresholds that have 61.8% AT richness (8–9% stripping thresholds) and three that have 61.9% AT richness (10–12% stripping thresholds) and three with 62% AT richness (13–15%).

One hypothesis for the overall decrease seen in AT richness across the dataset (from all gaps remaining to no gaps remaining) may be due to poly-A and/or poly-T sequences in the matrices. Duvall et al. (2020) noted that a relatively substantial portion of the gapped positions removed per stripping threshold corresponded to these poly-A and T sequences, and as the stripping thresholds increased in the percentage of gapped positions removed, the more of these ambiguous poly-A and T regions were removed. Likewise, Illumina kits have a documented bias in AT rich regions (Burke et al., 2016a,b).

Mean pairwise identity across the dataset behaves the opposite of AT richness. The highest mean pairwise identity across stripping thresholds occurs when no gaps remain in an alignment, and the lowest mean pairwise identity occurs when all gaps remain in an alignment. As gaps are increasingly removed from the alignment matrices, fewer ambiguously aligned regions remain and therefore, mean pairwise identity increases.

Total number of gaps per stripping threshold alignment were catalogued in 10,000 bp segments and noted. Gaps fluctuated greatly across the 10,000 bp segments in each stripping threshold alignment. For example, in the 4% stripping threshold alignment numbers of gaps per

segment range from 10,899 between bases 1–10,000; 13,376 gaps between segment bases 10,001–20,000; to 7,452 gaps between segment bases 20,001–30,000. The fluctuations in gaps per segment are most noticeable when all gaps remain in the stripping threshold alignment. Here we see numbers of gaps ranging in the hundreds of thousands to millions of gaps among the segments (Fig. 2). This variability has shown to have a substantial effect on topological inconsistencies as many programs used for nucleotide alignment are unable to unambiguously interpret regions with complex distributions of gapped positions, which may then interfere with phylogenetic inference.

The topologies of each stripping threshold were extensively compared and catalogued here by tribe or supertribe level (Fig. 1, Supplemental Table 3), and relationships summarized by taxonomic grouping (see below for in depth text descriptions of specific relationships). In the case of Poeae, relationships were also compared at the subtribe level. Relationships were also compared and contrasted between the assessments of Kellogg (2015), Soreng et al. (2017), and Tkach et al. (2020). While this general taxonomic overview falls in line with the more recent classification done by Soreng et al. (2017), we do recognize minor differences between assessments.

4.5.1. Brachyelytreae Ohwi

Represented in its entirety by one genus and three species (Soreng et al., 2017), and in this study, one species represents the tribe (*Brachyelytrum aristosum* (Michx.) P. Beauv. ex Trel.). The Brachyelytreae has remained sister to all other Pooideae in all plastid phylogenetic analyses of the subfamily (Kellogg, 2015; Soreng et al., 2015, 2017; Saarela et al., 2015, 2018). As the earliest diverging of Pooideae, its
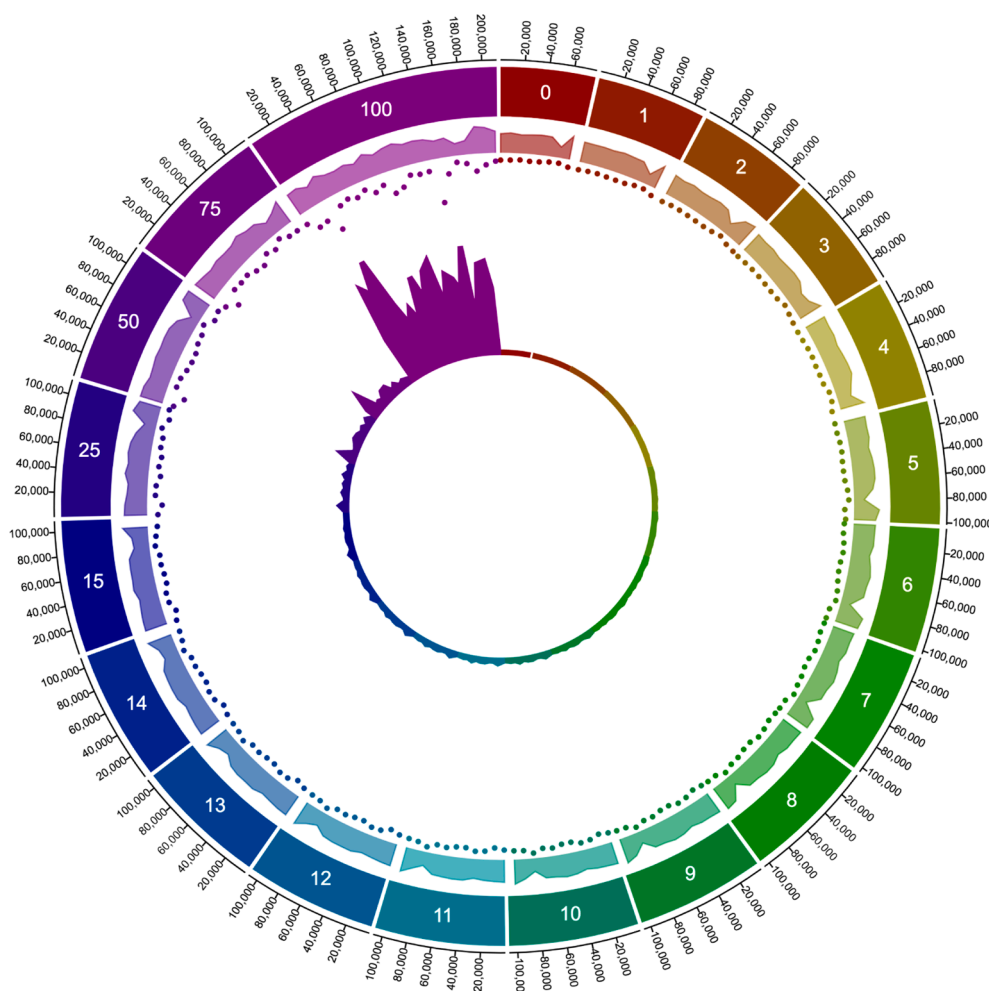
Fig. 2. Circa plot of data pertaining to alignment gap analysis. Outermost ring: legend, length of alignment in bp. Second ring (moving toward the center): bar length is proportional to the length of the alignment; the number label = stripping threshold. Third ring: histogram of % AT richness across the alignment (note this scale runs from zero to 100 percent, and is reflected by the lowest percentages closest to the center of the figure, and the highest percentages closest to the second ring). Fourth ring: dot plot of mean pairwise identity (%). Innermost ring: total number of gaps per 10,000 bp segment. Note the extreme variation in the 100% (all gaps included) stripping threshold.

position is maximally supported at all stripping thresholds.

4.5.2. Nardeae W.D.J.Koch and Lygeeae J.Presl (Nardodae Soreng)

Both *Nardus stricta* L. and *Lygeum spartum* L. belong to the supertribe Nardodae, and are maximally supported as sister taxa in our phylogenomic analyses as well as Soreng et al. (2017) and Saarela et al. (2018). In Kellogg (2015), Lygeeae is combined with Nardeae.

4.5.3. Phaenospermateae Renvoize & Clayton

Along with Brachyelytreae and Nardodae, Phaenospermateae is another of the early diverging tribes of Pooideae. The only representative of the Phaenospermateae (*Phaenosperma globosum* Munro ex Benth.) is present in this study and its position sister to Nardodae is maximally supported across all stripping thresholds. Previous studies (Saarela et al., 2018; Schubert et al., 2019) have also returned this position. Both Kellogg (2015) and Soreng et al. (2015) included Duthieeae as a synonym in the Phaenospermateae which then contained eight genera and 14 spp. in the previous circumscription. However, the revised Soreng et al. (2017) study separated Duthieeae from the Phaenospermateae. Additionally, there is evidence of hybridization between *Phaenosperma* and two or more tribes (Blaner et al., 2014; Hochbach et al., 2015).

4.5.4. Meliceae Link ex Endl. and Brylkinieae tateoka (Melicodae Soreng)

In Melicodae, *Brylkinia* F.Schmidt (Brylkinieae) is sister to Meliceae. All relationships in Melicodae are maximally supported (100% MLBV) across all gap analysis alignment stripping thresholds. Brylkinieae was placed in Meliceae in Kellogg (2015), considered as a possible subtribe in Meliceae in the previous Soreng et al. (2015) classification that was later updated in the Soreng et al. (2017) assessment, which places Brylkinieae as sister to Meliceae.

4.5.5. Stipeae Dumort. and Ampelodesmeae Tutin (Stipodae L.Liu)

The Stipeae, along with the Ampelodesmeae, form the supertribe Stipodae. Interestingly, the monophyly of Stipeae is interrupted by *Ampelodesmos mauritanicus* (Poir.) T.Durand & Schinz in our analyses and Romaschenko et al. (2012), and although the Ampelodesmeae is a member of Stipodae, the position of *A. mauritanicus* is placed within the core group of Stipeae, but poorly supported regardless of stripping threshold. This position is in the clade formed by *Oryzopsis asperifolia* Michx. and *Stipa roylei* (Nees) Duthie. In the alignment gap analysis, this position regularly rotated between (*A. mauritanicus*, (*Oryzopsis*, *S. roylei*)) and (*S. roylei*, (*A. mauritanicus*, *Oryzopsis*)) or (*Oryzopsis*, (*A. mauritanicus*, *S. roylei*)). Suspected hybridization between Ampelodesmeae and Stipeae was first described by Romaschenko et al. (2014) and is noted in Soreng et al. (2015, 2017), while also being listed in the "Worldwide Classification of Poaceae (Gramineae)" at www.tropicos.org, albeit representatives of Duthieeae were included in their analysis and are absent here. Additional hybridization of *Ampelodesmos* between other tribes (Duthieeae or Phaenospermateae) is evidenced in shared nuclear genes (Romaschenko et al., 2014). However, Kellogg (2015) includes *Ampelodesmos* in the Stipeae, which is consistent with our results. Yet, morphologically, the *Ampelodesmos* spikelet differs substantially from Stipeae and cannot be reconciled within the Stipeae as a result, and therefore, maintaining Ampelodesmeae as a separate tribe, but also a member of Stipodae respects the conflict between morphological and molecular results. Alternative hypotheses of this relationship include the possibility that these tribes diverged rather recently, or that mutation rates have decelerated. In all, it appears that these tribal representatives are quite closely related.

There also seems to be no particular identifying pattern to the placement of *A. mauritanicus* in Stipeae, as its position as sister to either *S. roylei* or *Oryzopsis* continued to alternate throughout the gap analysis, and there was no definite trend in relation to stripping thresholds, and topological alternation occurred between thresholds 25%, 50%, and 75%, 100%.

Several taxa within Stipeae in our tree (Fig. 1) are non-monophyletic, while others are monophyletic (this is well noted in several traditional circumscriptions) some of which can be attributed to nomenclatural changes, and others which have been revised to resolve the non-monophyletic relationships (Romaschenko et al., 2012, 2014). In general, the genus *Stipa* L. is monophyletic (*sensu stricto*) but also non-monophyletic (*sensu lato*) with consideration of those species that have been renamed (Columbus & Smith, 2010; Baldwin et al., 2012; Peterson et al., 2019): i.e. synonymous names, *Eriocoma hymenoides* (Roem. & Schult.) Rydb. = *Oryzopsis hymenoides* (Roem. & Schult.) Ricker ex Piper = *Stipa hymenoides* (Roem. & Schult.) Ricker ex Piper; *Jarava ichu* Ruiz & Pav. = *Stipa ichu* (Ruiz & Pav.) Kunth; *Orthoraphium roylei* Nees = *Stipa roylei* (Nees) Duthie; (https://www.ipni.org). Other members of Stipeae are recognized as monophyletic (i.e. 'Nassella' clade', core group of *Stipa*) when all gaps are removed. However, there is a lack of resolution across analyses (noted as "NR"—not resolved in Supplemental Fig. 1) for three nodes within the core *Stipa* representatives. Stipeae is also split into two smaller clades, to which *Patis racemosa* (Sm.) Romasch., P.M. Peterson & Soreng is sister. Romaschenko et al. (2012, 2014), note that *Patis* Ohwi, a genus that was previously "resurrected," is not a particularly close relative of the remaining Stipeae; Kellogg (2015) also notes the unique morphological trait of short rhizomes present in *Patis*.

Kellogg (2015) acknowledges the complexities between *Nassella* (Trin.) E.Desv. and *Amelichloa* Arriaga & Barkworth and that there is an argument for combining both genera, Soreng et al (2015, p. 122, 2017, p. 284) noted *Amelichloa* is "{nested within *Nassella*, but an intergeneric hybrid origin has not been ruled out}". This description is echoed in topological position shifts between *Amelichloa*, *Nassella hyalina* (Nees) Barkworth, and *Nassella neesiana* (Trin. & Rupr.) Barkworth across stripping thresholds. Here, *Amelichloa* is placed sister to the *Nassella* clade when all gaps are removed from the alignment, in agreement with Romaschenko et al. (2012).

4.5.6. Diarrheneae Tateoka ex C.S.Campb.

Only one species represents the Diarrheneae in this analysis: *Diarrhena obovata* (Gleason) Brandenburg. There are five total species and two genera (two *Diarrhena* and three *Neomolinia* Honda) in Diarrheneae. Here, Diarrheneae is sister to Brachypodieae + (Triticodae + Poodae) and is maximally supported (100% Maximum Likelihood Bootstrap Value: MLBV) at this position when all gaps remain in the alignment. However, as gaps are removed, the support decreases from 88.6% MLBV (1% stripping threshold) to 54% MLBV (0% stripping threshold/all gaps removed). Additionally, topology across the stripping thresholds is inconsistent. Among stripping thresholds ranging from 1 − 100% (all gaps present) stripping thresholds, the relationship of Diarrheneae and Brachypodieae is described as above. However, when all gaps are removed, Diarrheneae is sister to Brachypodieae only (Fig. 1; Supplemental Table 3).

4.5.7. Brachypodieae Harz.

The Brachypodieae forms a clade of the four representative species of *Brachypodium* P.Beauv. All relationships are supported at 100% (MLBV) across all gap stripping threshold analyses.

4.5.8. Littledaleeae Soreng & J.I.Davis, Bromeae Dumort., and Triticeae Dumort. (Triticodae T.D.Macfarl. & L.Watson)

The Triticodae supertribe forms a clade comprised of Littledaleeae, Bromeae, and Triticeae. However, Triticeae is paraphyletic with Bromeae, while Littledaleeae remains monophyletic as sister to the Triticodae. Within Triticeae, evidence of the *Triticum* L.-*Aegilops* L. hybridizing complex is minimal despite its prevalence among studies of this tribe which have noted reticulate relationships between both genera (Mason-Gamer & Kellogg, 1996; Petersen et al., 2006; Kawahara, 2009); it is likely that increased sampling would expose more evidence of hybridization/reticulation. Here, we see one clear instance of non-monophyly within Triticeae (*Triticum* species): *T. monococcum* subsp.

urartu (Thumanjan ex Gandilyan) Á.Löve & D.Löve and *T. monococcum* L. were found to be sister to the core *Aegilops* clade (Fig. 1).

Among the remaining genera included in the Triticeae, *Elymus* L. is split between Triticeae (*Elymus libanoticus* (Hack.) Melderis; a synonym of: *Pseudoroegneria tauri* subsp. *libanotica* (Hack.) Á.Löve) and Agrostidinae (*Elymus antiquus* (Nevski) Tzvelev; synonym: *Roegneria antiqua* (Nevski) B. S. Sun), but here we have genetic evidence that *E. antiquus* is likely a misidentified specimen. A comparison of pairwise alignments returned 99.9% identity between *E. antiquus* and *Agrostis stolonifera* L.*,* only 92.8% between *E. libanoticus* and *A. stolonifera,* and only 93.0% between *E. antiquus* and *E. libanoticus.* Our "*E. antiquus*" plastome has greater nucleotide identity with Agrostidinae than with the Triticeae, suggesting that the two accessions identified as *Elymus* in our study are not truly congeneric representatives of *Elymus*. *E. antiquus* is almost certainly a misidentified or contaminated specimen (See GRIN for seed accession used: (https://npgsweb.ars-grin.gov/gringlobal/accession detail.aspx?id=1066862).

Non-monophyletic relationships are also found among representatives of *Pseudoroegneria* (Nevski) Á.Löve*,* where *P. spicata* is synonymous of *Elymus spicatus* (Pursh) Gould. which is, in turn, synonymous of *Agropyron spicatum* (Pursh) Scribn. & J.G. Sm. However, *P. gracillima* (Nevski) Á.Löve is an unresolved species name (https://www.theplan tlist.org) and interestingly, is nested as (*P. gracillima*, (*Thinopyrum pycnanthun* [Godr.] Barkworth*, Elytrigia lolioides*)) while (*P. spicata, Connorochloa tenuis* [Buchanan] Barkworth, S.W.L.Jacobs & H.Q.Zhang) and also *Elymus libanoticus* are sister to this clade. Although these genera are non-monophyletic, this topology is not altogether unexpected based on nomenclatural changes. Many genera in Triticeae have synonyms in *Elymus* (see example above for *Pseudoroegneria*). The position of *P. gracillima* contributes to the non-monophyly of the genus. As *P. gracillima* is an accepted name we must consider alternate hypotheses for its position. These hypotheses include, 1) USDA misidentification of seed stock, and more likely 2) reclassification of *Pseudoroegneria* member species may have contributed to these unexpected relationships. Furthermore, Mason-Gamer et al. (2010a,b) identified evidence of hybridization events between *Pseudoroegneria, Hordeum* L. and *Elymus*, and even suggested that *Pseudoroegneria* and *Hordeum* are potential progenitors of the polyploids present among *Elymus* species. Multiple studies also explore reticulation among *Elymus* (Mason-Gamer et al., 2010a,b; Fan et al., 2013).

4.5.9. Poodae L.Liu, tribe Poeae R.Br.

Many of the relationships described in Orton et al. (2019) are confirmed in this study. There are 48 additional Poeae species included in this study expanding on the sampling in Orton et al. (2019), which focused on the membership of Group 1 and 2 Poeae species. This split between Groups 1 and 2 first noted in Soreng et al. (1990), is also seen in this analysis. Of the additional 48 species, 27 are newly published (Supplemental Table 1). There are 8 Poeae subtribes classified in the PPAM clade by Soreng et al. (2017): Coleanthinae Rouy, Poinae Dumort., Miliinae Dumort., Phleinae Dumort., Beckmanniinae Nevski, Cinninae Caruel, Alopecurinae Dumort., and Ventenatinae Holub ex L.J. Gillespie, Cabi & Soreng. Tkach et al. (2020) introduced newer subtribes to the PPAM clade: Brizochloinae Röser & Tkach and Avenulinae Röser & Tkach. Furthermore, Kellogg (2015) classified a number of these subtribes, as well as others that are recognized in Soreng et al. (2017), within closely related subtribes. These differences between the Kellogg (2015) and Soreng et al. (2017) classifications are noted. Recent work by Tkach et al. (2020) is discussed as well.

4.5.10. Brizinae Tzvelev

Brizinae (sensu Soreng et al., 2017) is maximally supported as sister to Echinopogoninae Soreng when all gaps remain in the alignment. However, when no gaps remain in the alignment support falls to 96% (MLBV) at the node uniting Brizinae and Echinopogoninae + Agrostidinae. Including additional species to represent both Brizinae and

Echinopogoninae in future research may clarify and better support these relationships. Brizinae is included in Agrostidinae in the Kellogg (2015) classification, and Echinopogoninae is not recognized in either Kellogg (2015) or Soreng et al. (2015), and was first introduced in Soreng et al. (2017). In the latter classification, Agrostidinae, Echinopogoninae and Brizinae (and Calothecinae, not included in the present study) are included in supersubtribe Agrostidodinae Soreng.

4.5.11. Agrostidinae Fr.

As elaborated in Orton et al. (2019), there are non-monophyletic relationships among representatives of Agrostidinae. In the 2019 analysis, Orton et al. found the monophyly of *Agrostis* L. species with the single exception that *Polypogon fugax* Nees ex Steud. (synonym: *Nowodworskya fugax* (Nees ex Steud.) Nevski) was included in this clade. Here, with expanded taxa included in this analysis, we have replicated this likely reticulate relationship between *Polypogon* Desf. and *Agrostis*. However, additional taxonomic sampling and inclusion of nuclear data will likely provide more definite assessments. In the Kellogg (2015) classification, the tribe Agrostidinae includes both subtribes Brizinae and Calothecinae (Calothecinae was not included in Kellogg's 2015 assessment, but the associated genera were included in Agrostidinae), which are recognized in Soreng et al. (2015, 2017) as independent subtribe designations and also members of the Agrostidodinae supersubtribe (see Brizinae above).

4.5.12. Aveninae J. Presl.

Recently, Barberá et al. (2017, 2019, 2020) transferred some groups of *Trisetum* to the genus *Koeleria* based on morphological and molecular evidence, and proposed a new genus (*Sibirotrisetum* Barberá, Soreng, Romasch., A. Quintanar & P.M. Peterson) while resurrecting another (*Acrospelion* Besser) to further reflect the relationships in Aveninae. Here we transfer *Trisetum glaciale* to *Acrospelion* based on the results of Barberá et al. (2017, 2018, 2020).

Acrospelion glaciale (Bory) Barberá, Soreng & Quintanar, **comb. nov.** *Avena glacialis* Bory, Ann. Gén. Sci. Phys. 3: 6 (1820) [basionym]. Type: Spain. Granada, Sierra Nevada, Aug. 1837, *P. E. Boissier 186* (neo-: G-00176302 [designated by Barberá et al., 2018, p. 381]; isoneotypes: B, BM, G, GH, GOET, JE, K, L, LE, M, MA, MO, MPU, MW, NY, O, P, RO, US, W).

4.5.13. Miliinae Dumort. & Phleinae Dumort.

Between our stripping thresholds from zero gaps remaining in the alignment to 25%, (*Milium effusum* L. + *Phleum alpinum* L.) is sister to the *Poa* clade. However, in our thresholds of 50%, 75%, and all gaps removed, the representatives of Miliinae and Phleinae are instead sister to Beckmanniinae Nevski. This differs from Soreng et al. (2017), in which Phleinae is sister to a clade comprised of Beckmanniinae plus three other subtribes (Cinninae, Alopecurinae, and Ventenatinae). Miliinae is sister to a clade comprised of all of the latter subtribes plus Poinae Dumort. Miliinae and Phleinae are unresolved in the multi-locus plastid phylogeny of Tkach et al. (2020).

4.5.14. Beckmanniinae Nevski

The lack of resolution between the sister relationship of Beckmanniinae and (*Arctagrostis* Griseb. + Cinninae) + (Alopecurinae + Ventenatinae) is potentially due to the inclusion of *Arctagrostis* (*incertae sedis* in the Poodinae). Soreng et al. (2017) classifies Beckmanniinae as a member of the Poodinae supersubtribe, while Kellogg (2015) places *Beckmannia* Host as a member of subtribe Poinae.

4.5.15. Airinae Fr.

Nomenclatural changes have clarified some relationships in the Airinae. The renaming of *Deschampsia flexuosa* (L.) Trin. to *Avenella flexuosa* (L.) Drejer has since resolved a suspected paraphyletic relationship among Aristaveninae F.Albers & Butzin and Airinae among our sampled taxa. Both Holcinae Dumort. and Aristaveninae are placed in

Airinae by Kellogg (2015). Our phylogenomic results are somewhat in disagreement with this placement given the more recent circumscriptions included in Tkach et al. (2020).

4.5.16. Aristaveninae F.Albers & Butzin

The Aristaveninae is represented in this study by a single species: *Deschampsia antarctica* E.Desv., which is supported as sister to *Scolochloa festucacea* (Willd.) Link L. + Loliodinae Soreng at 99% MLBV. The Orton et al. (2019) study also determined *D. antarctica* was sister to *S. festucacea* + Loliodinae.

4.5.17. Holcinae Dumort. & Helictochloinae Röser & Tkach

Holcus lanatus L. (Holcinae) is sister to *Helictochloa hookeri* (Scribn.) Romero Zarco (Helictochloinae) (86% MLBV), but due to a lack of additional representatives from this sub-tribe, an assessment of monophyly is limited. However, in Orton et al. (2019) this same topology was recovered in an analysis of Group 1 and Group 2 Poeae species. Additionally, Kellogg (2015) placed Holcinae within Airinae and Soreng et al. (2017) placed *Helictochloa* in Airinae. *Helictochloa hookeri* is included in its own subtribe (with *Molineriella* Rouy) as noted in Tkach et al. (2020), but no sister relationship with Airinae or Holcinae was detected.

4.5.18. Scolochloinae Tzvelev

Scolochloa festucacea (Willd.) Link is sister to the ten other representatives of the Loliinae subtribe, and nested within the supersubtribe Loliodinae at maximal support. This position, which differs from Soreng et al. (2017), was newly discovered in the Orton et al. (2019) study on Poeae Group 1 and 2 relationships, and reinforced here with similar sampling and maximal support. Previous relationships based on plastid loci (Soreng et al., 2017), placed Scolochloinae as sister to Airinae and Holcinae, falling outside of Loliodinae entirely. It appears that contrasting signals from complete plastomes versus select loci have produced substantially different topologies with regard to the placement of Scolochloinae. With the additional taxa included in this study, the evidence supports tentatively placing Scolochloinae within Loliodinae, pending further assessment of the subtribe, to include more of its representatives with plastome and nuclear level data. Analyses of nrDNA place *Scolochloa* Link. within Group 1 species (sister to Phalaridinae Fr. as noted in Tkach et al. (2020). This indicates a potential reticulate origin of *Scolochloa* that cannot be confirmed with plastid data alone.

4.5.19. Loliinae Dumort

Loliinae is split between the narrow-leaved clade (*Festuca ovina* L.*, Festuca alopecuros* Schousb.*, Festuca* sp., with *Castellia* Tineo as sister to the narrow-leaved *Festuca* representatives) and the remaining broad-leaved clade (*Lolium* L. spp., and *Drymochloa sylvatica* (Pollich) Holub). See also Tkach et al. (2020).

4.5.20. Incertae sedis in Poodinae Soreng & L.J.Gillespie

One species included in this study has been identified in Soreng et al. (2017) as *incertae sedis* in the Poodinae: *Arctagrostis latifolia* (R. Br.) Griseb. Our results place *A. latifolia* as sister to *Cinna arundinacea* L. in Cinninae, also with maximal support. Morphological assessments of *Arctagrostis* Griseb. and *Cinna* L. have indicated strong similarity, with each commonly having single-flowered spikelets (Hoffmann et al., 2013). This assessment is consistent with our phylogenomic results. However, additional sampling is necessary to assess a more accurate relationship and phylogenomic placement, as *C. arundinacea* L. is the only representative of Cinninae included in this study.

With plastome level data, *A. latifolia* is strongly supported in its position. Both Soreng et al. (2017) and Kellogg (2015), as well as Gillespie et al. (2010) place *Arctagrostis* in the HSAQN clade.

4.6. Divergence time estimation (BEAST2)

A Bayesian Evolutionary Analysis Sampling Trees (BEAST2) divergence analysis explored the context of divergence among the species included in this study. Paleoclimate and paleobiogeographic information were also informally investigated in a review and interpretation of previous literature, and discussed here.

We explored potential distinct patterns of speciation and divergence to determine if they coincided with major climatic events (Pimentel et al., 2017). We also investigated likely dispersal routes based on early continental arrangements and drift (Clayton, 1975).

Bremer (2002) and Bouchenak-Khelladi et al. (2010) estimated that the earliest grasses likely originated in the Gondwana region of Pangea circa 66–75 Ma, and subsequent prehistoric continental drift allowed for diverse distributions across broad climate regimes for these earliest grasses (Brown & Smith, 1972; Clayton, 1975, 1981). However, it is also estimated that regions of Gondwana began separating in the early Jurassic period (Cox, 1992; Jokat et al., 2003) eliminating potential overland routes of dispersal for early grasses long before their hypothesized origins, yet modern distributions allude to deeper prehistoric colonization of diverse habitats, suggesting earlier origins of Poaceae than first assessed. Many grasses with cosmopolitan distributions have small, easily dispersed diaspores (Cheplick, 1998). In this study, the spikelet clade is estimated to have diverged from the Anomochlooideae between 105.8 and 109.9 Ma, similar to the results of Burke et al. (2016b), which estimated the origins of the spikelet clade to be roughly 100 Ma. This evidence places the origins of Poaceae deeper into the Cretaceous than many earlier estimates.

Our results pertaining to Pooideae origins are in partial agreement with Schubert et al. (2019) where it was estimated that Pooideae diverged roughly 61 – 77 Ma, yet our timeline indicates a slightly older Cretaceous origin for Pooideae circa 76.3 – 93.3 Ma (Campanian – Turonian ages) (Fig. 3).

The atmosphere of the late Cretaceous contained more than double the CO_2 levels of modern-day (Andrews et al., 1995; DeConto et al., 1998, 1999), and temperatures were steadily increasing by this time. However, evidence suggests that there were isolated south pole glaciation events occurring during the Turonian age, which may have contributed to periods of constant or decreased temperatures in more southern latitudes (Bornemann et al., 2008). While modern Pooideae are comprised of cool season grasses, early lineages had only limited cold tolerance (Schubert et al., 2019). We speculate that later cooling trends of the Paleogene near the Eocene/Oligocene transition between 33 and 35 Ma, which exerted evolutionary pressures as a result of climate cooling, contributed to the later diversifications seen at the supertribe levels (e.g., Poodae, Nardodae, Stipodae) (Minaya et al., 2017; Pimentel et al., 2017; Schubert et al., 2019).

Nearly all Pooideae supertribes had diverged by 30 Mya, excluding Triticodae which is estimated to have diverged around 25 Ma. This younger age may be an artifact of phylogenomic signal from the wheat cultivar species skewing the age younger due to more recent selective breeding for desired genes, or also due to the inclusion of annuals which have high mutation rates and short generation times (Kilian et al., 2009).

Interestingly, grasses likely flourished during this time of greater cooling as they are particularly robust and effective in dispersing long distances, establishing or colonizing diverse habitats, as well as having the ability to adapt to rapidly changing environments, which makes them successful invasive species (Linder et al., 2018; Minaya et al., 2017). In particular, the cool season grasses, in which many lineages evolved frost tolerance, would have navigated the cooling climate more readily than the tropical varieties of grasses and other plants (Linder et al., 2017; Zhong et al., 2018; Schubert et al., 2019). This is reflected in many modern distributions of pooid grasses with ranges primarily in the northern hemisphere and some species inhabiting southern temperate regions (Soreng et al., 2015, 2017; Zhong et al., 2018; Watson et al., 1992 onwards).

Pooideae is a cosmopolitan subfamily, but predominantly distributed in North America and temperate Eurasia (Watson et al., 1992 onwards;
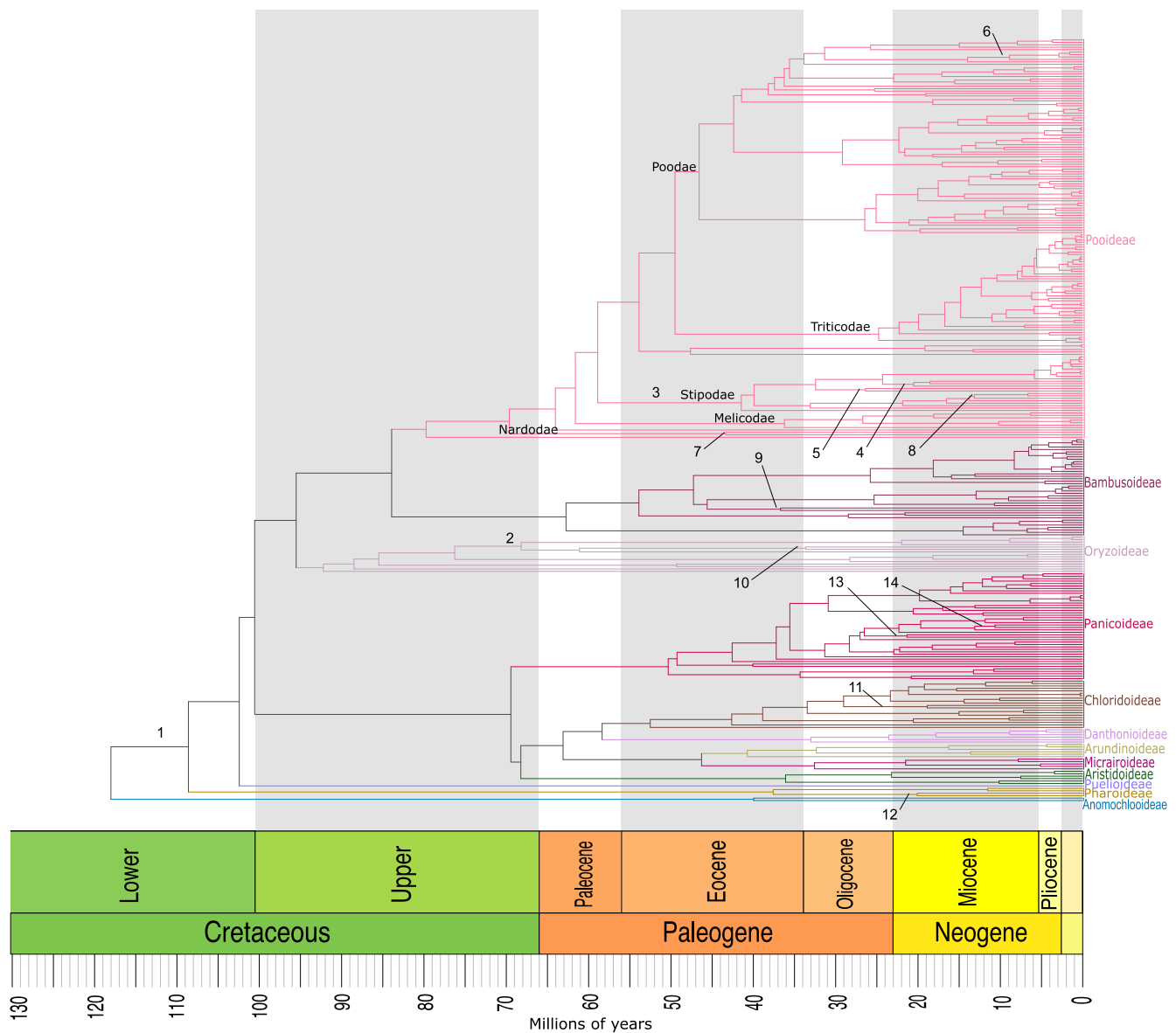
Fig. 3. Chronogram constructed from BEAST divergence estimation analysis. Major pooid clades and fossil calibrations are indicated. Refer to Table 1 for information on fossil calibration points.

Kellogg, 2015; Soreng et al., 2017; Zhong et al., 2018), with largely open habitat adapted species. There are a number of species exhibiting diverse habitats ranging from mountainous open areas to sandy coastal soils to species that thrive in shade tolerant savannas or woodlands (Watson et al., 1992 onwards). It is likely that most early pooid grasses dispersed northward through the southern region of Laurasia in Pangea and some pushed into southern temperate regions from initial ancestral origins in Gondwana (Brown & Smith, 1972; Minaya et al., 2017; Schubert et al., 2019).

Substantial continental movement continued through the Paleogene, and evidence of ancient dispersal routes for pooid grasses has brought about more recent and pointed studies aimed at exploring dispersal at the tribal or generic level (Blattner, 2006; Inda et al., 2008; Birch et al., 2014). Based on this research, and known prehistoric continental drift, it can be assumed that in general, many of the ancient pooids later migrated north into what is modern Europe and through areas now found in the Mediterranean and southwestern Asian corridor (once portions of Laurasia). Ancient pooids also expanded into temperate southern regions of what is now sub-Saharan Africa during the Miocene - Pliocene (Blattner, 2006; Inda et al., 2008; Minaya et al., 2017).

Eventually, the northward expanding groups moved to Australasia from Eurasian regions and separately to the New World through distribution along migratory routes into (what is now) Asia (Minaya et al., 2017). During the Pliocene, some New World bound pooids began migrating via land bridges into North America and continued expanding into areas of what is now Central and South America before transitioning into Australasian regions through long-distance dispersal among intermediate ancient island chains (Blattner, 2006; Jakob et al., 2007; Inda et al., 2008; Birch et al., 2014; Wen et al., 2016; Minaya et al., 2017). As continued continental drift led to dispersal barriers and vicariance, many groups began to diversify in secluded habitats creating disjunct but cosmopolitan distributions (Blattner, 2006).

5. Conclusion

Pooideae exhibit many complex and often reticulate relationships among and within member tribes and subtribes. Here we demonstrated that large-scale datasets using plastome scale information provide added resolution to many sparsely represented groups through the addition of genetic data. There still remains a deep deficit with regard to Poaceae

plastome level data available in sequence repositories. However, this study has explored the many benefits of utilizing the maximum level of genetic information that exists in plastomes.

Sampling strategies, as well as including or excluding alignment gaps, also greatly affect the quality of datasets and resulting analytical outcomes. Likewise, alignment gaps are a byproduct of sampling when introduced into a dataset. Here we have presented evidence that including alignment gaps can produce alternate topologies, which likely introduces artifacts from ambiguously aligned low-complexity regions of DNA, and should be avoided unless absolutely necessary. However, in comparing phylogenies produced from multiple stripping thresholds to explore the effect of alignment gaps on these analyses, we have seen that known/suspected reticulate relationships of hybridizing complexes behave in unexpected or unusual ways. This suggests that the underlying evolutionary complexities of biparentally inherited organismal histories are at least somewhat detectable, even in the phylogenies of uniparentally inherited plastomes.

Finally, the divergence estimation analysis determined a late Cretaceous origin for Pooideae at approximately 76.3–93.4 Ma. While this is slightly older than other recent estimates, it is in line with the hypothesis that divergences among lineages of grasses may be older than commonly estimated, as there is a lack of fossil evidence.

Future directions seek to include greater sampling in underrepresented taxonomic groups, particularly Duthieeae and Calothecinae, and include nuclear data to compare/contrast phylogenetic signals of both plastid and nucleus, as well as further explore possible hybridizing complexes.

CRediT authorship contribution statement

Lauren M. Orton: Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. **Patricia Barberá:** Validation, Resources, Writing - review & editing, Funding acquisition. **Matthew P. Nissenbaum:** Methodology, Writing - review & editing. **Paul M. Peterson:** Validation, Resources, Writing - review & editing. **Alejandro Quintanar:** Validation, Resources, Writing - review & editing, Funding acquisition. **Robert J. Soreng:** Validation, Resources, Writing - review & editing. **Melvin R. Duvall:** Conceptualization, Methodology, Validation, Formal analysis, Investigation, Resources, Writing - original draft, Writing - review & editing, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2021.107110.

References

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* 19 (6), 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Andrews, J.E., Tandon, S.K., Dennis, P.F., 1995. Concentration of carbon dioxide in the Late Cretaceous atmosphere. J. Geol. Soc. 152 (1), 1–3.

Attigala, L., Wysocki, W.P., Duvall, M.R., Clark, L.G., 2016. Phylogenetic estimation and morphological evolution of Arundinarieae (Bambusoideae: Poaceae) based on plastome phylogenomic analysis. Mol. Phylogenet. Evol. 101, 111–121.

Baldwin, B.G., Goldman, D.H., Keil, D.J., Patterson, R., Rosatti, T.J. (Eds.), 2012. The Jepson manual: vascular plants of California. Univ. of California Press.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Pyshkin, A.V., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19 (5), 455–477.

Barberá, P., Soreng, R.J., Peterson, P.M., Romaschenko, K., Quintanar, A., Aedo, C., 2020. Molecular phylogenetic analysis resolves *Trisetum* (Poaceae: Pooideae: Koeleriinae) polyphyletic: Evidence for a new genus, *Sibirotrisetum* and resurrection of *Acrospelion*. J. System. Evol. 58 (4), 517–526.

Barberá, P., Quintanar, A., Peterson, P.M., Soreng, R.J., Romaschenko, K., Aedo, C., 2019. New combinations, new names, typifications, and a new section, sect. *Hispanica*, in *Koeleria* (Poeae, Poaceae). *Phytoneuron* 2019 (46), 1–13.

Barberá, P., Romero-Zarco, C., Aedo, C., 2018. Taxonomic Revision of *Trisetum* sect. *Trisetum* (Poaceae: Pooideae: Aveninae) from Eurasia and North Africa1. Ann. Mo. Bot. Gard. 103 (3), 350–392.

Barberá, P., Romero-Zarco, C., Aedo, C., 2017. Taxonomic revision of *Trisetum* section *Acrospelion* (Poaceae: Pooideae: Aveninae) from Eurasia. Syst. Bot. 42 (4), 754–781.

Bell, M.A., Lloyd, G.T., 2015. strap: an R package for plotting phylogenies against stratigraphy and assessing their stratigraphic congruence. Palaeontology 58 (2), 379–389.

Birch, J.L., Cantrill, D.J., Walsh, N.G., Murphy, D.J., 2014. Phylogenetic investigation and divergence dating of *Poa* (Poaceae, tribe Poeae) in the Australasian region. Bot. J. Linn. Soc. 175 (4), 523–552.

Blaner, A., Schneider, J., Röser, M., 2014. Phylogenetic relationships in the grass family (Poaceae) based on the nuclear single copy locus topoisomerase 6 compared with chloroplast DNA. Syst. Biodivers. 12 (1), 111–124.

Blattner, F.R., 2006. Multiple intercontinental dispersals shaped the distribution area of *Hordeum* (Poaceae). New Phytol. 169 (3), 603–614.

Bornemann, A., Norris, R.D., Friedrich, O., Beckmann, B., Schouten, S., Damsté, J.S.S., Wagner, T., 2008. Isotopic evidence for glaciation during the Cretaceous supergreenhouse. Science 319 (5860), 189–192.

Bouchenak-Khelladi, Y., Verboom, G.A., Savolainen, V., Hodkinson, T.R., 2010. Biogeography of the grasses (Poaceae): a phylogenetic approach to reveal evolutionary history in geographical space and geological time. Bot. J. Linn. Soc. 162 (4), 543–557.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.H., Xie, D., Drummond, A.J., 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10 (4), e1003537.

Bremer, K., 2002. Gondwanan evolution of the grass alliance of families (Poales). Evolution 56 (7), 1374–1387.

Brown, W.V., Smith, B.N., 1972. Grass evolution, the Kranz syndrome, 13C/12C ratios and continental drift. Nature 239 (5371), 345–346.

Bruen, T.C., Philippe, H., Bryant, D., 2006. A simple and robust statistical test for detecting the presence of recombination. Genetics 172 (4), 2665–2681.

Burke, S.V., Grennan, C.P., Duvall, M.R., 2012. Plastome sequences of two New World bamboos—*Arundinaria gigantea* and *Cryptochloa strictiflora* (Poaceae)—extend phylogenomic understanding of Bambusoideae. Am. J. Bot. 99 (12), 1951–1961.

Burke, S.V., Lin, C.S., Wysocki, W.P., Clark, L.G., Duvall, M.R., 2016a. Phylogenomics and plastome evolution of tropical forest grasses (*Leptaspis*, *Streptochaeta*: Poaceae). Front. Plant Sci. 7, 1993.

Burke, S.V., Wysocki, W.P., Zuloaga, F.O., Craine, J.M., Pires, J.C., Edger, P.P., Duvall, M.R., 2016b. Evolutionary relationships in Panicoid grasses based on plastome phylogenomics (Panicoideae; Poaceae). BMC Plant Biol. 16 (1), 140.

Cahoon, A.B., Sharpe, R.M., Mysayphonh, C., Thompson, E.J., Ward, A.D., Lin, A., 2010. The complete chloroplast genome of tall fescue (*Lolium arundinaceum*; Poaceae) and comparison of whole plastomes from the family Poaceae. Am. J. Bot. 97 (1), 49–58.

Carlsen, M.M., Fér, T., Schmickl, R., Leong-Škorničková, J., Newman, M., Kress, W.J., 2018. Resolving the rapid plant radiation of early diverging lineages in the tropical Zingiberales: pushing the limits of genomic data. Mol. Phylogenet. Evol. 128, 55–68.

Caruccio, N., 2011. Preparation of next-generation sequencing libraries using Nextera™ technology: simultaneous DNA fragmentation and adaptor tagging by in vitro transposition. High-Through. Next Gen. Sequen. Meth. Appl. 2011, 241–255. https://doi.org/10.1007/978-1-61779-089-8_17.

Cheplick, G.P., 1998. Seed dispersal and seedling establishment in grass populations. In: Population biology of grasses. Cambridge University Press, Cambridge, pp. 84–105.

Christenhusz, M.J., Byng, J.W., 2016. The number of known plants species in the world and its annual increase. Phytotaxa 261 (3), 201–217.

Christin, P.A., Spriggs, E., Osborne, C.P., Strömberg, C.A., Salamin, N., Edwards, E.J., 2014. Molecular dating, evolutionary rates, and the age of the grasses. Syst. Biol. 63 (2), 153–165.

Clayton, W.D., 1975. Chorology of the genera of Gramineae. Kew Bull. 30, 111–132.

Clayton, W.D., 1981. Evolution and distribution of grasses. Ann. Mo. Bot. Gard. 68, 5–14.

Columbus, J.T., Smith Jr, J.P., 2010. Nomenclatural changes for some grasses in California and the *Muhlenbergia* clade (Poaceae). Aliso: J. System. Evol. Bot. 28 (1), 65–67.

Cotton, J.L., Wysocki, W.P., Clark, L.G., Kelchner, S.A., Pires, J.C., Edger, P.P., Duvall, M. R., 2015. Resolving deep relationships of PACMAD grasses: a phylogenomic approach. BMC Plant Biol. 15 (1), 178.

Cox, K.G., 1992. Karoo igneous activity, and the early stages of the break-up of Gondwanaland. Geological Society, London, Special Publications 68 (1), 137–148.

Cox, M.P., Peterson, D.A., Biggs, P.J., 2010. SolexaQA: at-a-glance quality assessment of Illumina second-generation sequencing data. BMC Bioinf. 11 (1), 485. https://doi. org/10.1186/1471-2105-11-485.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Meth. 9 (8), 772. https://doi.org/10.1038/ nmeth.2109.

DeConto, R.M., Hay, W.W., Bergengren, J.C., 1998. Modeling Late Cretaceous climate and vegetation. Zentralblatt für Geologie und Paläontologie Teil 1, 1996(11/12), 1433–1444.

DeConto, R.M., Hay, W.W., Thompson, S.L., Bergengren, J., 1999. Late Cretaceous climate and vegetation interactions: cold continental interior paradox. Special Papers-Geol. Soc. Am. 391–406.

Duvall, M.R., Fisher, A.E., Columbus, J.T., Ingram, A.L., Wysocki, W.P., Burke, S.V., Kelchner, S.A., 2016. Phylogenomics and plastome evolution of the chloridoid grasses (Chloridoideae: Poaceae). Int. J. Plant Sci. 177 (3), 235–246.

Duvall, M.R., Burke, S.V., Clark, D.C., 2020. Plastome phylogenomics of Poaceae – alternate topologies depend on alignment gaps. Bot. J. Linnean Soc. (in press).

Fan, X., Sha, L.N., Dong, Z.Z., Zhang, H.Q., Kang, H.Y., Wang, Y., Zheng, Y.L., 2013. Phylogenetic relationships and Y genome origin in *Elymus* L. sensu lato (Triticeae; Poaceae) based on single-copy nuclear Acc1 and Pgk1 gene sequences. Mol. Phylogenet. Evol. 69 (3), 919–928.

Felsenstein, J., 2005. PHYLIP (phylogeny inference package) Distributed by the author. Version 3. Seattle: Department of Genome Sciences. University of Washington.

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the nextgeneration sequencing data. Bioinformatics 28 (23), 3150–3152. https://doi. org/10.1093/bioinformatics/bts565.

Gallaher, T.J., Adams, D.C., Attigala, L., Burke, S.V., Craine, J.M., Duvall, M.R., Clark, L. G., 2019. Leaf shape and size track habitat transitions across forest–grassland boundaries in the grass family (Poaceae). Evolution 73 (5), 927–946.

Gillespie, L.J., Soreng, R.J., Paradis, L.M., Bull, R.D., 2010. Phylogeny and reticulation in Poinae subtribal complex based on nrITS, ETS, and *trn*TLF data. In: Seberg, O., Petersen, G., Barfod, A.S., Davis, J.I. (Eds.), Diversity, Phylogeny, and Evolution in the Monocotyledons. Aarhus University Press, Aarhus, pp. 589–617.

Hand, M.L., Spangenberg, G.C., Forster, J.W., Cogan, N.O., 2013. Plastome sequence determination and comparative analysis for members of the *Lolium-Festuca* grass species complex. *G3 (Bethesda): Genes, Genomes*. Genetics 3 (4), 607–616.

Heibl, C., Cusimano, N., Krah, F. S., Heibl, M.C., 2019. Package 'ips'.

Hochbach, A., Schneider, J., Röser, M., 2015. A multi-locus analysis of phylogenetic relationships within grass subfamily Pooideae (Poaceae) inferred from sequences of nuclear single copy gene regions compared with plastid DNA. Mol. Phylogen. Evol. 87, 14–27. https://doi.org/10.1016/j.ympev.2015.03.010.

Hoffmann, M.H., Schneider, J., Hase, P., Röser, M., 2013. Rapid and recent world-wide diversification of bluegrasses (*Poa*, Poaceae) and related genera. PLoS ONE 8 (3), e60061.

Huson, D.H., 1998. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics (Oxford, England) 14 (1), 68–73.

Iles, W.J., Smith, S.Y., Gandolfo, M.A., Graham, S.W., 2015. Monocot fossils suitable for molecular dating analyses. Bot. J. Linn. Soc. 178 (3), 346–374.

Inda, L.A., Segarra-Moragues, J.G., Müller, J., Peterson, P.M., Catalán, P., 2008. Dated historical biogeography of the temperate Loliinae (Poaceae, Pooideae) grasses in the northern and southern hemispheres. Mol. Phylogenet. Evol. 46 (3), 932–957.

Jakob, S.S., Ihlow, A., Blattner, F.R., 2007. Combined ecological niche modelling and molecular phylogeography revealed the evolutionary history of *Hordeum marinum* (Poaceae)—niche differentiation, loss of genetic diversity, and speciation in Mediterranean Quaternary refugia. Mol. Ecol. 16 (8), 1713–1727.

Jokat, W., Boebel, T., König, M., Meyer, U., 2003. Timing and geometry of early Gondwana breakup. J. Geophys. Res. Solid Earth 108 (B9).

Jones, S.S., Burke, S.V., Duvall, M.R., 2014. Phylogenomics, molecular evolution, and estimated ages of lineages from the deep phylogeny of Poaceae. Plant Syst. Evol. 300 (6), 1421–1436.

Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30 (4), 772–780.

Kawahara, T., 2009. Molecular phylogeny among *Triticum-Aegilops* species and of the tribe Triticeae. Breeding science 59 (5), 499–504.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Thierer, T., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12), 1647–1649.

Kellogg, E.A., 2015. Flowering plants, monocots, Poaceae. In: Kubitski, K. (Ed.), The families and genera of vascular plants, vol. 10. Springer International, Cham, pp. 978–983. https://doi.org/10.1007/978-3-319-15332-2.

Kilian, B., Özkan, H., Pozzi, C., Salamini, F., 2009. Domestication of the Triticeae in the Fertile Crescent. In: Genetics and Genomics of the Triticeae. Springer, New York, NY, pp. 81–119.

Leseberg, C.H., Duvall, M.R., 2009. The complete chloroplast genome of *Coix lacryma-jobi* and a comparative molecular evolutionary analysis of plastomes in cereals. J. Mol. Evol. 69 (4), 311–318.

Linder, H.P., Lehmann, C.E., Archibald, S., Osborne, C.P., Richardson, D.M., 2018. Global grass (Poaceae) success underpinned by traits facilitating colonization, persistence and habitat transformation. Biol. Rev. 93 (2), 1125–1144.

Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. J. 17 (1), 10–12. https://doi.org/10.14806/ej.17.1.200.

Mason-Gamer, R.J., Kellogg, E.A., 1996. Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). Syst. Biol. 45 (4), 524–545.

Mason-Gamer, R.J., Burns, M.M., Naum, M., 2010a. Phylogenetic relationships and reticulation among Asian *Elymus* (Poaceae) allotetraploids: analyses of three nuclear gene trees. Mol. Phylogenet. Evol. 54 (1), 10–22.

Mason-Gamer, R.J., Burns, M.M., Naum, M., 2010b. Reticulate evolutionary history of a complex group of grasses: phylogeny of *Elymus* StStHH allotetraploids based on three nuclear genes. PLoS ONE 5 (6), e10989.

Miller, M.A., Pfeiffer, W., Schwartz, T., 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In: Gateway computing environments workshop (GCE). IEEE, pp. 1–8. http://dx.doi.org/10.1109/GCE.2010.5676129.

Minaya, M., Hackel, J., Namaganda, M., Brochmann, C., Vorontsova, M.S., Besnard, G., Catalán, P., 2017. Contrasting dispersal histories of broad-and fine-leaved temperate Loliinae grasses: range expansion, founder events, and the roles of distance and barriers. J. Biogeogr. 44 (9), 1980–1993.

Mitchell, N., Lewis, P.O., Lemmon, E.M., Lemmon, A.R., Holsinger, K.E., 2017. Anchored phylogenomics improves the resolution of evolutionary relationships in the rapid radiation of Protea L. Am. J. Bot. 104 (1), 102–115.

Morris, L.M., Duvall, M.R., 2010. The chloroplast genome of *Anomochloa marantoidea* (Anomochlooideae; Poaceae) comprises a mixture of grass-like and unique features. Am. J. Bot. 97 (4), 620–627.

Nogueira da Silva, L.N., Essi, L., Iganci, J.R.V., Souza-Chies, T.T.D., 2020. Advances in the phylogeny of the South American cool-season grass genus *Chascolytrum* (Poaceae, Pooideae): A new infrageneric classification. Bot. J. Linn. Soc. 192 (1), 97–120.

Orton, L.M., Burke, S.V., Wysocki, W.P., Duvall, M.R., 2017. Plastid phylogenomic study of species within the genus *Zea*: rates and patterns of three classes of microstructural changes. Curr. Genet. 63 (2), 311–323.

Orton, L.M., Burke, S.V., Duvall, M.R., 2019. Plastome phylogenomics and characterization of rare genomic changes as taxonomic markers in plastome groups 1 and 2 Poeae (Pooideae; Poaceae). PeerJ 7, e6959.

Petersen, G., Seberg, O., Yde, M., Berthelsen, K., 2006. Phylogenetic relationships of *Triticum* and *Aegilops* and evidence for the origin of the A, B, and D genomes of common wheat (*Triticum aestivum*). Mol. Phylogenet. Evol. 39 (1), 70–82.

Peterson, P.M., Romaschenko, K., Soreng, R.J., Valdés Reyna, J., 2019. A key to the North American genera of Stipeae (Poaceae: Pooideae) with descriptions and taxonomic names for species of *Eriocoma, Neotrinia, Oloptum*, and five new genera: *Barkworthia,* ×*Eriosella, Pseudoeriocoma, Ptilagrostiella*, and *Thorneochloa*. PhytoKeys 126, 89–125. https://doi.org/10.3897/phytokeys.126.34096.

Pilger, R., 1954. Pflanzengeschichte und Pflanzengeographie. Botanische Jahrbücher für Systematik 76 (3), 313.

Pimentel, M., Escudero, M., Sahuquillo, E., Minaya, M.Á., Catalán, P., 2017. Are diversification rates and chromosome evolution in the temperate grasses (Pooideae) associated with major environmental changes in the Oligocene-Miocene? PeerJ 5, e3815.

Piot, A., Hackel, J., Christin, P.A., Besnard, G., 2017. One-third of the plastid genes evolved under positive selection in PACMAD grasses. Planta 1–12.

Poinar, G.J., Alderman, S., Wunderlich, J., 2015. One hundred million year old ergot: psychotropic compounds in the Cretaceous. Palaeodiversity 8, 13–19.

Posada, D., 2008. jModelTest: phylogenetic model averaging. Mol. Biol. Evol. 25 (7), 1253–1256. https://doi.org/10.1093/molbev/msn083.

Prasad, V., Strömberg, C.A., Alimohammadian, H., Sahni, A., 2005. Dinosaur coprolites and the early evolution of grasses and grazers. Science 310 (5751), 1177–1180.

Quintanar, A., Castroviejo, S., Catalán, P., 2007. Phylogeny of the tribe Aveneae (Pooideae, Poaceae) inferred from plastid trnT-F and nuclear ITS sequences. Am. J. Bot. 94 (9), 1554–1569. https://doi.org/10.3732/ajb.94.9.1554.

Rambaut, A., 2014. FigTree 1.4. 2 software. Institute of Evolutionary Biology, Univ. Edinburgh.

Romaschenko, K., Peterson, P.M., Soreng, R.J., Garcia-Jacas, N., Futorna, O., Susanna, A., 2012. Systematics and evolution of the needle grasses (Poaceae: Pooideae: Stipeae) based on analysis of multiple chloroplast loci, ITS, and lemma micromorphology. Taxon 61 (1), 18–44.

Romaschenko, K., Garcia-Jacas, N., Peterson, P.M., Soreng, R.J., Vilatersana, R., Susanna, A., 2014. Miocene-Pliocene speciation, introgression, and migration of *Patis* and *Ptilagrostis* (Poaceae: Stipeae). Mol. Phylogenet. Evol. 70, 244–259.

Saarela, J.M., Liu, Q., Peterson, P.M., Soreng, R.J., Paszko, B., 2010. Phylogenetics of the grass ' Aveneae-type plastid DNA clade'(Poaceae: Pooideae, Poeae) based on plastid and nuclear ribosomal DNA sequence data. In: Seberg, O., Petersen, G., Barfod, A.S., Davis, J.I. (Eds.), Diversity, Phylogeny, and Evolution in the Monocotyledons. Aarhus Univ. Press, Aarhus, pp. 557–588.

Saarela, J.M., Wysocki, W.P., Barrett, C.F., Soreng, R.J., Davis, J.I., Clark, L.G., Duvall, M.R., 2015. Plastid phylogenomics of the cool-season grass subfamily: clarification of relationships among early-diverging tribes. AoB Plants 7.

Saarela, J.M., Bull, R.D., Paradis, M.J., Ebata, S.N., Peterson, P.M., Soreng, R.J., Paszko, B., 2017. Molecular phylogenetics of cool-season grasses in the subtribes Agrostidinae, Anthoxanthinae, Aveninae, Brixinae, Calothecinae, Koeleriinae, and Phalaridinae (Poaceae: Pooideae: Poeae: Poeae chloroplast group 1). Phytokeys. https://doi.org/10.3897/phytokeys.87.12774.

Saarela, J.M., Burke, S.V., Wysocki, W.P., Barrett, M.D., Clark, L.G., Craine, J.M., Peterson, P.M., Soreng, R.J., Vorontsova, M.S., Duvall, M.R., 2018. A 250 plastome

phylogeny of the grass family (Poaceae): topological support under different data partitions. PeerJ 6, e4299. https://doi.org/10.7717/peerj.4299.

Schneider, J., Doring, E., Hilu, K.W., Röser, M., 2009. Phylogenetic structure of the grass subfamily Pooideae based on comparison of plastid *mat*K gene-30*trn*K exon and nuclear ITS sequences. Taxon 58, 405–424. https://doi.org/10.1002/tax.582008.

Schneider, J., Winterfeld, G., Hoffmann, M.H., Roeser, M., 2011. Duthieeae, a new tribe of grasses (Poaceae) identified among the early diverging lineages of subfamily Pooideae: molecular phylogenetics, morphological delineation, cytogenetics and biogeography. Syst. Biodivers. 9 (1), 27–44.

Schubert, M., Marcussen, T., Meseguer, A.S., Fjellheim, S., 2019. The grass subfamily Pooideae: Cretaceous-Palaeocene origin and climate-driven Cenozoic diversification. Glob. Ecol. Biogeogr. 28 (8), 1168–1182.

Soreng, R.J., Davis, J.I., Doyle, J.J., 1990. A phylogenetic analysis of chloroplast DNA restriction site variation in Poaceae subfam. Pooideae. Plant System. Evol. 172 (1–4), 83–97. https://doi.org/10.1007/BF00937800.

Soreng, R.J., Peterson, P.M., Romaschenko, K., Davidse, G., Zuloaga, F.O., Judziewicz, E. J., Morrone, O., 2015. A worldwide phylogenetic classification of the Poaceae (Gramineae). J. System. Evol. 53 (2), 117–137.

Soreng, R.J., Peterson, P.M., Romaschenko, K., Davidse, G., Teisher, J.K., Clark, L.G., Zuloaga, F.O., 2017. A worldwide phylogenetic classification of the Poaceae (Gramineae) II: An update and a comparison of two 2015 classifications. J. System. Evol. 55 (4), 259–290.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30 (9), 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Stöver, B.C., Müller, K.F., 2010. TreeGraph 2: combining and visualizing evidence from different phylogenetic analyses. BMC Bioinf. 11 (1), 7. https://doi.org/10.1186/1471-2105-11-7.

Stromberg, C.A., McInerney, F.A., 2011. The Neogene transition from C3 to C4 grasslands in North America: assemblage analysis of fossil phytoliths. Paleobiology 37 (1), 50–71.

Tkach, N., Schneider, J., Döring, E., Wölk, A., Hochbach, A., Nissen, J., Röser, M., 2020. Phylogenetic lineages and the role of hybridization as driving force of evolution in grass supertribe Poodae. Taxon 62 (2), 234–277.

Van Rossum, G., 1995. Python tutorial. Technical Report CS-R9526. Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

Warren, D., Geneva, A., Lanfear, R., 2017. RWTY (R We There Yet): an R package for examining convergence of Bayesian phylogenetic analyses, volume 34. R package version 1.0.2. Available at https://CRAN.R-project.org/package=rwty. http://dx.doi.10.1093/molbev/msw279.

Watson, L., Macfarlane, T.D., Dallwitz, M.J., 1992 onwards. The grass genera of the world: descriptions, illustrations, identification, and information retrieval; including synonyms, morphology, anatomy, physiology, phytochemistry, cytology, classification, pathogens, world and local distribution, and references. Version: 15th November 2019. delta-intkey.com.

Wen, J., Nie, Z.L., Ickert-Bond, S.M., 2016. Intercontinental disjunctions between eastern Asia and western North America in vascular plants highlight the biogeographic importance of the Bering land bridge from late Cretaceous to Neogene. J. System. Evol. 54 (5), 469–490.

Wysocki, W.P., Clark, L.G., Kelchner, S.A., Burke, S.V., Pires, J.C., Edger, P.P., Mayfield, D.R., Triplett, J.K., Columbus, J.T., Ingram, A.L., Duvall, M.R., 2014. A multi-step comparison of short-read full plastome sequence assembly methods in grasses. Taxon 63 (4), 899–910. https://doi.org/10.12705/634.5.

Wysocki, W.P., Clark, L.G., Attigala, L., Ruiz-Sanchez, E., Duvall, M.R., 2015. Evolution of the bamboos (Bambusoideae; Poaceae): a full plastome phylogenomic analysis. BMC Evol. Biol. 15 (1), 50.

Zerbino, D.R., Birney, E., 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18 (5), 821–829.

Zhao, L., Zhang, N., Ma, P.F., Liu, Q., Li, D.Z., Guo, Z.H., 2013. Phylogenomic analyses of nuclear genes reveal the evolutionary relationships within the BEP clade and the evidence of positive selection in Poaceae. PLoS ONE 8 (5), e64642.

Zhong, J., Robbett, M., Poire, A., Preston, J.C., 2018. Successive evolutionary steps drove Pooideae grasses from tropical to temperate regions. New Phytol. 217 (2), 925–938.

Table 1 References

Dugas, D.P., Retallack, G.J., 1993. Middle Miocene fossil grasses from Fort Ternan, Kenya. J. Paleontol. 67 (1), 113–128.

Elias, M.K., 1942. Tertiary prairie grasses and other herbs from the High Plains, vol. 41. Geological Society of America.

Jacobs, B.F., Kabuye, C.H., 1987. A middle Miocene (12· 2 my old) forest in the East African Rift Valley, Kenya. J. Human Evol. 16 (2), 147–155.

Jiménez-Moreno, G., Fauquette, S., Suc, J.P., 2010. Miocene to Pliocene vegetation reconstruction and climate estimates in the Iberian Peninsula from pollen data. Rev. Palaeobot. Palynol. 162 (3), 403–415.

Juchniewicz, K., 1975. Flora kopalna Turowa kolo Bogatyni w swietle analizy nablonkowey. Pr. Muz. Ziemi 24, 65–134.

MacGinitie, H.D., 1953. Fossil plants of the Florissant beds, Colorado, vol. 599. Carnegie institution of Washington.

Prasad, V., Strömberg, C.A.E., Leaché, A.D., Samant, B., Patnaik, R., Tang, L., Sahni, A., 2011. Late Cretaceous origin of the rice tribe provides evidence for early diversification in Poaceae. Nat. Commun. 2 (1), 1–9.

Strömberg, C.A., 2005. Decoupled taxonomic radiation and ecological expansion of open-habitat grasses in the Cenozoic of North America. Proc. Natl. Acad. Sci. 102 (34), 11980–11984.

Thomasson, J.R., 1978. Observations on the characteristics of the lemma and palea of the late Cenozoic grass *Panicum elegans*. Am. J. Bot. 65 (1), 34–39.

Thomasson, J.R., 1980. *Paleoeriocoma* (Gramineae, Stipeae) from the Miocene of Nebraska: Taxonomic and phylogenetic significance. System. Bot. 233–240.

Thomasson, J.R., 1985. Miocene fossil grasses: Possible adaptation in reproductive bracts (lemma and palea). Ann. Mo. Bot. Gard. 843–851.

Walther, H., 1974. Ergänzung zur Flora von Seifhennersdorf/Sachsen. II. Teil. Abh. Staatl. Mus. Mineral. Geol. Dresden 21, 143–185.