Summer 7-21-2019

Molecular and Morphological Variation among Populations of *Pediomelum tenuiflorum* (Pursh) A.N. Egan (Fabaceae) in Nebraska, USA

Chauncey L. Kellar Jr.
University of Nebraska at Omaha, ckellar@unomaha.edu

Roxi Kellar
University of Nebraska at Omaha, rkellar@unomaha.edu

Molecular and Morphological Variation among Populations of *Pediomelum tenuiflorum* (Pursh) A.N. Egan (Fabaceae) in Nebraska, USA

Chauncey L. Kellar and P. Roxanne Kellar

Biology Department, University of Nebraska at Omaha, 6001 Dodge St., Omaha, NE 68182, USA

Corresponding author — P. Roxanne Kellar, University of Nebraska at Omaha, 6001 Dodge St., Omaha, NE 68182; *email* rkellar@unomaha.edu, *tel* 402-554-2840

Abstract

Individuals of *Pediomelum tenuiflorum*, "wild alfalfa", from disjunct populations in Nebraska vary extensively in their overall gestalt. Those in the western and central part of the state have a very slender growth habit, with thin stems and few, small flowers; whereas, those in the southeast have a very robust growth habit with heavy-looking stems and many tightly clustered flowers. For nearly 200 years, taxonomists have alternated between splitting *P. tenuiflorum* into two species, with the many-flowered morphotype named *P. floribundum*, and lumping all the morphological variants into one species as they are now. In this study, we investigated morphological and molecular characters that could be used to clarify taxonomic classifications of these morphotypes. We measured 10 morphological characters on 51 specimens and sequenced nearly 300,000 nucleotide characters on the Illumina platform from three cellular genomes in seven samples of *Pediomelum* plus an outgroup taxon. Results revealed six significantly different morphological characters but ambiguous evolutionary histories of the plastid and mitochondrial genomes in *P. tenuiflorum*. Our complete plastid genomes and genes and noncoding regions of the mitochondrial genome may be used as a foundation for studying the evolutionary histories of these genomes. Additionally, we identified seven highly variable genomic regions in the chloroplast genome upon which a molecular phylogenetic investigation on an expanded set of samples from across the species' geographic distribution can be conducted to further define the taxonomic placements of *P. tenuiflorum* and *P. floribundum*.

Keywords: chloroplast, *floribundum*, genome, next-generation sequencing, phylogenetics, plastid, *Psoralea*, *Psoralidium*, taxonomy, systematics

doi 10.32873/unl.dc.tnas.39.17 ; https://doi.org/10.32873/unl.dc.tnas.39.17

Introduction

Pediomelum tenuiflorum (Pursh) A.N. Egan is a flowering plant species in Fabaceae and is distributed primarily across the Great Plains region of North America (Great Plains Flora Association 1986). The species was first collected along the Missouri River in an area now known as Lyman County, South Dakota in 1804 by Meriwether Lewis during the historic Lewis and Clark expedition (Lewis and Clark 1983). Frederick Pursh (1814) published the first description of the species, which he named *Psoralea tenuiflora* Pursh. The first specimen from Nebraska was collected in an area now known as Hall or Buffalo County in 1820 by Edwin James during the Long Expedition (Kaul, Sutherland, and Rolfsmeier 2012). In Nebraska, there are morphologically distinct populations of *P. tenuiflorum* in the central and western parts of the state that have little or no overlap (i.e., the populations are not close enough to cross-breed) with individuals in the southeastern part of the state (Kaul, Sutherland, and Rolfsmeier 2012, USDA, NRCS, 2019). Specimens such as those in SE Nebraska were originally named *Psoralea floribunda* Nutt. (Torrey and Gray 1838), but the taxon was later synonymized with *Psoralidium tenuiflorum* (Grimes 1990). Our study was inspired by lifetime plant taxonomists in Nebraska who noted in their description of *Psoralidium tenuiflorum* in *The Flora of Nebraska*, "Some authors have assigned plants with crowded, showy flowers (6–8 mm) to another species, *Psoralidium floribundum* (Nutt.) Rydb., or to a separate variety (as in *Atlas* and *FGP*). In Nebraska such plants are strictly from the s-e part and show little intergradation with the more northern and western plants. Both Grimes (1990: *Mem. N.Y. Bot. Gard.* 61:37) and Barneby (1989: *Intermountain Flora* 3B: 22) consider *P. floribundum* to be a simple synonym of *P. tenuiflorum*, but further study may show it to be worthy of recognition at the specific or infraspecific level" (Kaul, Sutherland, and Rolfsmeier 2012, p. 473).

Since it was first validly published in 1814, *Pediomelum tenuiflorum* has undergone multiple taxonomic studies and revisions. It was first split into two species, *Psoralea tenuiflora* (Pursh 1814) and *Psoralea floribunda* (Torrey and Gray 1838) and recognized as such by Vail (1894). The split was based on the following morphological differences: 1) *P. tenuiflora* is nearly glabrous, whereas *P. floribunda* is canescent; 2) *P. tenuiflora* is noticeably gland-dotted, whereas *P. floribunda* is not glandular; 3) *P. tenuiflora* has infrequent, very small flowers, whereas *P. floribunda* has abundant flowers, sometimes numbering 40–50 on a raceme; and 4) *P. floribunda* has more robust stems when compared to the slender stems of *P. tenuiflora* (Torrey and Gray 1838, Vail 1894). Vail (1894, p. 97) noted that, "… in the bulk of the material in the herbaria examined they can be easily separated." In 1895, Rydberg demoted *P. floribunda* to a variety of *P. tenuiflorum* – *Psoralea tenuiflora floribunda* (Nutt.) Rydberg – noting that, "… it can be easily distinguished, but sometimes, especially in the central portion of the state, it grades into the true *P. tenuiflora*." (Rydberg 1895, p. 55). However, in 1919, Rydberg separated out a new genus, *Pediomelum* Rydberg, re-elevated *Psoralea tenuiflora floribunda* to species status, *Psoralidium floribundum* (Nutt.) Rydberg, and transferred *Psoralea tenuiflora* to *Psoralidium tenuiflorum* (Pursh) Rydberg. The rest of the 20th century saw varied uses and combinations of names utilizing *Psoralea, Psoralidium,* and *Pediomelum* (e.g., Great Plains Flora Association, Barker, and Barkley 1977, Ward 1983, Great Plains Flora Association 1986), but since 1990, authors have primarily followed the circumscription by Grimes who supported *Psoralidium tenuiflorum* as a species, with *P. batesii* Rydberg, *P. bigelovii* Rydberg, *P. floribundum*, and *P. obtusilobum* (Torrey & A. Gray) Rydberg as synonyms (Grimes 1990).

In 2009, *Psoralidium tenuiflorum* was transferred to *Pediomelum tenuiflorum* (Egan and Reveal 2009) based on a phylogeny of North American members of Psoraleeae inferred from eight plastid and nuclear DNA markers (Egan and Crandall 2008). Their phylogeny included two specimens of *P. tenuiflorum* (Egan & Egan 192 and Egan & Egan 194) collected in central Texas in short/mixed grass prairie habitats similar to that of central and western Nebraska. No specimens of *P. tenuiflorum* resembling those from SE Nebraska were included in the Egan and Crandall (2008) analysis. Furthermore, the authors noted that, due to relatively recent speciation events, molecular variation within *Pediomelum* can be difficult to detect. This is evidenced in their phylogeny by lower posterior probabilities among splitting events leading to *Pediomelum tenuiflorum*, among others (see Egan and Crandall 2008, Fig. 3).

As massively parallel (or next-generation) sequencing becomes more reliable and cost effective, systematists have the opportunity to include larger numbers of, and potentially more variable, genetic regions in alignments for phylogenetic inference. However, unlike Sanger sequencing (Sanger and Coulson 1975), which requires only two matching, high-quality base calls for high confidence, massively parallel sequencing requires a minimum sequencing depth of 30× for plastome assemblies and 40× for high-quality nuclear ribosomal DNA (nrDNA) assemblies (Straub *et al*. 2012). But, larger numbers of parsimony-informative, nucleotide characters can be particularly valuable in recently diverged or slowly evolving groups (e.g. Steele *et al*. 2010). Another advantage to using genome survey sequences (GSS; Steele *et al*. 2012) is the chance to recover genetic regions from three, independently evolving, molecular lines of evidence – nuclear, chloroplast, and mitochondrial genomes (e.g., Steele and Pires 2011). Sequences from differing organellar genomes can be combined to infer a total molecular evidence phylogeny. Once these regions have been sequenced and aligned, researchers can also identify particularly variable regions and design primers to include a broader taxonomic group with focused Sanger sequencing.

In this study, we examined and compared morphological characters of *Pediomelum tenuiflorum* individuals from nine sites across Nebraska (Fig. 1), plus multiple herbarium specimens. We extracted DNA from samples in *Pediomelum*, conducted massively parallel sequencing on the Illumina (Illumina, Inc., San Diego, CA, USA) platform, mined Illumina reads from Kellar *et al*. (2015) and Ahrendsen *et al*. (2016), aligned sequences with some of those from Egan and Crandall (2008), and inferred phylogenies from multiple datasets to address the question, Is *Pediomelum floribundum* a species separate from *P. tenuiflorum*?

Materials and Methods

Sampling

Collection sites

Specimens of *Pediomelum tenuiflorum* were collected from nine locations across Nebraska (NE), U.S.A. (Fig. 1). In western/central NE, short/mixed grass and Sandhills

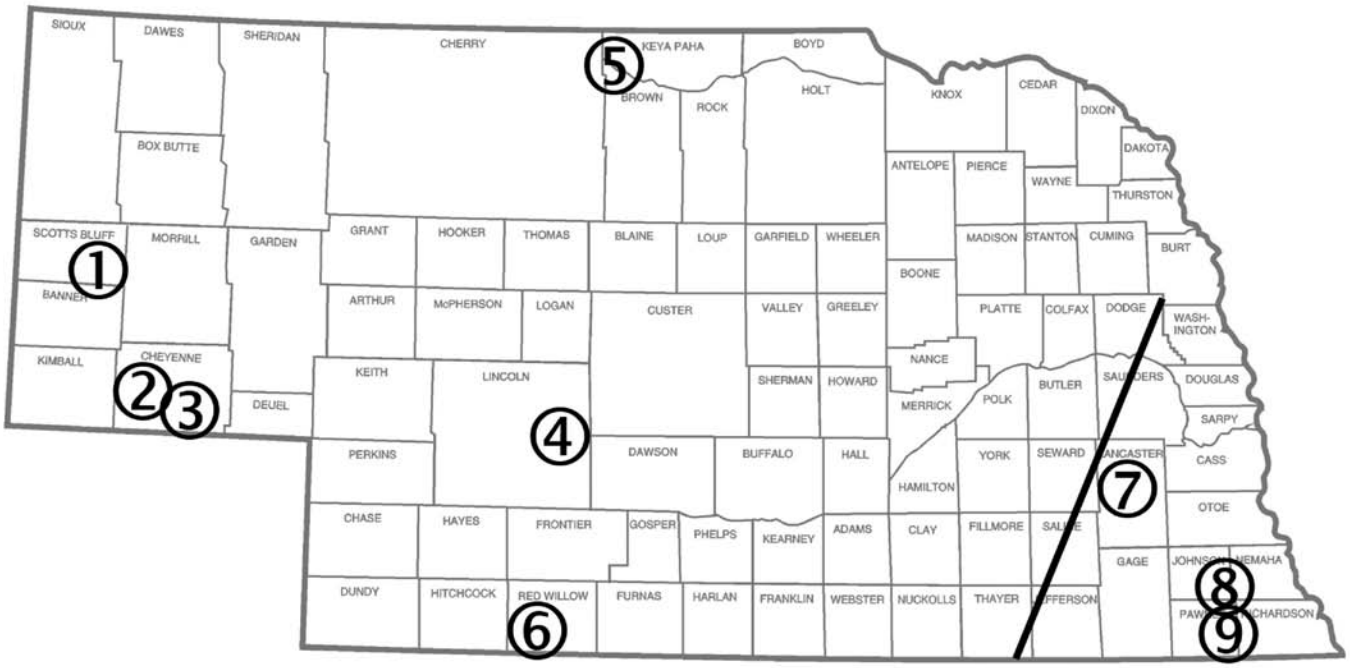

Figure 1. **Map of *Pediomelum tenuiflorum* collection sites in Nebraska, U.S.A.** 1) Buffalo Creek Wildlife Management Area (WMA); 2) Sidney, NE; 3) I-80 at Sidney, NE; 4) Jeffrey Reservoir; 5) Niobrara Valley Preserve; 6) McCook, NE; 7) Nine Mile Prairie; 8) Twin Oaks State WMA; 9) Table Rock State WMA. The heavy diagonal line shows the line of demarcation between western/central and southeastern Nebraska morphotypes. Base map retrieved from: https://nebraskalegislature.gov/about/counties.php

prairie sites included (number corresponds to the site number in Fig. 1; Co. = County): 1) Buffalo Creek Wildlife Management Area (WMA; 41º 42′ 28″ N, 103º 37′ 13″ W), Scottsbluff Co., 2) north side of Old Post Rd., 1 km (0.62 mi) west of Upland Pkwy in Sidney, NE (41º 07′ 02″ N, 102º 57′ 40″W), Cheyenne, Co., 3) I-80 at Sidney, NE, Upland Pkwy exit ramp (41º 06′ 49″ N, 102º 56′ 51″ W), Cheyenne Co., 4) Jeffrey Reservoir (40º 57′ 14″ N, 100º 23′ 50″ W), Lincoln Co., 5) Niobrara Valley Preserve (42º 47′ 21″ N, 100º 02′ 0″ W), Keya Paha Co., and 6) McCook, NE (40º 11′ N, 100º 37′ W), Red Willow Co. In southeastern Nebraska, mixed/tall grass prairie sites included: 7) Nine Mile Prairie (40º 52′ 03″ N, 96º 48′ 57″ W), Lancaster Co., 8) Twin Oaks State WMA (40º 19′ 13″ N, 96º 07′ 59″ W), Johnson Co., and 9) Table Rock State WMA (40º 10′ 49″ N, 96º 04′ 01″ W), Pawnee Co. The heavy diagonal line on the map in Figure 1 delineates the two morphotypes. Southeast of this line corresponds to the Loess Hills and Glacial Drift Hills regions of Nebraska (Chapman *et al*. 2001) as well as the region of Nebraska that receives ≥ 32 inches of rainfall per year (worldatlas, n.d.).Collections were deposited into OMA herbarium. In addition, leaflets were collected and dried over silica gel for DNA extractions.

Specimens included in phylogenetic analyses

Phylogenetic analyses included 32 taxa, incorporating markers downloaded from GenBank (Benson *et al*. 2018) for 25 species generated by Egan and Crandall (2008), new chloroplast and mitochondrial regions assembled from Illumina reads sequenced by Kellar *et al*. (2015) and Ahrendsen *et al*. (2016), and new Illumina sequences from two new samples of *Pediomelum tenuiflorum* from Nebraska. DNA was extracted from two *P. tenuiflorum* samples – one from Red Willow Co. (Sutherland 8317) and one from Pawnee Co. Nebraska (Sutherland 8144) (Fig. 2). These individuals were selected for being good representatives of the two morphotypes that are the focus of this study. Raw Illumina sequence reads for additional *Pediomelum* species and outgroups that were sequenced by Kellar *et al*. (2015) and Ahrendsen *et al*. (2016) were further mined for this work, meaning we extracted new sequences that were not used in the previous studies. Using current taxonomic names, the additional species included (collection number in parentheses): 1) *Pediomelum argophyllum* (Pursh) Grimes (Jones 1027), 2) *Pediomelum digitatum* (Nutt. ex Torr. & A. Gray) Isely. (Jones 1029), 3) *Pediomelum esculentum* (Pursh) Rydb. (Ahrendsen 14), 4)

Figure 2. ***Pediomelum tenuiflorum* morphotypes.** Collections used as representatives of west (A; Sutherland 8317) and east (B; Sutherland 8144) morphotypes in phylogenetic analyses. Photographs by PRK.

Pediomelum tenuiflorum (Jones 1026), and 5) *Ladeania lanceolata* (Pursh) A. N. Egan & Reveal (Ahrendsen 24). One additional sample of *Pediomelum digitatum* (Jones 1022) that was sequenced by Ahrendsen *et al*. (2016) but was not included in their study, is included here. Therefore, DNA sequences from eight samples were mined, processed, and analyzed from our collections, and new sequences were uploaded to GenBank.

The eight specimens that were Illumina-sequenced were included in an 8-taxon dataset for phylogenetic analysis. A 32-taxon dataset was assembled from seven of these eight specimens (Jones 1022 was not included because there were two other *P. digitatum* representatives), plus DNA sequences for 25 specimens downloaded from GenBank. ITS1, 5.8S, and ITS2, plus five chloroplast regions for the 25 species were downloaded from GenBank (accession numbers listed in Appendix 1) and aligned with our sequences. Two species in *Rupertia*, three specimens in *Ladeania*, and two species in *Orbexilum* were used as outgroups for the 32-taxon dataset because these genera are most closely related to *Pediomelum*, and they were used as outgroups in the Egan and Crandall (2008) analysis. *Ladeania lanceolata* was used as the outgroup in the 8-taxon dataset because it occurs in the sister group to *Pediomelum*, along with *Rupertia*, (Egan and Crandall 2008), and it grows in Nebraska.

Measurements and statistical analyses of morphological characters

Morphological characters of 51 samples were measured — 26 specimens from west/central NE and 25 specimens from southeast NE and included those listed above plus collections from the OMA herbarium (Table 1). Ten quantitative morphological characters were measured as follows: 1) average stem diameter (average of the largest and

Table 1. Fifty-one specimens (26 specimens from west/central NE and 25 specimens from southeast NE) on which 10 quantitative morphological characters were measured. Those from Southeastern NE were historically called *P. floribunda*. All specimens are housed in OMA herbarium.

County in Central/Western NE	*Collector*	*Collector No.*	*Collection Date*	*Site No. on Fig. 1*
Scottsbluff	C. Kellar & R. Kellar	52	24-Jul-2014	1
Scottsbluff	C. Kellar & R. Kellar	53	24-Jul-2014	1
Scottsbluff	C. Kellar & R. Kellar	66	4-Jul-2015	1
Scottsbluff	C. Kellar & R. Kellar	67	4-Jul-2015	1
Scottsbluff	C. Kellar & R. Kellar	64	4-Jul-2015	1
Scottsbluff	C. Kellar & R. Kellar	65	4-Jul-2015	1
Scottsbluff	Bragg	--	24-May-1992	N/A
Cheyenne	C. Kellar & R. Kellar	62	4-Jul-2015	2
Cheyenne	C. Kellar & R. Kellar	63	4-Jul-2015	2
Cheyenne	C. Kellar & R. Kellar	51	24-Jul-2014	3
Lincoln	C. Kellar & R. Kellar	50	24-Jul-2014	4
Lincoln	C. Kellar & R. Kellar	60	3-Jul-2015	4
Lincoln	C. Kellar & R. Kellar	61	3-Jul-2015	4
Red Willow	Sutherland	8317	27-Jul-2013	6
Keya Paha	C. Kellar & R. Kellar	8	7-Jul-2013	5
Keya Paha	C. Kellar & R. Kellar	7	7-Jul-2013	5
Keya Paha	C. Kellar & R. Kellar	10	7-Jul-2013	5
Keya Paha	C. Kellar & R. Kellar	11	7-Jul-2013	5
Keya Paha	Jones	1026	5-Jun-2012	5
Sioux	Bray	84	15-Jun-2001	N/A
Sioux	Bray	41	7-Jul-1992	N/A
Sioux	Sutherland	2557	25-Jul-1969	N/A
Lincoln	Sutherland	1879	20-Jun-1968	N/A
Dawes	Hill	PR-196	4-Aug-2001	N/A
Keith	Reitz	--	25-Oct-1981	N/A
Keith	Cooper	83	5-Jun-2004	N/A

County in Southeastern NE	*Collector*	*Collector No.*	*Collection Date*	*Site No. on Fig. 1*
Lancaster	C. Kellar & R. Kellar	47	13-Jul-2014	7
Lancaster	C. Kellar & R. Kellar	48	13-Jul-2014	7
Lancaster	C. Kellar & R. Kellar	56	7-Jun-2015	7
Lancaster	Ahrendsen	28	18-Jun-2013	7
Lancaster	Wendel	160	29-Jun-1992	N/A
Lancaster	Williams	92-66	5-Jul-1992	N/A
Lancaster	Sharpe	--	1-Jun-1967	N/A
Lancaster	Sharpe	--	7-Jun-1967	N/A
Lancaster	Lee	83	25-Jun-2001	N/A
Johnson	C. Kellar & R. Kellar	45	10-Jul-2014	8
Johnson	C. Kellar & R. Kellar	46	10-Jul-2014	8
Pawnee	D. Sutherland & R. Harms	8144	15-Jun-2013	8
Pawnee	C. Kellar & R. Kellar	40	10-Jul-2014	9
Pawnee	C. Kellar & R. Kellar	41	10-Jul-2014	9
Pawnee	C. Kellar & R. Kellar	42	10-Jul-2014	9
Pawnee	C. Kellar & R. Kellar	44	10-Jul-2014	9
Pawnee	C. Kellar & R. Kellar	58	13-Jun-2015	9
Pawnee	C. Kellar & R. Kellar	59	13-Jun-2015	9
Douglas	Lamphere	1196	26-Jun-1993	N/A
Thayer	Kaul	1845	7-Jul-1968	N/A
Butler	Sutherland	1700	9-Jun-1968	N/A
Gage	Fogell	48	25-May-1999	N/A
Barton Co., MO	Folstead	12800	7-Jun-1953	N/A

Note: "--" indicates there was no collector number on the specimen.

smallest stem on the specimen), 2) average leaflet width, 3) average leaflet length, 4) average peduncle diameter, 5) average corolla depth, 6) average banner width, 7) average number of flowers per inflorescence, 8) average fruit length, 9) average fruit width, and 10) average internode length (between peduncles). Stem diameters were measured using a digital micrometer in millimeters (mm). All other measurements were made using a metric ruler under a microscope and recorded in mm. For those measurements for which an average is reported, three measurements on the plant were made and averaged. For example, on one individual, for leaflet length, the average-sized mature leaf was selected from visual observation, then the three leaflets of the trifoliate leaf were measured and averaged.

For each specimen, each measurement (or average measurement) is independent of measurements from other specimens. Each average was averaged, and the mean was plotted on a bar graph. One standard error was calculated for each dataset and is shown on the bar graph. The means of west and east measurements were compared using two-sample *t*-tests. If west and east measurements were significantly different, the *P*-value was indicated on the bar graph. Additionally, we tested each character for normality, and then conducted a discriminant analysis using Minitab 18 (Minitab, LLC 2019) to determine how well the morphological characters of the two morphotypes separated into two groups, and which characters most contributed to separation.

DNA extraction, sequencing, and sequence assembly

Total DNA was extracted from ca. 20 mg silica-dried leaf tissue using the IBI Genomic DNA Mini Kit (IBI Scientific, Peosta, Iowa, USA). Multiple extractions were conducted to obtain a minimum of 12 µg of DNA for massively parallel sequencing. Samples were sent to the University of Missouri (MU) DNA Core Facilities for production of sequencing libraries and sequencing. Samples were run on the Illumina Hi-Seq platform at 12 samples per lane of single-pass, 100-bp reads.

Short-read sequence files of 100-bp reads for each sample were downloaded to the desktop software program Geneious (version 10.0.9; Biomatters [www.geneious.com]). The complete plastid genome for *Pediomelum tenuiflorum* (Jones 1026) was assembled using reference-based mapping, using the reference genome *Cicer arietinum* (Fabaceae; GenBank accession no. NC_011163; Jansen *et al*. 2008). In several cases, reference-based assembly left gaps in the consensus sequence that required multiple iterations of *de novo* assembly in which the ends of sequences were built from short-read scaffolding until the consensus sequences matched a known region. *De novo* assemblies were run multiple times until all gaps were filled and the boundaries of the inverted repeat (IR) regions were identified. The remainder of the plastid genomes were assembled using *P. tenuiflorum* as the reference genome. In a similar fashion, mitochondrial regions were assembled, initially using *Vigna angularis* (Fabaceae; GenBank accession no. NC_021092) as the reference genome, and subsequently using *P. tenuiflorum*. Nuclear ribosomal (nrDNA) internal transcribed spacers 1 & 2 (ITS1 & ITS2; including 5.8S) were assembled using *P. tenuiflorum* downloaded from GenBank (Accession no. EF517839; Egan and Crandall 2008) to assemble these regions for the other taxa. Average sequence coverage for each region was recorded.

Sequence alignment, data partitions, and phylogenetic analysis

Two datasets were assembled for phylogenetic analysis. The first dataset of 32 taxa included nrDNA ITS1, 5.8S, and ITS2 and five plastid regions (*matK*, *rpoB-trnC, trnD-trnY, trnL-F,* and *trnS-trnG*). *trnL-F* included a portion of the *trnL* intron, *trnL*-UAA exon 2, and the *trnL-trnF* intergenic spacer (IGS). The second dataset of eight taxa included complete chloroplast genomes, multiple coding and noncoding mitochondrial regions, and nrDNA ITS1, 5.8S, and ITS2. The plastid and nuclear regions were concatenated in the 32-taxon dataset. For the 8-taxon dataset, phylogenies were inferred from sequences from each organelle separately, and then were concatenated for a total molecular evidence phylogenetic analysis. In Geneious, all datasets were aligned using the sequence alignment tool MAFFT (v. 7.017; with default algorithm, scoring matrix: 200PAM / k = 2, gap open penalty: 1.53, and offset value 0.123; Katoh *et al*. 2002 & 2013). Sequence alignments were uploaded to Dryad (www.datadryad.org) and all newly assembled gene sequences used for phylogenetic analyses were uploaded to GenBank (Benson *et al*. 2018). We ran AICc through CIPRES Science Gateway (www.phylo.org) to determine the best models of evolution using AICc. Phylogenetic analyses were conducted using RAxML, version 8 (Stamatakis 2014), accessed through Geneious.

Maximum parsimony (MP) analyses were performed in PAUP* 4.0b10 (Swofford 2003). For each dataset,

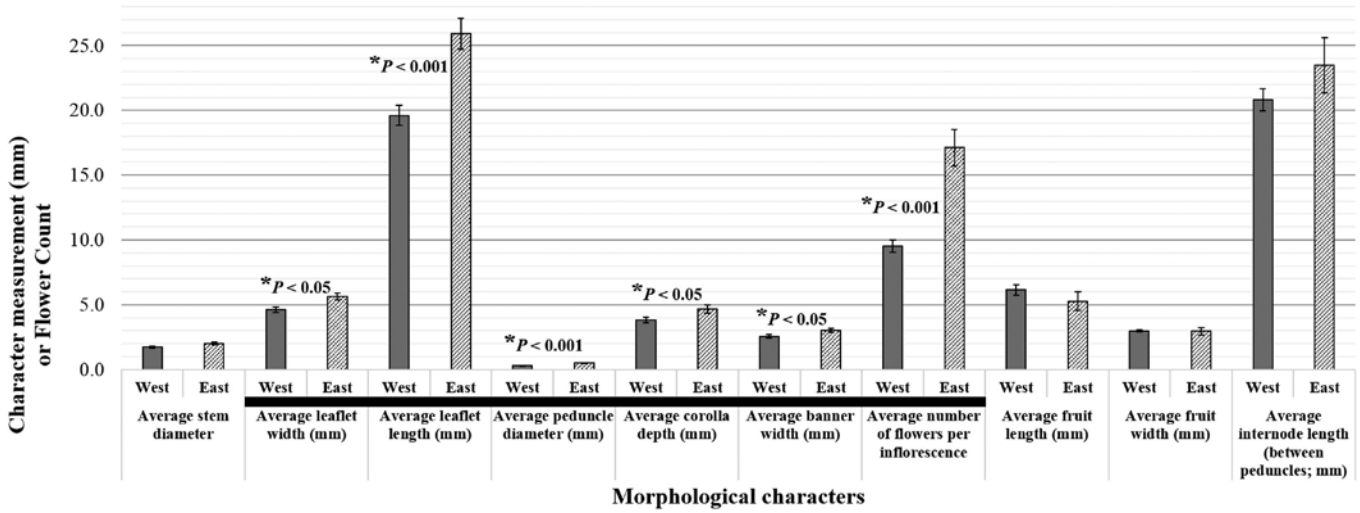

Figure 3. Measurements of morphological characters. Bar graph revealing morphological character measurements compared among *Pediomelum tenuiflorum* populations in western/central (West; solid gray bar) Nebraska and southeastern (East; hatched bar) Nebraska, U.S.A. *P*-values are indicated for (and the horizontal solid black bar highlights) the six measurements that were significantly different. Only nine specimens from West and six specimens from East had fruits. *n* = 51

heuristic searches were conducted using 100 random addition replicates with tree-bisection-reconnection (TBR) branch swapping with the maximum number of trees set to 5,000. Support for internal nodes was assessed using bootstrap analysis (Felsenstein 1985) of 1000 replicates with one random addition per replicate. Maximum likelihood (ML) analyses were conducted using RAxML, version 8 (Stamatakis, 2014) accessed through Geneious. ML analyses used rapid hill-climbing starting with a random tree. Likelihood scores of the optimal trees generated by RAxML were calculated in PAUP* 4.0b10 (Swofford 2003). Nonparametric bootstrap (BS) analyses on 1000 replicates were performed by RAxML in Geneious, and BS consensus trees were constructed in PAUP*.

Results

Morphological characters

Ten quantitative morphological characters were measured on 51 *Pediomelum tenuiflorum* specimens (Table 1); 26 from western/central NE and 25 from southeastern NE (Fig. 3). The solid black bar in Figure 3 indicates the six measurements that were significantly different between the two populations (Table 2). The ranges of all west measurements overlap with east measurements. Only nine specimens from West and six specimens from East Nebraska had fruits.

The morphological character data were tested for normality. Only average leaflet length was normally

Table 2. Measurements of six morphological characters that are significantly different between *Pediomelum tenuiflorum* samples from western/central Nebraska (West) and *P. tenuiflorum* samples from southeastern Nebraska (East). *n* = 51

Character	*Mean*		*Median*		*Range*		*P-value*
	West	*East*	*West*	*East*	*West*	*East*	
average leaflet width (mm)	4.63	5.62	4.6	5.3	3.1–7.5	3.6–8.7	*P* < 0.05
average leaflet length (mm)	19.60	25.91	19.8	26.0	13.2–26.3	8.7–35.3	*P* < 0.001
average peduncle diameter (mm)	0.30	0.49	0.3	0.5	0.1–0.6	0.2–0.8	*P* < 0.001
average corolla depth (mm)	3.81	4.69	4.0	4.7	2.5–5.3	3.0–7.7	*P* < 0.05
average banner width (mm)	2.57	3.03	2.7	3.0	1.9–3.4	1.9–4.2	*P* < 0.05
average number of flowers per inflorescence (no. of flowers)	9.51	17.12	9.7	16.7	4.0–15.3	7.0–31.0	*P* < 0.001

distributed. All other characters were binomially distributed; therefore, the characters were transformed using the natural logarithm (ln), and then were tested for normality. Ln average stem diameter, ln average leaflet width, ln average peduncle diameter, and ln average corolla depth were normally distributed. Average banner width, average number of flowers per inflorescence, and average internode length were not normally distributed, even when transformed. However, it is not considered fatal to the discriminant analysis if the data are not normally distributed (www.statsoft.com). Data on the fruits were not included in the discriminant analysis because many specimens did not have fruits.

The discriminant analysis revealed that 83.7% of the 51 *Pediomelum* specimens were placed in the correct group (West or East). The West group had the largest linear discriminant function (LDF) for internode length (0.243) as compared to the East group (0.136). The East group had the largest LDF for ln average stem diameter (2.249) as compared to the West group (0.072).

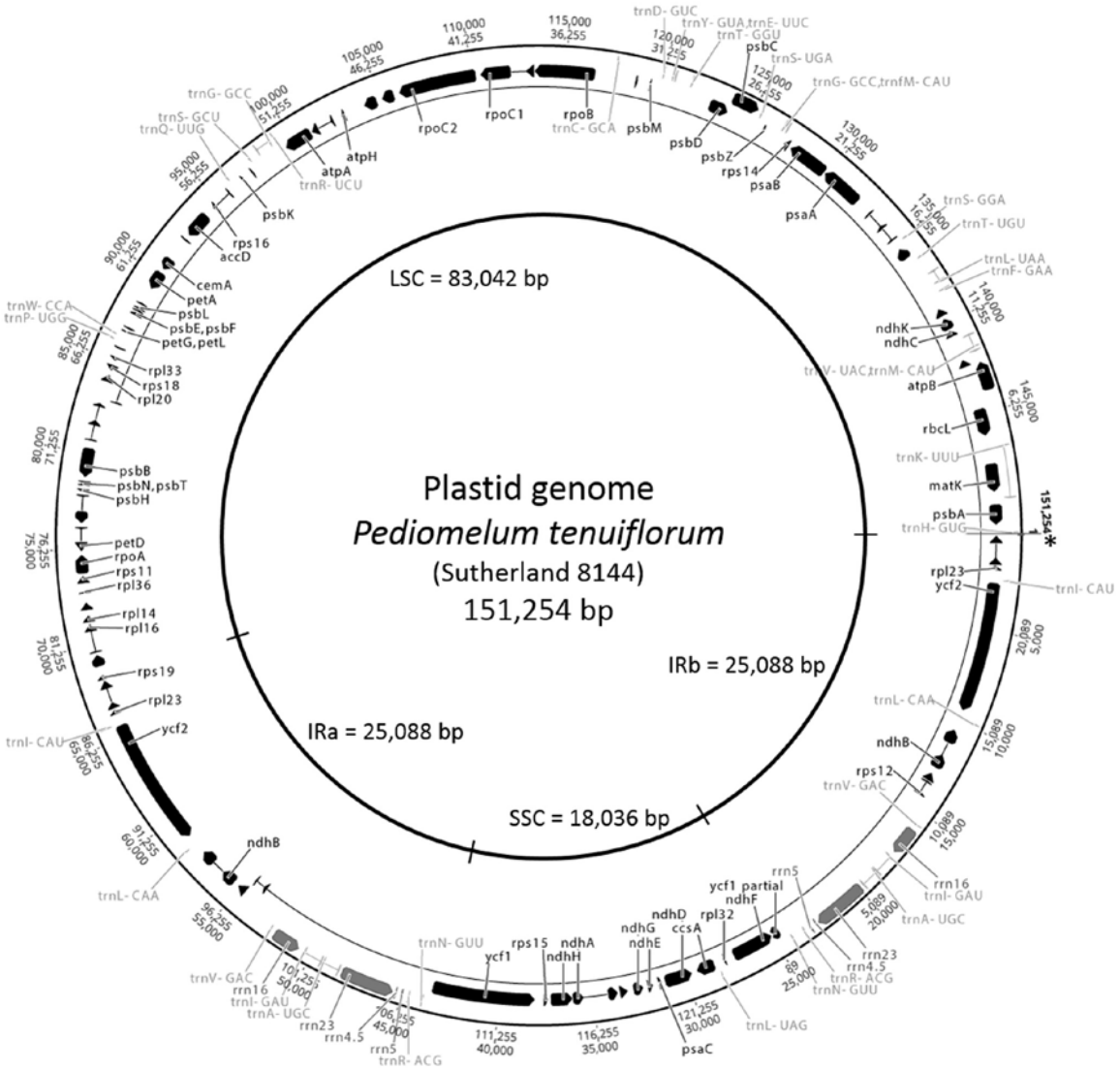

Figure 4. Complete plastid genome of *Pediomelum tenuiflorum*. The outer solid black circle represents the nucleotide sequence and is numbered beginning with *trnH*-GUG (marked with "*" at the 3-o'clock position). Coding regions are annotated with wide black arrows in the direction of transcription. Exon annotations are connected with lines. tRNAs and rRNAs are labeled and marked with gray arrows. The inner solid black circle marks boundaries of the quadripartite regions, labeled: LSC = Large single copy region; SSC = Small single copy region; IR = inverted repeat; bp = nucleotide base pairs. (Figure exported from Geneious, version 10.0.9, Biomatters [www.geneious.com] and modified in Microsoft Power Point.)

Table 3. Results of Illumina sequencing complete plastid genomes, nrDNA (ITS1, 5.8S, and ITS2), and multiple coding and noncoding regions of the mitochondrial genome of seven *Pediomelum* samples and one *Ladeania* species. "×" is the average number of raw reads per nucleotide position.

			Illumina sequence coverage (×)		
Species	*Collector and collection no.*	*No. of single-pass reads*	*Complete plastid genome*	*nrDNA*	*mtDNA*
Ladeania lanceolata	Ahrendsen 24	6,767,556	462.0	516.9	27.5
Pediomelum argophylum	Jones 1027	23,374,762	840.1	1659.3	71.9
Pediomelum digitatum	Jones 1022	15,953,739	997.0	1587.8	65.0
Pediomelum digitatum	Jones 1029	8,748,759	332.3	823.1	23.3
Pediomelum esculentum	Ahrendsen 14	28,478,757	964.7	7540.3	82.0
Pediomelum tenuiflorum	Jones 1026	21,127,997	781.7	1406.1	84.5
Pediomelum tenuiflorum	Sutherland 8144	27,004,367	1650.6	1689.2	98.3
Pediomelum tenuiflorum	Sutherland 8317	22,387,778	773.8	2179.1	115.3

Molecular characters

Illumina sequence recovery and assembly

Complete plastid genomes (e.g., Fig. 4), nrDNA, and multiple coding and noncoding regions of the mitochondrial genome (listed in Appendix 2) were recovered from Illumina sequencing of seven *Pediomelum* and one outgroup (*Ladeania*) species (Table 3). Illumina files contained 6.8 million to over 28 million short-read sequences. Referenced-based assembly resulted in the following datasets (average depths of nucleotide coverage): complete plastid genomes (332.3–1650.6×), nrDNA (516.9–7540.3×), and mtDNA (23.3–115.3×). All coverage depths far exceeded those estimated by Straub *et al*. (2012) to result in high confidence consensus sequences, except coverage of mtDNA for *L. lanceolata* (27.5×) and *P. digitatum* (Jones 1029; 23.3×), though these are not alarmingly low. The portions of coding, noncoding, and RNA sequences varied among the genomic compartments (Table 4). Overall, noncoding regions made up nearly 57% of the sequences that were recovered and compared. All new sequences that were recovered from Illumina sequencing were uploaded to Genbank (Accession numbers in Appendix 2).

Alignment and Phylogenetic analysis of the 32-taxon dataset

For the 32-taxon dataset, the five plastid markers – *matK*, *rpoB-trnC*, *trnD-trnY*, *trnL-F*, and *trnS-trnG* – and nrDNA ITS1 & ITS2 that were downloaded from Genbank for 25 *Pediomelum* species and outgroups (Appendix 1) were concatenated and aligned with the same DNA regions extracted from Illumina sequence files for the samples listed in Table 3. *Pediomelum digitatum* (Jones

Table 4. Quantities and portions of coding, noncoding, and RNA sequences that were recovered from Illumina sequencing for the various genomic compartments and were concatenated into alignments for phylogenetic analyses.

	Genomic compartment			
	Plastid	*nrDNA*	*mtDNA*	*TOTAL*
Average length (bp)	126,122	766	161,981	290,144
Genes (no. / % of total)	76 / 53.2%	--	32 / 27.3%	108 / 38.5%
rRNA (no. / % of total)	4 / 3.6%	1 / 43%	3 / 3.3%	7 / 3.5%
tRNA (no. / % of total)	30 / 1.8%	--	15 / 0.7%	45 / 1.2%
noncoding (% of total)	41.4%	57.0%	68.7%	56.8%
Alignment length (bp)	127,037	775	162,332	290,144
Pairwise identity	99.0%	98.0%	99.6%	99.3%
Parsimony-informative characters (bp)	385	4	296	685

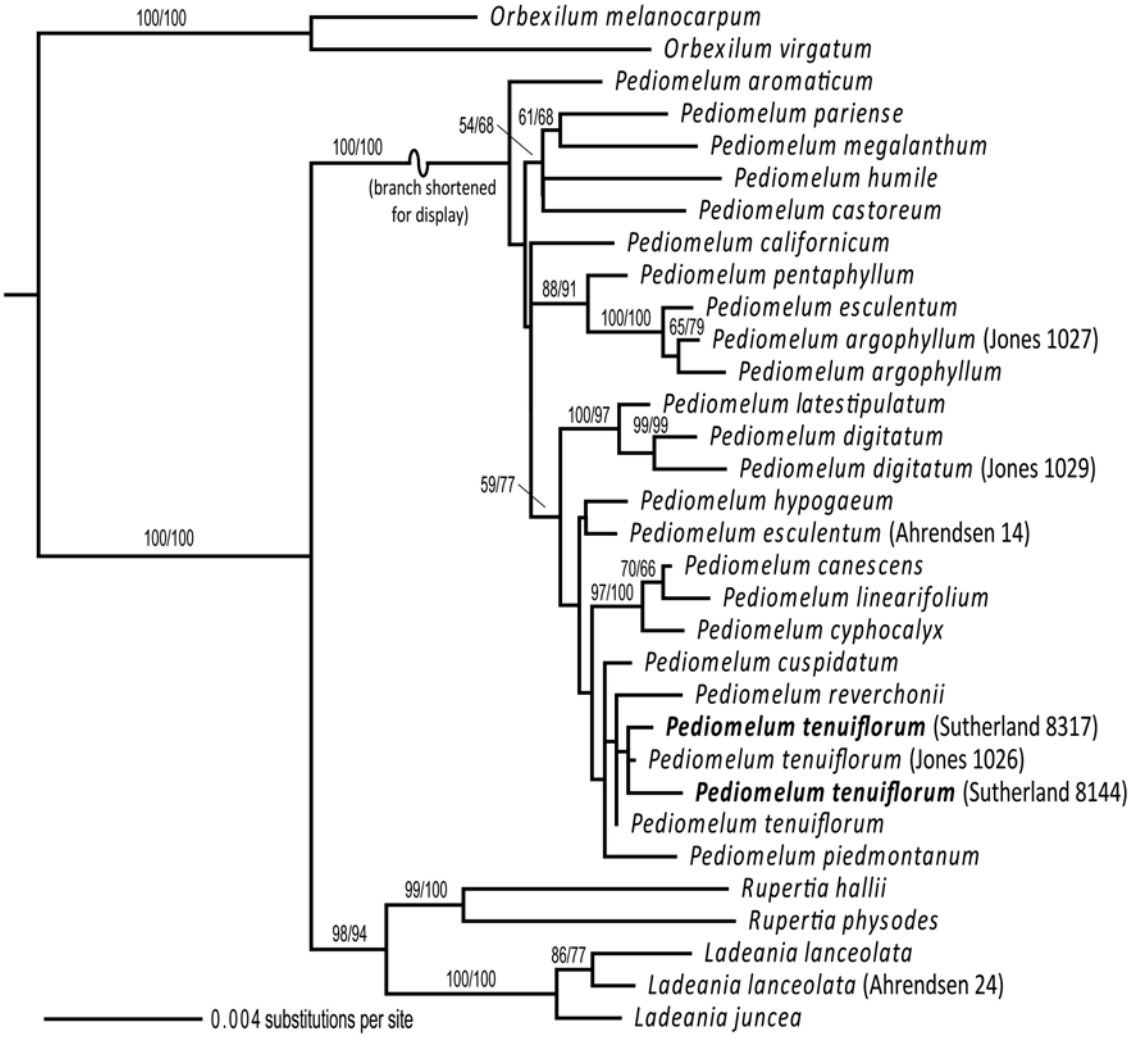

Figure 5. 32-taxon, six-marker phylogeny of *Pediomelum*. Maximum likelihood (ML) phylogeny (–ln L = 15470.35) for 24 samples (19 species) of *Pediomelum*, and seven outgroups, estimated from five plastid regions (*matK*, *rpoB-trnC*, *trnD-trnY*, *trnL-F*, and *trnS-trnG*) and nrDNA ITS1 & ITS2. Phylogeny is congruent with one of 29 most parsimonious trees. Sequences from Kellar-lab samples are noted with collection numbers, otherwise, sequence data were downloaded from Genbank (Appendix 1). Names in bold text are the samples from western/central (Sutherland 8317) and southeastern (Sutherland 8144) Nebraska. Numbers above branches indicate MP/ML bootstrap support values resulting from 1000 replicates.

1022) was excluded because the dataset included two other *P. digitatum*. The alignment is available from the Dryad Digital Repository: (doi:10.5061/dryad.ns8kp8j ; https://doi.org/10.5061/dryad.ns8kp8j).

The AICc algorithm selected the TIM+Γ model of evolution; however, we used GTR+CAT because it is more rigorous, and CAT improves the analysis (Geneious version 10.0.9; Biomatters [www.geneious.com]). The data were partitioned by gene (Appendix 3). The 32-taxon phylogeny (Fig. 5; Kellar-lab samples identified with collection numbers) had moderate (> 75) to high (> 90) MP and ML bootstrap scores (BS) for some clades; however, branch support among *P. tenuiflorum* samples was < 50 and collapsed in strict consensus trees. Each Kellar-lab sample is sister to the downloaded sample of the same species, except *P. esculentum*. The new Kellar-lab sample of *P. esculentum* (Ahrendsen 14) is sister to *P. hypogaeum* (Egan & Egan 209) in Figure 5 but with BS < 50. The ML tree was congruent with one of 29 most parsimonious trees (see Table 5 for tree statistics).

Table 5. Tree statistics for the two datasets: 1) 32-taxon, six regions and 2) 8-taxon, 3 genomic regions. bp = base pairs; p.i.c. = parsimony-informative characters; CI = consistency index (excluding uninformative characters); RI = retention index.

Dataset	*32-taxon, 6 regions*	*8-taxon, 3 genomes*
Alignment length (bp)	7,575	290,144
No. of p.i.c.s	279	685
No. of equally parsimonious trees	29	1
tree length	766	3331
CI	0.6903	0.6841
RI	0.8535	0.6039
–ln L	15470.35	424047.82
Figure no.	5	6

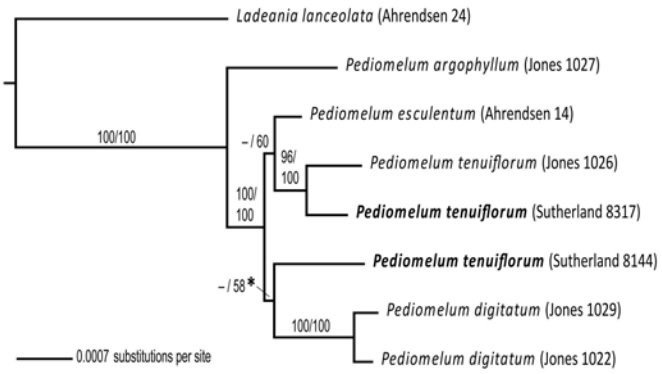

Figure 6. Eight-taxon, three-genomic region phylogeny of *Pediomelum*. Maximum likelihood (ML) phylogeny (–ln L = 424047.82) for seven *Pediomelum* samples (four species) and the outgroup *Ladeania lanceolata*, estimated from complete plastid genomes, nrDNA regions, and multiple mtDNA regions (described in Table 4). Congruent with the single most parsimonious tree, except *P. esculentum* grouped with the clade marked with "*" with BS = 64. Names in bold text are the samples from western/central (Sutherland 8317) and southeastern (Sutherland 8144) Nebraska. Numbers above branches indicate MP/ML bootstrap (BS) support values resulting from 1000 replicates. "–" indicates that this relationship did not occur in the MP tree.

Alignment and Phylogenetic analysis of eight-taxon dataset

Complete plastid genomes, nrDNA regions, and mtDNA regions for eight samples were aligned separately, and phylogenies were inferred for each organelle. Then, the three alignments were concatenated into an alignment for a total molecular evidence phylogenetic analysis. Pairwise identity in this dataset was very high: cpDNA = 99.0%, nrDNA = 98.0%, mtDNA = 99.6%. In the nrDNA alignment, there was only one nucleotide character that differed among the *Pediomelum tenuiflorum* samples. The alignments are available from the Dryad Digital Repository: (doi:10.5061/dryad.ns8kp8j ; https://doi.org/10.5061/dryad.ns8kp8j). In the chloroplast genome alignment, parsimony-informative characters (p.i.c.s) were scarce and were not consistent in informing sister relationships; however, seven noncoding regions were identified as having the highest concentrations of p.i.c.s as follows (noncoding region, pairwise identity, length [bp]): *ycf3*–*psaA*, 97.9%, 825 bp; *trnG*-GCC–*psbZ*, 94.9%, 734 bp; *psbD*–*trnT*, 98.8%, 1172 bp; *trnD*-GUC–*petN* (incl. *psbM*), 99.4%, 1187 bp; *petN*–*trnC*-GCA, 98.3%, 1071 bp; *atpI*–*atpH*, 98.0%, 1083 bp; *rpl20*–*rps12* (incl rpl20), 99.6%, 1227 bp. Parsimony-informative characters in the mitochondrial sequence alignment were not concentrated in any identifiable regions as they were in the chloroplast genome.

The AICc algorithm selected various models of evolution for the cpDNA (TVM+I+Γ), nrDNA (TRN+G), and mtDNA (TIM+I+Γ) datasets. We chose to use the GTR model of evolution because it is more rigorous, and added +CAT for improved performance. The data were partitioned by gene (Appendix 4), and the phylogeny resulting from the ML analysis (Figure 6; Table 5) is congruent with the single most parsimonious tree with one exception. *Pediomelum tenuiflorum* (Sutherland 8144) from eastern Nebraska grouped with the two *P. digitatum* samples (BS = 58). MP BS also supports this grouping but with *P. esculentum* sister to the two *P. digitatum* samples (BS = 58) and *P. tenuiflorum* (Sutherland 8144) sister to that clade (BS = 64).

We also inferred phylogenies from individual organellar datasets (Fig. 7 from cpDNA and mtDNA; Figure A1 in Appendix from nrDNA). The nrDNA phylogeny (Fig. A1) has very little resolution, but agrees with the cpDNA tree in the placement of *Pediomelum esculentum* sister to *P. digitatum* (BS = 64/63). The three *P. tenuiflorum* samples are in a clade (BS = 100/100). The ML analysis estimated from the mitochondrial (mtDNA) alignment alone (Fig. 7 (A)) strongly supports a clade containing the eastern (Sutherland 8144) and one of the western (Jones 1026) samples of *P. tenuiflorum* (MP/ML BS = 100/100). In the MP analysis, this clade is sister to the two samples of *P. digitatum* (BS = 92), with this clade sister to a clade containing the other western *P. tenuiflorum* sample (Sutherland 8317) and *P. esculentum* (BS = 94). However, in the ML analysis, the clade containing the eastern (Sutherland 8144) and one of the

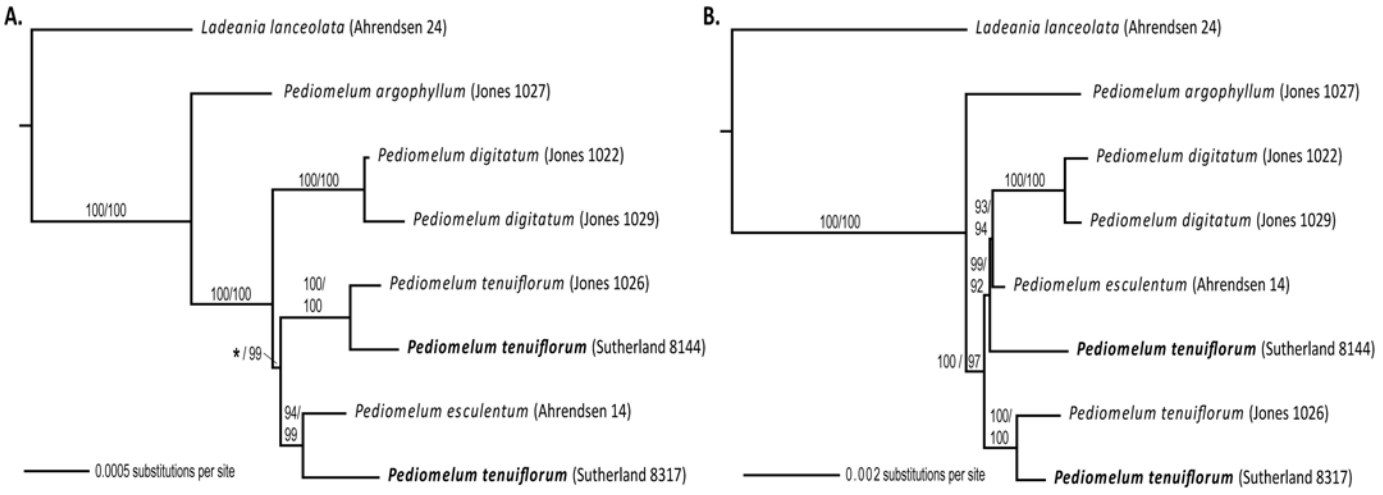

Figure 7. Comparative chloroplast and mitochondrial *Pediomelum* phylogenies. Maximum likelihood phylogenies for seven *Pediomelum* samples (four species) and the outgroup *Ladeania lanceolata*, estimated from (A) chloroplast genomes and (B) mitochondrial genes and noncoding regions. Name in bold text is the sample of *P. tenuiflorum* from southeastern NE. Branch marked with an "*" does not occur in the most parsimonious tree. Numbers above branches indicate MP/ML bootstrap (BS) support values resulting from 1000 replicates.

western (Jones 1026) samples of *P. tenuiflorum* is sister to a clade containing the other western *P. tenuiflorum* sample (Sutherland 8317) and *P. esculentum* (BS = 99), with this clade sister to the two samples of *P. digitatum* (BS = 100). In contrast, ML and MP analyses estimated from the chloroplast sequences (cpDNA) alone (Fig. 7 (B)) strongly supported the eastern *P. tenuiflorum* sample (Sutherland 8144) sister to a clade containing *P. esculentum* and the two *P. digitatum* samples (MP/ML BS = 99/92), with that clade sister to a clade containing the two western *P. tenuiflorum* samples (Sutherland 8317 and Jones 1026) (BS = 100/97).

Discussion

Our investigation into the potential split of *Pediomelum tenuiflorum* morphotypes into two species revealed significant differences in morphological characters (Fig. 3, Table 2). Individuals in the southeastern Nebraska populations are more robust with larger leaflets, larger flowers, and greater numbers of flowers per inflorescence, hence the original specific epithet *floribunda* assigned by Nuttall (Torrey and Gray 1838). The means of these measurements were significantly different, but the ranges of most measurements overlapped (Table 2). The vestiture of above-ground vegetative characters is not consistent with Nuttall's observations. Individuals from both populations are predominantly canescent, with only the abaxial leaflet surfaces primarily glabrous. Additionally, the leaflets of both morphotypes are gland-dotted. The number of flowers per node, discussed by Rydberg (1919), was also not consistently distinct among the two populations. Although not all morphological characters were significantly different, morphotypes from distinct populations seem clear from field observations (Fig. 2), and results of the discriminant analysis indicated that there is an 83.7% chance that a new plant collection would be placed into the correct morphotype group (West or East) based on these characters.

Molecular variation among *Pediomelum* samples is low, as is often seen with closely related species, making phylogenetic distinction difficult with only a few genomic markers. Even with 279 parsimony-informative characters in the 32-taxon phylogeny, many relationships are unsupported (BS < 50; Fig. 5). One notable placement is that of our sample of *P. esculentum* (Ahrendsen 14), which was separated by at least two moderately to strongly supported branches from the *P. esculentum* (Egan & Egan 216; *P. esculentum* 59) in Egan and Crandall (2008; Fig. 3) that we downloaded from Genbank. The two *P. esculentum* samples in Egan and Crandall (2008) were separated in their tree (Fig. 3) as well, suggesting that further study of this species is needed.

Massively-parallel sequencing identified additional genetic variation among species. Although the pairwise identity of sequences in the eight-taxon alignment was 99.3% (meaning differences totaled only 0.7%), with an overall length of 290,144 bp, this resulted in over 2,000 varying nucleotide characters, 685 bp of which were parsimony-informative (Tables 4 and 5). It is well-known that, in plants, nrDNA evolves at a faster rate than plastid genomes, and mitochondrial genomes evolve very slowly. Therefore, it is notable that there is only one nucleotide difference in the 434 bp ITS1+5.8S+ITS2 alignment among the three *Pediomelum tenuiflorum* samples – one character that differs between the west and east Nebraska samples.

The goal of combining genetic sequences from three cellular compartments for phylogenetics is to obtain a total molecular evidence tree, which may reveal the evolutionary history of a taxon or group of taxa. This is in contrast to a phylogeny inferred from just one gene that may only reveal the evolutionary history of that gene or a phylogeny inferred from a genome (e.g., chloroplast or mitochondrial) that may only reveal the evolutionary history of that genome. Concatenating sequences that may have differing histories can reduce some of the "noise" between genomes, but it can also mask past hybridization events or gene-transfer. Our analysis based on concatenated sequences from multiple genomic lineages resulted in a phylogeny with a wide range of branch support values (Fig. 6). *Pediomelum tenuiflorum* from southeastern Nebraska (Sutherland 8144) is clearly separated from the two *P. tenuiflorum* samples from western/central Nebraska (Sutherland 8317 and Jones 1026). However, the ML bootstrap support values grouping these samples with nearest relatives are rather low (BS = 60 and BS = 58). The phylogeny estimated from nrDNA (ITS1, 5.8S, and ITS2) is not very informative (Figure A1 in Appendix). This is not surprising given the pairwise identity for the nrDNA alignment was 98%.

To investigate this further, we estimated one phylogeny using only the chloroplast sequences and another using only the mitochondrial sequences (Fig. 7). These phylogenies were not congruent, suggesting different evolutionary histories for these two organelles. This lack of congruence may explain the low support for phylogenetic placement of the three *P. tenuiflorum* samples in the tree estimated from all three genomic compartments. Concatenating the alignments from three different organellar genomes may be confounding the question of evolutionary relationships of the specimens, begging further investigation into the evolutionary histories of the chloroplast and mitochondrial organelles.

The chloroplast genomes (e.g., Fig. 4) for the seven *Pediomelum* samples plus *Ladeania lanceolata* have gene orders similar to the gene order found in chickpea (*Cicer arietinum* [Fabaceae]; Jansen *et al*. 2008), including the large (ca. 50 kilo-bp) inversion, reversing the gene order between *rbcL* and *rps16*, and the internal stop codon in *ndhB*. In contrast to the *Cicer arietinum* plastid genome, however, our samples included both copies of the inverted repeat (IR). Additionally, the plastid genomes sequenced in our study included *rps16* exon 1 and exon 2 between *trnQ*-UUG and *accD*, and our genomes included three exons of *clpP*. The *Pediomelum* and *Ladeania* genomes included a total of 109 genes, grouped as follows: 76 protein-coding genes, four rRNA genes, and 29 tRNA genes.

Our work contributes both morphological and molecular data to taxonomic and phylogenetic questions within the genus *Pediomelum*. Sequencing of eight whole chloroplast genomes revealed the seven noncoding regions containing the most concentrated parsimony-informative characters (*ycf3–psaA*, *trnG–psbZ*, *psbD–trnT*, *trnD–petN* (incl. *psbM*), *petN–trnC, atpI–atpH*, *rpl20–rps12*). Only two of these regions (*psbD–trnT* and *atpI–atpH*) match the most variable noncoding regions found by Shaw et al. (2007) in the well-known *Tortoise and the Hare III* investigation of the best noncoding regions for phylogenetic analysis. The seven markers we identified may be used in focused Sanger sequencing to elucidate evolutionary relationships across *Pediomelum*. Primers can be designed using the plastid genome sequences as well as the mitochondrial genetic sequences. Additionally, consensus sequences from Illumina sequencing may be used as reference sequences for further reference-based assemblies from additional massively parallel sequencing.

According to the United States Department of Agriculture, Natural Resources Conservation Service (www.plants.usda.gov), the distribution of *Pediomelum tenuiflorum* ranges from Arizona to western Illinois and from central Texas north through Wyoming and into Montana. Although the morphotypes of *P. tenuiflorum* do not overlap in Nebraska, there may be regions of overlap elsewhere. Expanded sampling and additional sequencing will provide the opportunity to delineate populations of varying morphotypes, identify regions of overlap and/or hybridization, and explore the evolutionary histories of the mitochondrial and chloroplast organelles. Morphological measurements of specimens from an expanded distribution and inclusion of more samples that include fruits

may also contribute to a better understanding of variation among the populations of *P. tenuiflorum*.

In conclusion, some of the characters studied here may blur the lines between the western and eastern Nebraska populations of *Pediomelum tenuiflorum*, such as the contradictory phylogenies estimated from the plastid and mitochondrial datasets or the overlap in measurements of morphological characters. The conflicting results may indicate that *P. floribundum* does not exist as a species that should be separated from *P. tenuiflorum*; i.e., the two populations interbreed in unidentified locations, and the differing morphologies result from varying environmental factors. Conversely, these incongruous data may indicate an unclarified and complicated history of gene/genome evolution that can only be elucidated with both morphological and molecular investigations of many more samples from populations across the distribution of *P. tenuiflorum*. Despite the ultimate fate of *P. tenuiflorum* and/or *P. floribundum* (i.e., to split or not to split), our study provides a solid foundation of results and sequences upon which an expanded investigation can build.

Acknowledgments — The authors thank J.L. Peters and S. Jaksich for reviews of early versions of the manuscript, Dr. C. Rauter for help with statistics, two anonymous reviewers for many helpful suggestions, and D. Sutherland (UNO) for review comments, consistent scholarly support, help with collections, and taxonomic assistance. We also thank the MU Core Sequencing facility and the following granting institutions: Nebraska Native Plant Society and UNO Fund for Undergraduate Scholarly Experiences (FUSE).

Web Resources

Dryad Digital Repository
https://datadryad.org/

GenBank, NCBI (National Center for Biotechnology Information) https://www.ncbi.nlm.nih.gov/genbank/

Minitab 18 (Minitab, LLC 2019) www.minitab.com

United States Department of Agriculture, Natural Resources Conservation Service, http://plants.usda.gov

Literature Cited

Ahrendsen DL, Aust SK, and Kellar PR. (2016) Biodiversity assessment using next-generation sequencing: comparison of phylogenetic and functional diversity between Nebraska grasslands. *Plant Systematics and Evolution*, 302: 89–108.

Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, Pruitt KD, and Sayers EW. (2018) GenBank. *Nucleic Acids Research,* 46: D41–D47.

Chapman SS, Omernik JM, Freeouf JA, Huggins DG, McCauley JR, Freeman CC, Steinauer G, Angelo RT, and Schlepp RL (2001) *Ecoregion Download Files by State - Region 7* Retrieved from https://www.epa.gov/eco-research/ecoregion-download-files-state-region-7

Egan AN and Crandall KA. (2008) Incorporating gaps as phylogenetic characters across eight DNA regions: Ramifications for North American Psoraleeae (Leguminosae). *Molecular Phylogenetics and Evolution,* 46: 532–546.

Egan AN and Reveal JL. (2009) A New Combination in *Pediomelum* and a New Genus, *Ladeania*, from Western North America (Fabaceae, Psoraleeae) *Novon*, 19(3): 310–314.

Felsenstein J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39: 783–791.

Great Plains Flora Association, Barker WT, and Barkley TM. (1977) *Atlas of the Flora of the Great Plains* (Iowa State University Press, Aimes, IA, USA).

Great Plains Flora Association. (1986) *Flora of the Great Plains* (University Press of Kansas: Lawrence, KS, USA).

Grimes JW. (1990) A revision of the New World species of Psoraleeae (Leguminosae: Papilionoideae). *Memoirs of The New York Botanical Garden*. 61: 1–113.

Jansen RK, Wojciechowski MF, Sanniyasi E, Lee SB and Daniell H. (2008) Complete plastid genome sequence of the chickpea (*Cicer arietinum*) and the phylogenetic distribution of *rps12* and *clpP* intron losses among legumes (Leguminosae) *Molecular Phylogenetics and Evolution,* 48(3): 1204–1217.

Katoh K, Misawa K, Kuma K, and Miyata T. (2002) MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*. 30: 3059–3066.

Katoh K and Standley DM. (2013) MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Molecular Biology and Evolution* 30(4): 772–780.

Kaul RB, Sutherland DM, and Rolfsmeier SB. (2012) *The Flora of Nebraska*, second edition (Lincoln, Nebraska, USA: School of Natural Resources, University of Nebraska–Lincoln).

Kellar PR (Steele), Ahrendsen DL, Aust SK, Jones AR, and Pires JC. (2015) Biodiversity comparison among phylogenetic diversity metrics and between three North American prairies. *Applications in Plant Sciences* 3(7): 1400108.

Lewis M and Clark W. University of Nebraska–Lincoln. Center for Great Plains Studies, and American Philosophical Society. (1983) *The journals of the Lewis and*

Clark Expedition (G. Moulton & T. Dunlay, Eds.) (University of Nebraska Press: Lincoln, NE, USA)

Minitab 18 (Minitab, LLC 2019) www.minitab.com

Pursh FT. (1814) *Flora Americae Septentrionalis*, v.2. p. 475. (White, Cochran, and Co., London).

Rydberg PA and The Botanical Seminar, University of Nebraska. (1895) Flora of Nebraska: Part 21. Rosales. *Faculty Publications in the Biological Sciences*, 35.

Rydberg PA. (1919) Fabaceae, tribe Psoraleeae. Part 1. *North American Flora*, 24: 1–64.

Sanger F and Coulson AR. (1975) A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94(3): 441–446.

Shaw J, Lickey EB, Schilling EE, and Small RL. (2007). Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. *American Journal of Botany* 94: 275–288.

Stamatakis A. (2014) RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* 30(9):1312-3. https://doi.org/10.1093/bioinformatics/btu033

Steele PR, Gilbert LE, Friar L, and Jansen RK. (2010) Molecular systematics of the Neotropical genus *Psiguria* (Cucurbitaceae). *American Journal of Botany*, 97(1): 156–173.

Steele PR and Pires JC. (2011) Biodiversity assessment: state-of-the-art techniques in phylogenomics and species identification. *American Journal of Botany*, 98(3): 415–425.

Steele PR, Hertweck KL, Mayfield D, McKain MR, Leebens-Mack J, and Pires JC. (2012) Quality and quantity of data recovered from massively parallel sequencing: Examples in Asparagales and Poaceae. *American Journal of Botany*, 99(2): 330–348.

Straub SCK, Parks M, Weitemier K, Fishbein M, Cronn RC, Liston A. (2012) Navigating the tip of the genomic iceberg: next-generation sequencing for plant systematics. *American Journal of Botany*, 99(2): 349–364.

Swofford DL. (2003) PAUP*: Phylogenetic analysis using parsimony (* and other methods), version 4.0b10. (Sinauer, Sunderland, Massachusetts, USA).

Torrey J and Gray A. (1838) *A flora of North America: containing abridged descriptions of all the known indigenous and naturalized plants growing north of Mexico, arranged according to the natural system*, v.1. p. 300. (Wiley & Putnam: New York, USA).

USDA, NRCS. (2019) The PLANTS Database (http://plants.usda.gov, 6 March 2019). National Plant Data Team, Greensboro, NC 27401-4901 USA.

Vail AM. (1894) A Study of the Genus Psoralea in America. *Bulletin of the Torrey Botanical Club* 21(3): 91–119.

Ward DE. (1983) Chromosome Counts from New Mexico and Southern Colorado. *Phytologia*, 54(5): 302–308.

worldatlas (n.d.) *Nebraska Precipitation Map* Retrieved from https://www.worldatlas.com/webimage/countrys/namerica/usstates/weathermaps/neprecip.htm

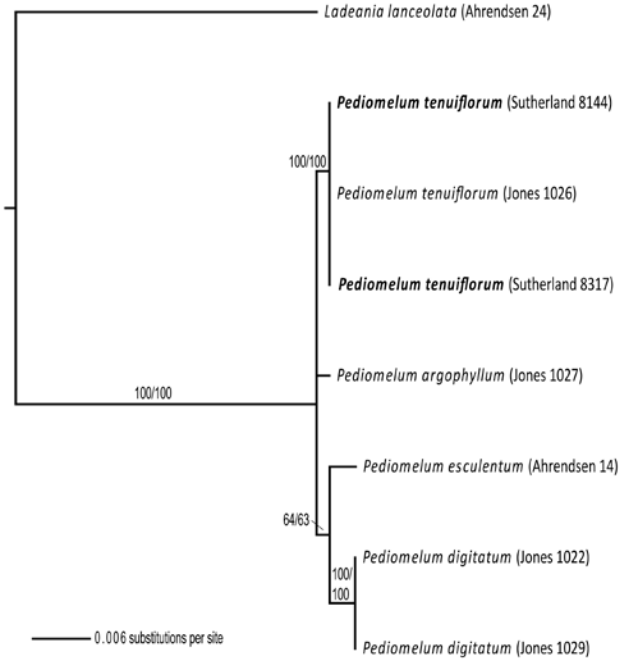

Figure A1. Eight-taxon, nrDNA phylogeny of *Pediomelum*. Maximum likelihood (ML) phylogeny (–ln L = 1294.66) for seven *Pediomelum* samples (four species) and the outgroup *Ladeania lanceolata*, estimated from ITS1, 5.8S, and ITS2. Phylogeny is congruent with the single most parsimonious tree. Names in bold text are the samples from western/central (Sutherland 8317) and southeastern (Sutherland 8144) Nebraska. Numbers above branches indicate MP/ML bootstrap (BS) support values resulting from 1000 replicates.

Appendix 1. Accession numbers for nrDNA (ITS1, 5.8S, and ITS2) plus five chloroplast regions for 25 species that were downloaded from GenBank.

Species	GenBank Accession Number					
	nrDNA	*matK*	*rpoB-trnC*	*trnD-trnY*	*trnL-F*	*trnS-trnG*
Ladeania juncea	EF517889	EF549986	EF549802	EF543489	EF543399	EF549899
Ladeania lanceolata	EF517891	EF549941	EF549804	EF543444	EF543401	EF549906
Orbexilum melanocarpum	EF517849	EF549948	EF549763	EF543452	EF543344	EF549859
Orbexilum virgatum	EF517834	EF549937	EF549750	EF543456	EF543349	EF549844
Pediomelum argophyllum	EF517892	EF549988	EF549805	EF543492	EF543377	EF549901
Pediomelum aromaticum	EF517875	EF549972	EF549788	EF543476	EF543385	EF549885
Pediomelum californicum	EF517876	EF549973	EF549789	EF543477	EF543386	EF549886
Pediomelum canescens	EF517900	EF549996	EF549813	EF543500	EF543410	EF549909
Pediomelum castoreum	EF517877	EF549974	EF549790	EF543478	EF543387	EF549887
Pediomelum cuspidatum	EF517874	EF549971	EF549787	EF543475	EF543384	EF549884
Pediomelum cyphocalyx	EF517842	EF549943	EF549757	EF543446	EF543381	EF549852
Pediomelum digitatum	EF517878	EF549975	EF549791	EF543479	EF543388	EF549888
Pediomelum esculentum	EF517879	EF549976	EF549792	EF543480	EF543389	EF549889
Pediomelum humile	EF517907	EF550000	EF549819	EF543507	EF543417	EF549916
Pediomelum hypogaeum	EF517880	EF549977	EF549793	EF543481	EF543390	EF549890
Pediomelum latestipulatum	EF517881	EF549978	EF549794	EF543482	EF543391	EF549891
Pediomelum linearifolium	EF517873	EF549970	EF549786	EF543474	EF543383	EF549883
Pediomelum megalanthum	EF517844	EF549945	EF549758	EF543447	EF543357	EF549854
Pediomelum pariense	EF517885	EF549982	EF549808	EF543486	EF543395	EF549895
Pediomelum pentaphyllum	EF517896	EF549983	EF549809	EF543487	EF543406	EF549896
Pediomelum piedmontanum	EF517901	EF549997	EF549814	EF543501	EF543411	EF549910
Pediomelum reverchonii	EF517887	EF549984	EF549800	EF543488	EF543397	EF549897
Pediomelum tenuiflorum	EF517839	EF549940	EF549755	EF543443	EF543354	EF549849
Rupertia hallii	EF517856	EF549953	EF549770	EF543458	EF543366	EF549866
Rupertia physodes	EF517863	EF549960	EF549777	EF543465	EF543373	EF549873

Appendix 2. Genbank accession numbers for all new sequences recovered from Illumina sequencing of seven *Pediomelum* samples and one species in *Ladeania*.

Species (voucher)	*Pediomelum digitatum* (Jones 1022)	*Pediomelum tenuiflorum* (Jones 1026)	*Pediomelum argophyllum* (Jones 1027)	*Pediomelum digitatum* (Jones 1029)	*Pediomelum tenuiflorum* (Sutherland 8317)	*Ladeania lanceolata* (Ahrendsen 24)	*Pediomelum esculentum* (Ahrendsen 14)	*Pediomelum tenuiflorum* (Sutherland 8144)
Chloroplast genome	MN115424	MN115425	MN115426	MN115427	MN115428	MN115429	MN115430	MN115431
ITS 1 + 5.8S + ITS 2	MN058285	MN058281	MN058286	MN058287	MN058282	MN058288	MN058283	MN058284
ccmB, nad1, nad6	MN066647	MN066648	MN066649	MN066650	MN066651	MN066652	MN066653	MN066654
trnF, trnP	MN058441	MN058442	MN058443	MN058444	MN058445	MN058446	MN058447	MN058448
trnP-trnM IGS	MN058473	MN058474	MN058475	MN058476	MN058477	MN058478	MN058479	MN058480
trnP-trnM IGS2	MN058481	MN058482	MN058483	MN058484	MN058485	MN058486	MN058487	MN058488
trnP-trnM IGS3	MN058489	MN058490	MN058491	MN058492	MN058493	MN058494	MN058495	MN058496
matR, nad1	MN066655	MN066656	MN066657	MN066658	MN066659	MN066660	MN066661	MN066662
nad4 (4 exons)	MN066679	MN066680	MN066681	MN066682	MN066683	MN066684	MN066685	MN066686
nad2 (3 exons)	MN066671	MN066672	MN066673	MN066674	MN066675	MN066676	MN066677	MN066678
atp9-1	MN058305	MN058306	MN058307	MN058308	MN058309	MN058310	MN058311	MN058312
nad5 (1 exon)	MN066687	MN066688	MN066689	MN066690	MN066691	MN066692	MN066693	MN066694
nad5 (2 exons)	MN066695	MN066696	MN066697	MN066698	MN066699	MN066700	MN066701	MN066702
nad5, rps1	MN066703	MN066704	MN066705	MN066706	MN066707	MN066708	MN066709	MN066710
rrn5, rrnS, trnW-cp	MN058401	MN058402	MN058403	MN058404	MN058405	MN058406	MN058407	MN058408
trnY, trnN-cp	MN058505	MN058506	MN058507	MN058508	MN058509	MN058510	MN058511	MN058512
nad2 (2exons)	MN066663	MN066664	MN066665	MN066666	MN066667	MN066668	MN066669	MN066670
nad2-atp8 IGS	MN058345	MN058346	MN058347	MN058348	MN058349	MN058350	MN058351	MN058352
atp8, nad3, rps12	MN058297	MN058298	MN058299	MN058300	MN058301	MN058302	MN058303	MN058304
atp1	MN058289	MN058290	MN058291	MN058292	MN058293	MN058294	MN058295	MN058296
ccmFn	MN058329	MN058330	MN058331	MN058332	MN058333	MN058334	MN058335	MN058336
trnG, trnQ	MN058497	MN058498	MN058499	MN058500	MN058501	MN058502	MN058503	MN058504
nad4L, atp4, rps10, cox1, nad1, rps3, rpl16	MN115416	MN115417	MN115418	MN115419	MN115420	MN115421	MN115422	MN115423
mttB	MN115408	MN115409	MN115410	MN115411	MN115412	MN115413	MN115414	MN115415
rpl16-rrnL IGS	MN058353	MN058354	MN058355	MN058356	MN058357	MN058358	MN058359	MN058360
rpl16-rrnL IGS2	MN058361	MN058362	MN058363	MN058364	MN058365	MN058366	MN058367	MN058368
rrnL, trnfM-1	MN058409	MN058410	MN058411	MN058412	MN058413	MN058414	MN058415	MN058416
rpl5, rps14, cob	MN058385	MN058386	MN058387	MN058388	MN058389	MN058390	MN058391	MN058392
atp6	MN115400	MN115401	MN115402	MN115403	MN115404	MN115405	MN115406	MN115407
nad9	MN058369	MN058370	MN058371	MN058372	MN058373	MN058374	MN058375	MN058376
trnK-TTT	MN058465	MN058466	MN058467	MN058468	MN058469	MN058470	MN058471	MN058472
rps4	MN058393	MN058394	MN058395	MN058396	MN058397	MN058398	MN058399	MN058400
trnC-GCA-1	MN058425	MN058426	MN058427	MN058428	MN058429	MN058430	MN058431	MN058432
trnC-trnH-cp IGS	MN058417	MN058418	MN058419	MN058420	MN058421	MN058422	MN058423	MN058424
trnH-cp	MN058449	MN058450	MN058451	MN058452	MN058453	MN058454	MN058455	MN058456
trnH-cp-atp9 IGS	MN058457	MN058458	MN058459	MN058460	MN058461	MN058462	MN058463	MN058464
apt9-2	MN058313	MN058314	MN058315	MN058316	MN058317	MN058318	MN058319	MN058320
nad7 (5exons)	MN058377	MN058378	MN058379	MN058380	MN058381	MN058382	MN058383	MN058384
ccmFc	MN058321	MN058322	MN058323	MN058324	MN058325	MN058326	MN058327	MN058328
cox3	MN058337	MN058338	MN058339	MN058340	MN058341	MN058342	MN058343	MN058344
trnC-GCA-2	MN058433	MN058434	MN058435	MN058436	MN058437	MN058438	MN058439	MN058440

Appendix 3. Data partitions for the 32-taxon maximum likelihood (ML) analysis.

DNA, 18S = 1–10
DNA, ITS1 = 11–228
DNA, 58S = 229–392
DNA, ITS2 = 393–630
DNA, 26S = 631–644
DNA, trnK1 = 645–1439
DNA, matK = 1440–2969
DNA, trnK2 = 2970–3257
DNA, rpoB = 3258–4406
DNA, trnDY = 4407–5918
DNA, trnLF = 5919–6924
DNA, trnSG = 6925 – 7575

Appendix 4. Data partitions for the eight-taxon, three-genomic region ML analysis.

DNA, trnH = 1-75
DNA, IGS01 = 76-409
DNA, psbA = 410-1471
DNA, IGS02 = 1472-1752
DNA, trnK1 = 1753-1787
DNA, trnKI1 = 1788-2106
DNA, matK = 2107-3636
DNA, trnKI2 = 3637-4391
DNA, trnK2 = 4392-4428
DNA, IGS03 = 4429-5110
DNA, rbcL = 5111-6538
DNA, IGS04 = 6539-7341
DNA, atpB = 7342-8834
DNA, atpE = 8835-9239
DNA, IGS05 = 9240-9376
DNA, trnM = 9377-9450
DNA, IGS06 = 9451-9636
DNA, trnV1 = 9637-9673
DNA, trnVI = 9674-10274
DNA, trnV2 = 10275-10311
DNA, IGS07 = 10312-10574
DNA, ndhC = 10575-10937
DNA, IGS08 = 10938-10980
DNA, ndhK = 10981-11619
DNA, IGS08 = 11620-11734
DNA, ndhJ = 11735-12211
DNA, IGS09 = 12212-12937
DNA, trnF = 12938-13010
DNA, IGS10 = 13011-13434
DNA, trnL1 = 13435-13481
DNA, trnLI = 13482-13995
DNA, trnL2 = 13996-14031
DNA, IGS11 = 14032-15020
DNA, trnT = 15021-15093
DNA, IGS12 = 15094-15379
DNA, rps4 = 15380-15985
DNA, IGS13 = 15986-16202
DNA, trnS = 16203-16290
DNA, IGS14 = 16291-16548
DNA, ycf3E1 = 16549-16674
DNA, ycf3I1 = 16675-17365
DNA, ycf3E2 = 17366-17593
DNA, ycf3I2 = 17594-18374
DNA, ycf3E3 = 18375-18527
DNA, IGS15 = 18528-19353
DNA, psaA = 19354-21606
DNA, IGS16 = 21607-21631
DNA, psaB = 21632-23836
DNA, IGS17 = 23837-23953
DNA, rps14 = 23954-24256
DNA, IGS18 = 24257-24430
DNA, trnfM = 24431-24504
DNA, IGS19 = 24505-24643
DNA, trnG = 24644-24714
DNA, IGS20 = 24715-25460
DNA, psbZ = 25461-25649
DNA, IGS21 = 25650-26014
DNA, trnS = 26015-26107
DNA, IGS22 = 26108-26292
DNA, psbC = 26293-27661
DNA, psbD = 27662-28723
DNA, IGS23 = 28724-29908
DNA, trnT = 29909-29980
DNA, IGS24 = 29981-30800
DNA, trnE = 30801-30873
DNA, IGS25 = 30874-30939
DNA, trnY = 30940-31023
DNA, IGS26 = 31024-31456
DNA, trnD = 31457-31530
DNA, IGS27 = 31531-31964
DNA, psbM = 31965-32069
DNA, IGS28 = 32070-32720
DNA, petN = 32721-32816
DNA, IGS29 = 32817-33903
DNA, trnC = 33904-33974
DNA, IGS30 = 33975-35011
DNA, rpoB = 35012-38224
DNA, IGS31 = 38225-38250
DNA, rpoC1E1 = 38251-38682
DNA, rpoC1I = 38683-39494
DNA, rpoC1E2 = 39495-41117
DNA, IGS32 = 41118-41333
DNA, rpoC2 = 41334-45494
DNA, IGS33 = 45495-45773
DNA, rps2 = 45774-46484
DNA, IGS34 = 46485-46741
DNA, atpI = 46742-47485
DNA, IGS35 = 47486-48603

DNA, atpH = 48604-48849
DNA, IGS36 = 48850-49311
DNA, atpFE1 = 49312-49456
DNA, atpFI = 49457-50193
DNA, atpFE2 = 50194-50606
DNA, IGS37 = 50607-50675
DNA, atpA = 50676-52208
DNA, IGS38 = 52209-52422
DNA, trnR = 52423-52494
DNA, IGS39 = 52495-52748
DNA, trnG1 = 52749-52797
DNA, trnGI = 52798-53495
DNA, trnG2 = 53496-53523
DNA, IGS40 = 53524-54074
DNA, trnS = 54075-54161
DNA, IGS41 = 54162-54328
DNA, psbI = 54329-54439
DNA, IGS42 = 54440-54923
DNA, psbK = 54924-55109
DNA, IGS43 = 55110-55607
DNA, trnQ = 55608-55679
DNA, IGS44 = 55680-56044
DNA, rps16E1 = 56045-56084
DNA, rps16I = 56085-56974
DNA, rps16E2 = 56975-57204
DNA, IGS45 = 57205-57723
DNA, accD = 57724-59133
DNA, IGS46 = 59134-59368
DNA, psaI = 59369-59473
DNA, IGS47 = 59474-60690
DNA, cemA = 60691-61380
DNA, IGS48 = 61381-61571
DNA, petA = 61572-62540
DNA, IGS49 = 62541-63413
DNA, psbJ = 63414-63536
DNA, IGS50 = 63537-63679
DNA, psbL = 63680-63796
DNA, IGS51 = 63797-63818
DNA, psbF = 63819-63938
DNA, IGS52 = 63939-63947
DNA, psbE = 63948-64199
DNA, IGS53 = 64200-64854
DNA, petL = 64855-64950
DNA, IGS54 = 64951-65104
DNA, petG = 65105-65218
DNA, IGS55 = 65219-65370
DNA, trnW = 65371-65444
DNA, IGS56 = 65445-65656
DNA, trnP = 65657-65730
DNA, IGS57 = 65731-66003
DNA, psaJ = 66004-66138
DNA, IGS58 = 66139-66578
DNA, rpl33 = 66579-66779
DNA, IGS59 = 66780-67011
DNA, rps18 = 67012-67324
DNA, IGS60 = 67325-67565
DNA, rpl20 = 67566-67958
DNA, IGS61 = 67959-68792
DNA, rps12E1 = 68793-68906
DNA, IGS62 = 68907-69120
DNA, clpPE3 = 69121-69348
DNA, clpPI2 = 69349-70047
DNA, clpPE2 = 70048-70339
DNA, clpPI1 = 70340-71041
DNA, clpPE1 = 71042-71112
DNA, IGS63 = 71113-71546
DNA, psbB = 71547-73073
DNA, IGS64 = 73074-73249
DNA, psbT = 73250-73357
DNA, IGS65 = 73358-73420
DNA, psbN = 73421-73552
DNA, IGS66 = 73553-73668
DNA, psbH = 73669-73890
DNA, IGS67 = 73891-74023
DNA, petBI = 74024-74864
DNA, petBE2 = 74865-75506
DNA, IGS68 = 75507-75703
DNA, petDI = 75704-76460
DNA, petDE2 = 76461-76935
DNA, IGS69 = 76936-77155
DNA, rpoA = 77156-78160
DNA, IGS70 = 78161-78234
DNA, rps11 = 78235-78651
DNA, IGS71 = 78652-79046
DNA, rpl36 = 79047-79160
DNA, IGS72 = 79161-79651
DNA, rps8 = 79652-80056
DNA, IGS73 = 80057-80326
DNA, rpl14 = 80327-80695
DNA, IGS74 = 80696-80819
DNA, rpl16E2 = 80820-81218
DNA, rpl16I = 81219-82289
DNA, IGS75 = 82290-82454
DNA, rps3 = 82455-83105
DNA, IGS76 = 83106-83533
DNA, rps19 = 83534-83812
DNA, IGS77 = 83813-83866
DNA, rpl2E2 = 83867-84300
DNA, rpl2I = 84301-85022
DNA, rpl2E1 = 85023-85413
DNA, IGS78 = 85414-85431
DNA, rpl23 = 85432-85713
DNA, IGS79 = 85714-86157
DNA, trnI = 86158-86231
DNA, IGS80 = 86232-86318
DNA, ycf2 = 86319-93201
DNA, IGS81 = 93202-93600
DNA, trnL = 93601-93681

DNA, IGS82 = 93682-94294
DNA, ndhBE2 = 94295-95050
DNA, ndhBI = 95051-95741
DNA, ndhBE1 = 95742-96518
DNA, IGS83 = 96519-96823
DNA, rps7 = 96824-97291
DNA, IGS84 = 97292-97345
DNA, rps12E3 =97346-97371
DNA, rps12I = 97372-97903
DNA, rps12E2 = 97904-98135
DNA, IGS85 = 98136-99473
DNA, trnV = 99474-99545
DNA, IGS86 = 99546-99766
DNA, rrn16 = 99767-101257
DNA, IGS87 = 101258-101548
DNA, trnI1 = 101549-101585
DNA, trnII = 101586-102538
DNA, trnI2 = 102539-102573
DNA, IGS88 = 102574-102637
DNA, trnA1 = 102638-102675
DNA, trnAI = 102676-103495
DNA, trnA2 = 103496-103530
DNA, IGS89 = 103531-103688
DNA, rrn23 = 103689-106499
DNA, IGS90 = 106500-106600
DNA, rrn45 = 106601-106704
DNA, IGS91 = 106705-106908
DNA, rrn5 = 106909-107029
DNA, IGS92 = 107030-107286
DNA, trnR = 107287-107361
DNA, IGS93 = 107362-108004
DNA, trnN = 108005-108076
DNA, IGS94 = 108077-108438
DNA, ycf1 = 108439-113853
DNA, IGS95 = 113854-114344
DNA, rps15 = 114345-114617
DNA, IGS96 = 114618-114717
DNA, ndhH = 114718-115900
DNA, ndhAE1 = 115901-116452
DNA, ndhAI = 116453-117788
DNA, ndhAE2 = 117789-118328
DNA, IGS97 = 118329-118403
DNA, ndhI = 118404-118889
DNA, IGS98 = 118890-119151
DNA, ndhG = 119152-119682
DNA, IGS99 = 119683-119911
DNA, ndhE = 119912-120217
DNA, IGS100 = 120218-120510
DNA, psaC = 120511-120756
DNA, IGS101 = 120757-120886
DNA, ndhD = 120887-122383
DNA, IGS102 = 122384-122662
DNA, ccsA = 122663-123640
DNA, IGS103 = 123641-123726
DNA, trnL = 123727-123806
DNA, IGS104 = 123807-124139
DNA, rpl32 = 124140-124301
DNA, IGS105 = 124302-124757
DNA, ndhF = 124758-127037
DNA, 18S = 127038-127120
DNA, ITS1 = 127121-127331
DNA, 58S = 127332-127495
DNA, ITS2 = 127496-127729
DNA, 26S = 127730-127812
DNA, mtNC1 = 127813-129148
DNA, ccmB = 129149-129769
DNA, mtNC2 = 129770-130442
DNA, nad1 = 130443-130829
DNA, mtNC3 = 130830-134191
DNA, nad6 = 134192-134810
DNA, mtNC4 = 134811-135130
DNA, trnFmt = 135131-135204
DNA, mtNC5 = 135205-135457
DNA, trnPmt = 135458-135532
DNA, mtNC6 = 135533-142784
DNA, trnMmt = 142785-142858
DNA, mtNC7 = 142859-143766
DNA, trnEmt = 143767-143838
DNA, mtNC8 = 143839-145949
DNA, nad1b = 145950-149408
DNA, mtNC9 = 149409-152790
DNA, trnImt = 152791-152872
DNA, mtNC10 = 152873-158088
DNA, nad4 = 158089-158549
DNA, nad4I1 = 158550-159970
DNA, nad4b = 159971-160575
DNA, nad4I2 = 160576-163788
DNA, nad4c = 163789-164211
DNA, nad4I3 = 164212-166799
DNA, nad4d = 166800-166888
DNA, mtNC11 = 166889-167067
DNA, nad2 = 167068-167228
DNA, nad2I1 = 167229-169817
DNA, nad2b = 169818-170389
DNA, nad2I2 = 170390-171868
DNA, nad2c = 171869-172057
DNA, mtNC12 = 172058-175178
DNA, atp9-1 = 175179-175403
DNA, mtNC13 =175404-176232
DNA, nad5 = 176233-177448
DNA, mtNC14 = 177449-178307
DNA, nad5b = 178308-178537
DNA, mtNC15 = 178538-178738
DNA, rps1mt = 178739-179359
DNA, rrn5mt = 179360-179475
DNA, mtNC16 = 179476-179641
DNA, rrnS = 179642-181655
DNA, mtNC17 = 181656-183421

DNA, trnWmt = 183422-183495
DNA, mtNC18 = 183496-190718
DNA, trnYmt = 190719-190801
DNA, mtNC19 = 190802-191186
DNA, trnNmt = 191187-191258
DNA, mtNC20 = 191259-192605
DNA, nad2d = 192606-192758
DNA, mtNC21 = 192759-194007
DNA, nad2e = 194008-194399
DNA, mtNC22 = 194400-197810
DNA, atp8 = 197811-198293
DNA, mtNC23 = 198294-198462
DNA, nad3 = 198463-198867
DNA, rps12mt = 198868-199245
DNA, atp1mt = 199246-200771
DNA, mtNC25 = 200772-200884
DNA, mtNC26 = 200885-202562
DNA, nad5d = 202563-202957
DNA, mtNC27 = 202958-203874
DNA, nad5e = 203875-204021
DNA, mtNC28 = 204022-204961
DNA, ccmFn = 204962-206698
DNA, mtNC29 = 206699-209850
DNA, trnGmt2 = 209851-209922
DNA, mtNC30 = 209923-213782
DNA, trnQmt2 = 213783-213854
DNA, mtNC31 = 213855-215346
DNA, nad4L = 215347-215648
DNA, mtNC32 = 215649-215836
DNA, atp4mt = 215837-216430
DNA, mtNC33 = 216431-217613
DNA, rps10 = 217614-220804
DNA, mtNC34 = 220805-221016
DNA, cox1 = 221017-222600
DNA, mtNC35 = 222601-224064
DNA, nad1c = 224065-225784
DNA, mtNC36 = 225785-227328
DNA, rps3mt = 227329-227402
DNA, rps3l = 227403-228783
DNA, rps3mt2 = 228784-230398
DNA, rpl16mt = 230399-230804
DNA, mtNC37 = 230805-241687
DNA, rrnL = 241688-244853
DNA, mtNC38 = 244854-245249
DNA, trnfMmt2 = 245250-245323
DNA, rpl5 = 245324-245892
DNA, rps14mt = 245893-246195
DNA, mtNC40 = 246196-247116
DNA, cob = 247117-248289
DNA, mtNC41 = 248290-249797
DNA, atp6 = 249798-250515
DNA, mtNC42 = 250516-250935
DNA, nad9 = 250936-251508
DNA, mtNC43 = 251509-258172
DNA, trnKmt = 258173-258245
DNA, mtNC44 = 258246-258728
DNA, rps4mt = 258729-259772
DNA, mtNC45 = 259773-261159
DNA, trnCmt = 261160-261232
DNA, mtNC46 = 261233-265036
DNA, trnHmt = 265037-265110
DNA, mtNC47 = 265111-268173
DNA, atp9mt2 = 268174-268398
DNA, mtNC48 = 268399-268563
DNA, nad7E1 = 268564-268706
DNA, nad7I1 = 268707-269610
DNA, nad7E2 = 269611-269679
DNA, nad7I2 = 269680-271006
DNA, nad7E3 = 271007-271473
DNA, nad7I3 = 271474-272532
DNA, nad7E4 = 272533-272776
DNA, nad7I4 = 272777-274806
DNA, nad7E5 = 274807-275068
DNA, mtNC49 = 275069-275251
DNA, ccmFcmt = 275252-280669
DNA, mtNC50 = 280670-282254
DNA, cox3 = 282255-283052
DNA, mtNC51 = 283053-284441
DNA, trnCmt2 = 284442-284512
DNA, mtNC52 = 284513-290144