UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EVOLUTION AND TAXONOMY OF THE DRYADOIDEAE

A THESIS

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

Master of Science

By

Leann Monaghan
Norman, Oklahoma
2024

EVOLUTION AND TAXONOMY OF THE DRYADOIDEAE

A THESIS APPROVED FOR THE
SCHOOL OF BIOLOGICAL SCIENCES

BY THE COMMITTEE CONSISTING OF

Dr. Abigail Moore, Chair

Dr. Richard Lupia

Dr. Lara Souza

Table of Contents

Abstract

Dryadoideae is one of three subfamilies within Rosaceae (the rose family) whose genera largely preside in the western part of North America. It contains mountain mahogany (*Cercocarpus*), mountain misery (*Chamaebatia*), cliff-rose (*Purshia*), and mountain avens (*Dryas*); many of these taxa house nitrogen-fixing bacteria belonging to the genus *Frankia*. There is also a long fossil record of these plants dating back to the Tertiary period. Morphological variation within Cercocarpus is of particular interest due to the presence of many high-quality leaf fossils. The precise relationships among and within the genera within Dryadoideae are not well-understood due to limitations in previous genetic data used and their conflicting phylogenetic hypotheses. The aim of this study is to resolve phylogenetic relationships within the subfamily, with an emphasis on Cercocarpus and Purshia, using hundreds of highly conserved and low-copy nuclear loci through target-capture (HybSeq) sequencing using the Angiosperm 353 baits and to morphologically compare leaves of extant and extinct species of *Cercocarpus*. Overall, in both genetic and morphological analyses, *C. mojadensis* and *C. pringlei* are most closely related. The subfamily Dryadoideae is strongly supported, and Purshia, Cercocarpus, and Chamaebatia were monophyletic. More data is needed to parse out how exactly the species in both *Cercocarpus* and *Purshia* are related. The morphological analyses offer some insight on how the *Cercocarpus* species relate to each other morphologically but would benefit from increased sampling

Introduction

The family Rosaceae, or the rose family, is a family that contains many economically important plants, from the roses planted in gardens to foods such as almonds, apples, and cherries. Rosaceae is divided into three subfamilies: Amygdaloideae, Rosoideae, and Dryadoideae (Xiang et al. 2017), the last being the focus of my thesis. Botanists still debate how exactly these subfamilies are related (Potter 2007, Xiang et al. 2017, Zhang et al. 2017). There are multiple fruit types in the Rosaceae. The Dryadoideae has achenes or aggregates of achenes and are a part of monophyletic, actinorhizal plant group, comprised of the orders Cucurbitales, Fagales, and Rosales (Chase et al. 2016), whose roots have nodules that house the nitrogen-fixing bacteria *Frankia* (Swensen et al. 1997*)*.

The subfamily Dryadoideae contains four genera. The genus *Dryas*, a small sub shrub, is found in arctic and alpine regions of Europe, Asia, and North America. It is used as fossil indicator of both the Younger Dryas and Older Dryas geological stages. The shrub *Chamaebatia* is an aromatic endemic of California and Baja California. *Purshia*, also a shrub, is found from British Columbia south throughout the western United States and into Mexico. The shrub *Cercocarpus* is found throughout the western United States and into Mexico. *Cercocarpus* and *Purshia* are highly valued as forage for large herbivores both domestic and wild.

The relationships within the Dryadoideae remain unresolved, but the subfamily itself is well supported (Xiang et al. 2017, Zhang et al. 2017, Chen et al. 2020). Several groups have looked at it, but they have come to different conclusions. While looking into the relationships in all of Rosaceae using complete chloroplast sequence data Zhang et al. (2017) found *Purshia* and *Cercocarpus* to be sister to each other, with *Chamaebatia* sister to them and *Dryas* sister to the

rest of the subfamily (all relationships, including Dryadoideae as a whole supported with 100% bootstrap support).

Xiang et al. (2017), using transcriptomes, could not resolve the relationships between *Cercocarpus*, *Chamaebatia*, and *Purshia.* They also found *Dryas* sister to the rest of the genera (with Dryadoideae as a whole and Dryadoideae without *Dryas* supported with 100% bootstrap support and the other relationships not supported). Chen et al. (2020), using 12 chloroplast regions, found *Cercocarpus* and *Chamaebatia* sister to each other and *Purshia* sister to both of them. *Dryas*, again, was found to be sister to the rest in the group (all relationships supported with 100% bootstrap support and posterior probabilities of 1.0).

The genus *Cercocarpus* itself has also been subject to much debate on how many species there are and how those are related. Martin (1950) believed there to be six species and 11 varieties based on morphological characteristics. Rydberg believed there to be 21 species (Rydberg, 1913) and Hendrickson et al. (2014) in *The Flora of North America* recognized eight species and six varieties.

Cercocarpus has been studied in detail by Vanden Heuvel and Linder (2001) using the Internal (ITS) and External (ETS) Transcribed Spacers of the nuclear ribosomal DNA. ITS proved not to be useful in delimiting the species as it was not variable enough, so they used ETS for their analyses. ETS, however, also had problems, as there were multiple copies of ETS found and even within copies the species were not monophyletic. Overall, none of the species which they sampled multiple individuals for were monophyletic.

Some work has been done to look at the genus *Cercocarpus* morphologically (Lis, 1992; Vanden Heuvel, 2002). Velasco de León and Cevallos-Ferriz (2000) compared some extant

species with fossil species including the species they described, † *C. mixteca*, as well as † *C. eastgatensis*, † *C. antiquus*, and † *C. ovatifolius*.

The objectives of this study are to: 1) Use targeted sequence capture with the Angiosperm353 kit (Johnson et al., 2018) to look at how all the genera in the Dryadoideae are related and to look at relationships between species of *Cercocarpus* and *Purshia.* 2) Compare current and extinct species of *Cercocarpus* with morphological analysis.

Methods

Sampling

A total of 34 plants were used for this project: 18 plants collected from the field, 3 from a botanical garden, and 13 from herbarium loans. A total of eight species and three varieties (Hendrickson et al. 2014) of the accepted species of *Cercocarpus* ,all five accepted species of *Purshia* and one debated species, one of three *Dryas species*, and one of two *Chamaebatia* species (Appendix 1). Outgroup genetic information was downloaded from NCBI SRI (https://www.ncbi.nlm.nih.gov/sra) (Appendix 1).

DNA Extraction

About 0.1g of dried leaf material was ground in a SPEX SamplePrep Genogrinder (Thomas Scientific, Swedesboro, NJ, USA). To break down more robust leaves like those of *Purshia*, sometimes a mortar and pestle were used. DNA was extracted from the ground leaves with a Qiagen Plant DNeasy Mini Kit (Qiagen Inc., Valencia, CA, USA). The manufacturer's instructions were followed with the following modifications: In step two, RNase was not added and tubes were incubated for 1 hour at 65 °C and vortexed every 20 min. In step four, tubes were

centrifuged for seven minutes. For steps eight and nine, 450 μL of AW2 buffer was added and then tubes were centrifuged for one minute. Tubes were then centrifuged for two minutes at 14,000 rpm to dry the membrane. Samples were eluted with 50 μL of buffer AE followed by an overnight incubation in a refrigerator. After the first elution, 50 μL of buffer AE was added again and the tubes were incubated at room temperature for 10 minutes.

Library Preparation

For library preparation, we used the NEB Next FS DNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA) following the manufacturer's protocol. The NEBNext® Multiplex Oligos for Illumina® (96 Index Primers) were used. During the library preparation, for inputs with adaptor ligated DNA greater than or equal to 100 ng the first bead selection was performed with 30 μL of beads and the second with 15 μL. The only changes to the protocol were that the beads were not allowed to dry after bead washes. For inputs with less than 100 ng of DNA the adaptor was diluted by 1:10. Polymerase chain reactions (PCR) were run for 16 cycles.

Hybridization

The myBaits Expert Angiosperms-353 v1 (Johnson et al., 2018; Daicel Arbor Biosciences, Ann Arbor, MI, USA) probe set was used. The manufacturer's protocol was followed. Hybridization temperature chosen was 65° C and PCR was run for 16 cycles. Samples were pooled together and then cleaned using a MinElute PCR Purification Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. Samples were then sent to the Oklahoma Medical Research Facility (OMRF, Oklahoma City, OK, USA) for sequencing.

Genetic analysis

Sequences were uploaded to the OU OSCER supercomputer. Trimmomatic version 0.39 was used (Bolger et al., 2014) to remove low quality base pairs and adapter sequences. HybPiper version 2.1.6 (Johnson et al., 2016) was used to map the reads to the target sequences from Angiosperm 353 and to obtain sequences of those regions from each individual. These sequences were then aligned with MAFFT version 7.429 (Katoh et al. 2002) and the alignments were pruned with trimAl version 1.2 (Capella-Gutierrez et al. 2009). Gene trees were made with RAxML version 8.2.11 (Stamatakis et al. 2014) and IQ-TREE version 2.2.2.6 (Minh et al. 2020) and coalescent trees were made with ASTRAL version 5.7.8 (Rabiee et al. 2019) using both RAxML and IQ-TREE gene trees. Concatenated alignments were made in Genius Prime 2024.0.4 (www.geneious.com) and RAxML was used to reconstruct concatenated trees, using 10000 bootstrap replicates. SVDquartets (Chifman & Kubatko, 2014, 2015) was run in PAUP version 4 3.99.169.0 (Swofford, 2003) with 10000 random quartets, a multispecies coalescent model, and 100 bootstrap replicates. Trees were viewed in FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

Scanning Electron Microscopy

One leaf from each species and variety of *Cercocarpus* sampled for genetic analyses was used for a total of 12 leaves. Leaf samples were mounted with carbon tape (Electron Microscopy Sciences, Hatfield, PA, USA), coated with iridium in a Quorum Q 150VES plus (Quorum, Laughton, East Sussex, England) sputter coater, and grounded with graphite conductive adhesive (Electron Microscopy Sciences, Hatfield, PA, USA). Leaves were then put under low vacuum in a TPS Quattro S environmental scanning electron microscope (ThermoFisher Scientific, Waltham, MA USA). Pictures were taken of both the adaxial and abaxial surfaces of the leaves.

Morphological Data

Characters were based on Velasco de León and Cevallos-Ferriz (2000) and my personal observations of the variation in trichomes and leaves between the different species (Appendix 2). Quantitative characters include leaf length, width, and length to width ratio. Binary characters include shape, margin type, apex type, base type, for both adaxial and abaxial: presence of substantially longer trichomes and shorter trichomes, if trichomes are short (0.3mm or less), long (0.3mm or more), straight, or curly (Figure 1). Leaf length was measured from the base to the apex using a small ruler to measure in millimeters. Width was measured at the widest point of the leaf. Length to width ratio was the length divided by the width. Shape was classified as linear, ovate, rhombic, obovate, lanceolate, or oblanceolate. Margin was marked as crenate, entire, or serrate. Apex was classified as obtuse, acute, or acuminate. Base was classified as cuneate or acuminate. Straight or curly was based off of human hair curl types. Anything in types one or two was considered straight and any trichome that would be classified in types three or four was considered curly (https://www.goodhousekeeping.com/beauty/hair/a32733411/curl-hair-types/) (Figure 2). ImageJ (Rasband, 2018) was used to measure the trichomes. Three to five trichomes were measured, and the mean was taken of those measurements. The presence of substantially longer and shorter trichomes on the leaf was also recorded. Characteristics from fossil species were taken from the pictures published of them from Velasco de León and Cevallos-Ferriz (2000).

Herbarium specimens and leaves given from botanical gardens were observed for morphological data. SEM images were used to gather information on the trichomes. For each species, five random leaves were observed for each specimen. Each leaf was analyzed separately in the UPGMA analysis. A one was given to a trait if it was present and a zero if it was absent.

Traits in a given category were not mutually exclusive to account for morphological variability within each individual.

SEM analysis

A distance matrix was created from the morphology matrix to compare traits and variables between the species in R version 4.3.1 (R Core Team, 2018). This was accomplished by following the R script tutorial (Neto 1012). First, variables were defined as either binary or quantitative. The ade4 package version 1.7-22 (Chessel et al., 2007; Dray & Dufour, 2007) was used to create the distance matrix. To create a dendrogram to visualize the relationships between the species, the package clue version 0.3-65 (Hornik, 2023) was used. The tree was then changed into a .phylo file with the package picante version 1.8.2 (Kembel et al., 2010). The .phylo files were then changed into nexus files using the phytools version 2.1-1 (Revell, 2024) package and visualized with FigTree (http://tree.bio.ed.ac.uk/software/figtree/). Dendrograms were created with this data with the following combinations: fossil and extant species using only binary data; extant species using only binary data; extant species using all data; and extant and extinct species using all data with quantitative characters coded as missing for the fossils. This was done to see if the addition of fossil species changed the relationships between the extant species.

Results

Genetic Data

Of the 34 plants sampled, 20 of them had good enough material to be sequenced. Total number of reads varied with the highest from *Dryas* with 58866047 and the lowest from *Cercocarpus betuloides* with 588. The number of reads that mapped to the targets also varied with the most from *C. montanus* at 35001201 and the lowest from *Purshia mexicana* with 22.

Dryas also had the highest number of genes with 345 and *P. mexicana* had zero. The highest number of paralogs found was 29 from *Pyrus regelii*, but the majority of samples had zero to four (Appendix 1).

Phylogenetics

Astral: IQTree

The Astral analysis of 351 gene trees reconstructed with IQTree (Figure 3) showed that all genera were monophyletic and Dryadoideae as a whole was supported with quartet score of one. *Dryas* was sister to the rest of the genera in the subfamily with a quartet score of 0.74. Between the other three genera it showed that *Chamaebatia* and *Cercocarpus* were sister with a 0.76 quartet score. In the genus *Cercocarpus*, *C. montanus* was sister to the rest of the species (0.84). *Cercocarpus pringlei* and *C. traskiae* were the most closely related (0.76). In the genus *Purshia*, *P. stansburyana* was the unsupported (0.41) sister to the other species. One individual of *Purshia subintegra* was more closely related to an individual of *P. tridentata* from Oregon than the other *P. subintegra*. *Purshia ericifolia* was sister to this group of *P. subintegra* and Oregon *P. tridentata*. ChamaebatiaCA2 and ChamaebatiaCA3 were the same specimen and were most closely related to one another. The Amygdaloideae subfamily group containing *Prunus*, *Pyrus*, and *Crataegus* was well supported with a quartet score of 1 and was sister to the Dryadoideae. *Pyrus* and *Crataegus*, which are both in the tribe Maleae of the Amygdaloideae, were each other's closest relatives with a quartet score of one. *Artocarpus* and *Parartocarpus* from the Moraceae were each other's closest relatives (1) and Moraceae was sister to Rosaceae.

Astral: RAxML

For the Astral analysis of gene trees reconstructed with RAxML (Figure 4), all genera were monophyletic and Dryadoideae as a whole was supported with a quartet score of one. *Dryas* was sister to the rest of the genera (1). The genera *Cercocarpus* and *Chamaebatia* were sister, but with only a quartet score of 0.45. For *Cercocarpus, C. pringlei* and C. *mojadensis* were sister with a quartet score of one. *Cercocarpus montanus* and *C. traskiae* were sister, but with only with a quartet score of 0.42. Both individuals of *Purshia subintegra* were sister with a quartet score of 0.61. Both individuals of *Purshia tridentata* were not each other's closest relatives. Amygdaloideae was sister to Dryadoideae with a quartet score of one and the tribe Maleae was supported with a quartet score of one within that. Moraceae was sister to Rosaceae with a quartet score of one.

SVDquartets

The SVDquartets analysis of the concatenated data showed all genera were monophyletic (Figure 5). Dryadoideae as a whole was supported with a bootstrap value of 100. *Cercocarpus* was supported with a bootstrap value of 72, *Purshia* with 100, and *Chamaebatia* with 71. *Cercocarpus* and *Purshia* were sister with a bootstrap value of 47. *Cercocarpus traskiae* and *C. montanus* were sister, but only with a bootstrap value of 38. *Cercocarpus pringlei* and *C. mojadensis* were each other's closest relatives, but only with a bootstrap value of 66. *Chamaebatia* was sister to the other three genera with a bootstrap value of 100. Both individuals of *Purshia subintegra* were not each other's closest relatives. The same was true with both individuals of *P. tridentata*. Amygdaloideae was sister to Dryadoideae with a bootstrap value of 100 and the tribe Maleae was supported with a bootstrap value of 100 within that. Moraceae was sister to Rosaceae with a bootstrap value of 50.

RAxML Concatenation

The RAxML analysis of the concatenated data showed that all genera were monophyletic with *Purshia* being sister to the other genera (100) and Dryadoideae as a whole being supported with a 100 bootstrap value (Figure 6). *Dryas* and *Cercocarpus* were sister with a bootstrap value of 100. Within *Cercocarpus, C. mojadensis* and *C. pringlei* were most closely related to each other (100) and *C. montanus* was sister to the other three species (100). In *Purshia*, both individuals of *Purshia subintegra* were not each other's closest relatives. The same is true with both individuals of *P. tridentata*. Amygdaloideae was sister to Dryadoideae with a bootstrap value of 100 and the tribe Maleae was supported with a bootstrap value of 100 within that. Moraceae was sister to Rosaceae with a bootstrap value of 100.

Morphological

SEM

Overall, the adaxial surfaces of *Cercocarpus* leaves have lower trichome density than the abaxial surfaces. In the *Cercocarpus fothergilloides, C. mojadensis,* and *C. pringlei* group (based on the morphological dendrograms) (Figure 7), trichomes are intensely coiled on the abaxial side. The trichomes on the veins on this side are also intensely coiled for both *C. pringlei* and *C. mojadensis*, while *C. fothergilloides* has straight trichomes on the veins. All the trichomes on the abaxial sides in these species are curly.

Between *C. breviflorus* var. *breviflorus* (Figure 8b) and *C. breviflorus var. eximius* (Figure 8a), the leaves are virtually indistinguishable from each other except for the presence of long and short trichomes on the abaxial side of *C. breviflorus* var. *breviflorus*. Both have straight trichomes on the adaxial side, and both straight and curly trichomes on the abaxial side. All trichomes are long. *Cercocarpus montanus* (Figure 9b) has long, straight trichomes on both sides

and short trichomes only on the abaxial side. *Cercocarpus minutiflorus* and *C. macrophyllus* have short, straight trichomes on both sides (Figure 9a and c). The abaxial side of *C. macrophyllus* species is particularly interesting with the presence of densely pubescent pits. *Cercocarpus ledifolius* (Figure 10a) has long, curly trichomes on the abaxial side and short, curly trichomes on the adaxial side. *Cercocarpus intricatus* (Figure 9b) has short trichomes on the abaxial side and both short and long trichomes on the adaxial side. *Cercocarpus betuloides* (Figure 11b) has short, straight trichomes on the adaxial side. On the abaxial side trichomes, were short and straight and the long trichomes were curly. On the abaxial side, *Cercocarpus traskiae* (Figure 11a) has long straight, and curly, short trichomes. The adaxial side the trichomes are long and bot straight and curly.

Dendrograms

Using fossil and extant species using all binary data (Figure 12) all taxa are monophyletic except *C. breviflorus* var. *breviflorus*, *C. breviflorus* var. *eximius*, *C. minutiflorus*, and *Cercocarpus fothergilloides*, *C. mojadensis*, and *C. pringlei* cluster together. *Cercocarpus breviflorus* vars. *breviflorus* and *eximius* cluster together. *Cercocarpus minutiflorus* clusters with *C. betuloides* and *C. traskiae* and † *C. eastgatensis* cluster with † *C. mixteca*. *Cercocarpus macrophyllus* and *C. montanus* cluster together and *C. intricatus* and *C. ledifolius* cluster together.

Looking at extant species using all binary data, all taxa are monophyletic. Species cluster similarly to the dendrogram with the fossils included except *Cercocarpus macrophyllus* which was closer to *C. minutiflorus*, *C. betuloides*, and *C. traskiae* (Figure 13).

Using extant species using all data, all taxa were monophyletic and species that clustered together were the same as the other two dendrograms, except for *C. ledifolius, C. montanus, C. macrophyllus*, and *C. intricatus*. *Cercocarpus intricatus* was sister to the rest of the species and *C. ledifolius* was closer to *C. minutiflorus*, *C. betuloides*, and *C. traskiae*. (Figure 14)

Using extant and extinct species using all data, with the length/width characters coded as missing for the fossil species (Figure 15), all taxa were monophyletic except *C. minutiflorus*, which contained † *C. ovatifolius*. Species clustered similarly to all other analyses, except for *C. montanus* and *C. macrophyllus*, which were in clades by themselves. *Cercocarpus intricatus* clustered with † *C. antiquus* and † *C. eastgatensis* clustered with † *C. mixteca*.

Discussion

Phylogenetics

In the genus *Cercocarpus*, *C. mojadensis* and *C. pringlei* are sister in all analyses except for the Astral tree made with IQTree gene trees. This relationship is supported by the fact that *C. pringlei* is sometimes considered to be a variety of *C. mojadensis* (Schneider 1905, Rydberg 1913). *Cercocarpus mojadensis* is found as far north as Chihuahua and as far south as Oaxaca with the type locality in Coahulia. *Cercocarpus pringlei* is found only in the southern portion of the same range with it being as far south as Oaxaca and as far north as Querétaro with the type locality in Oaxaca. Both only occur as far west as Durango and Guerro.

Cercocarpus montanus and *C. traskiae* are sister in two of the four trees, but with low support values. *Cercocarpus traskiae* is sometimes considered to be a variety of *C. montanus* (Martin, 1950) and sometimes a variety of *C. betuloides* (Dunkle 1940).

For all trees, the two *Purshia tridentata* accessions were not sister. In all but one tree, the *P. tridentata* collected in Oregon was most closely related to *P. subintegra*. In three of the four trees, *P. ericifolia* was sister to the clade formed by *P. subintegra* and *P. tridentata* from Oregon. There is evidence of hybridization between *P. tridentata* and *P. stansburyana* (Stutz, 1964). Their ranges overlap in the state of Utah, Nevada, Colorado, New Mexico, Arizona, and California . *Purshia tridentata* occurs at slightly lower elevations than *P. stansburyana*, but can be found in the same places. The specimen of *P. tridentata* from Utah that was used for genetic analyses could, in fact, be a hybrid, which would explain why it is close to *P. stansburyana* in all of the phylogenetic trees.

All trees fully supported the monophyly of *Chamaebatia australis*. ChamaebatiaCA2 and ChamaebatiaCA3 were from DNA extracted from the same individual and are sister in all trees. In three of the four analyses, *Chamaebatia* is sister to *Cercocarpus*.

The genus *Dryas* is sister to the other genera in two of the four trees with quartet scores of 0.74 and 0.72. Other literature supports this idea (Chen et al. 2020, Xiang et al .2017, Zhang et al. 2017). In the SVDquartets tree, it is sister to *Purshia* and *Cercocarpus* with a bootstrap value of 70 and in the concatenation RAxML tree it is sister to only *Cercocarpus* with a bootstrap value of 100.

Two samples from Moraceae, were used to root the Rosaceae. Members from the other two subfamilies of Rosaceae, Amygaloideae and Rosoideae, were used to provide sufficient outgroups relative to Dryadoideae. Though, the Dryadoideae as a whole has a long branch, so that does make its' relationship to the rest of Rosaceae hard to determine. The extant Dryadoideae samples analyzed were strongly-supported as monophyletic from ASTRAL, RAxML, and SVDquartet analyses as previously inferred (Chen et al. 2020, Xiang et al. 2017,

Zhang et al. 2017). Within Dryadoideae, *Dryas* was recovered as sister to the other genera in all anakyses, except the concatenated RAxML analysis. It showed that *Chamaebatia,* was sister to the other three genera. As the *Dryas* sample has a particularly long branch length, its position between the coalescent and concatenation analyses may have shifted due to the difficulty in placing samples with highly divergent sequences.

For the genus of *Cercocarpus* the Astral analysis of the IQTree gene trees gave a 0.76 quartet score and the Astral analysis of the RAxML gene trees gave a quartet score of 0.45. The Concatenation RAxML anaysis had a bootstrap value of 66. The SVDquartets analysis shows *Cercocarpus* sister to *Purshia,* but also with poor suport a 47 bootstrap value. Zhang et al. (2017) found *Purshia* and *Cercocarpus* sister to each other. Chen et al. (2020) and Xiang et al. (2017) found that *Cercocarpus* and *Chamaebatia* were sister to each other. Overall all genera were monophyletic and well supported.

Morphology

In all dendrograms, *C. pringlei* and *C. mojadensis* clustered together and *C. pringlei*, *C. mojadensis*, and *C. fothergilloides* all cluster together. This lends support to the idea that *C. pringlei* could be a variety of *C. mojadensis* (Schneider, 1905 & Rydberg, 1913). *Cercocarpus brevifolius* var. *breviflorus* and *C. brevifolius* var. *eximius* always clustered together. One tree had the individual leaves of *C. brevifolius* vars. *eximius* and *breviflorus* intermixed. *Cercocarpus betuloides* and *C. traskiae* always grouped together. Somethimes *Cercocarpus traskiae* is sometimes considered either a variety of *C. betuloides* (Dunkle, 1940) or of *C. montanus* (Martin, 1950). *Cercocarpus minutiflorus* never clusters with *C. montanus* even though it is sometimes considered a variety of *C. montanus* (Martin, 1950). *Cercocarpus intricatus* is closest to *C. ledifolius* in two of the four trees. There is also evidance of hybridization between the two

(Walker and Turley, 1999). *Cercocarpus intricatus* has been considered a variety of *C. ledifolius* (Jones, 1880). Yet, in the other two trees *C. ledifolius* clustered with *C. minutiflorus*, *C. betuloides*, and *C. traskiae*.

In the dendrograms that contained the fossil specimens, both † *C. eastgatensis* and † *C. mixteca* clustered together. † *Cercocarpus antiquus* moved arround, being closer to *C. ledifolius* in one and clustered with *C. minutiflorus*, *C. traskiae*, and *C. betuloides* in others. † *Cercocarpus ovatifolius* was within *C. minutiflorus.* The phenogram created by Velasco de León and Cevallos-Ferriz (2000) was quite different. Their analysis showed that † *C. eastgatensis* and † *C. ovatifolius* clustered together and † *C. mixteca* clustered with *C. montanus* and *C. fothergilloides*. They also found that *C. mojadensis* clustered together with *C. minutiflorus* and *C. traskiae* and that *C. pringlei* and *C. macrophyllus* clustered together. Velasco de León and Cevallos-Ferriz (2000) examined some additional characters: number of secondary veins, angle of secondary veins, and type of leaf areols. In addition, they did not use length to width ratio or data on trichomes.

Conclusions

Overall, in both genetic and morphological analyses, *C. mojadensis* and *C. pringlei* are most closely related. The subfamily Dryadoideae is very well supported and all genera are monophylletic. More data is needed to parse out how exactly the species in *Cercocarpus* and *Purshia* are related to each other. The morphological analyses offer some insight on how the *Cercocarpus* species relate to each other morphologically, but more data is also recommended.

Future Directions

The next step is to sequence the remaining species of *Cercocarpus* and *Chamaebatia*. Many of the samples did not give good enough results to be used in the analyses, so they need to be resequenced. The species *Purshia plicata* was sampled, but did not give good enough genetic data also needs to be redone and added to genetic analyses. Although *Dryas* is a complex genus and not the focus of the study, there is whole-genome sequencing data available in the NCBI Sequence Read Archive, which could allow the target loci to be extracted and more species to easily be included.

The morphological analyses would also benefit from the addition more individuals per species and the addition of more morphological characters, such as the angle of the leaf vein used in Mortensen (1973) and Velasco de León and Cevallos-Ferriz (2000).

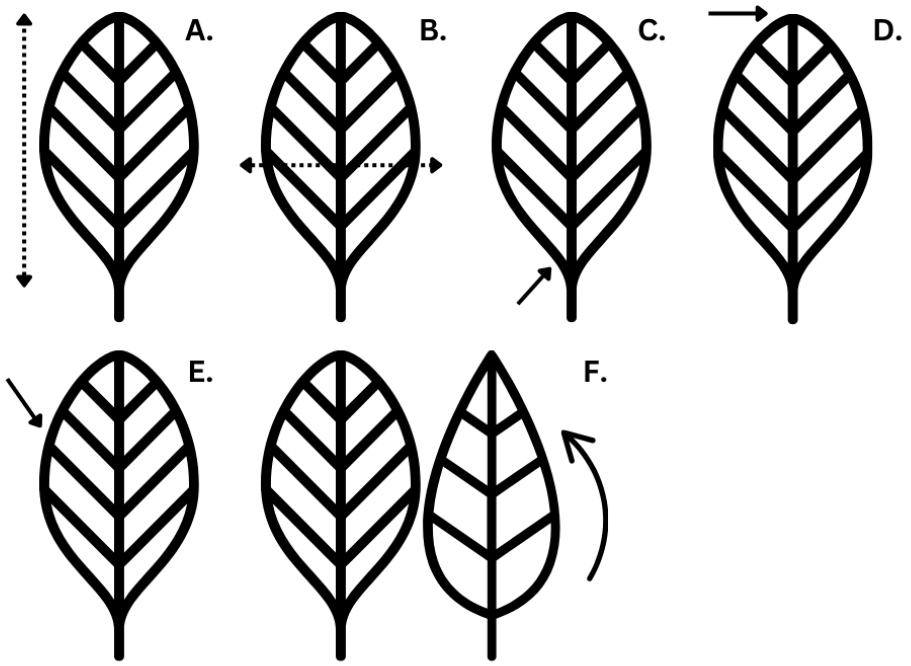

Figure 1. Diagram showing leaf characters. A) Length; B) Width; C) Base; D) Apex; E) Margin or Serration; F) Shape.

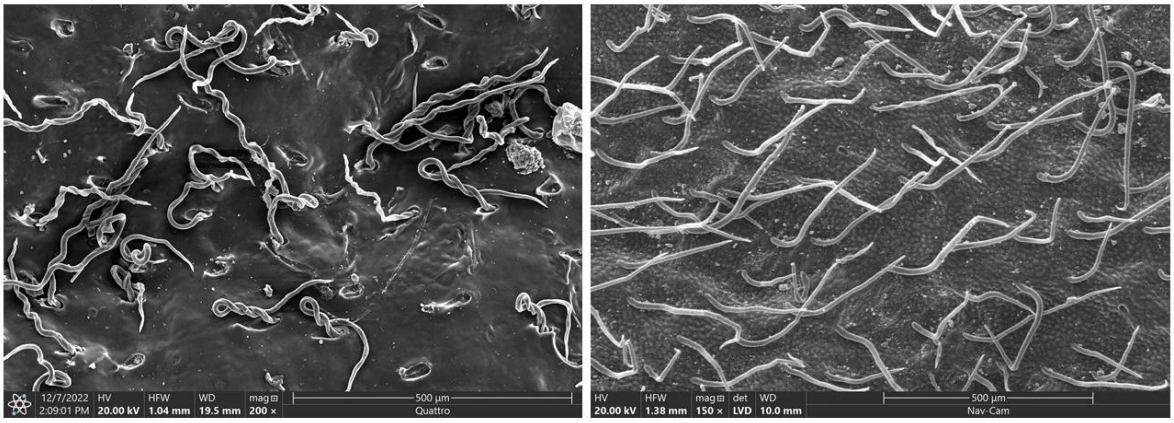

Figure 2. SEM pictures of *C. ledifolius* and *C. breviflorus* var. *breviflorus* adaxial surfaces showing curly and straight trichomes respectively.

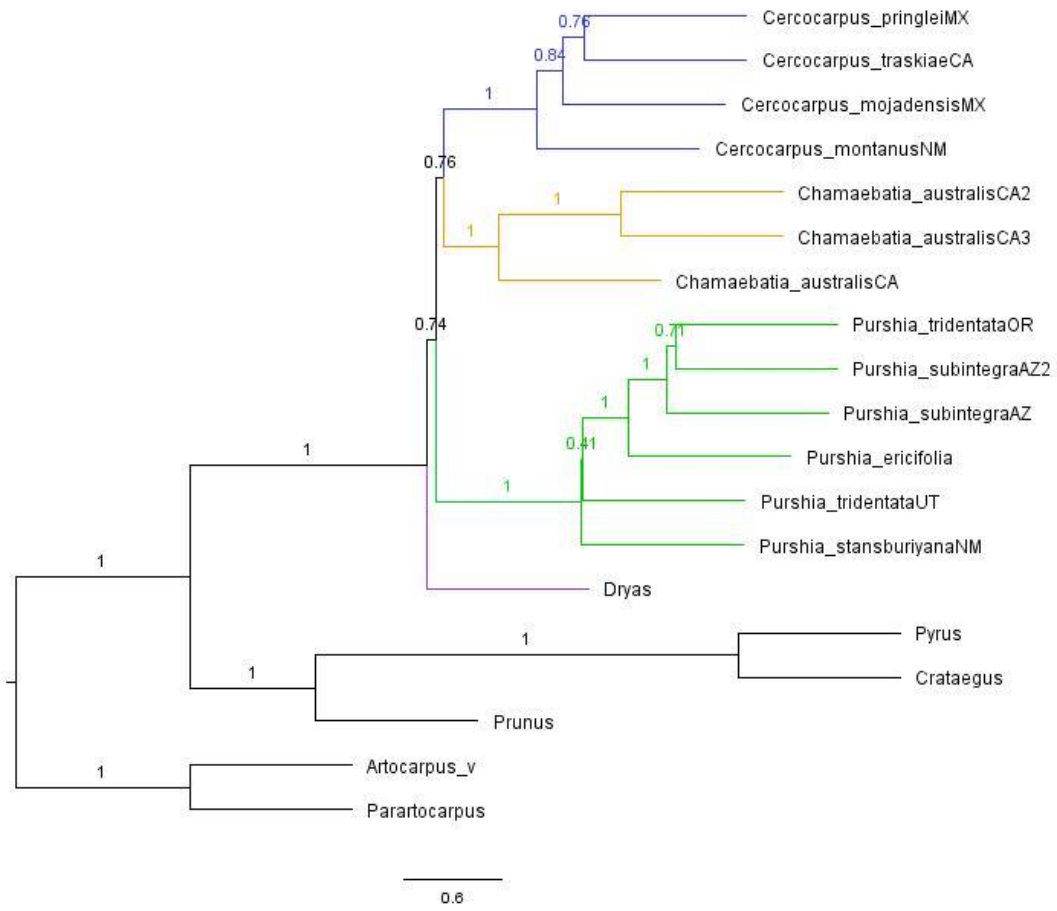

Figure 3. Astral analysis of IQTree gene trees. Numbers on the nodes signify quartet scores.

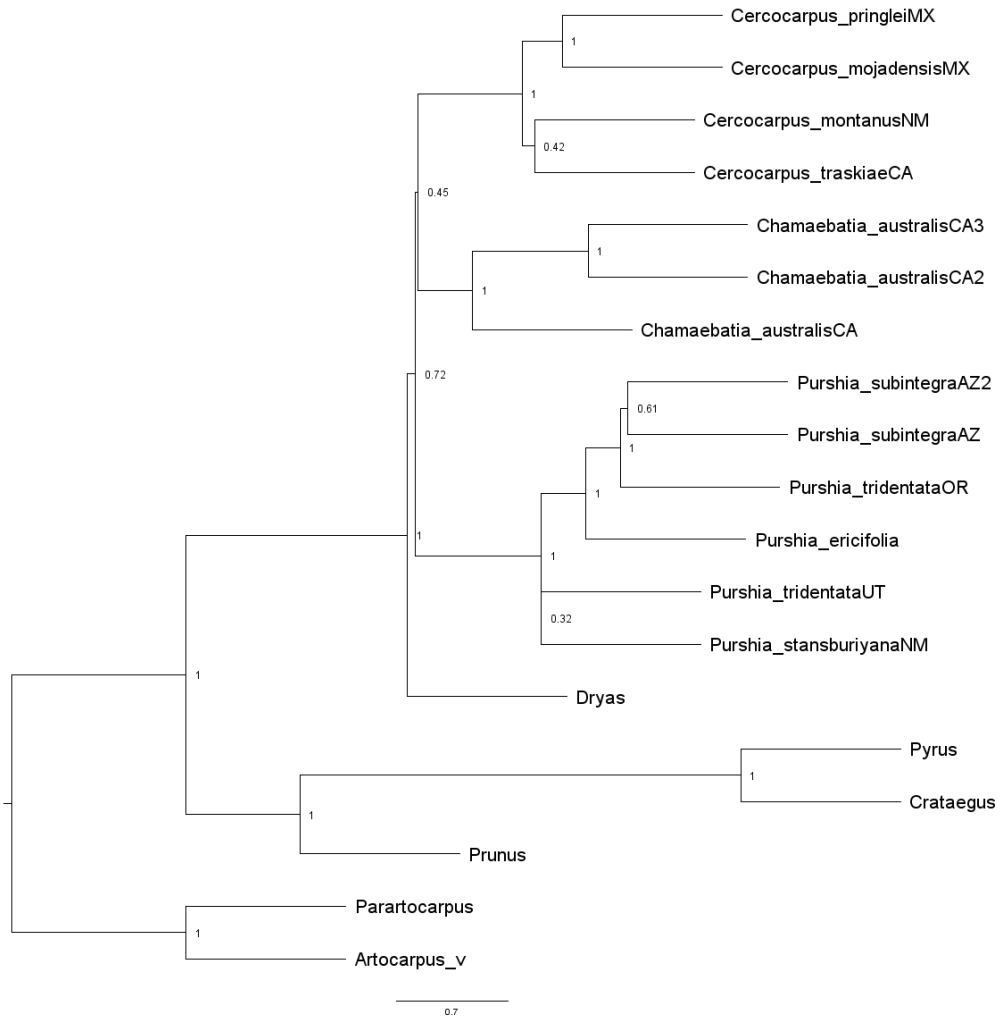

Figure 4. Astral analysis of RAxML gene trees. Numbers on the nodes signify quartet scores.

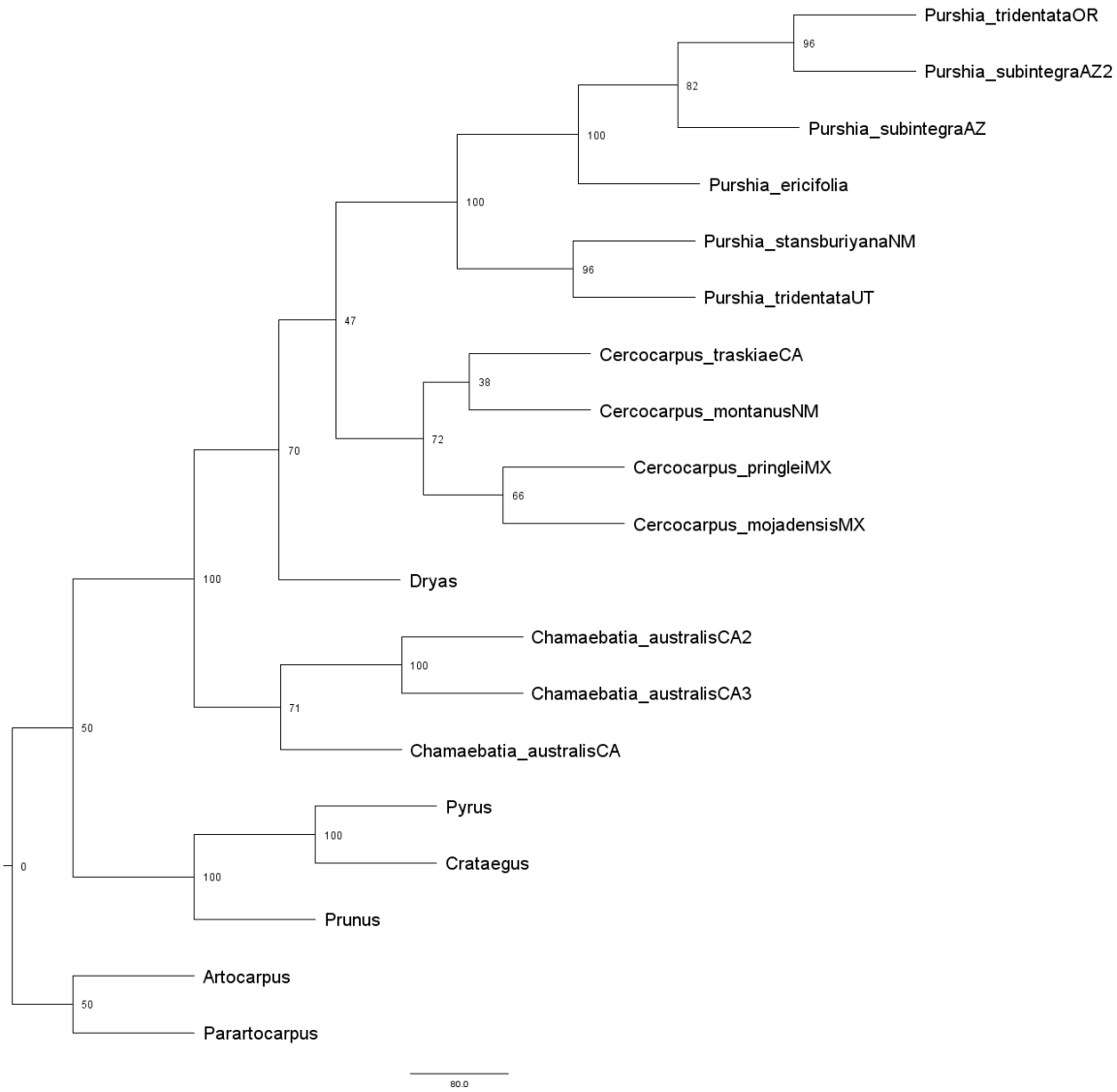

Figure 5. SVDquartets tree from concatenated data. Numbers at the nodes signify bootstrap values.

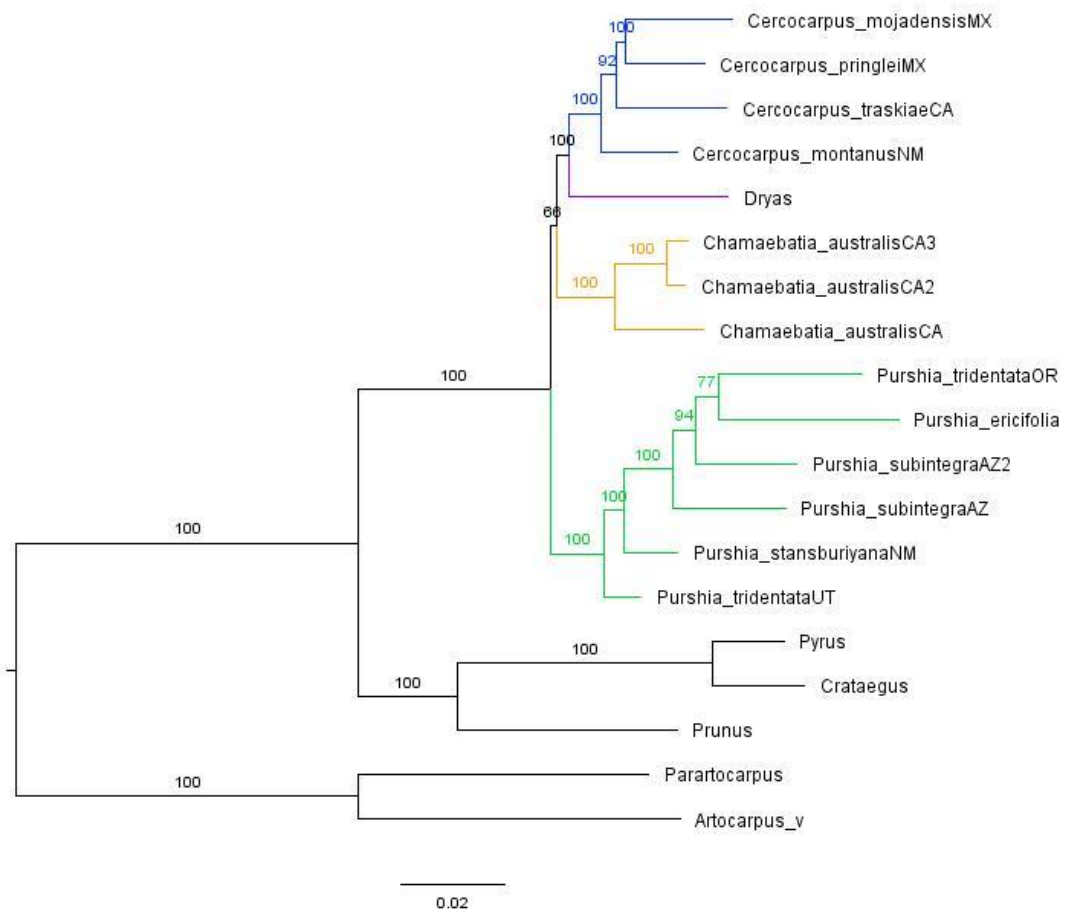

Figure 6. RAxML tree from concatenated data. Numbers at nodes indicate bootstrap support values.

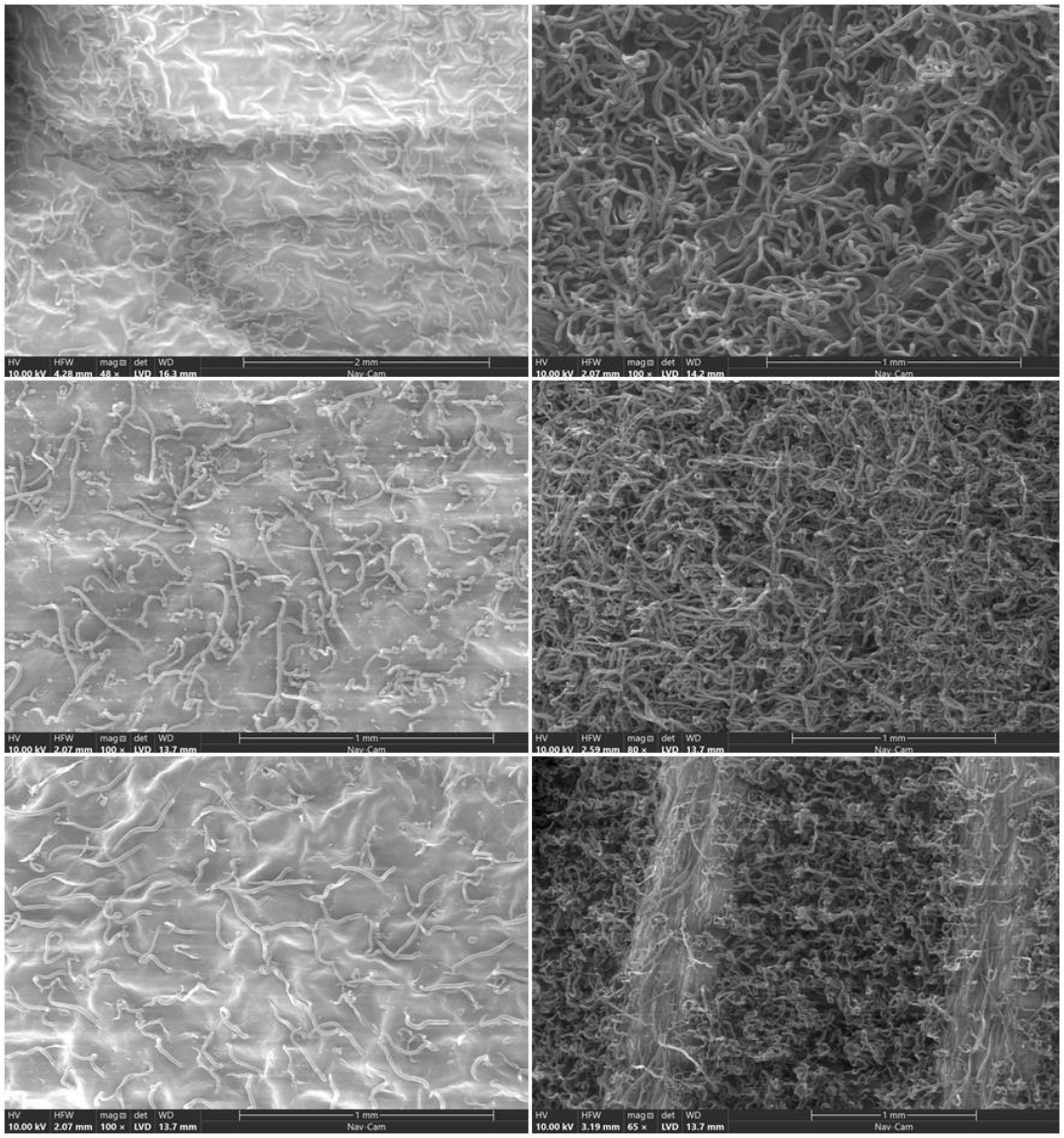

Figure 7. SEM images of (top to bottom) *Cercocarpus pringlei*, *C. mojadensis*, and *C. fothergilloides*. Images on the left are of the adaxial leaf surface and the right is abaxial surface.

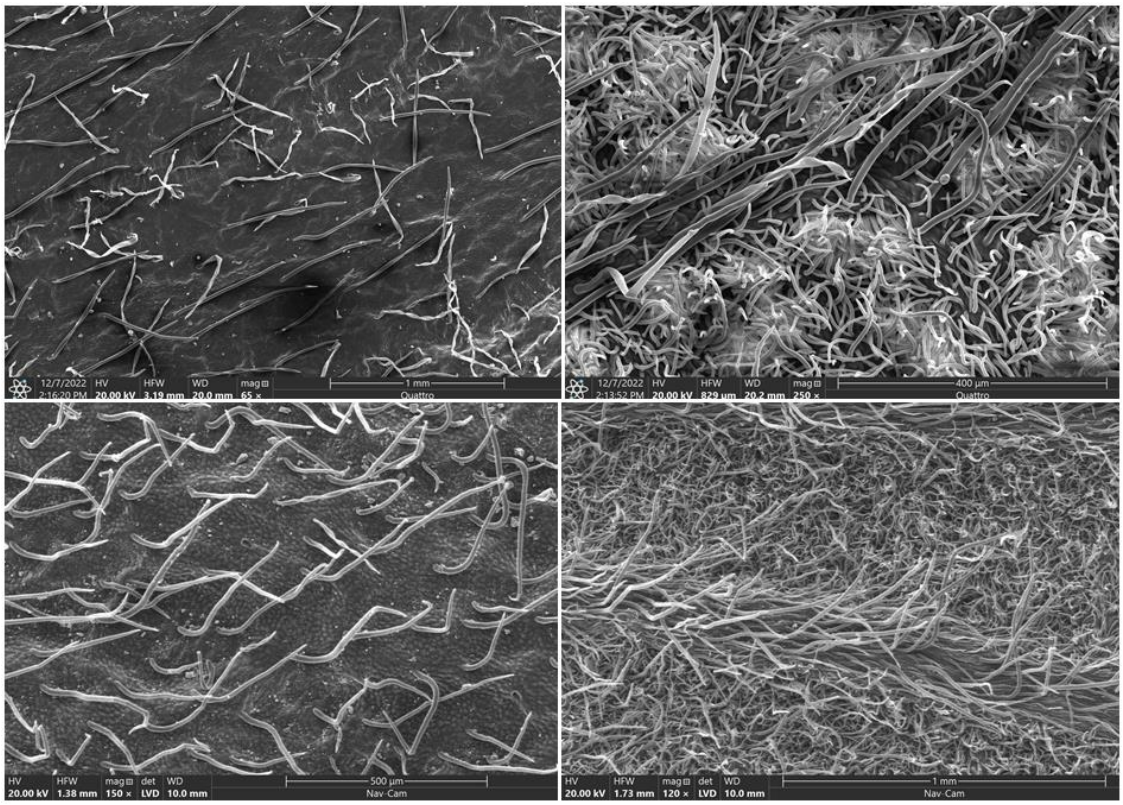

Figure 8. SEM images of leaf surfaces of *C. breviflorus* var. *eximius* (top) and *C. breviflorus* var. *breviflorus* (bottom). Images on the left are of the adaxial leaf surface and the right is abaxial surface.

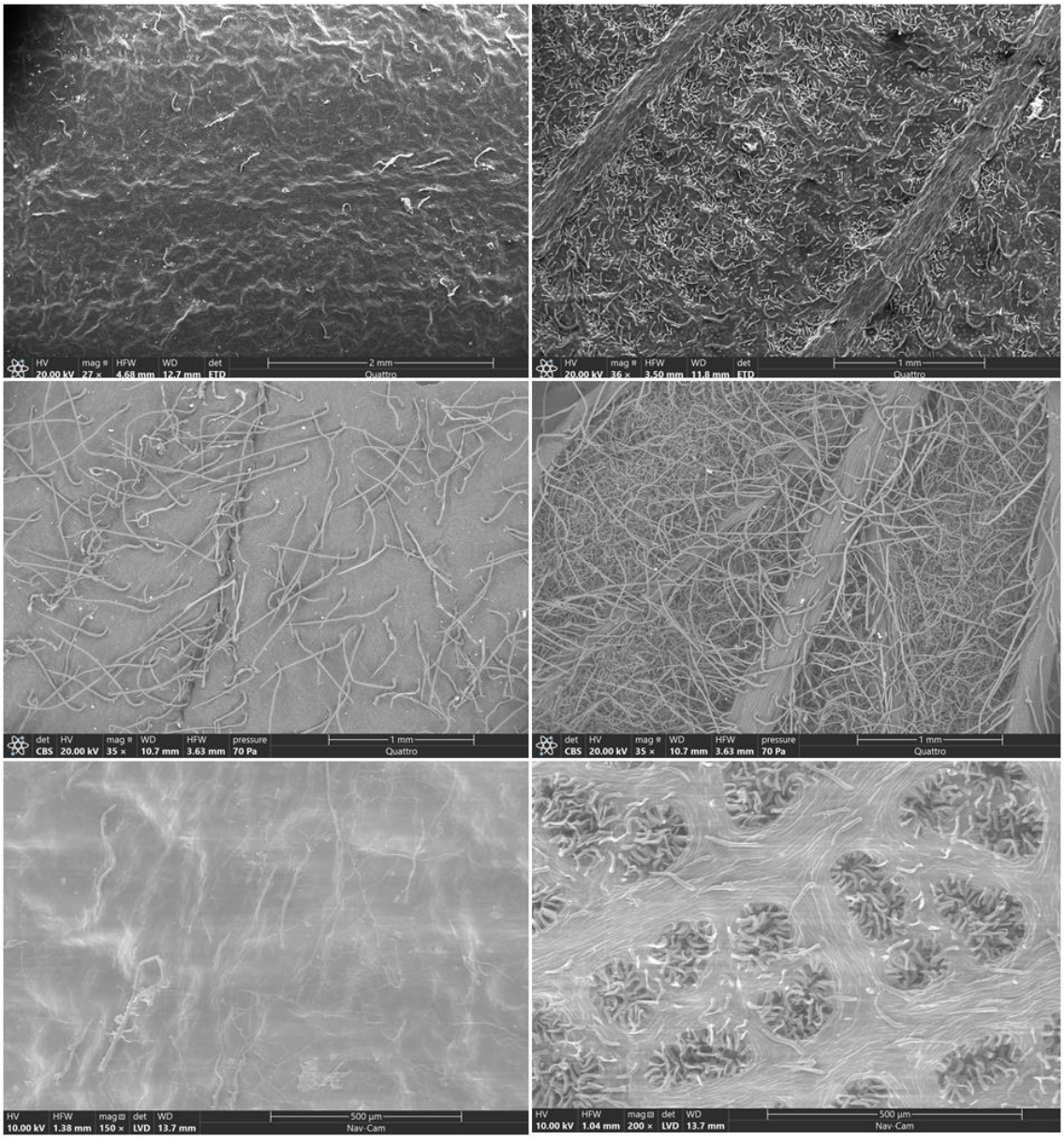

Figure 9. SEM images of leaf surfaces of (top to bottom) *C. minutiflorus*, *C. montanus*, and *C. macrophyllus.* Images on the left are of the adaxial leaf surface and the right is abaxial surface.

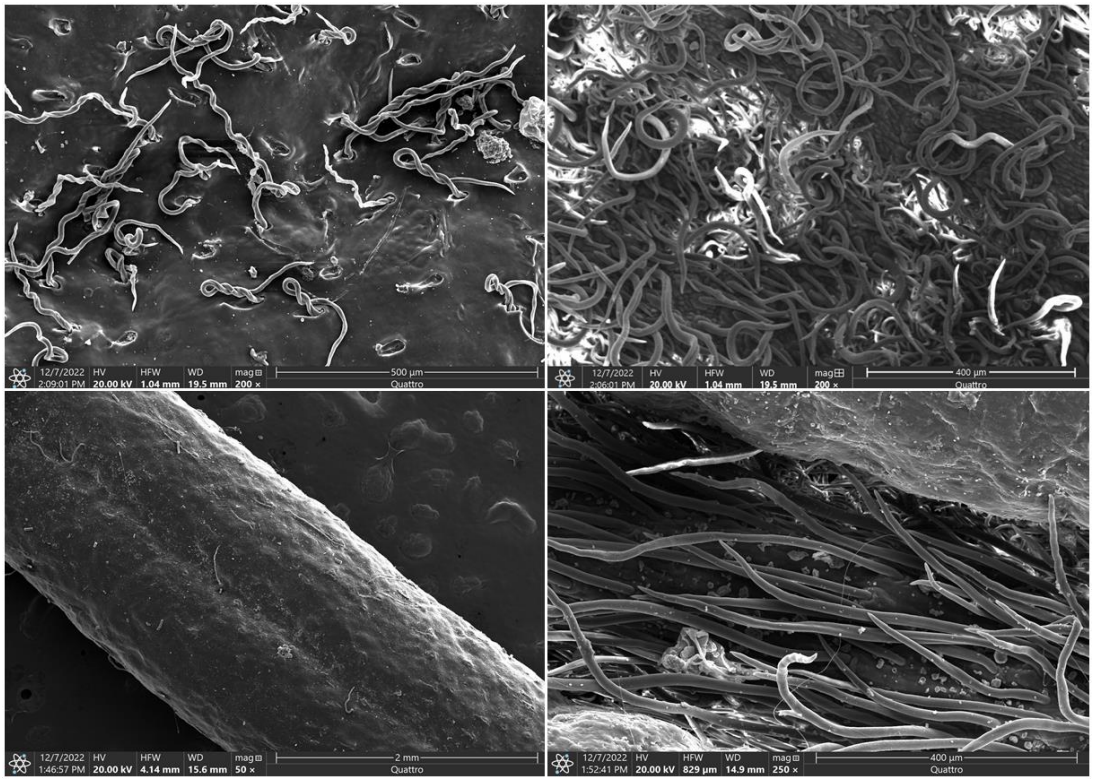

Figure 10. SEM images of leaf surfaces of *C. ledifolius* (top) and *C. intricatus* (bottom). Images on the left are of the adaxial leaf surface and the right is abaxial surface.

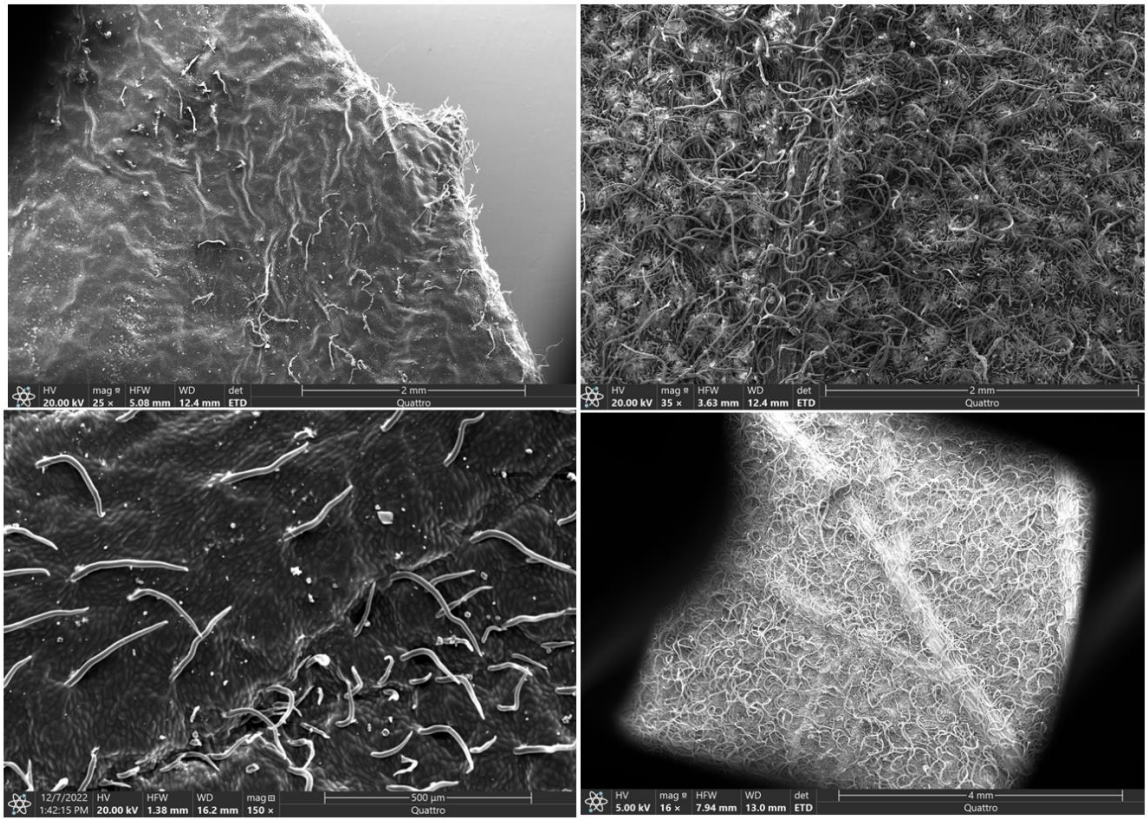

Figure 11. SEM images of leaf surfaces of *C. traskiae* (top) and *C. betuloides* (bottom). Images on the left are of the adaxial leaf surface and the right is abaxial surface.

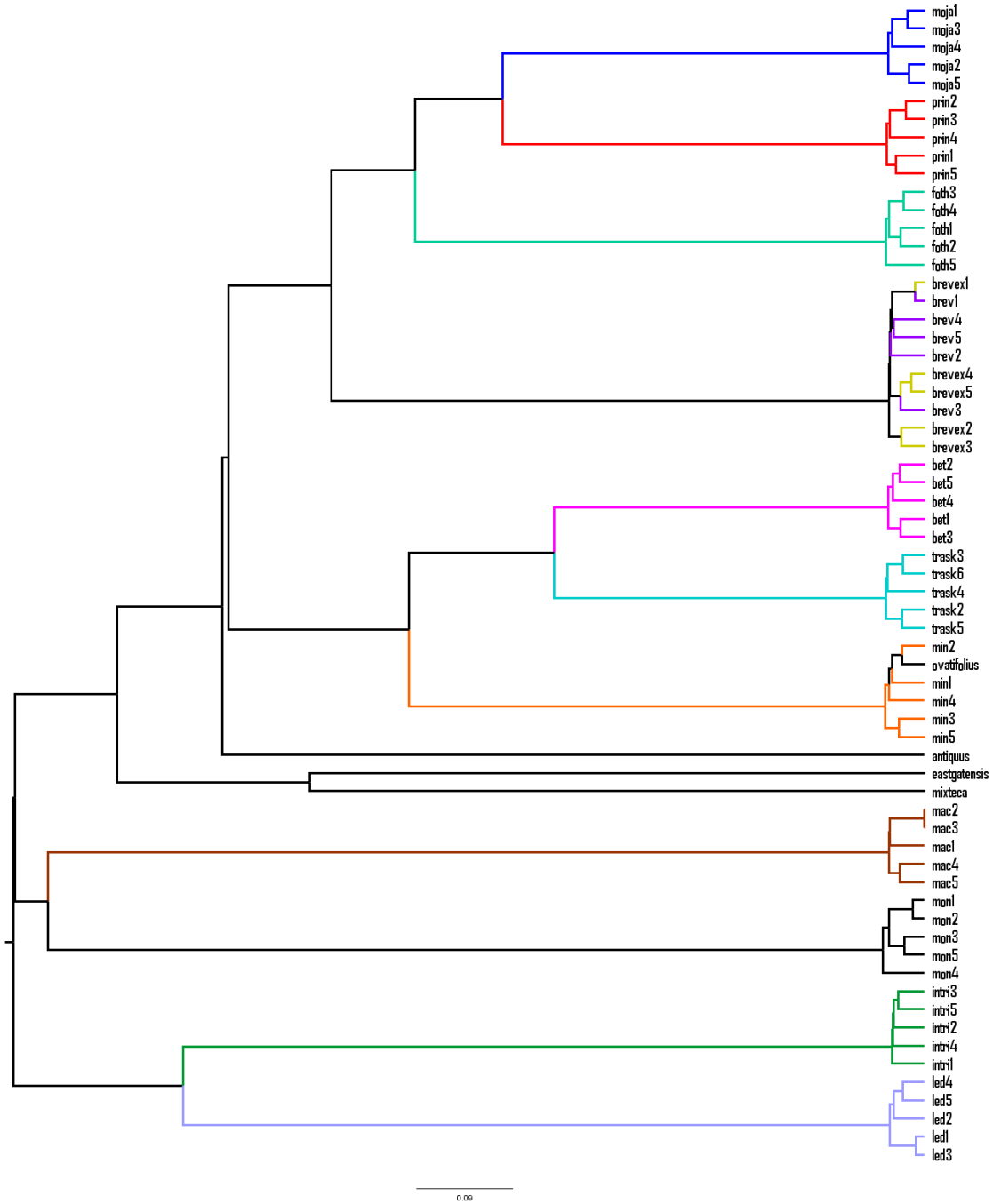

Figure 12. A UPGMA dendrogram including both fossil and extant species and binary characters only (excluding length, width, and length-width ratio). Leaves are named by an abbreviation of the species name and then their individual number. moja= *Cercocarpus mojadensis.* prin = *C. pringlei*. foth = *C. fothergilloides*. brevex = *C. breviflorus* var. *eximius*.

brev = *C. breviflorus* var. *breviflorus*. bet = *C. betuloides*. trask = *C. traskiae*. min = *C. minutiflorus*. ovatifolius= † *C. ovatifolius.* antiquus = † *C. antiquus*. eastgatensis = † *C. eastgatensis.* mixteca = † *C. mixteca*. mac = *C. macrophyllus*. mon = *C. montanus*. intri = *C. intricatus*. led = *C. ledifolius*.

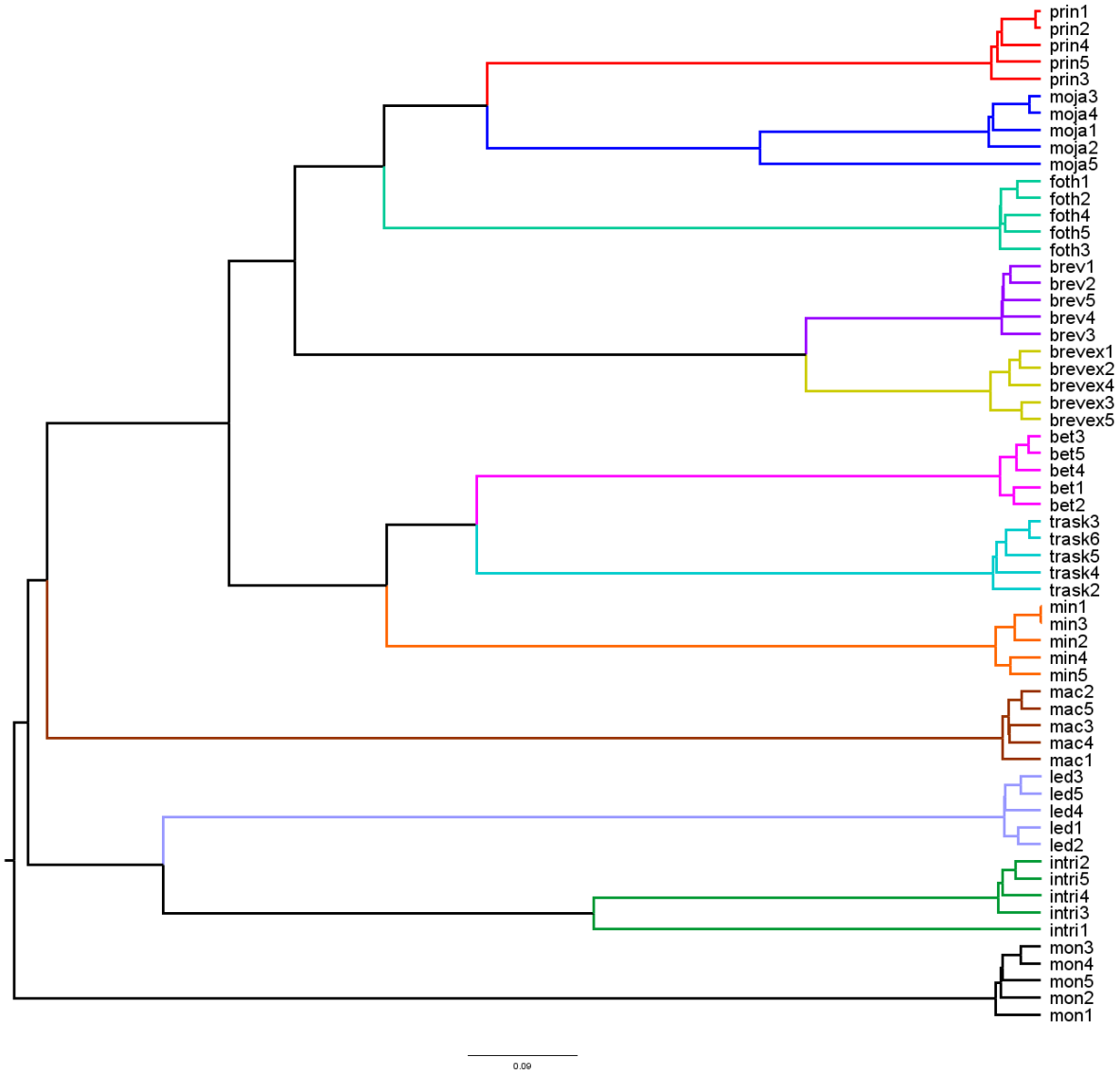

Figure 13. UPGMA dendrogram showing relationships between extant species using only binary data (all data except length, width, and length-width ratio). Leaves are named by an abbreviation of the species name and then their individual number. moja= *Cercocarpus mojadensis.* prin = *C. pringlei*. foth = *C. fothergilloides*. brevex = *C. breviflorus* var. *eximius*. brev = *C. breviflorus* var. *breviflorus*. bet = *C. betuloides*. trask = *C. traskiae*. min = *C. minutiflorus*. ovatifolius= † *C. ovatifolius.* antiquus = † *C. antiquus*. eastgatensis = † *C. eastgatensis.* mixteca = † *C. mixteca*. mac = *C. macrophyllus*. mon = *C. montanus*. intri = *C. intricatus*. led = *C. ledifolius*.

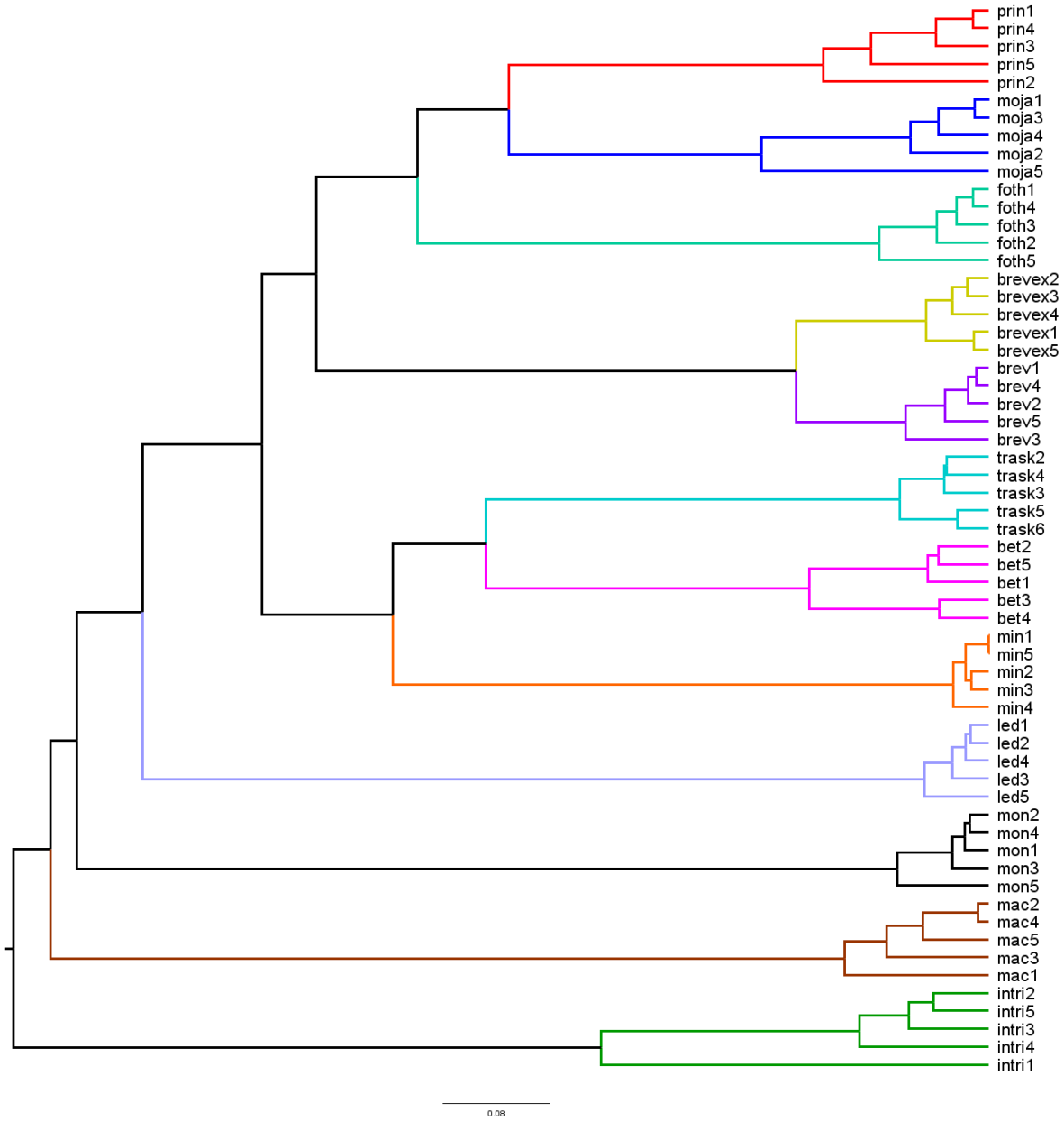

Figure 14. UPGMA dendrogram showing relationships between extant species using all data. Leaves are named by an abbreviation of the species name and then their individual number. moja= *Cercocarpus mojadensis.* prin = *C. pringlei*. foth = *C. fothergilloides*. brevex = *C. breviflorus* var. *eximius*. brev = *C. breviflorus* var. *breviflorus*. bet = *C. betuloides*. trask = *C. traskiae*. min = *C. minutiflorus*. ovatifolius= † *C. ovatifolius.* antiquus = † *C. antiquus*.

eastgatensis = † *C. eastgatensis.* mixteca = † *C. mixteca*. mac = *C. macrophyllus*. mon = *C. montanus*. intri = *C. intricatus*. led = *C. ledifolius*.

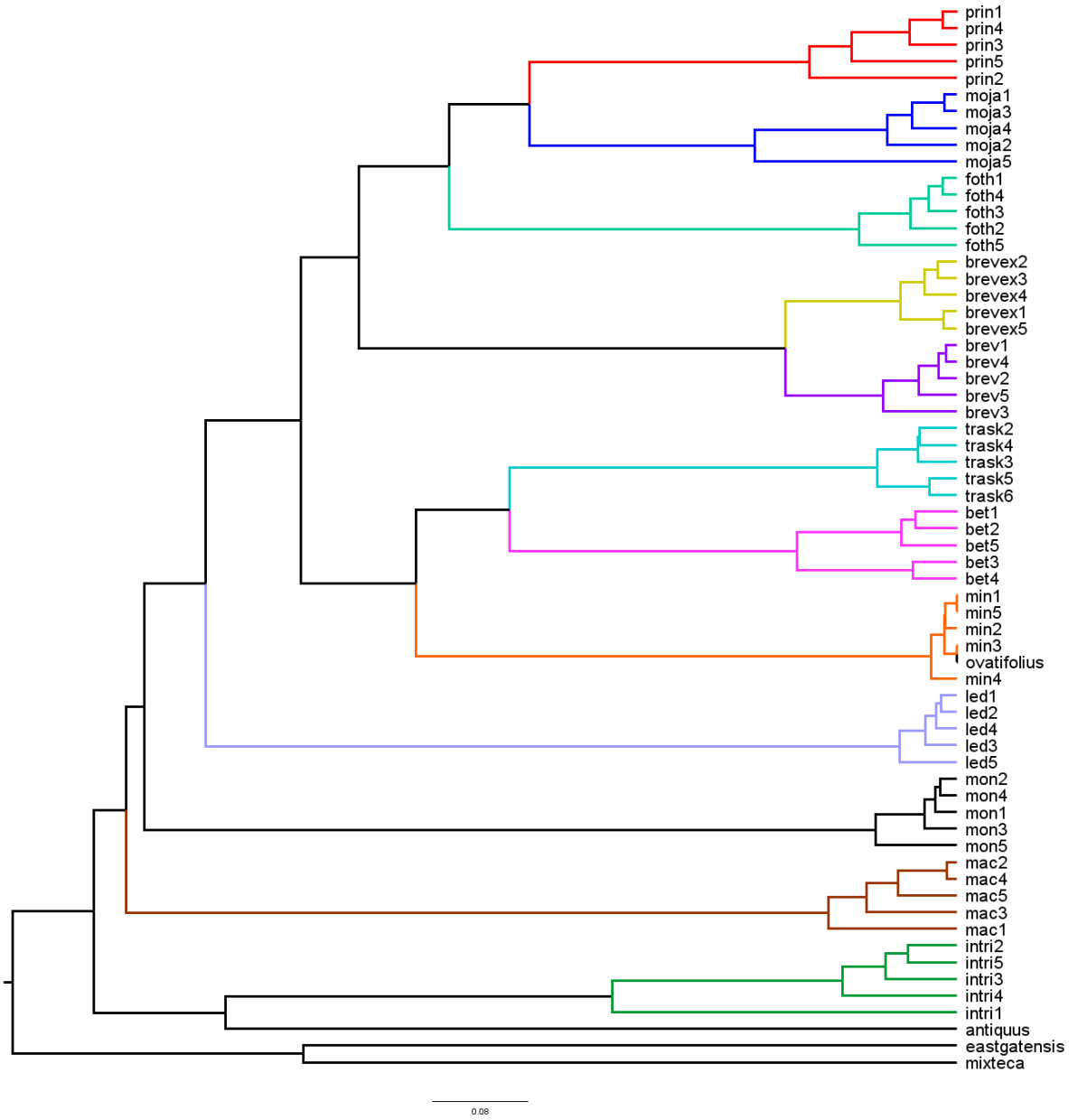

Figure 15. UPGMA dendrogram showing relationships between extant and extinct species using all data (with length, width, and length/width ratio coded as missing for fossil taxa). Leaves are named by an abbreviation of the species name and then their individual number. moja= *Cercocarpus mojadensis.* prin = *C. pringlei*. foth = *C. fothergilloides*. brevex = *C. breviflorus* var. *eximius*. brev = *C. breviflorus* var. *breviflorus*. bet = *C. betuloides*. trask = *C. traskiae*. min = *C. minutiflorus*. ovatifolius= † *C. ovatifolius.* antiquus = † *C. antiquus*. eastgatensis = † *C.*

eastgatensis. mixteca = † *C. mixteca*. mac = *C. macrophyllus*. mon = *C. montanus*. intri = *C. intricatus*. led = *C. ledifolius*.

Institution	Herbarium Number	Collector	Collection number	Genus	Species	Variety	Authority	State/Province	Country	Genbank SRA	Total Read Number	Number Reads Mapped	Number of Genes	Paralogs WarningLong	Citation	Morphology Code
OKL		Leann Monaghan, Abby Moore	16	Cercocarpus	montanus		Raf.	New Mexico	USA							mon1-5
OKL		Leann Monaghan, Abby Moore	20	Cercocarpus	montanus		Raf.	New Mexico	USA		49652506	35001201	20	5		
OKL		Leann Monaghan, Abby Moore	21	Cercocarpus	montanus		Raf.	New Mexico	USA							
OKL		Leann Monaghan, Abby Moore	23	Purshia	stansburyana		(Torr.) Henrickson	Arizona	USA							
OKL		Leann Monaghan, Abby Moore	26	Purshia	stansburyana		(Torr.) Henrickson	Arizona	USA		49354052	34186364	318	10		
OKL		Leann Monaghan, Abby Moore	28	Purshia	stansburyana		(Torr.) Henrickson	Arizona	USA							
OKL		Leann Monaghan, Abby Moore	30	Purshia	stansburyana		(Torr.) Henrickson	Arizona	USA							
OKL		Leann Monaghan, Abby Moore	36	Cercocarpus	intricatus		S. Watson	Utah	USA							
OKL		Leann Monaghan, Abby Moore	37	Purshia	stansburyana		(Torr.) Henrickson	Utah	USA							
OKL		Leann Monaghan, Abby Moore	39	Cercocarpus	intricatus		S. Watson	Utah	USA							
OKL		Leann Monaghan, Abby Moore	41	Purshia	stansburyana		(Torr.) Henrickson	Utah	USA							
OKL		Leann Monaghan, Abby Moore	43	Purshia	tridentata		(Pursh) DC.	Utah	USA							
OKL		Leann Monaghan, Mary Monaghan, Jackie Dougan	48	Dryas	drummondii		Richardson ex Hook.	Oregon	USA		58866047	34968070	345	1		
OKL		Leann Monaghan, Mary Monaghan, Jackie Dougan	53	Purshia	tridentata		(Pursh) DC.	Oregon	USA		38397404	26719901	225	0		
UC Botanical Garden at Berkeley	56.0907			Cercocarpus	traskiae		Eastw.	California	USA		14687720	8579909	342	1		
UC Botanical Garden at Berkeley	54.0829			Cercocarpus	minutiflorus		Abrams	California	USA							min1-5
UC Botanical Garden at Berkeley	65.1186	W. Roderick		Cercocarpus	betuloides		Nutt.	California	USA		588	63	1	0		bet1-5
OKL			2947	Cercocarpus	ledifolius		Nutt.				147452	941	1	0		
OKL			2548	Purshia	tridentata		(Pursh) DC.				19080148	14545399	344	20		
OKL		Marisa Szubryt & Abbot, JR.	2184	Cercocarpus	montanus		Raf.	Texas	USA							

ARIZ	393935	Culberto Angel Cruz-Espinosa and E Mzrtinez S.	382	Cercocarpus	pringlei		(C.K. Schneid.) Rydb.	Oaxaca	USA		14707002	6312855	342	4		pring1-5
TEX-LL	447641	Hendrickson	24638	Cercocarpus	fothergilloides	mojadednsis	(C.K. Schneid.) Henrickson	Coahuila	Mexico		58130	13926	117	0		moja1-5
BRY	636236	Robert L. Johnson, SM Clark, RC Mower	6079	Cercocarpus	brevifloru s		A. Gray	Arizona	USA							brev5
TEX	503840	Hendrickson	24816	Cercocarpus	macrophyllus		C.K. Schneid.	Nuevo Leon	Mexico							mac1-5
TEX	432786	Hendrickson and B. VandenHeuvel	23541	Cercocarpus	fothergilloides		C.K. Schneid.	Nuevo Leon	Mexico		64368	553	0	0		foth1-5
TEX	420951	Hendrickson	23348A	Purshia	mexicana		(D. Don) S.L. Welsh	Durango	Mexico		10960	22	0	0		
TEX	432699	Hendrickson and B. VandenHeuvel	23552	Purshia	plicata		(D. Don) Henrickson	Nuevo Leon	Mexico		289546	1717	3	0		
TEX	432812	Hendrickson and B. VandenHeuvel	23567	Purshia	ericifolia		(Torr. ex A. Gray) Henrickson	Coahuila	Mexico		42128446	31450744	193	1		
ASU	186213	Bob Denham		Purshia	subintegra		(Kearney) Henrickson	Arizona	USA		29608674	22652463	244	1		
ASU	210872	John L. Anderson		Purshia	subintegra		(Kearney) Henrickson	Arizona	USA		20709058	15916679	197	3		
UCR	UCR0042527	Jon P. Rebman		Chamaebatia	australis		(Brandege) Abrams	California	USA		233048	77383	199	2		
UCR	UCR0042517	Jon P. Rebman		Chamaebatia	australis		(Brandege) Abrams	California	USA		272058	105439	239	0		
UCR	UCR0042517	Jon P. Rebman		Chamaebatia	australis		(Brandege) Abrams	California	USA		106858	49048	229	0		
OKL		Abby Moore, J. Moore, B. Moore		Purshia	tridentata		(Pursh) DC.	Utah	USA		19080148	14545399	344	20		
LE	s.n.	A. V. Grebenjuk	192KAZ52 4-527	Pyrus	regelii			Turkistan	Kazakhstan	SRR12879558	1503914	107444	262	0	(Ufimov et al. 2021)	
	s.n.	S. Píšová	s.n	Prunus	tenella			Central Bohemia	Czech Republic	SRR12879560	1501078	29160	132	7	(Ufimov et al. 2021)	
TRT	12528	T. A. Dickinson	D1378	Crataegus	punctata			Ontario	Canada	SRR12879548	1596382	163090	317	1	(Ufimov et al. 2021)	
K	s.n.	Weiblen G.	1845	Artocarpus	vrieseanus					ERR5033888	1205220	42700	131	1	(Baker et al. 2021)	
K	K0012577 52	Regelado J. and Takeuchi, W.	1518	Parartocarpus	venenosus					ERR5033349	1146976	239687	333	29	(Baker et al. 2021)	

Appendix 1. Voucher table.

Code	Linear	Ovate	Rhombic	Elliptic	Obovate	Lanceolate	Oblanceolate	Length	widest_point	L_W_ratio	half_serration	1/4_serration	3/4_serration	Crenate_Margin	Entire_Margine	Serrate_Margin	Cuniate_Base	Acuminate_Base	Obtuse_Apex	Mucronate_Apex
brevex1	0	0	0	0	1	0	1	28	9	3.11	1	1	0	1	0	0	1	0	0	1
brevex2	0	0	0	0	1	0	1	18	6	3.00	1	1	0	1	0	0	1	0	0	1
brevex3	0	0	0	0	1	0	1	14.5	5	2.90	1	1	0	1	0	0	1	0	0	1
brevex4	0	0	0	0	1	0	1	19	8	2.38	1	1	0	1	0	0	1	0	0	1
brevex5	0	0	0	0	1	0	1	26	8	3.25	1	1	0	1	0	0	1	0	0	1
led1	0	0	0	1	0	1	1	18.5	5	3.70	0	0	0	0	1	0	1	0	0	0
led2	0	0	0	1	0	1	1	18	5.5	3.27	0	0	0	0	1	0	1	0	0	0
led3	0	0	0	1	0	1	1	17	4	4.25	0	0	0	0	1	0	1	0	0	0
led4	0	0	0	1	0	1	1	15	4.5	3.33	0	0	0	0	1	0	1	0	0	0
led5	0	0	0	1	0	1	1	11.5	5	2.30	0	0	0	0	1	0	1	0	0	0
intri1	1	0	0	0	0	1	0	19	1.5	12.67	0	0	0	0	1	0	1	0	0	0
intri2	1	0	0	0	0	1	0	13	1.5	8.67	0	0	0	0	1	0	1	0	0	0
intri3	1	0	0	0	0	1	0	14	1.9	7.37	0	0	0	0	1	0	1	0	0	0
intri4	1	0	0	0	0	1	0	21	1.8	11.67	0	0	0	0	1	0	1	0	0	0
intri5	1	0	0	0	0	1	0	15	1.5	10.00	0	0	0	0	1	0	1	0	0	0
brev1	0	0	0	0	1	0	1	14	4	3.50	1	1	0	1	0	0	1	0	0	1
brev2	0	0	0	0	1	0	1	14	3.5	4.00	1	1	0	1	0	0	1	0	0	1
brev3	0	0	0	0	1	0	1	19	10	1.90	1	1	0	1	0	0	1	0	0	1
brev4	0	0	0	0	1	0	1	16	4.5	3.56	1	1	0	1	0	0	1	0	0	1
brev5	0	0	0	0	1	0	1	14.4	5.5	2.62	1	1	0	1	0	0	1	0	0	1
mon1	0	0	1	0	1	0	0	35	12	2.92	1	1	0	0	0	1	0	1	0	0
mon2	0	0	1	0	1	0	0	34	10	3.40	1	1	0	0	0	1	0	1	0	0
mon3	0	0	1	0	1	0	0	36	9	4.00	1	1	0	0	0	1	0	1	0	0
mon4	0	0	1	0	1	0	0	37	11	3.36	1	1	0	0	0	1	0	1	0	0
mon5	0	0	1	0	1	0	0	21	8	2.63	1	1	0	0	0	1	0	1	0	0
bet1	0	1	1	1	1	1	1	39.5	25	1.58	1	1	0	1	0	0	1	0	0	0
bet2	0	1	1	1	1	1	1	32	23	1.39	1	1	0	1	0	0	1	0	0	0
bet3	0	1	1	1	1	1	1	58	33	1.76	1	1	0	1	0	0	1	0	0	0
bet4	0	1	1	1	1	1	1	56	38	1.47	1	1	0	1	0	0	1	0	0	0
bet5	0	1	1	1	1	1	1	33	18	1.83	1	1	0	1	0	0	1	0	0	0
trask2	0	1	0	1	1	0	0	22	17.5	1.26	1	1	0	1	0	0	1	0	0	0
trask3	0	1	0	1	1	0	0	27	21.5	1.26	1	1	0	1	0	0	1	0	0	0
trask4	0	1	0	1	1	0	0	29	17.5	1.66	1	1	0	1	0	0	1	0	0	0
trask5	0	1	0	1	1	0	0	34.5	28.6	1.21	1	1	0	1	0	0	1	0	0	0
trask6	0	1	0	1	1	0	0	32	25.6	1.25	1	1	0	1	0	0	1	0	0	0
min1	0	1	0	0	1	0	0	19	14	1.36	0	1	0	1	0	0	1	0	0	0
min2	0	1	0	0	1	0	0	17	13	1.31	0	1	0	1	0	0	1	0	0	0
min3	0	1	0	0	1	0	0	15	11.5	1.30	0	1	0	1	0	0	1	0	0	0
min4	0	1	0	0	1	0	0	12	12	1.00	0	1	0	1	0	0	1	0	0	0
min5	0	1	0	0	1	0	0	19	14	1.36	0	1	0	1	0	0	1	0	0	0
foth1	0	0	1	0	0	0	0	28	20	1.40	0	1	0	1	0	0	1	0	0	1
foth2	0	0	1	0	0	0	0	21	16	1.31	0	1	0	1	0	0	1	0	0	1
foth3	0	0	1	0	0	0	0	27.5	17.5	1.57	0	1	0	1	0	0	1	0	0	1
foth4	0	0	1	0	0	0	0	27	21.5	1.26	0	1	0	1	0	0	1	0	0	1
foth5	0	0	1	0	0	0	0	35	29	1.21	0	1	0	1	0	0	1	0	0	1
mac1	0	1	0	0	1	0	0	91	49	1.86	0	0	1	0	0	1	1	0	0	0
mac2	0	1	0	0	1	0	0	73	38	1.92	0	0	1	0	0	1	1	0	0	0
mac3	0	1	0	0	1	0	0	67.5	30.5	2.21	0	0	1	0	0	1	1	0	0	0
mac4	0	1	0	0	1	0	0	75	38	1.97	0	0	1	0	0	1	1	0	0	0
mac5	0	1	0	0	1	0	0	80	44	1.82	0	0	1	0	0	1	1	0	0	0
moja1	0	1	0	1	0	0	0	23	18.5	1.24	0	1	0	1	0	0	1	0	0	1
moja2	0	1	0	1	0	0	0	18	13.5	1.33	0	1	0	1	0	0	1	0	0	1
moja3	0	1	0	1	0	0	0	25	19.5	1.28	0	1	0	1	0	0	1	0	0	1
moja4	0	1	0	1	0	0	0	29	23.5	1.23	0	1	0	1	0	0	1	0	0	1
moja5	0	1	0	1	0	0	0	26.5	19	1.39	0	1	0	1	0	0	1	0	0	1
prin1	0	1	0	0	1	0	0	29.5	18.5	1.59	0	1	0	1	0	0	1	0	0	1
prin2	0	1	0	0	1	0	0	50	28	1.79	0	1	0	1	0	0	1	0	0	1
prin3	0	1	0	0	1	0	0	31	24	1.29	0	1	0	1	0	0	1	0	0	1
prin4	0	1	0	0	1	0	0	27	19	1.42	0	1	0	1	0	0	1	0	0	1
prin5	0	1	0	0	1	0	0	18	10	1.80	0	1	0	1	0	0	1	0	0	1
antiquus	0	1	0	0	0	1	0				1	1	0	1	0	0	1	0	0	0
eastgatensis	0	0	0	1	1	0	0				0	1	0	1	0	0	0	1	0	1
ovatifolius	0	1	0	0	1	0	0				0	1	0	0	0	0	1	0	0	0
mixteca	0	0	0	0	1	0	0				0	1	0	1	0	0	0	1	1	0

Appendix 2a. Morphology Matrix

Acute_Apex	Acuminate_Apex	Long_Trichomes_top	Long_Trichomes_bottom	Short_Trichomes_top	Short_Trichomes_bottom	Curly_Trichomes_top	Curly_Trichomes_bottom	Straight_Trichomes_top	Straight_Trichomes_botto	Guard_hairs_top	Guard_hairs_bottom
0	0	1	1	0	0	0	1	1	1	0	0
0	0	1	1	0	0	0	1	1	1	0	0
0	0	1	1	0	0	0	1	1	1	0	0
0	0	1	1	0	0	0	1	1	1	0	0
0	0	1	1	0	0	0	1	1	1	0	0
1	0	1	0	0	1	1	1	0	0	0	0
1	0	1	0	0	1	1	1	0	0	0	0
1	0	1	0	0	1	1	1	0	0	0	0
1	0	1	0	0	1	1	1	0	0	0	0
1	0	1	0	0	1	1	1	0	0	0	0
0	1	0	1	1	0	1	0	0	1	0	0
0	1	0	1	1	1	0	1	1	1	0	0
0	1	0	1	1	1	0	1	1	1	0	0
0	1	0	1	1	1	0	1	1	1	0	0
0	1	0	1	1	1	0	1	1	1	0	0
0	0	1	1	0	0	0	1	1	1	0	1
0	0	1	1	0	0	0	1	1	1	0	1
0	0	1	1	0	0	0	1	1	1	0	1
0	0	1	1	0	0	0	1	1	1	0	1
0	0	1	1	0	0	0	1	1	1	0	1
1	0	1	1	0	0	0	1	1	1	0	1
1	0	1	1	0	0	0	1	1	1	0	1
1	0	1	1	0	0	0	1	1	1	0	1
1	0	1	1	0	0	0	1	1	1	0	1
1	0	1	1	0	0	0	1	1	1	0	1
1	0	0	0	1	1	0	1	1	1	0	1
1	0	0	0	1	1	0	1	1	1	0	1
1	0	0	0	1	1	0	1	1	1	0	1
1	0	0	0	1	1	0	1	1	1	0	1
1	0	0	0	1	1	0	1	1	1	0	1
1	0	1	1	0	0	1	1	0	1	0	1
1	0	1	1	0	0	1	1	0	1	0	1
1	0	1	1	0	0	1	1	0	1	0	1
1	0	1	1	0	0	1	1	0	1	0	1
1	0	1	1	0	0	1	1	0	1	0	1
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
0	0	0	1	1	0	1	1	0	1	0	0
0	0	0	1	1	0	1	1	0	1	0	0
0	0	0	1	1	0	1	1	0	1	0	0
0	0	0	1	1	0	1	1	0	1	0	0
0	0	0	1	1	0	1	1	0	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
1	0	0	0	1	1	0	0	1	1	0	0
0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	1	1	1	1	0	0	0	0
0	0	0	0	1	0	1	1	0	0	0	0
0	0	1	0	0	1	1	1	0	0	0	0
0	0	1	0	0	1	1	1	0	0	0	0
0	0	1	0	0	1	1	1	0	0	0	0
0	0	1	0	0	1	1	1	0	0	0	0
0	0	1	0	0	1	1	1	0	0	0	0
0	1										
0	0										
1	0										
0	0										

Appendix 2b. Morphology Matrix continued

Literature Cited

Baker W.J, Bailey P., Barber V., Barker A., Bellot S., Bishop D., Botigué L.R., et al. 2021. A comprehensive phylogenomic platform for exploring the Angiosperm Tree of Life. *Systematic Biology* 71: 301–319.

Baker W.J., Dodsworth S., Forest F., Graham S.W., Johnson M.G., McDonnell A., Pokorny L., et al. 2021. Exploring ANGIOSPERMS353: An open, community toolkit for collaborative phylogenomic research on Flowering plants. *American Journal of Botany* 108: 1059–1065.

Bolger A.M., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. Aug 1;30(15):2114-20. doi: 10.1093/bioinformatics/btu170. Epub 2014 Apr 1. *PMID*: 24695404; PMCID: PMC4103590.

Capella-Gutierrez S., Silla-Martinez J.M, Gabaldon T. *Bioinformatics* 2009 25: 1972-1973. http://trimal.cgenomics.org/publications

Chase M.W., Christenhusz M.J.M., Fay M.F., Byang J.W., Judd W.S., Soltis D.E., Mabberly D.J., et al. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean society* 181(1): 1-20.

Chen X., Li J., Cheng T., Zhang W., Liu Y., Wu P., Yang X., et al. 2020. Molecular systematics of Rosoideae (Rosaceae). *Plant Systematics and Evolution* 306.

Dunkle, M. B. 1940. A Revision of the Channel Islands Forms of Cercocarpus. *Bulletin of the Southern California Academy of Sciences* 39(1): 2.

Chifman, J., L. Kubatko. 2014. Quartet inference from SNP data under the coalescent, *Bioinformatics* 30(23): 3317-3324.

Chifman, J., L. Kubatko. 2015. Identifiability of the unrooted species tree topology under the coalescent model with time-reversible substitution processes, site-specific rate variation, and invariable sites. *Journal of Theoretical Biology* 374: 35-47.

Chessel D., Dufour A., Thioulouse J. 2004. The ade4 Package – I: One-Table Methods. *R News* 4(1): 5–10. https://cran.r-project.org/doc/Rnews/.

Dray S., Dufour A. (2007). The ade4 Package: Implementing the Duality Diagram for Ecologists. *Journal of Statistical Software*: 22(4): 1–20. doi:10.18637/jss.v022.i04.

Hendrickson J., Vanden Heuvel B.D., American Journal of Botany Editorial Committee. 2014. Cercocarpus. *In* Flora of North America, 332–338. Oxford University Press, New York City, New York.

Hornik K., 2023. *clue: Cluster Ensembles*. R package version 0.3-65. https://CRAN.R-project.org/package=clue.

Johnson M.G., Gardner E. M., Liu Y., Medina R., Goffinet B., Shaw A.J., Zerega N.J.C., Wickett N.J. 2016. HybPiper: Extracting Coding Sequence and Introns for Phylogenetics from High-Throughput Sequencing Reads Using Target Enrichment. *Applications in Plant Sciences*. 4(7).

Johnson M.G., Pokorny K., Dodsworth S., Botigué L.R., Cowan R.S., Devault A., Eiserhardt W.L., et al. 2018. A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using K-medoids clustering. *Systematic Biology* 68: 594–606.

Jones M. E. 1880. *Botanical Gazette A Paper of Botanical Notes*. Botanical Gazette 5(12): 154.

Katoh K., Misawa K., Kuma K., Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research* 30(14): 3059-66.

Kembel S., Cowan P., Helmus M., Cornwell W., Morlon H., Ackerly D., Blomberg S., Webb C. 2010. Picante: R tools for integrating phylogenies and ecology. *Bioinformatics* 26: 1463–1464.

Lis, R.A. 1992. Leaf architectural survey of Cercocarpus (Rosaceae) and its systematic significance. *International Journal of Plant Sciences* 153: 258–272.

Martin, F.L. 1950. A revision of Cercocarpus. *Brittonia* 7: 91.

Mortenson T. 1973. Ecological variations in the leaf anatomy of selected species of Cercocarpus. *Aliso* 8: 19–48.

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., Von Haeseler A., Lanfear R. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evol*ution 37:1530-1534.

Neto J., 2012, Gower's Distance Modification (Generalization of Gower's distance). *R Functions.* Website. https://rfunctions.blogspot.com/2012/07/gowers-distance-modification.html.

Potter D., Eriksson T., Evans R.C., Oh S., Smedmark J.E., Morgan D.R, Kerr M., et al. 2007. Phylogeny and classification of Rosaceae. *Plant Systematics and Evolution* 266: 5–43.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.

Rabiee M., Sayyari E., Mirarab S. 2019. Multi-Allele Species Reconstruction Using ASTRAL. *Molecular Phylogenetics and Evolution* 130: 286–96.

Rasband W.S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.net/ij/, 1997-2018.

Revell L. 2024. phytools 2.0: an updated R ecosystem for phylogenetic comparative methods (and other things). *PeerJ* 12.

Rydberg A. 1913. North American flora. vol. 22: Part 5: Rosales. The New York Botanical Garden, New York, NY.

Schneider C. 1905. Illustriertes Handbuch der Laubholzkund. vol. 1: Part 3. 530. 1905.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312-3.

Stutz H.C., Thomas L.K. 1964. Hybridization and introgression in Cowania and Purshia. *Evolution* 18: 183.

Swensen S.M., Mullin B.C. 1997. The impact of molecular systematics on hypotheses for the evolution of root nodule symbioses and implications for expanding symbioses to new host plant genera. *Plant and Soil*. 194 (1/2): 185–192.

Swofford, D. L. 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4 3.99.169.0. Sinauer Associates, Sunderland, Massachusetts.

Vanden Heuvel, 2002, Molecular Systematics of Cercocarpus H.B.K. (Rosaceae), Dissertation for University of Texas at Austin

Vanden Heuvel B.D., Linder C.R. 2001. A Molecular Phylogeny for Cercocarpus H.B.K. (Rosaceae) Using the External Transcribed Spacer of the Nuclear Ribosomal Repeat. *USDA Forest Service Proceedings*.

Ufimov R., Zeisek V., Píšová S., Baker W.J., Fér T., Van Loo M., Dobeš C., Schmickl R. 2021. Relative performance of customized and universal probe sets in target enrichment: A case study in subtribe Malinae. *Applications in Plant Sciences* 9.

Velasco de León, P., Cevallos-Ferriz S.R.S. 2000. Leaves of Cercocarpus Mixteca n. sp. (Rosaceae) from Oligocene sediments, near Tepexi de Rodríguez, Puebla. *Review of Palaeobotany and Palynology* 111: 285–294.

Walker S.C., Turley D. 1999. Characteristics of Mountain Mahogany (Cercocarpus) Species and Hybrids in Utah Hybrid Zone. *USDA Forest Service Proceedings* RMRS-P-11.

Wickham H. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Xiang Y. Huang C.H., Hu Y., Wen J., Li S., Yi T., Chen H., et al. 2017. Evolution of Rosaceae fruit types based on nuclear phylogeny in the context of geological times and genome duplication. *Molecular Biology and Evolution* 34: 1026–1026.

Zhang S., Jin J., Chen S., Chase M.W., Soltis D.E., Li H. Yang J., et al. 2017. Diversification of Rosaceae since the late cretaceous based on plastid phylogenomics. *New Phytologist* 214: 1355–1367.

Acknowledgements

Microscopy data collection was performed at the Samuel Roberts Noble Microscopy Laboratory with the help of Dr. Preston Larson.

Thank you to Dr. Abby Moore, Dr. Rick Lupia, and Dr. Lara Souza for help developing my thesis and for guiding me along the way

Help in the field: Marisa Szubryt, Mary Monaghan, Jackie Dougan

Help in the lab: Marisa Szubryt, Maddie Medina

Help with morphological analyses: Claire Curry, Rick Lupia

[Some of] The computing for this project was performed at the OU Supercomputing Center for Education & Research (OSCER) at the University of Oklahoma (OU).

Sources of specimens: The Robert Bebb Herbarium (OKL), University of California Riverside Herbarium (UCR), UC Botanical Garden at Berkeley, The Arizona State University Vascular Plant Herbarium (ASU), Billie L. Turner Plant Resources Center at University of Texas at Austin (TEX-LL), Stanley L. Welsh Herbarium at Brigham Young University (BRY), and the University of Arizona Herbarium (ARIZ)