

RESEARCH ARTICLE

Complete chloroplast genome comparisons for *Pityopsis* (Asteraceae)E. Anne Hatmaker¹ , Phillip A. Wadl², Timothy A. Rinehart³, Jennifer Carroll⁴, Thomas S. Lane¹, Robert N. Trigiano^{1*}, Margaret E. Staton^{1*} , Edward E. Schilling^{5*}

1 Department of Entomology and Plant Pathology, University of Tennessee, Knoxville, Tennessee, United States of America, **2** U.S. Department of Agriculture, Agricultural Research Service, U.S. Vegetable Laboratory, Charleston, South Carolina, United States of America, **3** U.S. Department of Agriculture, Agricultural Research Service, Crop Production and Protection, Beltsville, Maryland, United States of America, **4** U.S. Department of Agriculture, Agricultural Research Service, Thad Cochran Southern Horticultural Laboratory, Poplarville, Mississippi, United States of America, **5** Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, Tennessee, United States of America

* Current address: Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee, United States of America

* rtrigian@utk.edu (RNT); mstaton1@utk.edu (MES); eschilling@utk.edu (EES)


 OPEN ACCESS

Citation: Hatmaker EA, Wadl PA, Rinehart TA, Carroll J, Lane TS, Trigiano RN, et al. (2020) Complete chloroplast genome comparisons for *Pityopsis* (Asteraceae). PLoS ONE 15(12): e0241391. <https://doi.org/10.1371/journal.pone.0241391>

Editor: Tzen-Yuh Chiang, National Cheng Kung University, TAIWAN

Received: June 30, 2020

Accepted: October 13, 2020

Published: December 28, 2020

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Data Availability Statement: Illumina reads are available through the NCBI SRA accession numbers SRR4457826–4457837. The chloroplast genome sequence for *Pityopsis falcata* is associated with NCBI accession KY045817.1. Chloroplast and nuclear alignments are available through figshare <https://doi.org/10.6084/m9.figshare.13260023.v2> and <https://doi.org/10.6084/m9.figshare.13260056.v1>, respectively.

Funding: This work was supported by the U.S. Department of Agriculture (USDA/MOA number

Abstract

Pityopsis includes several regionally and one federally endangered species of herbaceous perennials. Four species are highly localized, including the federally endangered *P. ruthii*. The genus includes several ploidy levels and interesting ecological traits such as drought tolerance and fire-dependent flowering. Results from previous cladistic analyses of morphology and from initial DNA sequence studies did not agree with one another or with the infrageneric taxonomic classification, with the result that infrageneric relationships remain unresolved. We sequenced, assembled, and compared the chloroplast (cp) genomes of 12 species or varieties of *Pityopsis* to better understand generic evolution. A reference cp genome 152,569 bp in length was assembled *de novo* from *P. falcata*. Reads from other sampled species were then aligned to the *P. falcata* reference and individual chloroplast genomes were assembled for each, with manual gapfilling and polishing. After removing the duplicated second inverted region, a multiple sequence alignment of the cp genomes was used to construct a maximum likelihood (ML) phylogeny for the twelve cp genomes. Additionally, we constructed a ML phylogeny from the nuclear ribosomal repeat region after mapping reads to the *Helianthus annuus* region. The chloroplast phylogeny supported two clades. Previously proposed clades and taxonomic sections within the genus were largely unsupported by both nuclear and chloroplast phylogenies. Our results provide tools for exploring hybridity and examining the physiological and genetic basis for drought tolerance and fire-dependent flowering. This study will inform breeding and conservation practices, and general knowledge of evolutionary history, hybridization, and speciation within *Pityopsis*.

Introduction

Pityopsis is a small genus of Asteraceae with its center of diversity in the southeastern United States [1]. The genus includes a wide variety of ploidy levels across species and a large

58-6404-1-637 received by RT, <https://www.usda.gov/>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

geographic range throughout southeastern North America, in Mexico and Central America, and in the Bahamas [2]. Notably, four species of *Pityopsis* are rare and of conservation concern: *P. ruthii* (listed as endangered federally), *P. flexuosa* (listed as endangered by the state of Florida), *P. falcata* (listed as endangered by the state of Connecticut and of special concern by the state of Rhode Island), and *P. pinifolia* (listed as threatened by the state of Georgia). *Pityopsis* has been the subject of several phylogenetic studies [3–5], but intrageneric relationships for all species and varieties in the genus have not been fully resolved, resulting in significant variation in the number of species recognized within *Pityopsis*. The genus includes many polyploid varieties and several ecologically adaptive traits such as fire-stimulated flowering [3,6] and drought-tolerance [7]. Studying species relationships often allows for better evolutionary understanding of traits.

Phylogenetic studies are conducted to clarify taxonomic relationships and classification [8]. They have proved useful for understanding plant-pathogen interactions [9] and community ecology [10]. Additionally, phylogenetic studies can translate to predictions of phenological response and adaptation in related species, especially adaptation in regard to climate change [11]. Phylogenies have additional use in studies focused on evolutionary history [12]. *Pityopsis* is an excellent candidate for such analysis as the genus includes species that vary for traits such as fire-adaptive flowering, as well as species with varying ploidy levels [4]. In *Pityopsis*, species distinctions are not well understood and require further resolution, which has been difficult due to the differing ploidy levels in the genus and apparent hybridization. For example, in *P. graminifolia* alone there are three ploidy levels present in different varieties of the species: diploid (*P. graminifolia* var. *graminifolia*), tetraploid (var. *latifolia*), and hexaploid (var. *tracyi*) [13]. Analyzing datasets with a range of ploidy levels creates difficulties when using biparental nuclear markers. However, with a well-supported phylogeny based on molecular markers, *Pityopsis* could be used to examine the evolution of adaptive traits and the role of hybridity in the evolution of polyploidy.

Nuclear microsatellites have been developed for two different *Pityopsis* species and chloroplast microsatellites have been developed for one species [14–16]. However, whole chloroplast (cp) genomes are lacking for all species in the genus. With the availability of next-generation sequencing, phylogenetic studies using entire cp genomes is becoming more reliable and common, especially for closely related species [17]. Chloroplast genome sequences have become a convenient way to find repetitive sequences and single nucleotide polymorphisms (SNPs) that could be used for further ecological and evolutionary studies, as well as clarifying taxonomy in genera with muddled history [18]. Many similar studies have been conducted on phylogenetic relationships within economically important plants, such as wheat, rice, and maize [19], strawberry [20], and cotton [21]. Using cp genomes to analyze the species relationships within *Pityopsis* allows further studies regarding past polyploid events to use a simplified system due to the haploid nature of chloroplasts, though only the maternal line is revealed in the case of species arising from hybridization events resulting in allopolyploidy.

Pityopsis includes seven species: *P. aspera* (Shuttlew. ex Small) Small, *P. falcata* (Pursh) Small, *P. flexuosa* (Nash) Small, *P. graminifolia* (Michx.) Nutt., *P. oligantha* (Chapm. ex Torr. & Gray) Small, *P. pinifolia* (Ell.) Nutt., and *P. ruthii* (Small) Small [15]. Both *P. aspera* and *P. graminifolia* have multiple varieties, some of which have previously been recognized as separate species [22]. *Pityopsis* is endemic to the eastern United States, and though *P. graminifolia* and *P. aspera* have a large range, other species in the genus are more localized, such as *P. ruthii* and *P. flexuosa*. All species are perennial and have yellow inflorescences, as indicated by the common name for plants in the genus, goldenaster [13].

The division of *Pityopsis* into sections remains unresolved. Semple and Bowers [13], divided the genus into two sections: section *Pityopsis* with *P. falcata*, *P. flexuosa*, *P. pinifolia*, and *P.*

ruthii, and section *Graminifoliae* with *P. aspera*, *P. graminifolia*, and *P. oligantha*. However, the phylogenetic analysis conducted by Gowe and Brewer [3] based on morphology divided the species into two clades that did not coincide with the sectional classification, and were referred to informally as the Falcata clade, which includes *P. falcata*, *P. flexuosa*, *P. graminifolia*, *P. pinifolia*, and *P. oligantha*, and the Aspera clade, which includes *P. aspera*, *P. adenolepis*, and *P. oligantha*. In contrast, a molecular study utilized sequences from chloroplast and nuclear regions of all seven species and concluded that two new clades should be named: Ruthii and Flexuosa [4]. Clade Ruthii includes *P. falcata*, *P. pinifolia*, *P. ruthii*, and *P. graminifolia* var. *latifolia*. Splitting the species *P. graminifolia*, clade Flexuosa includes *P. graminifolia* var. *aequilifolia*, *P. graminifolia* var. *tenuifolia*, and *P. graminifolia* var. *graminifolia*, as well as *P. aspera*, *P. adenolepis*, and *P. oligantha*. Both the 2005 and the 2008 studies include *P. adenolepis* as a separate species from *P. aspera* as per Clewell [22], although Nesom [2] considers them synonymous based on his interpretation of morphology. We have continued to use the taxonomic designations set forth by Semple and Bowers [13] as there is no agreement on naming of varieties or species even as recently as 2019 [2,23]. With little to no consensus between morphological and molecular studies, any information derived from molecular studies within the genus can only improve taxonomic resolution.

In this study, 12 *Pityopsis* chloroplast genomes were assembled, compared to other Asteraceae chloroplast genomes, and used to construct phylogenetic trees. To provide additional information from the biparentally inherited nuclear genome, we also constructed phylogenetic trees using the nuclear external transcribed spacer (ETS) region, which is highly conserved, to complement the chloroplast phylogenies and to add to our knowledge of hybridity and evolution of the *Pityopsis* genus. Although resolving the taxonomic problems of the genus is beyond the scope of the current study, here we present data on chloroplast genomes that will provide a foundation for future studies of *Pityopsis*.

Methods

Ethics statement for plant collection

Leaf tissue of seven species including seven varieties of *Pityopsis* was collected from the southeastern United States (Table 1). Leaf tissue from plants maintained in a greenhouse at the University of Tennessee was collected for *P. graminifolia* var. *tracyi*. This study used tissue collected in 2010 and 2013 and kept at -80°C from *P. ruthii* [14], *P. falcata*, and *P. graminifolia*

Table 1. State of tissue after collection, date, and location of *Pityopsis* individuals.

Species	Type of tissue	Year collected	Location
<i>P. aspera</i> var. <i>adenolepis</i>	Dried	2014	South Carolina
<i>P. aspera</i> var. <i>aspera</i>	Dried	2016	Florida
<i>P. falcata</i>	Dried	2010	Rhode Island
<i>P. flexuosa</i>	Dried	2015	Florida
<i>P. graminifolia</i> var. <i>aequilifolia</i>	Dried	2015	Florida
<i>P. graminifolia</i> var. <i>graminifolia</i>	Frozen	2014	South Carolina
<i>P. graminifolia</i> var. <i>latifolia</i>	Frozen	2013	Tennessee
<i>P. graminifolia</i> var. <i>tenuifolia</i>	Frozen	2014	South Carolina
<i>P. graminifolia</i> var. <i>tracyi</i>	Fresh	2014	Florida
<i>P. oligantha</i>	Dried	2015	Florida
<i>P. pinifolia</i>	Dried	2016	South Carolina
<i>P. ruthii</i>	Frozen	2010	Tennessee

<https://doi.org/10.1371/journal.pone.0241391.t001>

var. *latifolia* [15], respectively. Original plant material for *P. ruthii* and *P. graminifolia* var. *latifolia* was collected in Polk County, Tennessee under permits from the Tennessee Valley Authority (TE117405-2) and the United States Fish and Wildlife Service (TE134817-1). All other tissue was collected from public land which required no permit, under permit from the South Carolina Department of Natural Resources, or in coordination with United States Forest Service scientists. *Pityopsis aspera* var. *adenolepis*, *P. graminifolia* var. *tenuifolia*, and *P. graminifolia* var. *graminifolia* were collected from Florence and Darlington counties in South Carolina in 2014. Tissue for *P. graminifolia* var. *aequilifolia* was collected in Ocala National Forest in 2015. For *P. oligantha* and *P. flexuosa*, tissue was collected in 2014 and 2015 from Liberty and Wakulla counties, respectively, in Florida. Tissue was collected in 2016 for *P. aspera* var. *aspera* in Florida and *P. pinifolia* in the Peachtree Rock Heritage Preserve in Lexington County, South Carolina. Vouchers are available at the Florida State University Herbarium (FSU) for *P. oligantha* and *P. flexuosa* (Anderson 28905 and Anderson 28533, respectively).

Library construction and sequencing

Total genomic DNA (gDNA) was isolated using a DNeasy Plant Mini Kit (Qiagen, Valencia, CA) following manufacturer's protocol. Genomic DNA of all samples was cleaned and concentrated using the Zymo Genomic DNA Clean and Concentrator Kit (Zymo Research Corp., Irvine, CA). The libraries were prepared using the Nextera DNA Library Preparation Kit (Illumina, San Diego, CA). DNA was fragmented using transposase-mediated tagmentation and paired end sequenced using dual indexes. The Illumina MiSeq version 3 sequencing platform (Illumina, San Diego, CA) was used for 250 bp paired-end sequencing of the DNA. Three libraries were pooled for three runs and four pooled for one run. One run was discarded due to low-quality.

Sequence trimming and alignment

The sequence quality of all sequences was checked using FastQC [24] for kmer content, GC content, and average length of reads. Adaptors and low quality ends were trimmed using Trimmomatic v. 0.35 [25]. After trimming, quality was assessed again using FastQC, which showed that overall quality improved in all individuals. Using the program Bowtie2 [26], the data from all individuals was aligned against the chloroplast genome of *Helianthus annuus*, which was downloaded from NCBI (GenBank: DQ383815.1; downloaded November, 2015). *P. falcata* had the highest number of mapped reads after the first round of sequencing, and was therefore selected for *de novo* assembly of a reference cp genome.

Genome assembly and annotation

After mapping *P. falcata* reads to the *H. annuus* chloroplast genome to filter out genomic DNA, the *P. falcata* reads were assembled into a reference cp genome using the program ABySS v 1.5.2 [27], which is designed for short, paired-end reads. Gaps within the draft genome were closed using the "map to reference" option within Geneious 11.1.5 [28] using *P. falcata* reads and default parameters. Additionally, *P. falcata* reads were mapped back to the *P. falcata* reference cp genome to fill short gaps and call variants. Reads from each individual species and variety were mapped to the *P. falcata* reference to generate a consensus sequence, which served as a draft cp genome. The draft genomes were then gapfilled within Geneious [28] using the "map to reference" option. We also assembled a cp genome from a second *P. ruthii* individual for quality control using sequencing data from previous work [14] (S1–S3 Tables) and the same methodology.

The reference genome from *P. falcata* was annotated using DOGMA [29], which is specific to organelle genomes and also identifies tRNAs and rRNAs. The annotations were manually reviewed and edited within Geneious v. 11.1.5 [28]. Visualization of the genome annotation as a gene map was created using the program OGDraw [30].

Alignment and comparison

The *Pityopsis* and *H. annuus* cp genomes (after removal of the duplicate copy of the inverted region) were then aligned using Mauve [31]. Pairwise differences were calculated between all *Pityopsis* cp genomes and the outgroup *H. annuus* cp genome within Mauve [32]. The substitution model was chosen using the corrected Akaike information criterion (AICc) as calculated by JModelTest [33]. The maximum likelihood cp phylogenetic tree was built using RAxML 8.2.11 [34] GTR + GAMMA + I parameter within RAxML [34]. Bootstrap analysis was conducted using 1000 replicates. The consensus tree was drawn with a 10% burn-in and 50% support threshold.

Using Geneious [28], gDNA reads of all 12 *Pityopsis* individuals were mapped to the 9814 bp nuclear ribosomal repeat region of *H. annuus* (KF767534.1) including the 28S, 18S, and 5.8S genes. Consensus sequences were called and gDNA mapped back to the consensus sequences for all *Pityopsis* taxa. The nuclear ribosomal region of *H. annuus* was used as an outgroup and aligned along with the *Pityopsis* ribosomal regions using MUSCLE [35] with anchor optimization, and the substitution model selected using AICc as previously outlined. A maximum likelihood phylogeny was reconstructed in RAxML [34] using the GTR + GAMMA parameter with 1000 replicates for bootstrapping. The consensus tree was drawn with a 10% burn-in and 50% support threshold.

Results

Chloroplast genome sequencing, assembly, and annotation

Using the Illumina MiSeq sequencing platform, we sequenced gDNA and assembled cp genomes for 12 samples representing 7 species from *Pityopsis*, including 7 varieties (TtT 1). Illumina paired-end sequencing produced from 3,451,455 (*P. oligantha*) to 33,339,900 (*P. graminifolia* var. *aequilifolia*) reads per individual (S1 Table). Of these reads, 6,571 (*P. graminifolia* var. *aequilifolia*) to 199,621 (*P. ruthii*) reads mapped to the *H. annuus* reference cp genome, with 5–240 x coverage (S2 Table). *P. aspera* var. *aspera* had the highest number of basepairs mapped to the *Helianthus* reference (S2 Table), whereas *P. graminifolia* var. *aequilifolia* had the fewest mapped reads.

The reference *Pityopsis* cp genome from *P. falcata* is a single, circular chromosome, with a large single copy (LSC), small single copy (SSC), and two inverted repeat regions (IR) (Fig 1). The *P. falcata* reference was 152,683 bp in length; the LSC was 84,377 bp in length, the SSC was 20,144 bp in length, and the two IRs were 24,081 bp in length. 113 unique genes were identified: 29 transfer RNA (tRNA) genes, 4 ribosomal RNA (rRNA) genes, and 80 protein-coding genes. The IR regions each contained four rRNAs, seven tRNAs, and seven protein-coding genes. Genes directly related to photosynthesis accounted for 42% of all gene function. All assembled *Pityopsis* cp genomes shared synteny with one another and included the same gene features. No inversions or genome rearrangements were apparent in the *Pityopsis* cp genome when compared to each other or other Asteraceae species. The *Pityopsis* cp genome length (152,683 bp) was comparable to cp genomes of other Asteraceae species such as *Lactuca sativa* (lettuce) and *Jacobaea vulgaris*, although the LSC and SSC were longer than those of other species (Table 2). Asteraceae cp genomes contain approximately 114 genes according to Wang *et al.* [36]; we identified 113 genes from the *Pityopsis* cp genome. When including genes

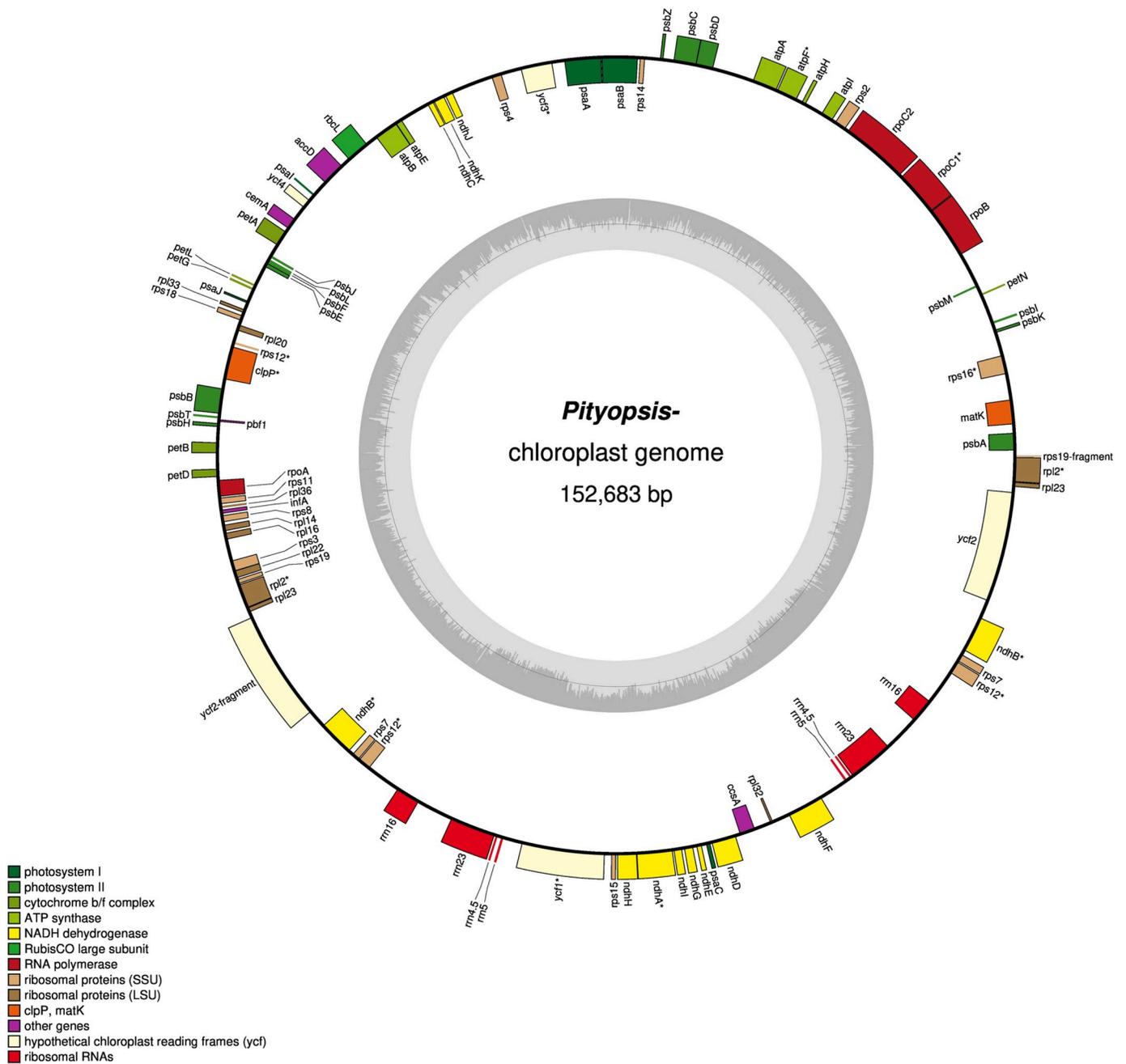


Fig 1. Visualization of *Pityopsis falcata* chloroplast gene map with annotations. The inner circle is GC content. Genes are color coded based on function as per the legend. Genes on the inside of the outer circle are minus (-) strand and genes on the outside of the outer circle are plus (+) strand.

<https://doi.org/10.1371/journal.pone.0241391.g001>

duplicated in the IRs, 131 genes were identified, of which 87 were protein-coding. This is five genes fewer than found (including duplicates) in *J. vulgaris* [37]. Functional groups of genes were all appropriately represented in the *Pityopsis* cp genomes as compared to those of *Aster spathulifolius* [38], with all anticipated protein-coding genes seen in *Pityopsis*. All photosynthesis system I and II genes expected in angiosperms were seen, as compared to the list from Wakasugi *et al.* [39].

Table 2. Comparison of the *Pityopsis falcata* chloroplast genome to other Asteraceae species.

	<i>Pityopsis falcata</i>	<i>Aster spathulifolius</i>	<i>Chrysanthemum indicum</i>	<i>Helianthus annuus</i>	<i>Jacobea vulgaris</i>	<i>Lactuca sativa</i>
Length (bp)	152,683	149,473	150,972	151,104	150,686	152,772
LSC (bp)	84,377	81,998	82,740	83,530	82,855	84,105
SSC (bp)	20,144	17,973	18,394	18,308	18,258	18,599
IR (bp)	24,081	24,751	24,971	24,633	25,000	25,034
No. genes	113	111	114	115	115	115
No. protein-coding genes	80	78	80	81	81	78
No. tRNAs	29	29	30	30	30	30
No. rRNAs	4	4	4	4	4	4
Duplicated genes	18	18	18	18	18	18

<https://doi.org/10.1371/journal.pone.0241391.t002>

We included a single IR in the Mauve alignment and pairwise analyses. Percent identity was higher between cp genomes than the nuclear sequences. Pairwise percent identity was calculated for all 12 *Pityopsis* cp genomes and the outgroup, *H. annuus* (Table 3). The most similar cp genomes based on percent identity were *P. graminifolia* var. *latifolia* and *P. aspera* var. *aspera* (Table 3). The taxa with the most similar nuclear ribosomal regions were *P. aspera* var. *adenolepis* and var. *aspera* (99.83%), *P. aspera* var. *adenolepis* and *P. graminifolia* var. *graminifolia* (99.83%), and *P. aspera* var. *adenolepis* and *P. graminifolia* var. *latifolia* (99.84%) (Table 3).

Phylogenetic analyses

We reconstructed two maximum likelihood (ML) phylogenies, one with cp genomes (Fig 2) and one using the nuclear ribosomal repeat region (Fig 3). Only branches with 50% bootstrap (BS) support or higher were included in the topology. The cp genome and nuclear ribosomal region ML phylogenies differ primarily in the placement of some varieties of *P. graminifolia* and the two varieties of *P. aspera*. A close relationship between *P. pinifolia* and *P. flexuosa* was supported in both phylogenies, with the chloroplast tree showing the two as sister species (BS > 98). A similar relationship between *P. graminifolia* var. *aequilifolia* and var. *tenuifolia*

Table 3. Pairwise alignment comparison of 12 *Pityopsis* species and varieties. Below diagonal is pairwise percent identity between chloroplast genomes calculated from a Mauve multiple sequence alignment. Above diagonal is pairwise percent identity between the short nuclear ribosomal region of the same species, calculated from a MUSCLE multiple sequence alignment.

	1	2	3	4	5	6	7	8	9	10	11	12	13
<i>Helianthus annuus</i>		77.72	77.73	77.66	77.69	77.75	77.69	77.74	77.72	77.69	77.76	77.73	77.62
<i>P. aspera</i> var. <i>adenolepis</i>	87.40		99.83	99.44	98.77	98.66	99.83	99.84	98.65	99.71	99.36	98.67	99.33
<i>P. aspera</i> var. <i>aspera</i>	87.47	99.65		99.33	98.66	98.56	99.72	99.90	98.53	99.72	99.23	98.55	99.21
<i>P. falcata</i>	87.41	99.56	99.66		98.78	98.71	99.44	99.36	98.71	99.28	99.17	98.73	99.49
<i>P. flexuosa</i>	87.49	99.61	99.70	99.62		99.11	98.76	98.71	99.12	98.65	98.94	99.15	98.76
<i>P. graminifolia</i> var. <i>aequilifolia</i>	87.53	99.73	99.86	99.75	99.76		98.64	98.57	99.74	98.48	99.12	99.72	98.65
<i>P. graminifolia</i> var. <i>graminifolia</i>	87.40	99.62	99.66	99.62	99.65	99.79		99.75	98.61	99.67	99.34	98.63	99.39
<i>P. graminifolia</i> var. <i>latifolia</i>	87.52	99.67	99.97	99.65	99.72	99.86	99.69		98.57	99.70	99.27	98.59	99.25
<i>P. graminifolia</i> var. <i>tenuifolia</i>	87.44	99.54	99.86	99.55	99.60	99.74	99.77	99.86		98.47	99.11	99.77	98.64
<i>P. graminifolia</i> var. <i>tracyi</i>	87.45	99.65	99.68	99.61	99.63	99.75	99.83	99.66	99.77		99.19	98.49	99.16
<i>P. oligantha</i>	87.51	99.75	99.77	99.73	99.77	99.86	99.72	99.78	99.66	99.78		99.13	99.06
<i>P. pinifolia</i>	87.51	99.60	99.72	99.63	99.86	99.77	99.68	99.72	99.65	99.67	99.75		98.62
<i>P. ruthii</i>	87.50	99.66	99.72	99.66	99.69	99.80	99.65	99.71	99.61	99.71	99.80	99.70	

<https://doi.org/10.1371/journal.pone.0241391.t003>

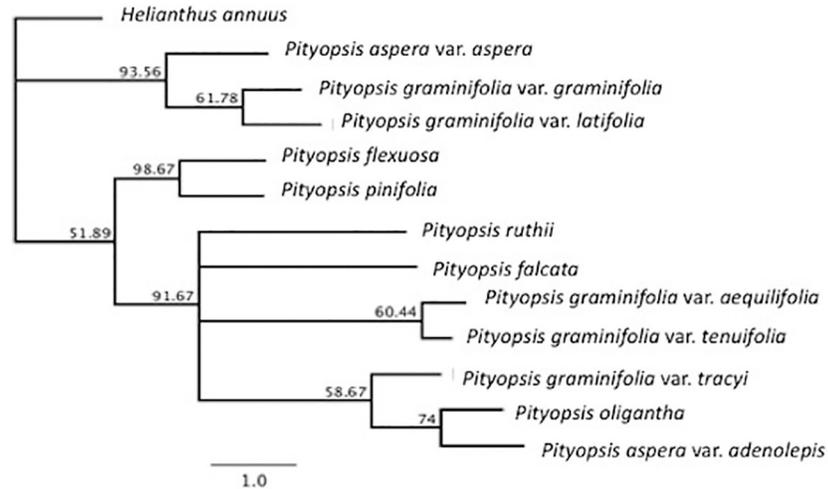


Fig 2. Phylogenetic relationships within *Pityopsis* using the maximum likelihood approach from whole chloroplast genomes, with *Helianthus annuus* as the outgroup. The tree was constructed using the GTR + GAMMA model within RAxML with 1000 replicates. A consensus tree was built using a 50% support threshold. Numbers above nodes are bootstrap values (> 50%).

<https://doi.org/10.1371/journal.pone.0241391.g002>

was moderately supported in the nuclear (BS = 80.6) but weakly supported in the chloroplast tree (BS = 61.8). *P. falcata* and *P. ruthii* were placed near one another (BS = 91.67) in the chloroplast tree, and the two were placed as sister species in the nuclear phylogenetic tree with strong support (BS = 97.6). The placement of the two *P. aspera* varieties is incongruent, with the chloroplast tree showing divergence and the nuclear tree supporting (BS = 89.2) a closer relationship. The placement of *P. oligantha* was also incongruent between the two trees.

Discussion

In this study, we examined relationships among *Pityopsis* species using whole genome sequencing to assemble and compare whole chloroplast genomes. All twelve complete *Pityopsis* cp

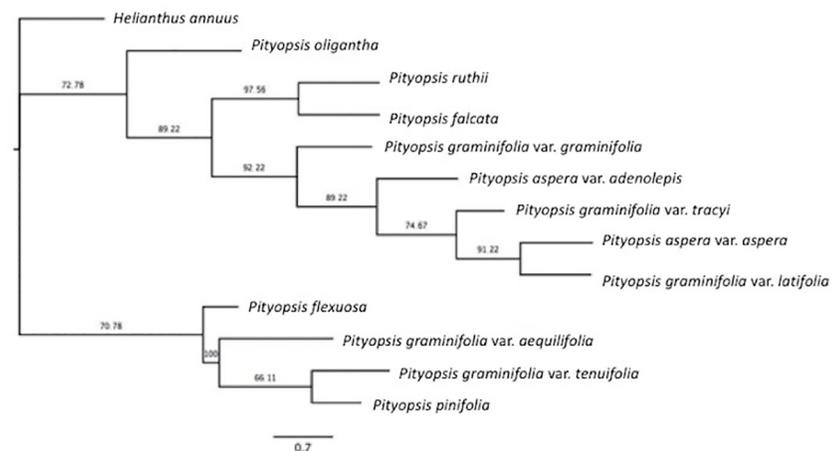


Fig 3. Phylogenetic tree constructed from nuclear ribosomal region of all sampled *Pityopsis* taxa with related species *Helianthus annuus* as an outgroup using the maximum likelihood approach. The tree was constructed using the GTR + GAMMA model within RAxML with 1000 replicates. A consensus tree was built using a 50% support threshold. Numbers above nodes are bootstrap values (> 50%).

<https://doi.org/10.1371/journal.pone.0241391.g003>

genomes displayed attributes common among angiosperm cp genomes, with quadripartite structure including the LSC, SSC, and a pair of inverted repeats (IRa and IRb). Although there were no genomic rearrangements apparent and gene order was maintained, the sizes of the cp genomes ranged from 152,558 to 152,747, suggesting small genetic differences.

The *Pityopsis* cp genome included 29 unique tRNA genes, comparable to *A. spathulifolius* (29) and *J. vulgaris* (29). Within Asteraceae, 29 tRNA genes per cp genome is typical [36,40]. The number of rRNA genes found in the IR of *Pityopsis* is consistent with the number found in several other Asteraceae species, including *A. spathulifolius* [38], *H. annuus* and *L. sativa* [40], and *J. vulgaris* [37]. The *ycf1* and *ndhH* genes in *Pityopsis* did not overlap, consistent with *H. annuus* and other species within Heliantheae, rather than overlapping as seen in Astereae species such as *A. spathulifolius* [38]. Additionally, the *ycf15* gene was present in *Pityopsis* cp genomes, a phenomenon that distinguishes *H. annuus* from *Chrysanthemum indicum*, *C. × morifolium*, and *Guizotia abyssinica*, in which *ycf15* is absent [36]. We do not know whether *ycf15* is co-transcribed with trnL-CAA and *ycf2*, as in *Camellia* [41], but all three genes are present in *Pityopsis* cp genomes. Due to the phylogenetic relationship between *Pityopsis* and *H. annuus*, we expected to see similarities with *Helianthus*, such as presence of *ycf15*, rather than similarities with the more distantly related genera such as *Chrysanthemum*.

The close relationship between *P. flexuosa* and the varieties of *P. graminifolia* seen in a previous study [4] was not evident in the whole cp genome phylogeny, or the nuclear ribosomal region phylogeny. Our findings are also not consistent with the division of the genus into the sections of *Graminifoliae* and *Pityopsis* proposed by Semple and Bowers [13]. In the nuclear and chloroplast phylogenetic trees section, *Pityopsis* is separated by species within *Graminifoliae*. Both sets of trees placed *P. ruthii* and *P. falcata* close together as sister species and did the same with *P. flexuosa* and *P. pinifolia*. The two species with the most disagreement between datasets are both tetraploids, *P. aspera* var. *adenolepis* and *P. oligantha*. This incongruence might be explained by the difference in inheritance for the two datasets, with the chloroplast inherited maternally and the ribosomal region biparentally; in allopolyploids, this means inheriting the ribosomal region from two different species. Resolution of the contributions to the incongruent placement of polyploid species within *Pityopsis* will require expanded sampling at the population level beyond the scope of the current study.

It is not confirmed whether *Pityopsis* polyploids are auto- or allopolyploids, though there is some evidence that allopolyploidy is the mechanism of genome duplication in *P. graminifolia* var. *latifolia*, and that polyploid varieties within *Pityopsis* may be allopolyploid hybrids of other species and varieties [4]. The incongruence between the nuclear and cp trees in the placement of *P. aspera* varieties suggest there has been cp transfer through hybridization involved, likely involving *P. aspera* var. *adenolepis* and *P. oligantha*. Our results do not provide evidence to refute Nesom's [2] categorization of the two varieties as separate species based on morphology and distribution: *P. aspera* and *P. adenolepis*. The differences between our nuclear and chloroplast phylogenies also support the hypothesis that *P. oligantha* is allopolyploid.

Incongruences between nuclear and chloroplast datasets support allopolyploidy in both *P. oligantha* and *P. aspera* var. *adenolepis*. *P. graminifolia* var. *tracyi* is also a possible allopolyploid, as the hexaploid is placed differently in the two phylogenies, with it placed in the same clade as *P. graminifolia* var. *latifolia* and both *P. aspera* varieties in the nuclear phylogenies but with *P. oligantha* and only *P. aspera* var. *adenolepis* in the chloroplast phylogeny. Our results support further investigation into the *P. graminifolia* complex, as the varieties were not placed together in either phylogeny, supporting a reorganization of the species, particularly for *P. graminifolia* var. *aequilifolia* and var. *tenuifolia* which were placed closer to one another than to other *P. graminifolia* varieties in both trees. Indeed, based on morphology and distribution, Nesom [2] names the varieties as distinct species from *P. graminifolia*: *P. aequilifolia* and the

more widespread *P. tenuifolia*. The cp genomes from all species and varieties of *Pityopsis* will provide information to future researchers interested in the genus or speciation in plants. Although our study presents the most complete molecular dataset for *Pityopsis* to date, sampling inconsistencies between morphological and molecular studies may contribute to taxonomic confusion.

The variation among chloroplast genomes of *Pityopsis* species provide a mechanism of distinguishing between species and varieties for use in future studies, as well as a broader of understanding diversity within the genus. We have assembled whole chloroplast genomes that will allow further study of individual species as well, opening possibilities for future work in chloroplast transcriptomics, furthering knowledge of variable regions within the chloroplast, and providing information for future studies of *Pityopsis* and Asteraceae.

Supporting information

S1 Table. Statistics from original genomic sequences for all *Pityopsis* individuals.
(DOCX)

S2 Table. Statistics of *Pityopsis* sequences mapped to the *Helianthus annuus* chloroplast reference genome using Bowtie2.
(DOCX)

S3 Table. Pairwise alignment comparison of 12 *Pityopsis* species and varieties. Below diagonal is pairwise percent identity between chloroplast genomes calculated from a Mauve multiple sequence alignment. Above diagonal is pairwise percent identity between the short nuclear ribosomal region of the same species, calculated from a MUSCLE multiple sequence alignment.
(XLSX)

Acknowledgments

We would like to thank Drs. Loran Anderson, Susan Carr, Joan Walker, and Eugene Wofford for their assistance in tissue collection, as well as T.J. Savereno. The research reported in this article is from the thesis of Hatmaker (2016) which is available online at http://trace.tennessee.edu/utk_gradthes/3744. Mention of trade names or commercial products in this article is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the University of Tennessee or the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer. Samples collected in Tennessee were collected under Tennessee Valley Authority Permit # TE117405-2 and U.S. Fish and Wildlife Service Permit # TE134817-1.

Author Contributions

Conceptualization: E. Anne Hatmaker, Phillip A. Wadl, Timothy A. Rinehart, Robert N. Trigiano, Margaret E. Staton.

Data curation: E. Anne Hatmaker, Thomas S. Lane, Edward E. Schilling.

Formal analysis: E. Anne Hatmaker, Thomas S. Lane, Edward E. Schilling.

Funding acquisition: Timothy A. Rinehart, Robert N. Trigiano.

Investigation: E. Anne Hatmaker, Phillip A. Wadl.

Methodology: E. Anne Hatmaker, Phillip A. Wadl, Jennifer Carroll, Thomas S. Lane, Robert N. Trigiano, Edward E. Schilling.

Project administration: Phillip A. Wadl, Timothy A. Rinehart, Robert N. Trigiano, Margaret E. Staton, Edward E. Schilling.

Resources: Timothy A. Rinehart.

Supervision: Phillip A. Wadl, Robert N. Trigiano, Edward E. Schilling.

Validation: E. Anne Hatmaker, Edward E. Schilling.

Visualization: E. Anne Hatmaker.

Writing – original draft: E. Anne Hatmaker.

Writing – review & editing: E. Anne Hatmaker, Phillip A. Wadl, Timothy A. Rinehart, Jennifer Carroll, Thomas S. Lane, Robert N. Trigiano, Margaret E. Staton, Edward E. Schilling.

References

1. Semple JC. Pityopsis Nuttall. in Flora of North America. Vol. 20. Asteraceae, Part 2. Astereae and Senecioneae. Committee FNAE, editor. New York: Oxford Univ. Press, New York.; 2006. 222–226 p.
2. Nesom GL. Taxonomic synopsis of pityopsis (asteraceae). Phytoneuron. 2019; 2019(1):1–31.
3. Gowe AK, Brewer JS. The evolution of fire-dependent flowering in goldenasters (Pityopsis spp.). J Torrey Bot Soc. 2005; 132:384–400.
4. Teoh V-H. Phylogeny, Hybridization and the Evolution of Fire-stimulated Flowering Within the Grass-leaved Goldenasters (Pityopsis, Asteraceae). University of Mississippi; 2008.
5. Costa CM. Molecular Phylogeny of the Goldenasters, Subtribe Chrysopsidinae (Asteraceae: Astereae), Based on Nuclear Ribosomal and Chloroplast Sequence Data. Towson University; 2014.
6. Gornish ES. Effects of density and fire on the vital rates and population growth of a perennial goldenaster. AoB Plants. 2013; 5.
7. Moore PA, Wadl PA, Skinner JA, Trigiano RN, Bernard EC, Klingeman WE, et al. Current knowledge, threats, and future efforts to sustain populations of *Pityopsis ruthii* (Asteraceae), an endangered southern Appalachian species¹. J Torrey Bot Soc. 2016;
8. Wan QH, Wu H, Fujihara T, Fang SG. Which genetic marker for which conservation genetics issue? Electrophoresis. 2004. <https://doi.org/10.1002/elps.200305922> PMID: 15274000
9. Gilbert GS, Webb CO. Phylogenetic signal in plant pathogen-host range. Proc Natl Acad Sci. 2007; 104(12):4979–83. <https://doi.org/10.1073/pnas.0607968104> PMID: 17360396
10. Vamوسي SM, Heard SB, Vamوسي JC, Webb CO. Emerging patterns in the comparative analysis of phylogenetic community structure. Vol. 18, Molecular Ecology. 2009. p. 572–92. <https://doi.org/10.1111/j.1365-294X.2008.04001.x> PMID: 19037898
11. Hoffmann A, Griffin P, Dillon S, Catullo R, Rane R, Byrne M, et al. A framework for incorporating evolutionary genomics into biodiversity conservation and management. Clim Chang Responses. 2015; 2(1):1.
12. Byrne M. Phylogeography provides an evolutionary context for the conservation of a diverse and ancient flora. Vol. 55, Australian Journal of Botany. 2007. p. 316–25.
13. Semple JC, Bowers F. Cytogeography of Pityopsis Nutt., the grass-leaved goldenasters (Compositae: Astereae). Rhodora. 1987;381–9.
14. Hatmaker EA, Staton ME, Dattilo AJ, Hadziabdic, Rinehart TA, Schilling EE, et al. Population Structure and Genetic Diversity Within the Endangered Species Pityopsis ruthii (Asteraceae). Front Plant Sci. 2018; 9(July):1–15.
15. Boggess SL, Wadl PA, Hadziabdic D, Scheffler BE., Windham AS, Klingeman WE, et al. Characterization of 12 polymorphic microsatellite loci of Pityopsis graminifolia var. latifolia. Vol. 6, Conservation Genetics Resources. 2014. p. 1043–5.
16. Wadl PA, Dattilo AJ, Scheffler BE, Trigiano RN. Development of microsatellite loci for the endangered species Pityopsis ruthii (Asteraceae). Am J Bot. 2011; 98(12):e342–5. <https://doi.org/10.3732/ajb.1100100> PMID: 22058180

17. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* 2009; 7(1):84. <https://doi.org/10.1186/1741-7007-7-84> PMID: 19954512
18. Huang H, Shi C, Liu Y, Mao SY, Gao LZ. Thirteen *Camellia* chloroplast genome sequences determined by high-throughput sequencing: Genome structure and phylogenetic relationships. *BMC Evol Biol.* 2014; 14(1):151. <https://doi.org/10.1186/1471-2148-14-151> PMID: 25001059
19. Matsuoka Y, Yamazaki Y, Ogihara Y, Tsunewaki K. Whole chloroplast genome comparison of rice, maize, and wheat: Implications for chloroplast gene diversification and phylogeny of cereals. *Mol Biol Evol.* 2002; 19(12):2084–91. <https://doi.org/10.1093/oxfordjournals.molbev.a004033> PMID: 12446800
20. Njuguna W, Liston A, Cronn R, Ashman TL, Bassil N. Insights into phylogeny, sex function and age of *Fragaria* based on whole chloroplast genome sequencing. *Mol Phylogenet Evol.* 2013; 66(1):17–29. <https://doi.org/10.1016/j.ympev.2012.08.026> PMID: 22982444
21. Xu Q, Xiong G, Li P, He F, Huang Y, Wang K, et al. Analysis of complete nucleotide sequences of 12 *Gossypium* chloroplast genomes: Origin and evolution of Allotetraploids. *PLoS One.* 2012; 7(8): e37128. <https://doi.org/10.1371/journal.pone.0037128> PMID: 22876273
22. Clewell AF. Guide to the vascular plants of the Florida panhandle. Tallahassee: University Presses of Florida, Florida State University; 1985. 605 p.
23. Semple JC, Jabbour F. Type specimens of *inula* (*pityopsis*) *graminifolia* (*asteraceae*: *astereae*). *Phyto-neuron.* 2019; 2019–22(April):1–9.
24. Andrews S. FastQC: A quality control tool for high throughput sequence data [Internet]. Babraham Bioinformatics; 2010. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
25. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
26. Langmead Ben, Salzberg SL. Bowtie2. *Nat Methods.* 2013; 9(4):357–9.
27. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. ABySS: A parallel assembler for short read sequence data. *Genome Res.* 2009; 19(6):1117–23. <https://doi.org/10.1101/gr.089532.108> PMID: 19251739
28. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* 2012; 28(12):1647–9. <https://doi.org/10.1093/bioinformatics/bts199> PMID: 22543367
29. Wyman SK, Jansen RK, Boore JL. Automatic annotation of organellar genomes with DOGMA. *Bioinformatics.* 2004; 20(17):3252–5. <https://doi.org/10.1093/bioinformatics/bth352> PMID: 15180927
30. Greiner S, Lehwark P, Bock R. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. *Nucleic Acids Res.* 2019; 47(W1):W59–64. <https://doi.org/10.1093/nar/gkz238> PMID: 30949694
31. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 2004; 14(7):1394–403. <https://doi.org/10.1101/gr.2289704> PMID: 15231754
32. Bateman A, Martin MJ, O'Donovan C, Magrane M, Alpi E, Antunes R, et al. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* 2017; 45(D1):D158–69. <https://doi.org/10.1093/nar/gkw1099> PMID: 27899622
33. Posada D. jModelTest: Phylogenetic model averaging. *Mol Biol Evol.* 2008; 25(7):1253–6. <https://doi.org/10.1093/molbev/msn083> PMID: 18397919
34. Stamatakis A. RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics.* 2014; 30(9):1312–3. <https://doi.org/10.1093/bioinformatics/btu033> PMID: 24451623
35. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32(5):1792–7. <https://doi.org/10.1093/nar/gkh340> PMID: 15034147
36. Wang M, Cui L, Feng K, Deng P, Du X, Wan F, et al. Comparative Analysis of Asteraceae Chloroplast Genomes: Structural Organization, RNA Editing and Evolution. *Plant Mol Biol Report.* 2015; 33(5):1526–38.
37. Doorduyn L, Gravendeel B, Lammers Y, Ariyurek Y, Chin-A-Woeng T, Vrieling K. The complete chloroplast genome of 17 individuals of pest species *Jacobaea vulgaris*: SNPs, microsatellites and barcoding markers for population and phylogenetic studies. *DNA Res.* 2011; 18(2):93–105. <https://doi.org/10.1093/dnares/dsr002> PMID: 21444340
38. Choi KS, Park SJ. The complete chloroplast genome sequence of *Aster spathulifolius* (Asteraceae); genomic features and relationship with Asteraceae. *Gene.* 2015; 572(2):214–21. <https://doi.org/10.1016/j.gene.2015.07.020> PMID: 26164759

39. Wakasugi T, Tsudzuki T, Sugiura M. The genomics of land plant chloroplasts: Gene content and alteration of genomic information by RNA editing. Vol. 70, Photosynthesis Research. 2001. p. 107–18. <https://doi.org/10.1023/A:1013892009589> PMID: 16228365
40. Timme RE, Kuehl J V., Boore JL, Jansen RK. A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: Identification of divergent regions and categorization of shared repeats. *Am J Bot.* 2007; 94(3):302–12. <https://doi.org/10.3732/ajb.94.3.302> PMID: 21636403
41. Shi C, Liu Y, Huang H, Xia EH, Zhang H Bin, Gao LZ. Contradiction between Plastid Gene Transcription and Function Due to Complex Posttranscriptional Splicing: An Exemplary Study of *ycf15* Function and Evolution in Angiosperms. *PLoS One.* 2013; 8(3):e59620. <https://doi.org/10.1371/journal.pone.0059620> PMID: 23527231