

RESEARCH ARTICLE

Plastome variation and phylogeny of *Taxillus* (Loranthaceae)Huei-Jiun Su^{1*}, Shu-ling Liang¹, Daniel L. Nickrent²**1** Department of Earth and Life Sciences, University of Taipei, Taipei, Taiwan, **2** Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States of America* hjsu@utapei.edu.tw

Abstract

Several molecular phylogenetic studies of the mistletoe family Loranthaceae have been published such that now the general pattern of relationships among the genera and their biogeographic histories are understood. Less is known about species relationships in the larger (> 10 species) genera. This study examines the taxonomically difficult genus *Taxillus* composed of 35–40 Asian species. The goal was to explore the genetic diversity present in *Taxillus* plastomes, locate genetically variable hotspots, and test these for their utility as potential DNA barcodes. Using genome skimming, complete plastomes, as well as nuclear and mitochondrial rDNA sequences, were newly generated for eight species. The plastome sequences were used in conjunction with seven publicly available *Taxillus* sequences and three sequences of *Scurrula*, a close generic relative. The *Taxillus* plastomes ranged from 121 to 123 kbp and encoded 90–93 plastid genes. In addition to all of the NADH dehydrogenase complex genes, four ribosomal genes, *infA* and four intron-containing tRNA genes were lost or pseudogenized in all of the *Taxillus* and *Scurrula* plastomes. The topologies of the plastome, mitochondrial rDNA and nuclear rDNA trees were generally congruent, though with discordance at the position of *T. chinensis*. Several variable regions in the plastomes were identified that have sufficient numbers of parsimony informative sites as to recover the major clades seen in the complete plastome tree. Instead of generating complete plastome sequences, our study showed that *accD* alone or the concatenation of *accD* and *rbcl* can be used in future studies to facilitate identification of *Taxillus* samples and to generate a molecular phylogeny with robust sampling within the genus.

OPEN ACCESS

Citation: Su H-J, Liang S-l, Nickrent DL (2021) Plastome variation and phylogeny of *Taxillus* (Loranthaceae). PLoS ONE 16(8): e0256345. <https://doi.org/10.1371/journal.pone.0256345>

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

Received: February 15, 2021

Accepted: August 5, 2021

Published: August 18, 2021

Copyright: © 2021 Su et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting Information](#) files.

Funding: This study was funded by MOST grants (106-2311-B-845-001-MY3 and 109-2311-B-845-001) from Ministry of Science and Technology, Taiwan to HJS.

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

Introduction

The family Loranthaceae is the largest in the sandalwood order Santalales with 76 genera and over 1000 species distributed worldwide [1]. Despite its ecological and ethnobotanical value, relatively little is known about species relationships within the larger genera. Nearly all genera have been represented as placeholder species in broad-scale phylogenetic work within Santalales and Loranthaceae [1–4], however, relatively few studies have examined interspecific relationships within a genus. Of the 76 Loranthaceae genera, 28 have 10 or more species [5]. Among these, very few published studies that included a robust sampling of species within the genus are of *Tristerix* [6], *Loranthus* [7] and *Dendropemon* [8]. For some genera, molecular

phylogenies have been published but only for a subset of species, such as *Psittacanthus* [9, 10] and *Taxillus* [11].

Plastomes in angiosperms contain highly conserved as well as variable regions [12], therefore these genomes have been widely used to address phylogenetic questions at various taxonomic levels [13]. Whole plastome sequences were used to examine relationships among families of the sandalwood order [14], and the recovered topology was congruent with previous multigene analyses [1, 3]. Although complete plastome sequences can be used to address species relationships within a mistletoe genus (e.g. *Loranthus*), there is interest in identifying which plastome regions are most effective in resolving relationships at lower taxonomic levels. Sometimes referred to as DNA “barcodes”, these regions can be targeted to allow implementation of a rapid and cost-effective method to identify specimens to the rank of species [15, 16]. It then follows that such gene regions are also useful in generating species level phylogenies. Although genome skimming [17] can be used to generate complete plastomes (as reported herein), this method is more costly and cannot always be used with degraded DNA that might be present in herbarium specimens or forensic samples. The rationale proposed here is to use genome skimming for a subset of species within a genus, determine which regions are appropriately variable, and then use those regions in species level analyses with more robust sampling. In such studies, more cost effective and rapid PCR methods can be employed to generate sequences across a non-trivial number of samples that broadly span the species diversity. This “depth then breadth” approach would allow more rapid advancement toward achieving species-level phylogenies for many genera of Loranthaceae and other angiosperms.

The genus treated in the present study is *Taxillus* Tiegh. which comprises 35–40 mistletoe species that are mainly distributed in tropical and temperate Asia. Although some molecular phylogenetic analyses have been published for the genus [11, 18–20] all included five or fewer species because their main focus was to report complete chloroplast genome (plastome) sequences. *Taxillus*, along with its close generic relative *Scurrula*, are the only two genera of Tribe Loranthaeae, Subtribe Scurrulinae [2, 21]. Recognition of their close relationship predates the molecular era [22–24] and indeed these authors expressed difficulty in distinguishing the two genera. Moreover, species relationships within the genera are unclear and no monograph exists for either. A list of 39 potentially good *Taxillus* species was compiled from information the Plant List (now World Flora Online, WFO) where 50 clear synonyms were removed, and six names listed as uncertain status (S1 Table). Both *T. lonicerifolius* (Hayata) S.T. Chiu and *T. rhododendricola* (Hayata) S.T. Chiu, were listed as synonyms of *T. nigrans* (Hance) Danser in Flora of China [25] and on WFO. In addition, *T. matsudai* (Hayata) Danser is considered a synonym of *T. caloreas* (Diels) Danser in these two works. These names were used in the Flora of Taiwan [26] and are here retained so that their relationships to the other species can be tested using molecular data.

Fifteen *Taxillus* species occur in mainland China, 7 in Taiwan, 14 in India, Sri Lanka, and/or Bangladesh, and three in Vietnam, Laos and/or Cambodia. One species (*T. yadoriki*) is known from Japan and the geographically most distant taxon is *T. wiensii* from Kenya. Both Danser (1931) and Barlow (1997) indicated that fruit features best distinguish *Taxillus* from *Scurrula*, where the latter has stipitate, obovoid or clavate fruits and the former has non-stipitate, ovoid or ellipsoid fruits. Features used in the Flora of China [25] that differentiate Chinese species include the presence or absence of pubescence on branches, leaves, calyx and corolla, the color and type of indumentum (e.g. tomentose vs. with stellate hairs), pedicel and corolla dimensions, typical vs. multilocellate anthers, and surface features of the fruit (granulose, verrucose, verruculose or scabrid). Problems with the branch indumentum and fruit surface features are that these characters are labile, often disappearing in mature structures. Another practical limitation is that fruits are not present on all herbarium specimens.

Of the ca. 39 potential species, complete plastome sequences from 15 were examined here, which included seven previously published and eight newly generated ones. Six of the endemic Taiwan species were sequenced and eight other species found in China and surrounding regions were also included. No sequences were obtained from *Taxillus* species from India. Because sampling is incomplete, the aim of this study was not to provide a definitive molecular phylogeny of *Taxillus* but to explore the genetic diversity present in the plastomes, locate genetically variable hotspots, and test these for their utility as potential barcodes. These selected candidate markers will be useful for sample identification and future studies aimed at resolving phylogenetic relationships among all species of *Taxillus*. The plastome phylogenies were further compared with the topologies inferred from nuclear and mitochondrial rDNA regions to assess subcellular genome congruence and screen for potential hybridization events.

Materials and methods

Ethics statement

Fresh *T. pseudochinensis* material were collected from Hengchun, Pingtung, Taiwan with permission from Hengchun Tropical Botanic Garden, Taiwan. The collection of other *Taxillus* fresh materials were carried out in compliance with relevant laws of Taiwan. Herbarium samples were accessed with approval from the Herbarium of Taiwan Forestry Research Institute (TAIF) and Research Center for Biodiversity, Academia Sinica, Taipei (HAST). No endangered species were used in this study.

Sample collection, DNA extraction and sequencing

26 *Taxillus* samples belonging to nine species were collected and detailed voucher specimen information is shown in [S2 Table](#). 21 of the samples were used for genome skimming. Total genomic DNA was extracted using the CTAB method [27] and sheared into fragments between 400 and 700 bp in size for library construction. Sequencing data were retrieved on a 300 bp paired-end Illumina Miseq platform at VYM Genome Research Center, Taiwan. In addition, paired-end genomic sequences of *Scurrula chingii* (SRX9233076), *Taxillus chinensis* (CNS0039580) and *T. sutchuenensis* (CNS0039577) were downloaded from publicly available databases and used in the analyses. Additional sequence data used in various phylogenetic analyses were retrieved from Genbank ([S3 Table](#)).

Acquisition of plastome, nuclear and mitochondria rDNA sequences

De novo assemblies of contigs were performed by the CLC Genomics Workbench software (v8.5.1, CLC Bio, Aarhus, Denmark). The adaptor sequences and low-quality bases of paired reads were removed before assembly using Trimmomatic (v0.39) [28]. Contigs containing plastid genes were selected and used for scaffold building. The plastid contigs were obtained by BLASTN searches against the plastomes of *Nicotiana tabacum* (NC_001879.2, e-value 1e-10) and *Schoepfia jasminodora* (NC_034228.1). About three plastid contigs were recovered from each of the sequenced samples and the orientation of the contigs was based on *N. tabacum*. The gaps between contigs were closed by iterative mapping of raw reads using BWA (v0.7.4) [29] and samtools [30]. The final assemblies were visually inspected by IGV [31] to ensure each plastome was fully covered by overlapping paired-end reads.

The preliminary annotations of complete plastomes were performed using GeSeq [32] using other land plant chloroplast genomes as reference. For intron-containing genes, the intron and exon boundaries were carefully examined by sequence comparisons with the sequences of *N. tabacum*. To explore the presence/absence of genes that were missing from the

annotation, gene sequences of selected eudicot species were used as queries for BLASTN searches (e-value $1e-10$). To verify the *clpP* pseudogenization event in the *T. nigrans* plastome (MH095982.1), RNA-seq data of *T. nigrans* (SRX2755388) was used and the *clpP* sequences were recovered in comparison to the respective nucleotide sequence [20].

In addition to the plastome sequences, the nuclear ribosomal DNA (nrDNA) cistron was constructed that included sequence spanning the small-subunit (SSU) to large-subunit (LSU) rDNA sequences, including the two internal transcribed spacers and 5.8S rDNA. Similarly, mitochondrial ribosomal DNA (mrDNA) sequences of the small- and large-subunit rDNA were obtained (these are not organized into a cistron). Both regions were obtained by using BLASTN (e-value $1e-10$) searches against sequences of other eudicots. NCBI Genbank accession numbers of newly obtained sequences can be found in [S2 Table](#).

Molecular phylogenetic analyses

Sequences of the full plastomes, the nrDNA cistron, and the mrDNA (concatenated SSU and LSU rDNA sequences) were aligned independently using MAFFT v7.222 [33]. The first plastome dataset containing 29 *Taxillus* sequences (13 species) as well as sequences for seven other Loranthaceae. *Schoepfia jasminodora* was included as outgroup based on previous studies [2, 3]. Sixty-three protein-coding genes (PCGs) and the four rRNA genes were collected and individually aligned using MUSCLE v3.8.31 [34]. This alignment was converted to aligned amino acids (codons) using PAL2NAL [35] and the concatenated alignment of PCGs and rRNA was used to infer a Maximum likelihood (ML) phylogeny. This and all subsequent ML phylogenetic trees were generated using RAxML v8.1.17 [36] with the GTRMMI model. One thousand rapid bootstrap replicates were used to evaluate support values of the ML trees. The second plastid dataset was constructed using 23 *Taxillus* full plastome sequences (11 species). The datasets of nrDNA and mrDNA were also constructed using the same sampling. The ML trees resulting from analyses of these alignments were used to compare topologies among the three subcellular compartment partitions using *Scurrula chingii* as outgroup. The Maximum parsimony (MP) analyses were conducted using the MPBoot software [37] with default parsimony ratchet search options and the bootstrap analyses included 1,000 bootstrap replicates using the same search options. Additional *Taxillus* from NCBI Genbank were not included because the nrDNA and mrDNA sequences were not available for all taxa.

In addition to sequences generated in this study, plastome sequences as well as sequences from the plastid intergenic spacers (IGS) *trnF-trnL* and *trnH-psbA* were included in another ML phylogenetic analysis ([S3 Table](#)). The sampling here included additional *Taxillus*, *Scurrula* and *Helixanthera* accessions [38–40] with the goal of testing monophyly of the species.

Morphological data

To investigate the taxonomic utility of morphological characters often used for identification of *Taxillus*, four characters were selected, and their character states plotted on the 67 gene ML phylogenetic tree. The four characters were leaf shape, the presence abaxial leaf hairs, the presence of corolla hairs, and fruit shape. These characters were chosen based on keys and descriptions published in regional floras [25, 26] as well as visual inspection of the plant samples used in this study.

Sequence variation, genetic distances and informative variables analyses

To assess the presence of divergent regions in the 29 *Taxillus* plastomes, the full MAFFT alignment and the sliding window analysis of nucleotide diversity was conducted using the software DnaSP v6 with a step size of 100 bp and window length of 500 bp [41]. To calculate

infrageneric genetic distance of the plastome and the nrDNA sequences, pairwise Kimura 2-parameter (K2P) distances between the sample pairs were obtained using MEGAX [42]. To evaluate sequence diversity of the various molecular regions in the *Taxillus* plastomes, 63 shared PCGs, four rRNA genes and seven intergenic spacers were analyzed to determine proportion of variable sites and the proportion of parsimony informative sites in these regions. Tests were also conducted to determine the phylogenetic utility of the seven variable intergenic spacers as well as five PCGs that are >1000 bp in length and have been proposed as potential barcodes. Each of the genetic regions was aligned using MAFFT and the diversity information was computed by the AMAS program [43]. The trees were generated using ML and bootstrap support was obtained using RaxML as described earlier. A simple phylogenetic utility test was also used to evaluate the trees using the baseline topology of the complete plastome ML tree as the standard for comparison. The monophyly of *Taxillus* clades I-IV contributed one point each, the monophyly of *Scurrula* one point, and the overall topology one point for a possible total of six points. Topological differences within clades were ignored for this analysis. *Taxillus chinensis* was not included because only one accession was used in some analyses (see below).

Development of *Taxillus*-specific barcodes

Five primer sets for three plastid regions (*accD*, *matK* and *trnL-trnF*) were identified and tested for PCR amplification efficiency. The specific primers were designed based on the 29 *Taxillus* plastome sequences (S4 Table). PCR amplifications were performed in 50 μ l reactions that each contained 25 μ l of 2 \times Taq Plus PCR MasterMix (Tiangen Biotech Co., Ltd.), 0.2 μ M of forward and reverse primers, 22 μ l sterile double distilled water (ddH₂O) and 20–50 ng *Taxillus* genomic DNA. Negative controls used only ddH₂O. The PCR conditions included a preheating at 94°C for 5 min, 35 cycles of 94°C for 30 s, annealing temperatures between 46 and 50°C for 50 s (see S4 Table), and elongation at 72°C for 55 s, followed by a final extension at 72°C for 5 min. Seven *Taxillus* species were tested and the amplicons (including negative controls) were visualized on 1.5% agarose gels. The PCR products of five additional *Taxillus* samples were purified and sequences were obtained from Sanger sequencing.

Results

General features of the *Taxillus* plastomes

The 21 complete plastomes of *Taxillus* assembled and reported here can be compared to the eight published plastome sequences of other *Taxillus* species and those of three species of *Scurrula* (S5 Table). Overall, the plastomes of *Taxillus* ranged in size from 121,305 bp (*T. chinensis*) to 123,074 bp (*T. liquidambaricola*) with GC contents ranging from 37.2 to 37.4%. Gene content and order of the 29 *Taxillus* plastomes are highly conserved, each containing a total of 90–93 unique and putative functional genes, including 61–63 PCGs, 4 rRNA and 25 or 26 tRNA genes. Compared to typical angiosperms plastomes that contain 113 unique genes, the *Taxillus* and *Scurrula* (as well as *Helixanthera*) plastomes have lost or contain pseudogenes for 11 NAD (P)H dehydrogenase (*ndh*) genes, four ribosomal protein genes (*rps15*, *rps16*, *rpl16* and *rpl32*) and four intron-containing tRNA genes (*trnK*-UUU, *trnG*-UCC, *trnV*-UAC, *trnI*-GAU) (S5 Table and S1 Fig). Following the major gene losses that occurred along the branch leading to all these accessions of Loranthaceae, loss or pseudogenization has occurred independently in a few lineages (S1 Fig), such as the pseudogenization of *trnA*-UGC in *T. matsudai* and *H. parastitica* and the independent losses of *infA* in the *Scurrula* clade, *T. nigrans* and *T. chinensis*.

The boundaries between large single copy (LSC) region, small single copy (SSC) region, and inverted repeats (IRs) showed a nearly identical pattern in the 13 *Taxillus* and the three *Scurrula* plastomes (S2 Fig). The junction of LSC/IR_B was located in the exon2 of *rpl2* and the

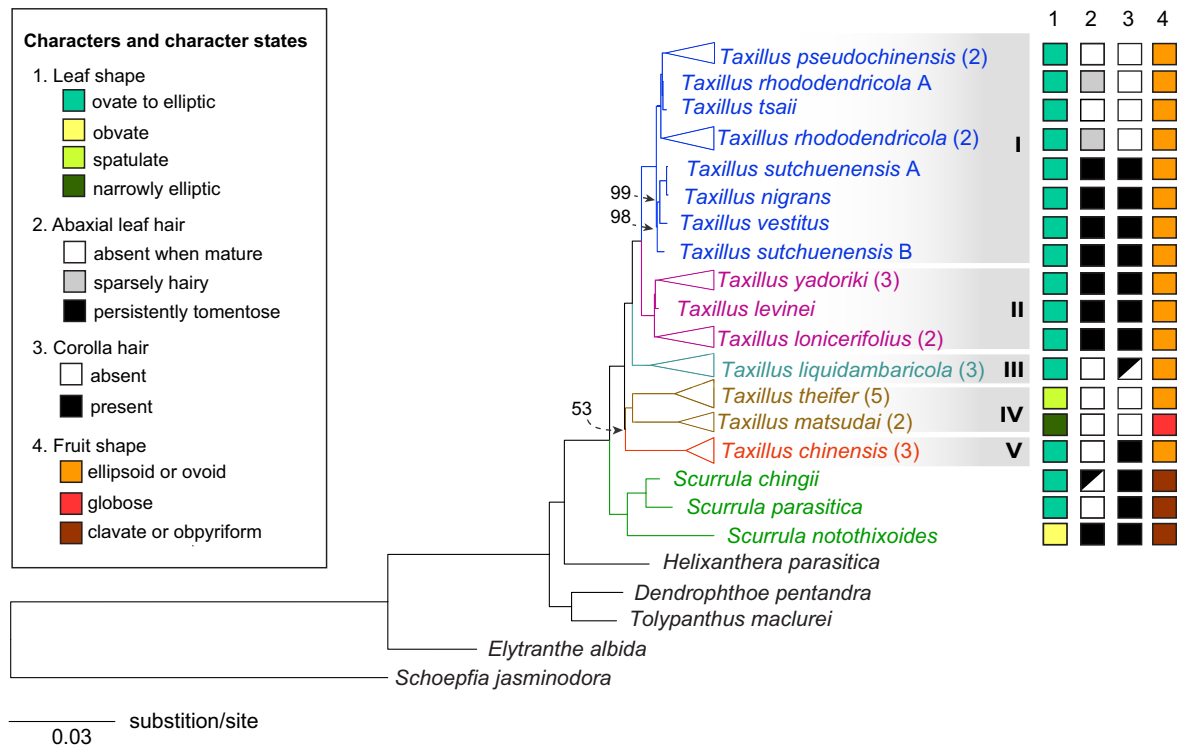


Fig 1. Phylogenetic tree of *Taxillus* and other Loranthaceae. This ML tree derives from analysis of the 67 gene sequence dataset (63 PCGs and four rRNAs). Only bootstrap values < 100% are labeled and *Schoepfia jasminodora* was used as the outgroup. Triangles at the branch tips represent multiple accessions of the same species and the number of accessions used is shown in parentheses. Four morphological characters and their states are shown to the right of the phylogenetic tree. The grey shading marks the five major clades of *Taxillus* (see text).

<https://doi.org/10.1371/journal.pone.0256345.g001>

junction SSC/IR_a was identified within *ycf1*. The junction of IR_b/SSC was found within *trnL-UAG* and a short portion of the *trnL-UAG* overlapped with the duplicated *ycf1* fragment. The junction of IR_a/LSC was identified within the *trnH-GUG* gene, except in *T. nigrans* whose *trnH-GUG* was not detected in the plastome.

Plastome, mitochondrial and nuclear gene trees

The phylogeny based on the 67 shared plastid genes included samples of 13 *Taxillus* species as well as *Scurrula* and five other genera of Loranthaceae (Fig 1). A monophyletic *Taxillus* was strongly supported (BS = 100%) and that clade was sister to a clade consisting of the three *Scurrula* accessions (BS = 100%). The *Taxillus* phylogenetic tree contained five distinct clades, the first four of which were strongly supported (100 BS): the first clade included *T. nigrans*, *T. pseudochinensis*, *T. rhododendricolus*, *T. sutchuenensis*, *T. tsaii*, and *T. vestitus* while the second contained *T. levinei*, *T. lonicerifolius* and *T. yadoriki*. A third clade was sister to the first two and contained three accessions of *T. liquidambaricola*. The fourth clade was composed of *T. theifer* and *T. matsudai*. The fifth clade was composed of three accessions of *Taxillus chinensis* whose topology as sister to the remaining clades was poorly supported (53% MLBS).

The four morphological characters and their states were plotted on the ML tree (Fig 1). Leaf shape was generally uninformative given that most *Taxillus* species were scored as ovate or elliptic. Although abaxial leaf hair presence showed some pattern across states for clades II-IV, clade I was polymorphic. Even less conformance to clade was seen for the presence of corolla hairs. Finally, fruit shape appears important in distinguishing *Taxillus* from *Scurrula* but has

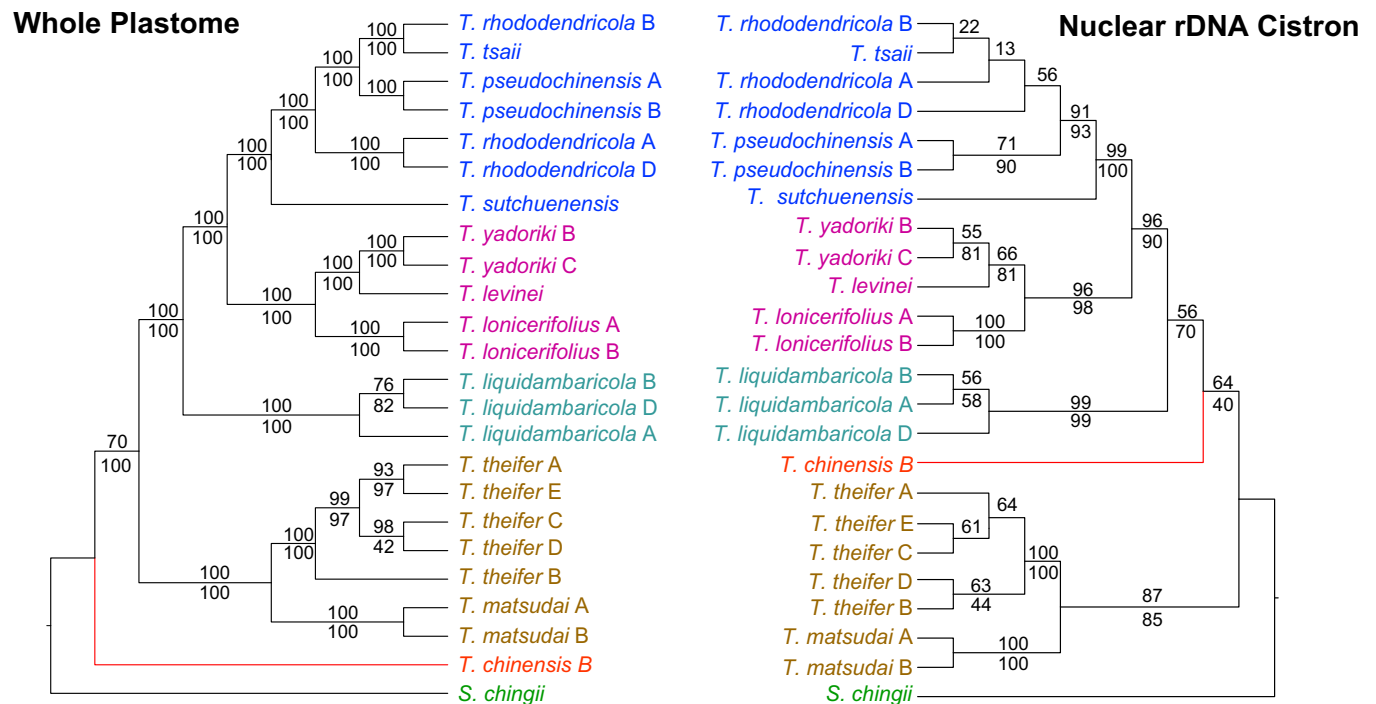


Fig 2. Comparison of plastome and nuclear rDNA ML gene trees for *Taxillus* species. A. ML tree derived from an analysis of the whole plastome dataset. This tree is generally congruent with the one obtained from mitochondrial rDNA sequences. The Maximum Likelihood Bootstrap values (MLBS) and Maximum Parsimony Bootstrap (MPBS) values are indicated above and below the branches, respectively. B. Tree inferred from the ML analysis of the nuclear rDNA sequences. MLBS and MPBS values are indicated above and below the branches, respectively.

<https://doi.org/10.1371/journal.pone.0256345.g002>

little value in differentiating *Taxillus* species. For both leaf and fruit shape, several autapomorphic states exist and, with the present sampling, contribute no cladistic information. Overall, these morphological features are insufficient to distinguish the five clades within *Taxillus*.

Trees resulting from ML and MP analyses of the mrDNA dataset were identical to the plastome tree in terms of topology of the five major clades; however, relationships among conspecific accessions of *T. liquidambaricola* and *T. theifer* differed slightly (S3 Fig). A comparison of the phylogenetic trees resulting from analyses of the plastome and the nrDNA cistron partitions is shown in Fig 2. The whole plastome tree is mostly congruent with the topology shown in Fig 1, differing only topology among accessions in Clade I. Although the nrDNA tree was less resolved than the plastome tree, clades I-III received strong bootstrap support (BS > 95%) and clade IV moderate support (87%). The plastome tree shows *T. chinensis* as sister to all the remaining *Taxillus* whereas in the nrDNA tree this taxon is sister to clades I-III, but with low support. The two topologies are similar in most respects, differing mainly in the interspecific relationships of clade I samples and the infraspecific relationships within the *T. liquidambaricola* and *T. theifer* clades.

Plastome nucleotide diversity and genetically informative regions

The nucleotide diversity values (π) for the 29 *Taxillus* plastome sequences (Fig 3) ranged from 0 to 0.0635 (*rpoB-trnC*) and the mean nucleotide diversity value was 0.0125, showing the sequences were generally conserved. Compared to the nucleotide diversity in the LSC and SSC regions, the two IR regions were the most conserved. The six most variable regions included five IGSS (*rpoB-trnC*, *rps4-trnT*, *accD-psaI*, *rpl14-rps3*, *ccsA-psaC*) and one PCG (*ycf1*). 67

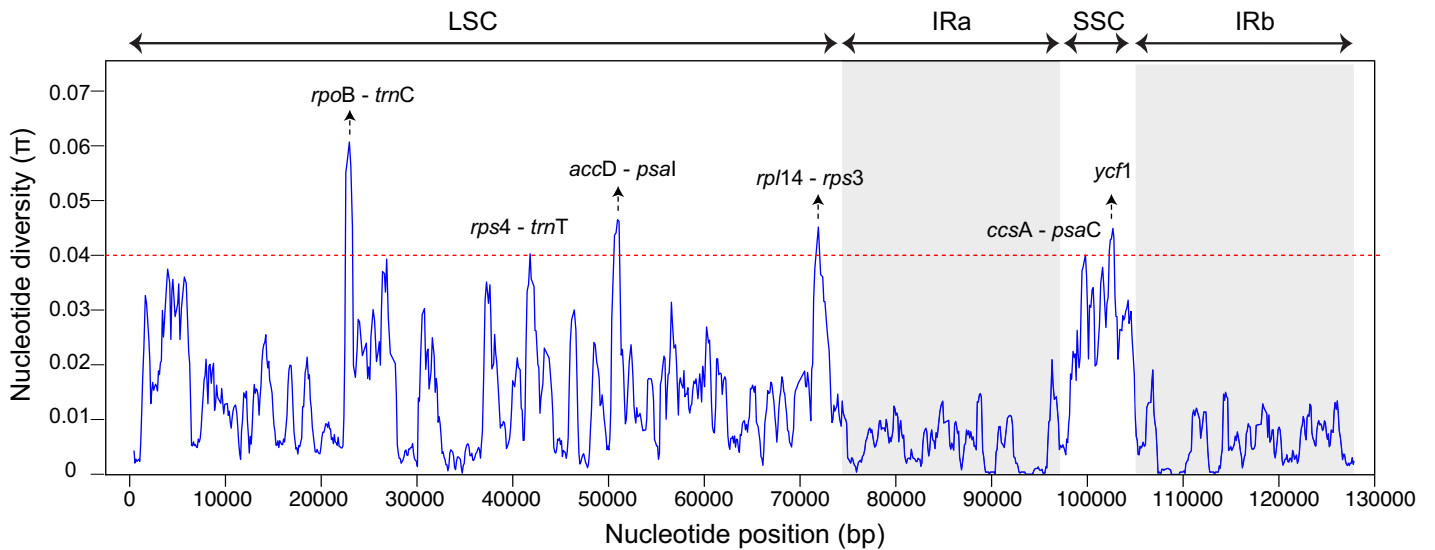


Fig 3. Nucleotide variability in 29 *Taxillus* plastomes. Sequence diversity was calculated using a sliding window analysis (window size = 500 bp, step size = 100 bp). The six most variable regions are labeled.

<https://doi.org/10.1371/journal.pone.0256345.g003>

plastid genes and seven IGS regions were arranged according to size and the proportion of parsimony informative sites and plotted on the same graph (Fig 4). Six of the seven IGS regions were less than 1000 bp in length yet all showed a high proportion of parsimony informative sites (8–25%). Although the PCGs were significantly longer, the proportion of parsimony informative sites was generally <5%.

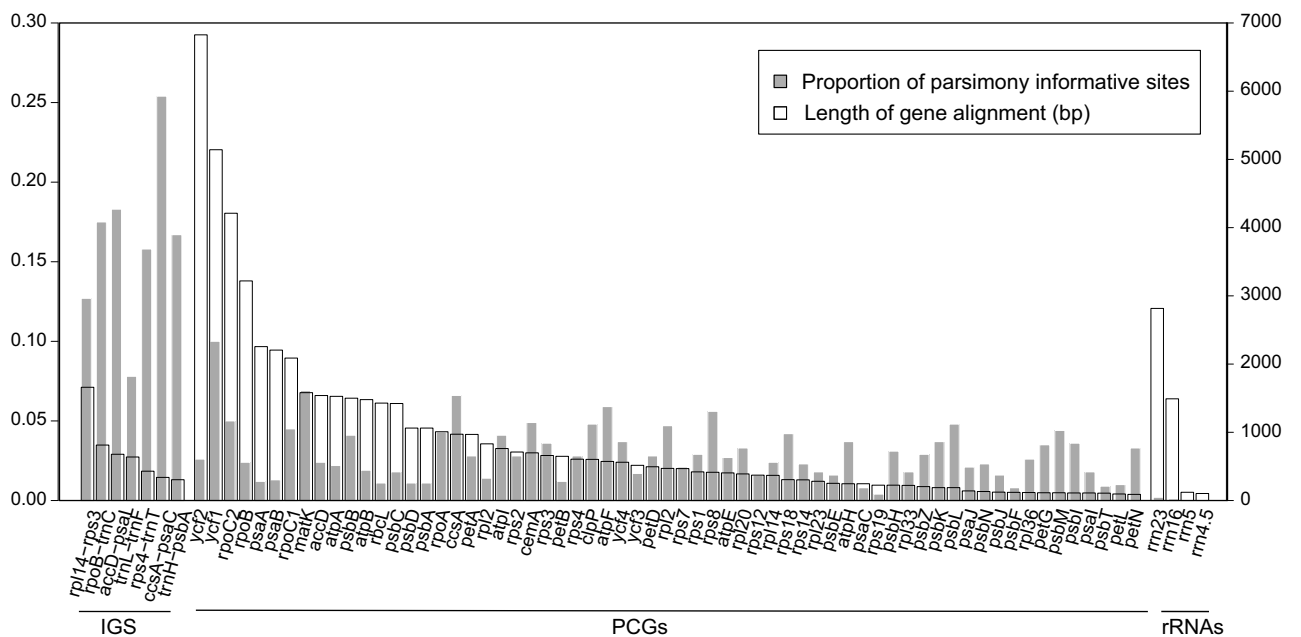


Fig 4. Proportion of parsimony informative sites and gene lengths for 29 *Taxillus* plastomes. Intergenic spacer regions (IGS), protein coding genes (PCG) and *rnm* genes (rRNA) are plotted by length (empty vertical bars, length from the sequence alignment) and proportion of parsimony informative sites (vertical filled bars).

<https://doi.org/10.1371/journal.pone.0256345.g004>

Table 1. Simple phylogenetic utility test scores^a for the 29 *Taxillus* plastid gene regions.

Gene tree ^c	Gene region	Length of alignment	Proportion of PI sites ^b	Clade I	Clade II	Clade III	Clade IV	Monophyly <i>Scurrula</i>	<i>Taxillus</i> topology	Total score
A	<i>rpl14-rps3</i>	1660	0.13	1	1	1	1	1	1	6
B	<i>rpoB-trnC</i>	813	0.18	1	1	1	0	1	0	4
C	<i>accD-psaI</i>	678	0.18	1	1	1	0	1	0	4
D	<i>trnL-trnF</i>	638	0.08	1	1	1	1	1	0	5
E	<i>rps4-trnT</i>	430	0.16	1	1	1	1	0	0	4
F	<i>ccsA-psaC</i>	339	0.25	1	1	0	0	1	0	3
G	<i>trnH-psbA</i>	305	0.17	0	0	0	0	0	0	0
H	<i>ycf1</i>	5142	0.10	1	1	1	1	1	0	5
I	<i>rpoc1</i>	2088	0.05	1	1	1	1	1	0	5
J	<i>matK</i>	1581	0.07	1	1	1	0	1	0	5
K	<i>accD</i>	1500	0.04	1	1	1	1	0	1	5
L	<i>rbcL</i>	1428	0.01	1	1	1	1	1	0	5
M	<i>rbcL+accD</i>	2913	0.03	1	1	1	1	1	1	6
N	<i>rbcL+matK</i>	3009	0.05	1	1	1	0	1	0	4
O	<i>matK+accD</i>	3066	0.06	1	1	1	0	1	0	4

^aTree score was made up of one point for each monophyletic clade I, II, III, IV, one point for monophyletic *Scurrula*, and one point for *Taxillus* clade topology matching overall plastome tree.

^b PI: Parsimonious informative.

^c See S4 Fig.

<https://doi.org/10.1371/journal.pone.0256345.t001>

The simple phylogenetic utility test results are presented for the IGSs and PCGs (Table 1 and S4 Fig). Trees generated from the seven IGS regions, including *trnH-psbA* that is commonly used as a barcode, show a range of topologies (S4 Fig). Among the IGSs, only the lengthy *rpl14-rps3* region matched the overall plastome topology (score of 6) and the shortest region (*trnH-psbA*) received a score of zero. Three of the IGSs (*rpl14-rps3*, *trnL-trnF* and *rps4-trnT*) supported monophyly of clade IV. After *rpl14-rps3* the next highest score (5) was *trnL-trnF* that did not recover the overall *Taxillus* topology. This IGS was examined further by including six additional sequences from NCBI Genbank (S5 Fig). A species not previously sampled, *Taxillus thibetensis*, emerged as part of clade I but weakly supported as sister to *T. sutchuenensis*. A major clade composed of clades I-III (but not IV) received high bootstrap support (87%) and all sequences of *T. chinensis* were strongly supported as monophyletic (100%). The taxon named *T. wiensii* from Kenya emerged outside the *Taxillus/Scurrula* clade (S5 Fig). As shown in S4 Fig, *trnH-psbA* performed poorly relative to other IGSs where species such as *Taxillus chinensis*, *T. liquidambaricola*, *T. theifer* and *T. matsudai* were resolved as monophyletic. For the PCGs, a score of five was obtained for all five genes, but only the *accD* topology received high nodal support for Clades I-IV (S4 Fig). This contrasts with *rbcL* where most nodes received weak support. As with four of the IGSs, *matK* did not yield a monophyletic clade IV and this was not corrected even when concatenated with *rbcL* or *accD* (S4 Fig). The highest score (6) among PCGs was obtained by concatenating *rbcL* with *accD*.

Genetic distances (K2P) among the plastome and nrDNA sequences of the 29 *Taxillus* and three *Scurrula* plastomes were compared (Fig 5). The intertaxon distances for their plastomes was about two-fold higher than those derived from nrDNA. The genetic distances between *Taxillus* and *Scurrula* plastomes ranged from 0.0200 to 0.0369, whereas the maximum distance within *Taxillus* (0.020) occurred between *T. chinensis* C to *T. theifer* F. For plastomes, the

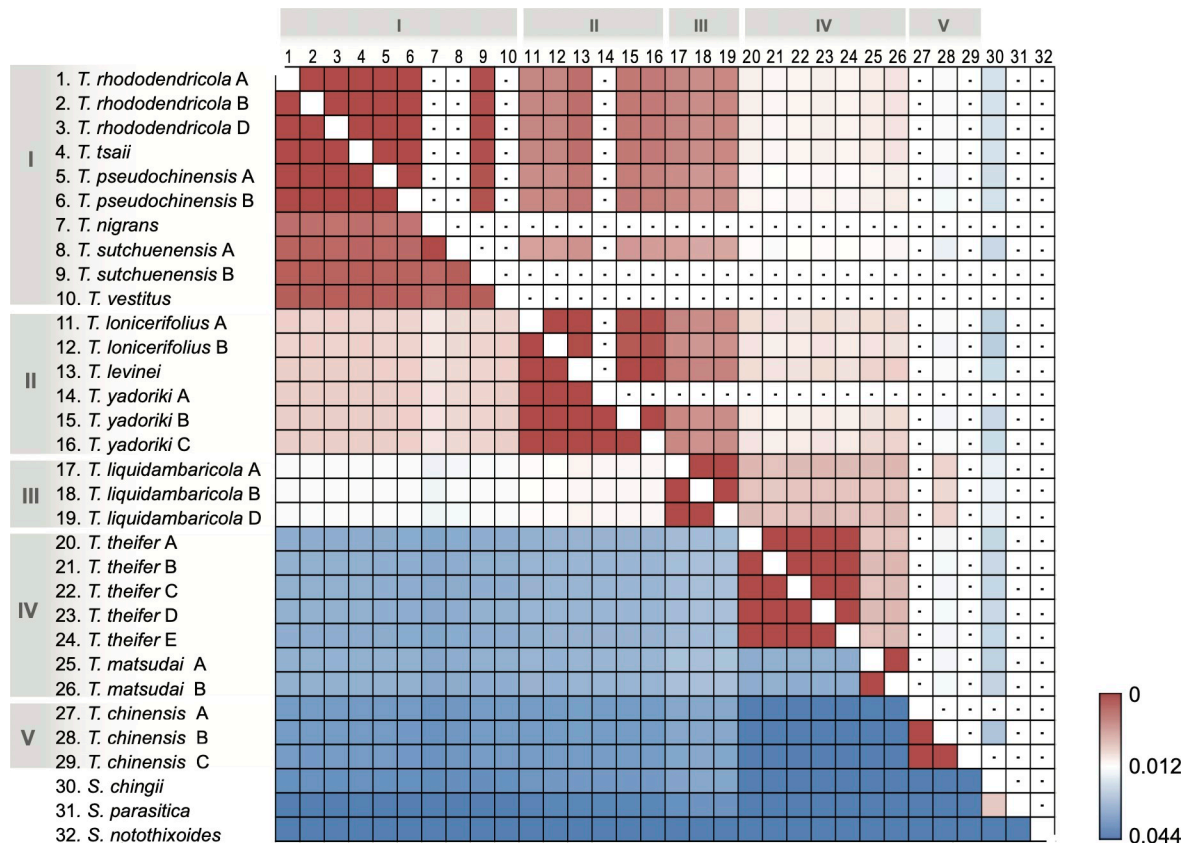


Fig 5. A heatmap of pairwise genetic distance values of the plastome and nuclear rDNA sequences. The upper right portion of the matrix contains genetic distances calculated from the nuclear rDNA sequence while the lower left portion contains genetic distance computed from whole plastome sequences. All the analyses were performed using the Kimura 2-parameter (K2P) model.

<https://doi.org/10.1371/journal.pone.0256345.g005>

lowest K2P distance (near-zero) was detected within samples of the same species and among *T. pseudochinensis*, *T. tsaii* and *T. rhododendricola* A. Similarly, for nrDNA sequences, K2P distances near zero were found among the same species samples as well as among *T. tsaii*, *T. rhododendricola* and *T. pseudochinensis*. The genetic distances among members of the same *Taxillus* clade were generally lower than 0.01.

The five specific primer sets for three genetic regions (*accD*, *matK* and *trnL-trnF*) successfully amplified their target fragments in all seven species. The fragments ranged from 562 bp to 750 bp in size and the results of agarose gel electrophoresis are shown in [S6 Fig](#). The Sanger sequences for the five additional *Taxillus* samples were incorporated into their respective gene alignments and the resulting ML gene trees of 34 *Taxillus* and three *Scurrula* sequences are shown in [S7 Fig](#). Conspecific *Taxillus* species were consistently placed together in the same clade and bootstrap values generally increased. However, only *T. liquidambaricola*, *T. theifer*, *T. matsudai* and *T. chinensis* were resolved as reciprocally monophyletic as opposed to Clade I and II species that were not.

Discussion

Plastome evolution in *Taxillus*

In the comparative analyses of *Taxillus* and *Scurrula* plastomes, we observed an overall high conservation in genome size, GC content, gene content, and IR/SC boundaries. The sizes of

the *Taxillus* plastomes (121–123 kbp) are similar to *Loranthus* [7] and near the mean value for the entire family (122.3 kbp, range 116–139 kbp from 45 values). For the non-mistletoe members of Santalales (excluding holoparasites), plastomes range in size from 118.7 kbp (*Schoepfia jasminodora*) to 156.8 kbp (*Ximenia americana*) with a mean of 138.6 kbp. Reductions in plastome size across the order have been discussed [14, 44, 45] and it appears that this process is occurring independently in several clades through gene loss and pseudogenization. As proposed in several plastome degradation models [46, 47], the first stage is pseudogenization or loss of *ndh* complex genes and this appears to be the case with all parasitic lineages in Santalales [46]. With the exception of *ndhB*, all *ndh* genes appear to have been lost in the common ancestor of Loranthaceae and Schoepfiaceae [46] (S1 Fig). Within Loranthaceae, further reductions involve *infA* and tRNA genes. It should be noted that *trnK*-UUU, *trnG*-UCC, *trnV*-UAC, and *trnI*-GAU (all group IIA intron-containing tRNA genes) have been lost or pseudogenized, but the intron processing gene *matK* remains present in these plastomes (S1 Fig), suggesting a gradual loss of intron-containing tRNA genes prior the loss of *matK* during plastome reduction. The complete loss of *matK* in Santalales has not been documented, however, its pseudogene is present in *Viscum* [48] (Viscaceae) and *Cecarria* [7] (Loranthaceae). The gene contents of the *Taxillus* plastomes are similar to the previous reports [11, 18–20, 41], but with two notable exceptions. The four genes including *ycf1*, *trnH*-GUG, *trnL*-UAG and *trnL*-UAA, were reported as missing in Li et al. (2017) [19], but following reannotation were found in *T. chinensis* and *T. sutchuenensis*. Moreover, *clpP* was reported as a pseudogene in *T. nigrans* by Zhao et al. (2019) [20], but an intact *clpP* sequence was recovered from the transcriptome data, therefore, this loss was not labeled on S1 Fig.

The nucleotide variability values (π) within the 29 *Taxillus* plastomes (0–6.4%, mean 1.3%) (Fig 3) were higher than the sequence divergence values reported for nine *Dalbergia* plastomes (0–4.4%, mean 0.86%) [49], 32 *Artemisia* plastomes (0–0.97%, mean 0.24%) [50], 40 *Populus* plastomes (0–0.71%, mean 0.36%) [51] and 65 *Picea* plastomes (mean 0.3%) [52]. The high variability within the *Taxillus* plastome sequences reflect a general molecular rate acceleration in the genomes of parasitic plants [53]. The fact that ca. 20% of the plastid genes have been lost or pseudogenized in *Taxillus* also supports the concept that plastome functional reduction is higher in heterotrophic than autotrophic plants [14, 46].

Phylogenetic and taxonomic considerations

The phylogenetic analysis of 13 species of *Taxillus* using complete plastome sequences (Fig 2) resulted in four well-supported clades with less support for the position of the single accession of *T. chinensis*. This topology showed some minor conflicts with the 67 plastid gene tree (Fig 1) and nrDNA tree (Fig 2). Taken together, we interpreted the full plastome topology in Fig 2 as the best reflection of the species phylogeny and thus used this as the standard for comparing individual plastomic regions to each other (S4 Fig). Conflicts between cytoplasmic organellar and nuclear phylogenies may indicate potential introgressive hybridization events [54], incomplete lineage sorting [55] or differences in dispersal ability of plastid vs. nuclear DNA [56, 57]. Given the minor phenotypic differences among *T. rhododendricola*, *T. tsaii* and *T. pseudochinensis* [58] and that their plastome and rDNA genetic distances are extremely low (0–0.0010) (Fig 5), it is possible that these three taxa are conspecific or a part of a closely related species complex. A broader sampling and further analyses are required to assess genetic diversity among these taxa.

Despite support for five *Taxillus* clades, the distribution of character states for the four morphological characters examined did not differentiate these clades. As already shown by Danser (1931) and Barlow (1991), fruit features best distinguish *Taxillus* from *Scurrula* but have little

value differentiating the 13 species of *Taxillus* included here. It is possible that expanding the number and type of morphological characters (quantitative and qualitative), as well as subjecting the data to multivariate analysis, could result in relationships congruent with the molecular phylogeny, but this remains to be determined.

The genetic data obtained here allows some insight into *Taxillus* names that were considered synonyms by World Flora Online (WFO). As mentioned above, *T. rhododendricola* (accessions A and B) were resolved within a clade containing *T. tsaii* and *T. pseudochinensis*. WFO synonymized *T. rhododendricola* with *T. nigrans*, but that species is genetically more distant from this taxon based on interspecific K2P plastome distances (Fig 5), thus at this time it seems prudent not to combine these taxa. Moreover, the genetic distance between *T. nigrans* and *T. sutchuenensis* is very low, suggesting these taxa are conspecific.

The ML trees shown in S5 Fig provides additional insights into *Taxillus* relationships. Using only *trnL-trnF* sequences, *T. thibetensis* is weakly supported as part of clade I. The added sequences of *T. tsaii*, *T. sutchuenensis*, and *T. chinensis* were placed in conspecific clades, thus their identification was likely accurate. The sequence of *T. wiensii* (voucher A. Robertson 7364) from coastal Kenya emerged as sister to *Taxillus* and *Scurrula*, not as part of any of the former's clades. Doubts about the placement of this taxon in *Taxillus* were expressed by Polhill & Wiens (1998) [59] who suggested that it is perhaps more similar to other African genera. Indeed, this view was supported by the analysis of ITS and *trnL-trnF* [60] where *T. wiensii* was not sister to *Scurrula* but associated with genera from subtribe Tapinanthinae.

The topology of the *trnH-psbA* tree shows that this region performed poorly compared with other IGSs (S4 Fig) which can likely be attributed to short sequence length (297 bp) that provided few characters. Here accessions identified as *T. chinensis* and *Scurrula chingii* occurred in two distinct clades each. Incomplete lineage sorting or misidentification of voucher specimens could also be involved, particularly for common and widespread species such as *T. chinensis* where this name is used for specimens collected by nonspecialists. Potential misidentification of *Taxillus* samples are also revealed in S8 Fig. The *T. thibetensis* accession (MH161427) was embedded within the *Scurrula* clade while the accession identified as *T. delavayi* (MH161427) was sister to the *Taxillus/Scurrula* clade. However, because voucher information for the two sequence accessions is not available, it is difficult to investigate this issue further. This problem highlights the importance of requiring readily accessible voucher data (preferably photographs) for any organism whose DNA sequences are deposited in public repositories.

Potential barcodes in *Taxillus*

The challenges in developing broad scale, easily implemented barcodes in plants have been numerous. For plastome genes, *rbcl* is easy to PCR amplify but has only “modest resolving power” across angiosperms [61]. In contrast, *matK* provides greater numbers of informative characters but finding PCR priming sites flanking this region that are “universally” conserved is near impossible. This results in workers designing taxon-specific primers, thus defeating the concept of universal barcodes. As shown in this study, better performance is sometimes achieved by concatenating plastome loci. As stated by Fazekas et al. [62], there is a limit to resolution no matter which region or regions are chosen, and this was empirically demonstrated for *Crocus* by Seberg and Petersen (2009) [63] where number of species identified reached a plateau at ca. 4000 bp. This problem is exacerbated in “taxonomically complex groups where species limits are often very narrowly defined” [61]. Although complete plastid genome sequencing has been proposed [64, 65], cost and other factors currently limit this method compared with PCR based barcoding methods. An intermediate approach has been taken by some

workers where a subset of taxa are subjected to complete plastome sequencing and then lineage-specific loci identified and used as DNA barcodes for species level identification and phylogenetics [15, 66–68]. This was the approach taken in the current study.

The comparison of PCGs and IGSs among the 29 *Taxillus* plastomes indicates that the latter generally display a higher proportion of parsimony informative sites than the former (Fig 4). One might then assume they are more phylogenetically useful; however, given the simple phylogenetic utility scoring system employed herein, the opposite conclusion is reached where scores were on average higher for PCGs, not IGSs. The reasons for this are complex owing to the interplay of a number of variables. These include: 1) taxonomic group, 2) taxon sampling, 3) length of the DNA region, 4) proportion of variable sites, 5) proportion of parsimony informative sites, and 6) the quality of the parsimony informative sites. A brief scan of the plant barcoding literature shows that different regions and combinations have been recommended depending upon what taxonomic rank is required. The *rbcL*+*matK* combination may be effective in placing a plant in a family but may not be effective in discriminating among species of different plant groups. In terms of sampling, Yonghua et al. (2010) [69] recommended *trnH-psbA* as a barcode locus, but only five species of *Taxillus* were included in their evaluation. Our more robust sampling shows that this marker is unsuitable for interspecific discrimination. As shown in Table 1 and S4 Fig, the lengths of the IGS and PCG regions examined varied from very long (*ycf1*) to very short (*trnH-psbA*). From a practical perspective, *ycf1* (ca. 5 kbp) is too long for routine PCR amplifications. The proportion of variable and parsimony informative sites within the PCGs and IGSs also varied where the latter were generally higher. But variability alone does not automatically translate into higher phylogenetic utility. For example, *trnH-psbA* is likely too variable to be effective for species discrimination. Although a site may be scored as parsimony informative, the synapomorphies may be for clades that are not present on the reference tree (here the complete plastome tree), i.e. the variation is homoplastic. The total number of parsimony informative sites (length x proportion) is similarly a nonintuitive measure of phylogenetic utility. Two of the IGSs (*rpoC-trnC* and *accD-psaI*) showed higher numbers of parsimony informative sites than *trnL-trnF*, yet the latter obtained a higher phylogenetic utility score (5) simply because it yielded a monophyletic clade IV.

All five of the examined PCGs received simple phylogenetic utility scores of five, but only *accD* recovered the overall *Taxillus* topology. *matK* also reconstructed the overall phylogeny except for a monophyletic clade IV. Only two of the regions tested received perfect scores of six: *rpl14-rps3* and the concatenated *rbcL + accD*. The former is promising as a barcode marker because its length allows easy PCR amplification with just two internal primers. Other longer regions, such as *ycf1* and the intron-containing gene *rpoC1* may also provide barcode loci if internal PCR primers were designed. The *rbcL + accD* combination appears to be the best barcode candidate for resolving species relationships in *Taxillus* (S4 Fig). Although of similar length and containing similar numbers of parsimony informative sites, the other two PCG combinations (*rbcL + matK*, *accD + matK*) did not recover the relationship between *T. thelifer* and *T. matsudai* in clade IV, again highlighting the complex interaction of variables.

Conclusions

This study reports plastome organization and phylogenetic analyses of the hemiparasitic mistletoe *Taxillus*, a genus whose species are difficult to differentiate based on morphological characters. The plastome phylogeny of 13 *Taxillus* and three *Scurrula* species showed well-resolved infrageneric relationships and the topology was largely congruent with the relationships based on mitochondrial and nuclear rDNA sequences. Maximum likelihood phylogenetic trees were built for individual intergenic spacer sequences and protein coding genes and

these were evaluated based on the complete plastome reference tree. From 12 examined regions/genes, the *rpl14-rps3* IGS and concatenated *rbcL* + *accD* were found to have the highest phylogenetic utility as measured by their similarity to the reference tree. Future work will involve acquiring sequences from additional samples of *Taxillus* and *Scurrula* and testing the utility of the barcodes identified in this study to differentiate species.

Supporting information

S1 Raw gel images. Original gel images of S6 Fig.
(PDF)

S1 Fig. Phylogenetic relationships in *Taxillus* with putative events of gene loss and pseudogenization plotted on the tree. The putative events of gene loss and pseudogenization are inferred based on the most parsimonious scenario. Arrows indicate gene losses (black) or pseudogenization (blue).
(EPS)

S2 Fig. Comparison of the IR boundaries among 13 *Taxillus* and other Lorantheaceae accessions.
(EPS)

S3 Fig. Comparison of mitochondrial rDNA MP and ML gene trees for *Taxillus* species. MPBS and MLBS values are indicated above the branches.
(EPS)

S4 Fig. Plastome barcode regions and the tree topologies. Trees containing 29 *Taxillus* and three *Scurrula* sequences generated from separate ML analyses of IGS (A-G) and PCG (H-O) regions. Numbers in parentheses indicate the length of the gene alignment, proportion of parsimony informative sites and score of relationship consistency. The *Helixanthera* sequences were used as outgroups on all trees. The four major clades of *Taxillus* as well as *Scurrula* are shown in different colors. Bootstrap values >70% are labeled above branches.
(EPS)

S5 Fig. The ML analysis of *T. wiensii* and other *Taxillus* using the *trnL-trnF* region.
(EPS)

S6 Fig. PCR amplification of target plastid loci for the *Taxillus* species. (A) *accD* region 1, (B) *accD* region 2, (C) *matK* region 1 (D) *matK* region 2 (E) *trnL-trnF*. The original raw images are provided as [S1 Raw gel images](#).
(TIFF)

S7 Fig. The *matK*, *accD* and *trnL-trnF* trees with increased sampling. Trees containing 34 *Taxillus* and three *Scurrula* sequences generated from separate ML analyses. Numbers in parentheses indicate the score of relationship consistency. Bootstrap values >70% are labeled above branches.
(EPS)

S8 Fig. Potential sample misidentification revealed by the phylogeny of *Taxillus* and other Lorantheaceae. Two potential misidentification accessions are shown in blue. The topology was generated from ML analysis of the 67 plastid gene sequence data (63 PCGs and four rRNAs). ML bootstrap values are shown above the branches and *Schoepfia jasminodora* was used as outgroup.
(EPS)

S1 Table. Taxonomic information for 39 potential *Taxillus* species as well as selected synonyms.

(PDF)

S2 Table. Sample voucher information and accession number of newly obtained sequences in Genbank.

(XLSX)

S3 Table. List of the published accessions used in the phylogenetic analysis.

(XLSX)

S4 Table. Specific primers designed for PCR amplification in selected *Taxillus* samples.

(XLSX)

S5 Table. Characteristics of complete plastomes in the *Taxillus* and *Scurrula* species.

(XLSX)

Acknowledgments

We are thankful to Mr. Chien-Hsun Yu at the Forestry Bureau of Taiwan and Mr. Ding-Hong Yeh at Taiwan Forestry Research Institute for their kind assistance in plant collections. The authors thank the staff of the herbaria TAIF and HAST that provided materials for this study and the Clinical and Industrial Genomic Application Development Service Center of National Core Facility for Biopharmaceuticals, Taiwan for NGS sequencing.

Author Contributions

Conceptualization: Huei-Jiun Su, Daniel L. Nickrent.

Data curation: Huei-Jiun Su, Daniel L. Nickrent.

Formal analysis: Huei-Jiun Su, Shu-ling Liang, Daniel L. Nickrent.

Funding acquisition: Huei-Jiun Su.

Investigation: Huei-Jiun Su, Shu-ling Liang.

Methodology: Huei-Jiun Su, Daniel L. Nickrent.

Project administration: Huei-Jiun Su.

Resources: Huei-Jiun Su.

Visualization: Huei-Jiun Su.

Writing – original draft: Huei-Jiun Su.

Writing – review & editing: Huei-Jiun Su, Shu-ling Liang, Daniel L. Nickrent.

References

1. Nickrent DL, Anderson F, Kuijt J. Inflorescence evolution in Santalales: integrating morphological characters and molecular phylogenetics. *Am J Bot.* 2019; 106(3):402–14. <https://doi.org/10.1002/ajb2.1250> PMID: 30856677.
2. Liu B, Le CT, Barrett RL, Nickrent DL, Chen Z, Lu L, et al. Historical biogeography of Loranthaceae (Santalales): Diversification agrees with emergence of tropical forests and radiation of songbirds. *Mol Phylogenet Evol.* 2018; 124:199–212. <https://doi.org/10.1016/j.ympev.2018.03.010> PMID: 29550535.
3. Su H-J, Hu J-M, Anderson FE, Der JP, Nickrent DL. Phylogenetic relationships of Santalales with insights into the origins of holoparasitic Balanophoraceae. *Taxon.* 2015; 64(3):491–506. <https://doi.org/10.12705/643.2>

4. Vidal-Russell R, Nickrent DL. Evolutionary relationships in the showy mistletoe family (Loranthaceae). *American Journal of Botany*. 2008; 95:1015–29. <https://doi.org/10.3732/ajb.0800085> PMID: 21632422
5. Nickrent DL. Parasitic angiosperms: How often and how many? *Taxon*. 2020; 69(1):5–27. <https://doi.org/10.1002/tax.12195>
6. Amico GC, Vidal-Russell R, Nickrent DL. Phylogenetic relationships and ecological speciation in the mistletoe *Tristerix* (Loranthaceae): the influence of pollinators, dispersers, and hosts. *Am J Bot*. 2007; 94(4):558–67. <https://doi.org/10.3732/ajb.94.4.558> PMID: 21636426.
7. Nickrent DL, Su H-J, Lin R-Z, Devkota MP, Hu J-M, Glatzel G. Examining the needle in the haystack: evolutionary relationships in the mistletoe genus *Loranthus* Jacq. (Loranthaceae). *Systematic Botany*. 2021; In press.
8. Caraballo-Ortiz MA. *Phylotranscriptomics, taxonomy, and ecology of the caribbean-endemic mistletoes Dendropemon* (Loranthaceae). University Park, Pennsylvania 16802: Pennsylvania State University; 2017.
9. Francisco-Ornelas J, Gándara E, Vásquez-Aguilar AA, Ramírez-Barahona S, Ortiz-Rodríguez AE, González C, et al. A mistletoe tale: postglacial invasion of *Psittacanthus schiedeanus* (Loranthaceae) to Mesoamerican cloud forests revealed by molecular data and species distribution modeling. *BMC Evolutionary Biology*. 2016; 16(78).
10. Ortiz-Rodríguez AE, Guerrero E. Y., Ornelas JF. Phylogenetic position of Neotropical *Bursera*-specialist mistletoes: the evolution of deciduousness and succulent leaves in *Psittacanthus* (Loranthaceae). *Botanical Sciences*. 2018; 96(3):443–61.
11. Guo X, Ruan Z, Zhang G. The complete plastome of *Taxillus vestitus* (Loranthaceae), a hemiparasitic plant. *Mitochondrial DNA Part B*. 2019; 4(2):3188–9. <https://doi.org/10.1080/23802359.2019.1667912> PMID: 33365912
12. Jansen RK, Cai Z, Raubeson LA, Daniell H, Depamphilis CW, Leebens-Mack J, et al. Analysis of 81 genes from 64 plastid genomes resolves relationships in angiosperms and identifies genome-scale evolutionary patterns. *Proc Natl Acad Sci U S A*. 2007; 104(49):19369–74. <https://doi.org/10.1073/pnas.0709121104> PMID: 18048330.
13. Daniell H, Lin CS, Yu M, Chang WJ. Chloroplast genomes: diversity, evolution, and applications in genetic engineering. *Genome Biol*. 2016; 17(1):134. <https://doi.org/10.1186/s13059-016-1004-2> PMID: 27339192.
14. Chen X, Fang D, Wu C, Liu B, Liu Y, Sahu SK, et al. Comparative plastome analysis of root- and stem-feeding parasites of Santalales untangle the footprints of feeding mode and lifestyle transitions. *Genome Biol Evol*. 2019. <https://doi.org/10.1093/gbe/evz271> PMID: 31845987.
15. Awad M, Fahmy RM, Mosa KA, Helmy M, El-Feky FA. Identification of effective DNA barcodes for *Triticum* plants through chloroplast genome-wide analysis. *Comput Biol Chem*. 2017; 71:20–31. <https://doi.org/10.1016/j.compbiolchem.2017.09.003> PMID: 28961510.
16. Teshome GE, Mekbib Y, Hu G, Li ZZ, Chen J. Comparative analyses of 32 complete plastomes of *Tef* (*Eragrostis tef*) accessions from Ethiopia: phylogenetic relationships and mutational hotspots. *PeerJ*. 2020; 8:e9314. <https://doi.org/10.7717/peerj.9314> PMID: 32596045.
17. Dodsworth S. Genome skimming for next-generation biodiversity analysis. *Trends Plant Sci*. 2015; 20(9):525–7. <https://doi.org/10.1016/j.tplants.2015.06.012> PMID: 26205170.
18. Cho W-B, Han E-K, Son DC, Lee J-H. The complete chloroplast genome sequence of *Taxillus yadoriki* (Loranthaceae): a hemi-parasitic evergreen shrub in East Asia. *Mitochondrial DNA Part B*. 2020; 5(3):3190–1. <https://doi.org/10.1080/23802359.2020.1806755> PMID: 33458100
19. Li Y, Zhou JG, Chen XL, Cui YX, Xu ZC, Li YH, et al. Gene losses and partial deletion of small single-copy regions of the chloroplast genomes of two hemiparasitic *Taxillus* species. *Sci Rep*. 2017; 7(1):12834. <https://doi.org/10.1038/s41598-017-13401-4> PMID: 29026168.
20. Zhao S, Li J, Ma R, Miao N, Mao Q, Mao K. Characterization of the complete chloroplast genome of *Taxillus nigrans*. *Mitochondrial DNA Part B*. 2019; 4(1):472–3. <https://doi.org/10.1080/23802359.2018.1536485>
21. Nickrent DL, Malecot V, Vidal-Russell R, Der JP. A revised classification of Santalales. *Taxon*. 2010; 59(2):538–58. WOS:000276617400019.
22. Danser BH. The Loranthaceae of the Netherlands Indies. *Bulletin du Jardin Botanique Buitenzorg*. 1931; 11(Series 3):233–519.
23. Barlow BA. Conspectus of the genera *Scurrula* L. and *Taxillus* Tieghem (Loranthaceae) in the Malesian region. *Blumea*. 1991; 36:63–85.
24. Barlow BA. Loranthaceae. *Flora Malesiana series 1*. 1997; 13:209–401.

25. Qiu H, Gilbert MG. Loranthaceae. In: Wu ZY, Raven PH, Hong DY, editors. Flora of China Vol5 (Ulmaceae through Basellaceae). 19: Science Press, Beijing, and Missouri Botanical Garden Press, St. Louis.; 2003.
26. Chiu ST. Loranthaceae. In: Huang TC, editors. Flora of Taiwan. Vol 2. 2 ed. Taipei, Taiwan: Editorial Committee of the Flora of Taiwan; 1996. p. 269–86.
27. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 1987; 19(1):11–5.
28. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014; 30:2114–2120. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404.
29. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009; 25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324> PMID: 19451168.
30. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics.* 2009; 25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943.
31. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative genomics viewer. *Nat Biotechnol.* 2011; 29(1):24–6. <https://doi.org/10.1038/nbt.1754> PMID: 21221095.
32. Tillich M, Lehwarck P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, et al. GeSeq—versatile and accurate annotation of organelle genomes. *Nucleic Acids Res.* 2017; 45(W1):W6–W11. <https://doi.org/10.1093/nar/gkx391> PMID: 28486635.
33. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013; 30(4):772–80. <https://doi.org/10.1093/molbev/mst010> PMID: 23329690.
34. 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.
35. Suyama M, Torrents D, Bork P. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 2006; 34(Web Server issue):W609–12. <https://doi.org/10.1093/nar/gkl315> PMID: 16845082.
36. 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.
37. Hoang DT, Vinh LS, Flouri T, Stamatakis A, von Haeseler A, Minh BQ. MPBoot: fast phylogenetic maximum parsimony tree inference and bootstrap approximation. *BMC Evol Biol.* 2018; 18(1):11. Epub 2018/02/03. <https://doi.org/10.1186/s12862-018-1131-3> PMID: 29390973.
38. Jiang D, Ma R, Li J, Mao Q, Miao N, Mao K. Characterization of the complete chloroplast genome of *Scurrula parasitica*. *Mitochondrial DNA Part B.* 2018; 4(1):247–8. <https://doi.org/10.1080/23802359.2018.1547135>
39. Yuan L-X, Wang J-H, Chen C-R, Zhao K-K, Zhu Z-X, Wang H-F. Complete chloroplast genome sequence of *Scurrula nothoxoides* (Loranthaceae): a hemiparasitic shrub in South China. *Mitochondrial DNA Part B.* 2018; 3(2):580–1. <https://doi.org/10.1080/23802359.2018.1471366> PMID: 33474249
40. Shin HW, Lee NS. Understanding plastome evolution in Hemiparasitic Santalales: Complete chloroplast genomes of three species, *Dendrotrophe varians*, *Helixanthera parasitica*, and *Macrosolen cochinchinensis*. *PLoS One.* 2018; 13(7):e0200293. <https://doi.org/10.1371/journal.pone.0200293> PMID: 29975758.
41. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. *Mol. Biol. Evol.* 2017; 34(12):3299–3302. <https://doi.org/10.1093/molbev/msx248> PMID: 29029172.
42. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol Evol.* 2018; 35(6):1547–9. <https://doi.org/10.1093/molbev/msy096> PMID: 29722887.
43. Borowiec ML. AMAS: a fast tool for alignment manipulation and computing of summary statistics. *PeerJ.* 2016; 4:e1660. <https://doi.org/10.7717/peerj.1660> PMID: 26835189.
44. Li M, Zhang Y, Li Y, Zhang L. The complete chloroplast genome of *Scurrula chingii* (W.C. Cheng) H.S. Kiu (Loranthaceae), a hemiparasitic shrub. *Mitochondrial DNA Part B.* 2021; 6(1):282–4. <https://doi.org/10.1080/23802359.2020.1863166> PMID: 33553646
45. Guo X, Liu C, Zhang G, Su W, Landis JB, Zhang X, et al. The complete plastomes of five hemiparasitic plants (*Osyris wightiana*, *Pyralaria edulis*, *Santalum album*, *Viscum liquidambaricolum*, and *V. ovalifolium*): comparative and evolutionary analyses within Santalales. *Front Genet.* 2020; 11:597. <https://doi.org/10.3389/fgene.2020.00597> PMID: 32612639.

46. Wicke S, Muller KF, dePamphilis CW, Quandt D, Bellot S, Schneeweiss GM. Mechanistic model of evolutionary rate variation en route to a nonphotosynthetic lifestyle in plants. *Proc Natl Acad Sci U S A*. 2016; 113(32):9045–50. <https://doi.org/10.1073/pnas.1607576113> PMID: 27450087.
47. Wicke S, Naumann J. Molecular evolution of plastid genomes in parasitic flowering plants. In: Chaw S-M, Jansen RK, editors. *Advances in Botanical Research*. 85: Elsevier; 2018. p. 315–47.
48. Petersen G, Cuenca A, Seberg O. Plastome evolution in hemiparasitic mistletoes. *Genome Biol Evol*. 2015; 7(9):2520–32. <https://doi.org/10.1093/gbe/evv165> PMID: 26319577.
49. Qin M, Zhu C-J, Yang J-B, Vatanparast M, Schley R, Lai Q, et al. Comparative analysis of complete plastid genome reveals powerful barcode regions for identifying wood of *Dalbergia odorifera* and *D. tonkinensis* (Leguminosae). *Journal of Systematics and Evolution*. 2020;n/a(n/a). <https://doi.org/10.1111/jse.12598>
50. Kim GB, Lim CE, Kim JS, Kim K, Lee JH, Yu HJ, et al. Comparative chloroplast genome analysis of *Artemisia* (Asteraceae) in East Asia: insights into evolutionary divergence and phylogenomic implications. *BMC Genomics*. 2020; 21(1):415. <https://doi.org/10.1186/s12864-020-06812-7> PMID: 32571207.
51. Zong D, Gan P, Zhou A, Zhang Y, Zou X, Duan A, et al. Plastome sequences help to resolve deep-level relationships of populus in the family Salicaceae. *Front Plant Sci*. 2019; 10:5. <https://doi.org/10.3389/fpls.2019.00005> PMID: 30723484.
52. Sullivan AR, Schiffthaler B, Thompson SL, Street NR, Wang XR. Interspecific plastome recombination reflects ancient reticulate evolution in *Picea* (Pinaceae). *Mol Biol Evol*. 2017; 34(7):1689–701. <https://doi.org/10.1093/molbev/msx111> PMID: 28383641.
53. Bromham L, Cowman PF, Lanfear R. Parasitic plants have increased rates of molecular evolution across all three genomes. *BMC Evol Biol*. 2013; 13:126. <https://doi.org/10.1186/1471-2148-13-126> PMID: 23782527
54. Tsitrone A, Kirkpatrick M, Levin DA. A model for chloroplast capture. *Evolution*. 2003; 57(8):1776–82. <https://doi.org/10.1111/j.0014-3820.2003.tb00585.x> PMID: 14503619.
55. Maddison WP, Knowles LL. Inferring phylogeny despite incomplete lineage sorting. *Syst Biol*. 2006; 55(1):21–30. <https://doi.org/10.1080/10635150500354928> PMID: 16507521.
56. Currat M, Ruedi M, Petit R, Excoffier L. The hidden side of invasions: massive introgression by local genes. *Evolution* 2008; 62:1908–20. <https://doi.org/10.1111/j.1558-5646.2008.00413.x> PMID: 18452573
57. Petit R, Excoffier L. Gene flow and species delimitation. *Trends in Ecology & Evolution*. 2009; 24:386–93. <https://doi.org/10.1016/j.tree.2009.02.011> PMID: 19409650
58. Chiu S-T. Notes on the genus *Taxillus* Van Tieghem (Loranthaceae) in Taiwan. *Taiwania*. 1996; 41(2):154–67.
59. Polhill R, Wiens D. *Mistletoes of Africa*. Richmond-Surrey, UK: The Royal Botanic Gardens Kew; 1998. 370 p.
60. Wilson CA, Calvin C. Character divergences and convergences in canopy-dwelling Loranthaceae. *Botanical Journal of the Linnean Society*. 2006; 150:101–13.
61. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. *PLoS One*. 2011; 6(5):e19254. <https://doi.org/10.1371/journal.pone.0019254> PMID: 21637336.
62. Fazekas AJ, Burgess KS, Kesanakurti PR, Graham SW, Newmaster SG, Husband BC, et al. Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE* 2008; 3:e2802. <https://doi.org/10.1371/journal.pone.0002802> PMID: 18665273
63. Seberg O, Petersen G. How many loci does it take to DNA barcode a *Crocus*? *PLoS ONE*. 2009; 4(2):e4598. <https://doi.org/10.1371/journal.pone.0004598> PMID: 19240801
64. Nock C, Waters D, Edwards M, Bowen S, Rice N, Cordeiro G, et al. Chloroplast genome sequences from total DNA for plant identification. *Plant Biotechnol Journal*. 2011; 9(3):328–33. <https://doi.org/10.1111/j.1467-7652.2010.00558.x> PMID: 20796245
65. Parks M, Cronn R, Liston A. Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biology*. 2009; 7(1):84. <https://doi.org/10.1186/1741-7007-7-84> PMID: 19954512
66. Magdy M, Ou L, Yu H, Chen R, Zhou Y, Hassan H, et al. Pan-plastome approach empowers the assessment of genetic variation in cultivated *Capsicum* species. *Hortic Res*. 2019; 6:108. <https://doi.org/10.1038/s41438-019-0191-x> PMID: 31645963.
67. Jiao L, Lu Y, He T, Li J, Yin Y. A strategy for developing high-resolution DNA barcodes for species discrimination of wood specimens using the complete chloroplast genome of three *Pterocarpus* species. *Planta*. 2019; 250(1):95–104. <https://doi.org/10.1007/s00425-019-03150-1> PMID: 30923906.

68. Song F, Li T, Burgess KS, Feng Y, Ge XJ. Complete plastome sequencing resolves taxonomic relationships among species of *Calligonum* L. (Polygonaceae) in China. *BMC Plant Biol.* 2020; 20(1):261. <https://doi.org/10.1186/s12870-020-02466-5> PMID: 32513105.
69. Yonghua L, Jinlan R, Shilin C, Jingyuan S, Kun L, Dong L, et al. Authentication of *Taxillus chinensis* using DNA barcoding technique. *Journal of Medicinal Plants Research.* 2010; 4(24):2706–9. <https://doi.org/10.5897/jmpr09.677>