MITOCHONDRIAL DNA PART B: RESOURCES
2024, VOL. 9, NO. 2, 272–276
https://doi.org/10.1080/23802359.2024.2316071

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

The complete chloroplast genome of *Nekemias hypoglauca* (Hance) J. Wen & Z. L. Nie 2014 (Family: Vitaceae) and its phylogenetic analysis

Guan-Hao He[a], Lei Zhang[a], Ying Meng[a], Jun Wen[b] and Ze-Long Nie[a]

[a]College of Biology and Environmental Sciences, Jishou University, Jishou, Hunan, China; [b]Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

ABSTRACT

Nekemias is a perennial woody vine with nine species that had been originally placed in *Ampelopsis*. These species of *Nekemias* are economically and medically important. Limited information is available on the genomic characteristics of the chloroplasts of this genus. *Nekemias hypoglauca* (Hance) J. Wen & Z. L. Nie 2014 contains 131 unique genes (86 protein-coding genes, 8 rRNAs, and 37 tRNAs). The complete chloroplast sequence contains 162,976 bp. The large single-copy region contains 89,291 bp; the small single-copy region contains 19,063 bp, and a pair of inverted repeat sequences is composed of 27,311 bp. There are 84 simple sequence repeat (SSR) loci in the complete chloroplast genome of *N. hypoglauca*, with mononucleotide, dinucleotide, trinucleotide, tetranucleotide and hexanucleotide SSRs of 58, 9, 6, 10 and 1, respectively. A total of 337 repeats were identified, including 172 forward repeats, three reverse repeats and 163 palindromic repeats. A phylogenetic analysis based on the complete genome data of the chloroplasts of 10 plant species indicated the monophyly of *Nekemias* and determined the phylogenetic relationships of *N. hypoglauca* in *Nekemias*. This study provides a reference for further studies on the taxonomy, identification, origin and evolution of *N. hypoglauca* and *Nekemias*.

ARTICLE HISTORY
Received 2 June 2023
Accepted 4 February 2024

KEYWORDS
Chloroplast genome; phylogenetic relationship; *Nekemias hypoglauca*

1. Introduction

Nekemias hypoglauca (Hance) J. Wen & Z. L. Nie 2014 is a perennial woody and vine species that is characterized by opposite tendril leaves, which are mostly forked to sometimes tridentate, and pinnately compound leaves. *N. hypoglauca* is primarily found in Jiangxi Province, Fujian Province and Guangdong Province, China. The genus was separated from *Ampelopsis* by Wen et al. (2014) with nine species that are disjunctly distributed in East Asia and North America. *N. hypoglauca* is rich in flavonoids and economically and medically important. With the advent of next-generation sequencing technologies that can measure data at the genome level, it provides important information to study the chloroplast genome of this genus. The chloroplast genomes of *N. cantoniensis* (Hook. & Arn.) J. Wen & Z.L. Nie, *N. grossedentata* (hand-mazz.) J. Wen & Z.L. Nie, and *N. arborea* (L.) J. Wen & Boggan have been reported (Gu et al. 2020; Liu et al. 2021; Luo et al. 2022). In this study, we sequenced and assembled the chloroplast genome of *N. hypoglauca*. In addition, we constructed a new phylogenetic tree based on the whole genome of *Nekemias* chloroplasts using a richer sample of taxonomic units, and the phylogeny of *Nekemias* appears to have a higher resolution than that described in previous studies.

2. Materials and methods

Fresh leaves were collected from Jiangxi Province, China (115°31′ 82 east longitude, 25°17′ 41 north latitude) (Figure 1) and dried using silica gel. A voucher specimen is stored in

Figure 1. *Nekemias hypoglauca* plant. This photo was taken by Ze-Long Nie with the author's permission for use.

CONTACT Ze-Long Nie niez@jsu.edu.cn College of Biology and Environmental Sciences, Jishou University, Jishou, Hunan, China.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2024.2316071.

the herbarium at Jishou University (https://www.jsu.edu.cn/, Ying Meng, mengying@jsu.edu.cn) with a voucher number of Chu211. The DNA was extracted using a modified CTAB procedure. The genomic DNA was fragmented using a Q800R2 ultrasonicator (QSonica, Newtown, CT, USA) that resulted in a final DNA fragment that was 400-500 bp long. Total DNA libraries were constructed for enrichment using an NEBNext Ultra II DNA Library Prep kit for Illumina (New England Biolabs, Ipswich, MA, USA). The library sample was sent to a high-throughput sequencing company for sequencing using Illumina HiSeq (San Diego, CA, USA) to obtain bidirectional reads that were 150 bp long.

The raw data was assembled using the GetOrganelle program (Jin et al. 2020) and annotated on GeSeq (https://chlorobox.mpimpgolm.mpg.de/geseq.html) (Tillich et al. 2017). Furthermore, a map of the chloroplast genome of *N. hypoglauca* was drawn using a web link (https://chlorobox.mpimpgolm.mpg.de/OGD-raw.html) (Lohse et al. 2007).

Simple sequence repeats (SSRs) were examined using the MISA online tool (https://webblast.ipk-gatersleben.de/misa/) (Beier et al. 2017) with the parameters of single, dinucleotides and trinucleotides to 10, 5 and 4 repeats, respectively. The parameters for tetranucleotides, pentanucleotides and hexanucleotides were then set to 3.

The online program REPuter (http://bibiserv.techfak.uni-bielefeld.de/reputer/) was used to identify repeat sequences in the chloroplast genome, which included four types of repeats. The detection parameters were set to a minimum repeat size of 30 bases and a Hamming distance of 3 (Kurtz et al. 2001). To evaluate the relative synonymous codon usage (RSCU) and the degree of codon bias, we calculated the codon bias index (CBI), codon adaptation index (CAI),

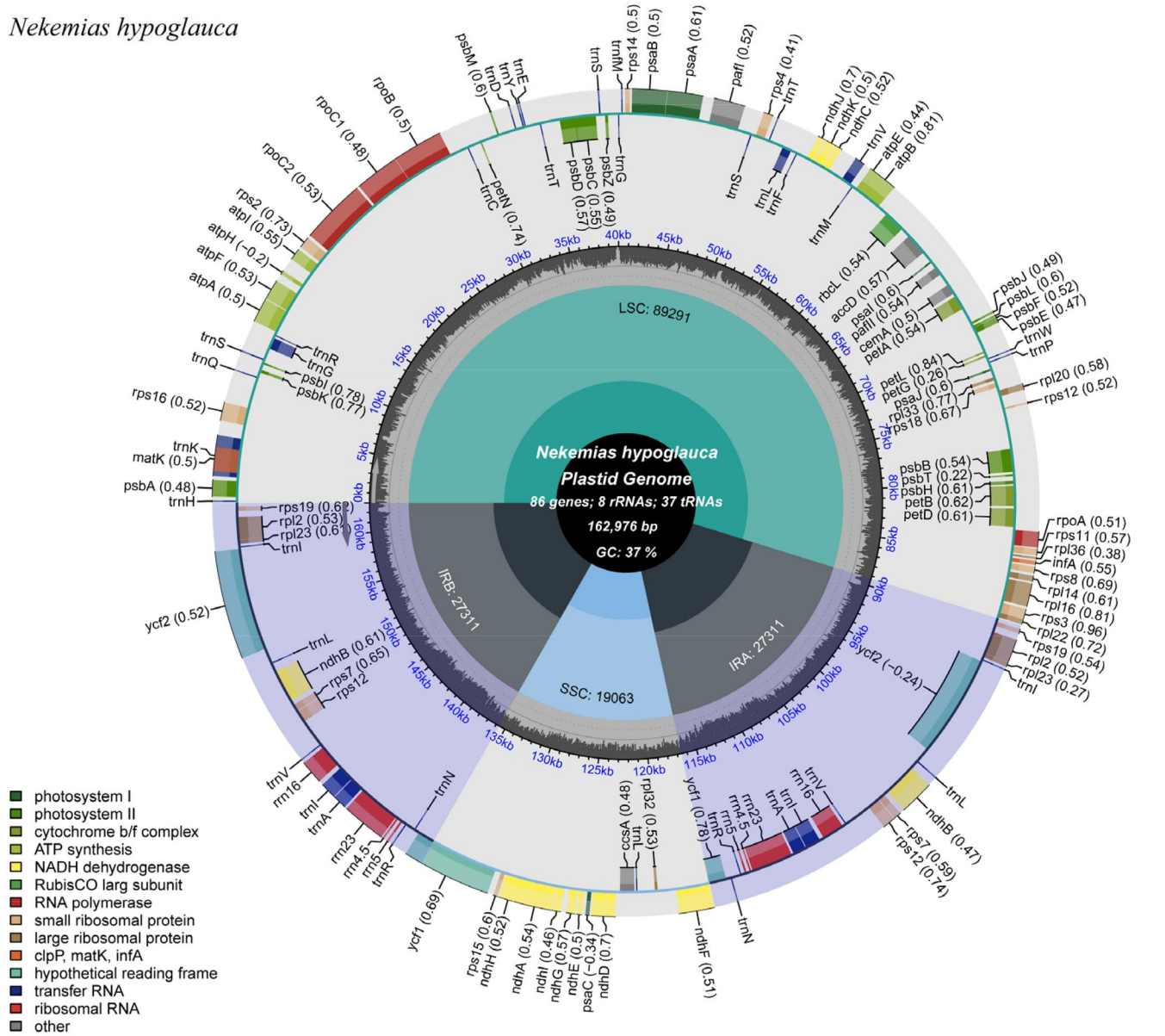

Figure 2. Map of the *Nekemias hypoglauca* chloroplast genome. Genes drawn outside the outer circle are transcribed counterclockwise, and genes drawn inside the outer circle are transcribed clockwise. Genes that are members of different functional groups are color-coded. The differently colored legends in the bottom left corner indicate genes with different functions. The dark gray inner circle indicates the GC content of the chloroplast genome and the presence of nodes in the LSC, SSC, IR regions. IR, inverted repeat; LSC, large single-copy; SSC, small single-copy.

frequency of optimal codons (FOP), effective numbers of codons (NC), and synonymous third codon position of the GC content (GC3) (Xie et al.2019; Sheikh-Assadi et al. 2022).

In addition to the newly sequenced *N. hypoglauca* chloroplast genome, nine other chloroplast genomes were obtained from NCBI to determine the phylogenetic position of *Nekemias* in the Ampelopsideae, elucidate the phylogenetic position of *N. hypoglauca* in the genus, and reconstruct the phylogenetic relationships within *Nekemias*. *Parthenocissus vitacea* was used as an outgroup. The sequences from 10 chloroplast whole genomes were compared using MAFFT V7.309 (Katoh and Standley 2013). A maximum likelihood (ML) phylogenetic inference was based on a supermatrix of 10 chloroplast genome assemblies using RAxML v8 in the GTRGAMMA model to assess the branching support (Nguyen et al. 2015).

3. Results and discussion

We assembled a closed circular chloroplast genome for *N. hypoglauca* that was 162,731 bp long with a typical quadripartite structure, with an average depth of 324.68 X (Supplementary Figure 1). This structure included a pair of inverted repeat regions (IRs) of 27,148 bp separated by a large single-copy region (LSC) of 89,371 bp and small single-copy regions (SSC) of 19,064 bp (Figure 2). The chloroplast genome contained 131 genes, including 86 protein-coding genes, 8 ribosomal RNA genes (rRNA) and 37 transfer RNA genes (tRNA). A total of 16 of these genes, including *atpF*, *ndhA*, *ndhB*, *petB*, *petD*, *rpl2*, *rpl16*, *rpoC1*, *rps16*, *rrn23*, *trnA-UGC*, *trnG-UCC*, *trnI-GAU*, *trnK-UUU*, *trnL-UAA*, and *trnV-UAC*, contained one intron, while the three genes *clpP1*, *pafI* and *rps12* contained two introns. A total of 19 genes contained introns that affect the RNA stability, regulation of gene expression, and selective splicing (Dinh et al. 2016; Peng et al. 2022). A gene structural analysis was conducted for *rps16*, *atpF*, *rpoC1*, *pafI*, *petB*, *petD*, *rpl16*, *rpl2*, *ndhA*, and *ndhB,* and it was difficult to annotate the genes (Supplementary Figure 2). The total GC content of this genome was 37.3%, with 35.4%, 31.7% and 42.5%, which corresponded to the LSC, SSC and IR regions, respectively.

The *N. hypoglauca* chloroplast genome contains 84 SSRs, which included 58 mononucleotide repeats, nine dinucleotide repeats, six trinucleotide repeats, 10 tetranucleotide repeats and one hexanucleotide repeat. The SSRs primarily appeared in the LSC region and were dominated by mononucleotide repeats. The most common SSRs were mononucleotide repeats that consisted of A and T. A total of 337 repeat sequences were identified in *N. hypoglauca*, including 172 forward repeats (F), three reverse repeats (R), and 163 palindromic repeats (P). The palindromic repeats and forward repeats were more common than the other types of repeats. The largest repeats were composed of two palindromic repeats of 164 bp and forward repeats. Larger and more complex repetitive sequences may significantly affect rearrangements of the chloroplast genome and divergence of the sequences (Cavalier-Smith 2002; Timme et al. 2007; Peng et al. 2022).

A total of 63 long tandem repeat sequences (LTRs) were identified in the chloroplast genome of *N. hypoglauca*. Most of these tandem repeats had repeat units between 10 and 30, and the copy number of the repeats ranged from two to three. The longest repeat unit was in the IR region (89,645–90,020 and 161,890–162,265) and was 103 bp long, while the shortest was in the LSC region (39,747–39,856) at only 2 bp.

An analysis of codon usage in the *N. hypoglauca* chloroplast genome was based on 54 CDS genes by calculating the relative synonymous codon usage (RSCU). A total of 21,747 codons were detected. Leucine and cysteine were the most and least frequently used amino acids with 2976 and 403

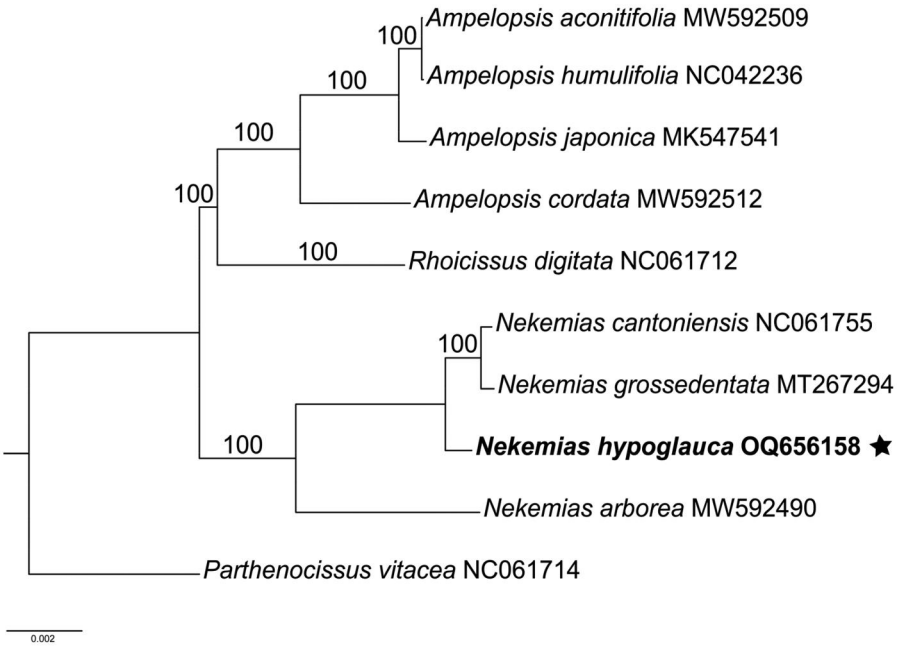

Figure 3. A ML tree based on the sequences of the chloroplast genomes of *Nekemias* and *Ampelopsis* with GenBank accession numbers. The following sequences were used: *Ampelopsis aconitifolia* (MW592509), *Ampelopsis humulifolia* (NC042236), *Ampelopsis japonica* (MK547541), *Ampelopsis cordata* (MW592512), *Rhoicissus digitata* (NC061712), *Nekemias cantoniensis* (NC061775) (Luo et al. 2022), *Nekemias grossedentata* (MT267294) (Gu et al. 2020), *Nekemias arborea* (MW592490) (Liu et al. 2021), and *Parthenocissus vitacea* (NC061714). The bootstrap support values are shown on the nodes. ML, maximum likelihood.

codons, respectively. An analysis of the RSCU showed 30 codons with RSCU values > 1, which indicated that there was some biased usage of these codons. Moreover, Met and Trp were encoded by a single codon (RSCU $= 1$) with no biased usage. In addition, among the codons with RSCU > 1, only the serine codon (UCC) terminated with C, while the remaining 23 codons terminated with either A or U, which suggested that there was a biased usage pattern of codons with A/U endings in the *N. hypoglauca* chloroplast genome.

An ML tree was constructed using the chloroplast genome sequences, including four species of *Nekemias*, four of *Ampelopsis*, and one of *Rhoicissus* with *Parthenocissus vitacea* as an outgroup (Figure 3). All the nodes had a high degree of support. *Nekemias* is located at the first diverged lineage of Ampelopsideae with *Ampelopsis* and *Rhoicissus* as sister groups, which is consistent with the findings of previous studies (Wen et al. 2007, 2014; Lu et al. 2018; Ma et al. 2021). *N. hypoglauca*, *N. cantoniensis*, and *N. grossedentata* form a branch, and *N. hypoglauca* is located at the base of this branch.

4. Conclusions

In this study, the complete chloroplast genome of *N. hypoglauca* was sequenced for the first time and had a total length of 162,731 bp. A total of 131 genes were annotated and composed of 86 PCGs, eight rRNA genes, and 37 tRNA genes. There were no significant differences in the size of genome and the content of genes between *N. hypoglauca* and the other chloroplast genomes in *Nekemias*. The species of *Nekemias* are divided into two branches with *N. arborea* from North America serving as the sister to the three species from East Asia. The intrageneric relationships are also consistent with those of previous studies (Nie et al. 2012). *N. hypoglauca* was the first diverged lineage in the East Asian clade. In conclusion, our results provide basic genetic resources to develop an identification of the species and study the phylogenetic relationships in the *Nekemias* genus.

Ethical approval

No permission was necessary in this study for the sample collection. *Nekemias hypoglauca* (Hance) J. Wen & Z. L. Nie is widely distributed in Fujian, Guangdong and Jiangxi provinces of China, and is not on the List of National Key Protected Wild Plants.

Authors' contributions

In this research, Guan-Hao He is the experimental designer and executor. He has completed the data analysis and the first draft of the paper. Lei Zhang and Ying Meng have contributed to the experimental design and the analysis of experimental results. Jun Wen and Ze-Long Nie have been responsible for supervising the experimental design, data analysis, and the writing and revision of the paper. The final version of the manuscript was read and approved by all authors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by grants from Natural Sciences Foundation of China[(32060055], Natural Sciences Foundation of Hunan Province [2019JJ40232 and 2019JJ40233], and the Hunan Provincial Innovation Foundation for Postgraduate [CX20221121].

Data availability statement

In support of the findings of this study, the genome sequence data are openly available in GenBank of the NCBI at https://www.ncbi.nlm.nih.gov/ under accession OQ656158. The associated BioProject, SRA, and BioSample numbers are PRJNA955411, SRR24184573, and SAMN34177316, respectively.

References

Beier S, Thiel T, Münch T, Scholz U, Mascher M. 2017. MISA-web: a web server for microsatellite prediction. Bioinformatics. 33(16):2583–2585. doi:10.1093/bioinformatics/btx198.

Cavalier-Smith T. 2002. Chloroplast evolution: secondary symbiogenesis and multiple losses. Curr Biol. 12(2):R62–R64. doi:10.1016/S0960-9822(01)00675-3.

Dinh SN, Sai TZT, Nawaz G, Lee K, Kang H. 2016. Abiotic stresses affect differently the intron splicing and expression of chloroplast genes in coffee plants (*Coffea arabica*) and rice (*Oryza sativa*). J Plant Physiol. 201:85–94. doi:10.1016/j.jplph.2016.07.004.

Gu L, Zhang N, Feng C, Yi Y, Yu ZW. 2020. The complete chloroplast genome of *Ampelopsis grossedentata* (Hand.-Mazz.) WT Wang (Family: Vitaceae) and its phylogenetic analysis. Mitochondrial DNA Part B Resour. 5(3):2423–2424. doi:10.1080/23802359.2020.1775508.

Jin JJ, Yu WB, Yang JB, Song Y, DePamphilis CW, Yi TS, Li DZ. 2020. GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biol. 21(1):241. doi:10.1186/s13059-020-02154-5.

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 30(4):772–780. doi:10.1093/molbev/mst010.

Kurtz S, Choudhuri JV, Ohlebusch E, Schleiermacher C, Stoye J, Giegerich R. 2001. REPuter: the manifold applications of repeat analysis on a genomic scale. Nucleic Acids Res. 29(22):4633–4642. doi:10.1093/nar/29.22.4633.

Liu B-B, Ma Z-Y, Ren C, Hodel RGJ, Sun M, Liu X-Q, Liu G-N, Hong D-Y, Zimmer EA, Wen J. 2021. Capturing single-copy nuclear genes, organellar genomes, and nuclear ribosomal DNA from deep genome skimming data for plant phylogenetics: a case study in Vitaceae. J Syst Evol. 59(5):1124–1138. doi:10.1111/jse.12806.

Lohse M, Drechsel O, Bock R. 2007. OrganellarGenomeDRAW (OGDRAW): a tool for the easy generation of high-quality custom graphical maps of plastid and mitochondrial genomes. Curr Genet. 52(5–6):267–274. doi:10.1007/s00294-007-0161-y.

Lu L, Cox CJ, Mathews S, Wang W, Wen J, Chen Z. 2018. Optimal data partitioning, multispecies coalescent and Bayesian concordance analyses resolve early divergences of the grape family (Vitaceae). Cladistics. 34(1):57–77. doi:10.1111/cla.12191.

Luo Y, Wang R, Li Q, Liu J, Deng Z. 2022. The complete chloroplast genome sequence of *Nekemias cantoniensis* (Hook. et Arn.) Planch 1873 (Vitaceae). Mitochondrial DNA Part B Resour. 7(8):1536–1539. doi:10.1080/23802359.2022.2109436.

Ma ZY, Nie ZL, Ren C, Liu XQ, Zimmer EA, Wen J. 2021. Phylogenomic relationships and character evolution of the grape family (Vitaceae). Mol Phylogenet Evol. 154:106948. doi:10.1016/j.ympev.2020.106948.

Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 32(1):268–274. doi:10.1093/molbev/msu300.

Nie ZL, Sun H, Manchester SR, Meng Y, Luke Q, Wen J. 2012. Evolution of the intercontinental disjunctions in six continents in the *Ampelopsis* clade of the grape family (Vitaceae). BMC Evol Biol. 12(1):17. doi:10.1186/1471-2148-12-17.

Peng J-Y, Zhang X-S, Zhang D-G, Wang Y, Deng T, Huang X-H, Kuang T-H, Zhou Q. 2022. Newly reported chloroplast genome of *Sinosenecio albonervius* Y. Liu & QE Yang and comparative analyses with other *Sinosenecio* species. BMC Genomics. 23(1):639. doi:10.1186/s12864-022-08872-3.

Sheikh-Assadi M, Naderi R, Kafi M, Fatahi R, Salami SA, Shariati V. 2022. Complete chloroplast genome of *Lilium ledebourii* (Baker) Boiss and its comparative analysis: lights into selective pressure and adaptive evolution. Sci Rep. 12(1):9375. doi:10.1038/s41598-022-13449-x.

Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R, Greiner S. 2017. GeSeq–versatile and accurate annotation of organelle genomes. Nucleic Acids Res. 45(W1):W6–W11. doi:10.1093/nar/gkx391.

Timme RE, Kuehl JV, Boore JL, Jansen RK. 2007. A comparative analysis of the *Lactuca* and *Helianthus* (Asteraceae) plastid genomes: identification of divergent regions and categorization of shared repeats. Am J Bot. 94(3):302–312. doi:10.3732/ajb.94.3.302.

Wen J, Boggan J, Nie ZL. 2014. Synopsis of *Nekemias* Raf., a segregate genus from *Ampelopsis* Michx.(Vitaceae) disjunct between eastern/southeastern Asia and eastern North America, with ten new combinations. PhytoKeys. 42(42):11–19. doi:10.3897/phytokeys.42.7704.

Wen J, Nie ZL, Soejima A, Meng Y. 2007. Phylogeny of Vitaceae based on the nuclear *GAI1* gene sequences. Can J Bot. 85(8):731–745. doi:10.1139/B07-071.

Xie D-F, Yu H-X, Price M, Xie C, Deng Y-Q, Chen J-P, Yu Y, Zhou S-D, He X-J. 2019. Phylogeny of Chinese *Allium* species in section Daghestanica and adaptive evolution of *Allium* (Amaryllidaceae, Allioideae) species revealed by the chloroplast complete genome. Front Plant Sci. 10:460. doi:10.3389/fpls.2019.00460.