MITOCHONDRIAL DNA PART B: RESOURCES
2025, VOL. 10, NO. 8, 753–757
https://doi.org/10.1080/23802359.2025.2535628

Taylor & Francis
Taylor & Francis Group

PLASTOME REPORT

OPEN ACCESS Check for updates

Complete sequence of the *Achillea ptarmica* chloroplast genome determined by long-read sequencing

Hiroki Natsume and Makoto Kusaba

Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, Japan

ABSTRACT

Achillea species, commonly referred to as yarrow, belong to the family Asteraceae and are typically perennial, outbreeding plants used in traditional medicine. The genus *Achillea* has undergone hybridization and polyploidization events, resulting in phylogenetically complex relationships. Considering the limited scope of genomic studies on this genus, a genomic analysis of a diploid *Achillea* species is important for advancing our understanding of its evolutionary biology and genetic diversity. In this study, we assembled the complete chloroplast genome sequence of a diploid *Achillea* species (*A. ptarmica*) using nanopore long-read sequencing technology. The assembled chloroplast genome (149,252 bp), with a GC content of 38%, was revealed to include a large single-copy region (82,498 bp), a small single-copy region (18,458 bp), and two inverted repeat regions (24,148 bp each). The genome contains 89 protein-coding genes, 8 rRNAs, and 37 tRNAs. A phylogenetic analysis clustered *A. ptarmica* with other *Achillea* species (*A. millefolium* and *A. wilsoniana*), positioning them within the same clade as the genus *Tanacetum*. The accuracy of nanopore sequencing was validated by comparing the results for 10 chloroplast genes with the corresponding Sanger sequencing results. Our findings provide valuable genetic resources for further taxonomic, evolutionary, and phylogenetic studies of the genus *Achillea*.

ARTICLE HISTORY
Received 26 December 2024
Accepted 11 July 2025

KEYWORDS
Asteraceae; nanopore; phylogenetic analysis

Introduction

Achillea species (yarrow) belong to the family Asteraceae and are characterized by aromatic and hairy leaves as well as flat clusters of small capitula in various colors, features that have made them popular choices for garden landscaping (Saeidnia et al. 2011). Beyond their ornamental appeal, *Achillea* species have long been utilized in traditional medicine because of their spasmolytic, anti-inflammatory, and antidiabetic properties (Sheidai et al. 2009; Saeidnia et al. 2011; Veryser et al. 2017). Members of this genus have a basic chromosome number of $x = 9$, and includes species with diploid (2x), tetraploid (4x), hexaploid (6x), and octoploid (8x) genomes (Sheidai et al. 2009). *Achillea* species have complex phylogenetic relationships because of hybridization and polyploidization during evolution (Guo et al. 2004). Among *Achillea* species, *A. millefolium* L. and *A. wilsoniana* (Heimerl ex Hand.-Mazz.) Hand.-Mazz. have had their chloroplast genomes fully sequenced. An analysis of *A. ptarmica* L. (1753), which is a diploid perennial herb (Figure 1), may provide additional information regarding intergeneric and intrageneric relationships of *Achillea* species.

Illumina short-read sequencing has been widely used for whole-genome sequencing of chloroplasts, but its limited read lengths make it challenging to assemble large inverted repeats (IRs) (Kolodner and Tewari 1979; Scheunert et al. 2020). Long-read sequencing methods, such as PacBio SMRT DNA sequencing and nanopore sequencing, may be better for assembling complete chloroplast genomes (Scheunert et al. 2020). However, the low accuracy of nanopore sequencing can sometimes pose a problem (Rang et al. 2018). We herein report the complete chloroplast genome sequence of *A. ptarmica* determined using nanopore sequencing. This study aimed to clarify the taxonomic and evolutionary characteristics of the genus *Achillea* and evaluate the feasibility of using nanopore sequencing to analyze chloroplast genomes.

Materials and methods

Plant materials

A. ptarmica leaf samples were collected at the Laboratory of Plant Chromosome and Gene Stock at Hiroshima University (34°40′17″ N, 132°71′67″ E). A voucher specimen (SH1002) is stored by the Laboratory of Plant Chromosome and Gene Stock at Hiroshima University (contact M. Kusaba, akusaba@hiroshima-u.ac.jp). Genomic DNA extracted from leaf samples is available from the National Bioresource Project Chrysanthemum (https://shigen.nig.ac.jp/chrysanthemum/). To minimize interference

CONTACT Makoto Kusaba akusaba@hiroshima-u.ac.jp Graduate School of Integrated Sciences for Life, Hiroshima University, Higashi-Hiroshima, Japan
Supplemental data for this article can be accessed online at https://doi.org/10.1080/23802359.2025.2535628.

Figure 1. Photographs of *Achillea ptarmica* taken by Hiroki Natsume in the Laboratory of Plant Chromosome and Gene Stock at Hiroshima University, Hiroshima, Japan (34°40′17″ N, 132°71′67″ E). (A) Capitula; the capitula are white and arranged in a corymbose inflorescence. The involucre is hemispherical. Ray florets are large, white, and arranged in two rows, while disk florets are also white, (B) Shoot. Leaves are elliptic to lanceolate-linear with fine serrations, and their bases partially clasp the stem.

from secondary products, genomic DNA was extracted from leaves after a 2-day incubation in darkness.

DNA extraction and sequencing

DNA for constructing a nanopore sequencing library was prepared as described by Naito (2023). Approximately 1.5 g leaf samples were ground in liquid nitrogen and then DNA was extracted using a NucleoBond HMW DNA kit (Macherey-Nagel, Düren, Germany). DNA fragments smaller than 10 kbp were removed using a Short Read Eliminator Kit (Pacific Biosciences, USA). DNA quality was assessed by 0.8% agarose gel electrophoresis, whereas the DNA concentration was determined using a Qubit fluorometer (Thermo Fisher, Waltham, MA, USA). A sequencing library was prepared using an SQK-LSK114 Ligation Sequencing DNA V14 kit (ONT, Oxford, UK) and then sequenced using an R10.4.1 flow cell.

Genome assembly and annotation

A complete chloroplast genome was assembled using ptGAUL (v.1.0.5) (Zhou et al. 2023), with the *A. wilsoniana* (NC_069982) chloroplast genome serving as a reference. The genome was annotated and visualized using CPGAVAS2 (Shi et al. 2019) and CPGView (Liu et al. 2023), respectively.

Phylogenetic analysis

Phylogenetic relationships were inferred on the basis of all protein-coding sequences encoded by 23 chloroplast genomes within the family Asteraceae. These sequences were obtained from GenBank and aligned with *A. ptarmica* sequences using MAFFT (v.7.0) (Katoh et al. 2019). A maximum-likelihood phylogenetic tree was constructed using RAxML (v.8.2.11) (Stamatakis 2014) and the default parameters of the GTRGAMMA model, with branch support evaluated using 100 bootstrap replicates.

Sequencing accuracy analysis

Ten chloroplast genes (*rpoC1*, *rps2*, *rps16*, *trnL*, *ycf2*, *accD*, *ndhF*, *petB*, *psbB*, and *rpoB*) containing sequence variations specific to *A. ptarmica* (compared with *A. wilsoniana* and *A. millefolium* sequences) were selected for Sanger sequencing. The generated sequences were compared with nanopore sequencing results to assess consistency and accuracy. Primers used for PCR amplifications and Sanger sequencing are listed in Supplementary Table S1.

Results

We obtained 68.1 Gb sequence data from *A. ptarmica* genomic DNA using a PromethION nanopore system. By assembling these sequences using ptGAUL, a 149,252 bp circular chloroplast genome was generated (Figure 2; GenBank accession PQ412534). The genome has an overall GC content of 38% and consists of a large single-copy region (82,498 bp), a small single-copy region (18,458 bp), and two IR regions (24,148 bp each). The chloroplast genome includes 89, 8, and 37 protein-coding, rRNA, and tRNA sequences, respectively (Figure 2). The chloroplast genome was sequenced at an average depth of 2970.8× coverage (Supplementary Figure 1). Cis-spliced genes and the trans-spliced gene *rps12* were identified in the chloroplast genome (Supplementary Figure S2). According to a phylogenetic analysis, *A. ptarmica* is closely related to *A. wilsoniana* and *A. millefolium*, consistent with their classification within the genus *Achillea*. Additionally, in the phylogenetic tree, *Achillea* species were clustered with two *Tanacetum* species (Figure 3), suggesting these genera are closely related. To evaluate the accuracy of the sequences obtained by nanopore sequencing, 10 polymorphic sites between *A. ptarmica* and *A. wilsoniana*/*A. millefolium* sequences, including nucleotide substitutions (five transitions and three transversions), one insertion (5 bp), and one deletion (5 bp), were validated *via* Sanger sequencing (Supplementary Figures S3 and S4). The comparison revealed that *A. ptarmica* sequence variations were consistent with Sanger sequencing data, confirming that the sequence variations were not due to nanopore sequencing errors.

Discussion and conclusion

This study assembled the *A. ptarmica* chloroplast genome using long-read nanopore sequencing data. The consistency in the sequences at 10 *A. ptarmica*-specific polymorphic sites obtained *via* nanopore sequencing and Sanger sequencing reflected the reliability of nanopore sequencing for assembling chloroplast genomes. The size, GC content, and gene

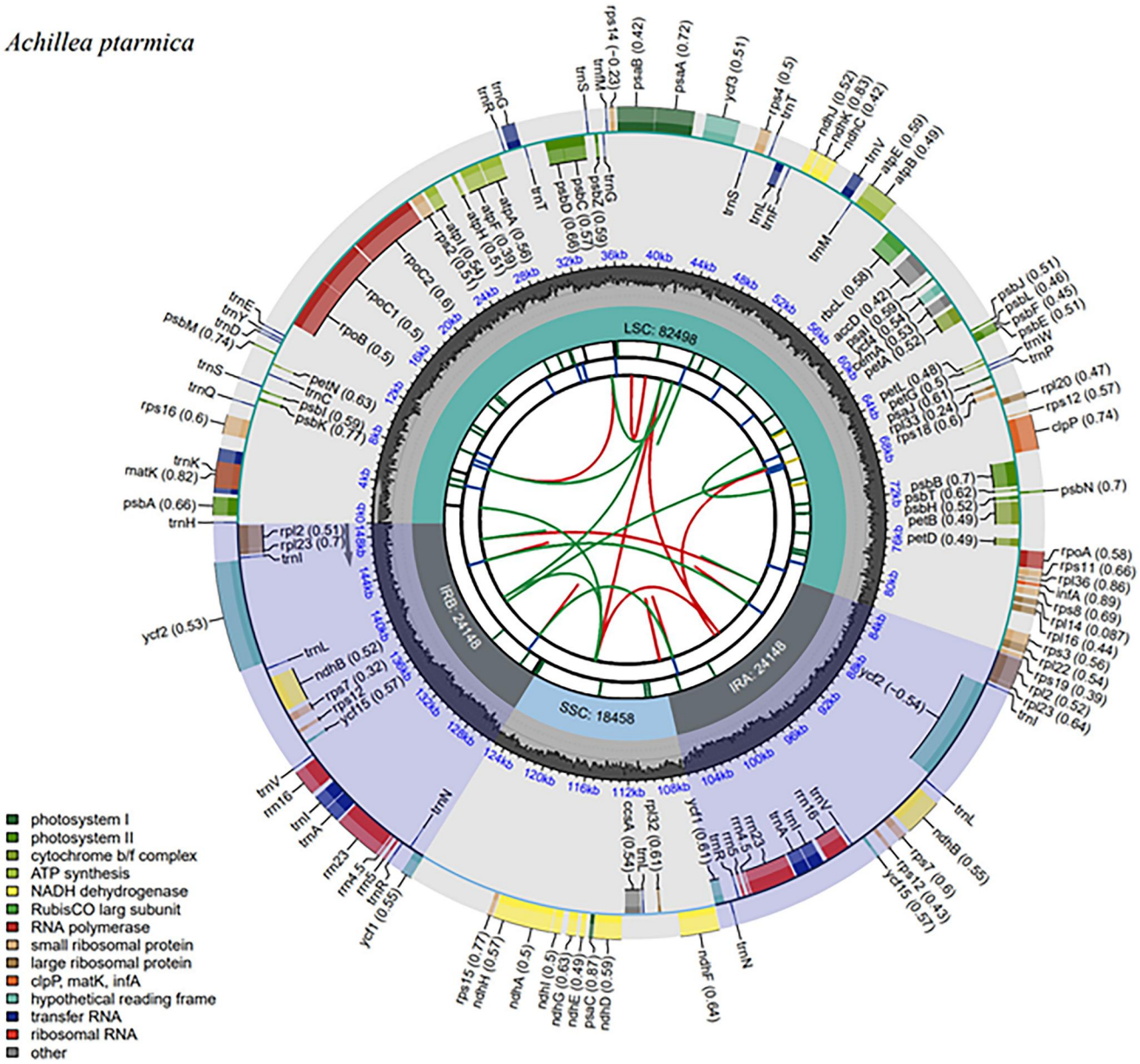

Figure 2. Schematic map of the overall features of the *Achillea ptarmica* chloroplast genome visualized by CPGView. The map contains six tracks arranged from the center outward. The first track presents dispersed repeats, consisting of direct and palindromic repeats, connected with red and green arcs, respectively. The second track presents long tandem repeats as short blue bars, whereas the third track presents short tandem repeats or microsatellite sequences as colored bars (black: complex repeat; green: repeat unit of 1; yellow: repeat unit of 2; purple: repeat unit of 3; blue: repeat unit of 4; orange: repeat unit of 5; red: repeat unit of 6). The fourth track includes small single-copy (SSC), inverted repeat (IRa and IRb), and large single-copy (LSC) regions. The fifth track presents the GC content along the genome, whereas the sixth track presents genes, which are color-coded according to their functional classification. Transcriptional directions for the inner and outer genes are clockwise and counterclockwise, respectively. Gene functional classifications are indicated in the bottom left corner.

composition of the *A. ptarmica* chloroplast genome were similar to those of the *A. millefolium* chloroplast genome (Liu et al. 2023). A phylogenetic analysis suggested that *A. millefolium* and *A. wilsoniana* are closely related species, whereas *A. ptarmica* is more distantly related. This is in accordance with the findings of an earlier study (Guo et al. 2004) that assigned *A. ptarmica* to an early-diverging lineage, distinct from the more recently evolved *A. millefolium* complex. Additionally, the clustering of *Achillea* species within the same clade as *Tanacetum* species is suggestive of a close relationship between the subtribes Matricariinae (containing *Achillea*) and Anthemidinae (containing *Tanacetum*), consistent with the results of a recent study (Manzo and Tomasello 2024).

In summary, we herein present a complete *A. ptarmica* chloroplast genome dataset, making it the third such dataset for the genus *Achillea*, along with a phylogenetic reconstruction of the genus *Achillea* and related genera. Furthermore, the study data highlight the reliability of nanopore sequencing for analyzing chloroplast genomes, supporting its application in evolutionary studies.

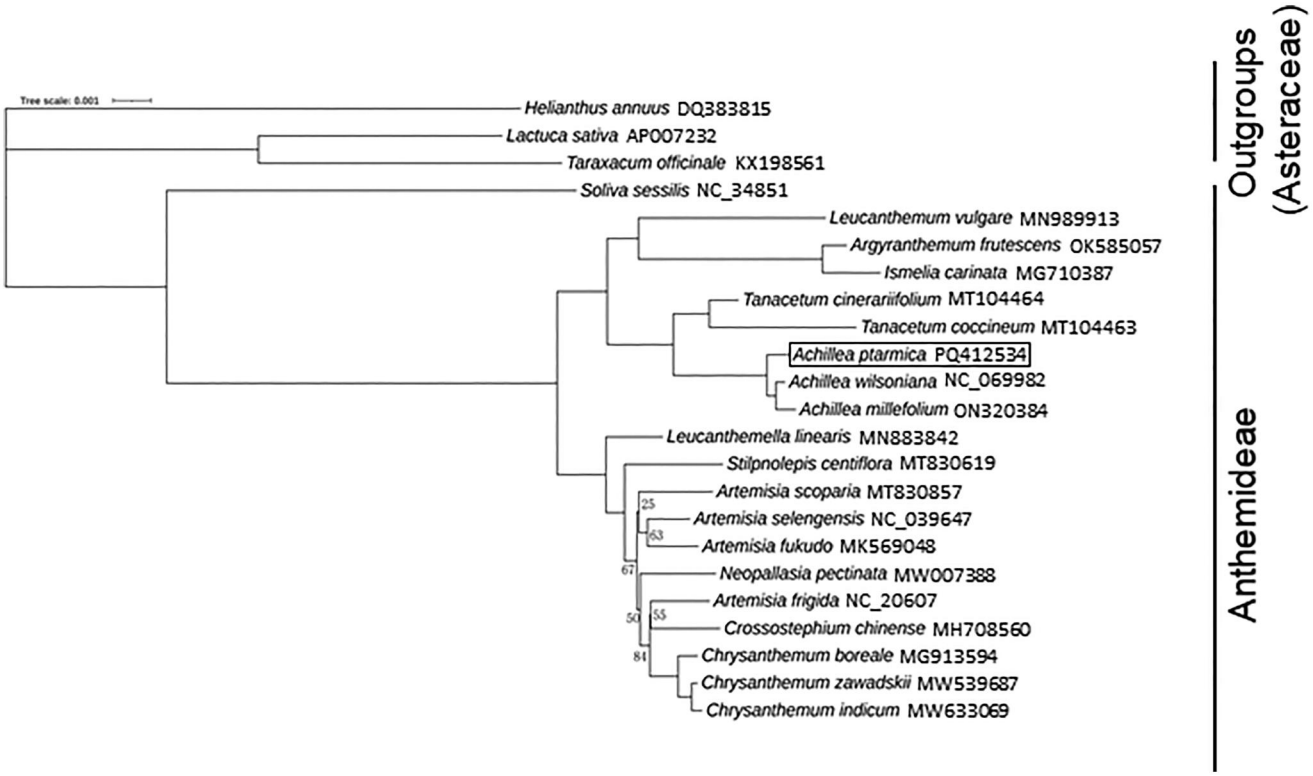

Figure 3. Maximum-likelihood phylogenetic tree of *Achillea ptarmica* and 22 additional species. The tree was constructed using the RAxML based on protein-coning genes. The dataset included 20 species from the tribe Anthemideae. Three species were used as outgroup. Bootstrap support values lower than 100 are displayed next to the corresponding nodes in the tree, while values of 100 are omitted for clarity. The following sequences were used: *Achillea millefolium* (ON320384; Liu et al. 2023), *Achillea ptarmica* (PQ412534), *Achillea wilsoniana* (NC_069982), *Argyranthemum frutescens* (OK585057; Zhao et al. 2022), *Artemisia frigida* (NC_20607; Liu et al. 2013), *Artemisia fukudo* (MK569048; Min et al. 2019), *Artemisia scoparia* (MT830857; Li et al. 2020), *Artemisia selengensis* (NC_039647; Peng et al. 2018), *Chrysanthemum boreale* (MG913594; Won et al. 2018), *Chrysanthemum indicum* (MW633069), *Chrysanthemum zawadskii* (MW539687; Baek et al. 2021), *Crossostephium chinense* (MH708560; Chen et al. 2019), *Ismelia carinata* (MG710387; Liu et al. 2018), *Leucanthemella linearis* (MN883842; Kim and Kim 2020), *Leucanthemum vulgare* (MN989913; Scheunert et al. 2020), *Neopallasia pectinata* (MW007388; Yu et al. 2021), *Soliva sessilis* (NC_34851; Vargas et al. 2017), *Stilpnolepis centiflora* (MT830619; Shi and Xie 2020), *Tanacetum cinerariifolium* (MT104464), *Tanacetum coccineum* (MT104463); *Helianthus annuus* (DQ383815; Timme et al. 2006), *Lactuca sativa* (AP007232; Kanamoto et al. 2006), *Taraxacum officinale* (KX198561; Zhang et al. 2017).

Acknowledgments

We thank Yumi Nagashima and Eriko Kaneda for their technical assistance as well as Kenji Taniguchi for helpful discussions. This work was conducted with the facilities in the Natural Science Center for Basic Research and Development (N-BARD) at Hiroshima University [NBARD-00109]. The first author (Hiroki Natsume) was previously known as Dong Qin. **HN**: Conceptualization, Investigation, Data curation, Writing – original draft, Writing – review & editing; **MK**: Conceptualization, Resources, Supervision, Writing – review & editing. All authors agree to be accountable for all aspects of this work.

Author contributions

CRediT: **Hiroki Natsume**: Data curation, Investigation, Methodology, Writing – original draft; **Makoto Kusaba**: Conceptualization, Resources, Supervision, Writing – review & editing.

Ethical approval

The plant material used in this study is classified as "Least Concern" on the IUCN Red List of Threatened Species, and the sampling site is not located in any protected area.

Disclosure of interest

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (Grant no. 24H00C13 to M.K.).

Data availability statement

The genome sequence data that support the findings of this study are openly available in GenBank of NCBI at [https://www.ncbi.nlm.nih.gov] (https://www.ncbi.nlm.nih.gov/) under the accession no. PQ412534. The associated BioProject, SRA, and Bio-Sample numbers are PRJNA1195354, SAMN45217404, and SRR31654399, respectively. The raw data of sanger sequencing available at figshare (https://doi.org/10.6084/m9.figshare.28021436.v1).

References

Baek J, Park S, Lee J, Min J, Park J, Lee GW. 2021. The complete chloroplast genome of *Chrysanthemum zawadskii* Herbich (Asteraceae) isolated in Korea. Mitochondrial DNA B Resour. 6(7):1956–1958. doi:10.1080/23802359.2021.1934148.

Chen T, Du Y, Zhao H, Dong M, Liu L. 2019. The complete chloroplast genome of *Crossostephium chinense* (Asteraceae), using genome skimming data. Mitochondrial DNA Part B Resour. 4(1):322–323. doi:10.1080/23802359.2018.1544038.

Guo Y, Friedrich E, Rosabelle S. 2004. Phylogeny and systematics of *Achillea* (Asteraceae-Anthemideae) inferred from nrITS and plastid *trnLF* DNA sequences. Taxon. 53(3):657–672. A. doi:10.2307/4135441.

Kanamoto H, Yamashita A, Asao H, Okumura S, Takase H, Hattori M, Yokota A, Tomizawa KI. 2006. Efficient and stable transformation of *Lactuca sativa* L. cv. Cisco (lettuce) plastids. Transgenic Res. 15(2):205–217. doi:10.1007/s11248-005-3997-2.

Katoh K, Rozewicki J, Yamada KD. 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 20(4):1160–1166. doi:10.1093/bib/bbx108.

Kim HT, Kim JS. 2020. The complete chloroplast genome sequence of *Leucanthemella linearis* (Matsum. ex Matsum.) Tzvelev (Asteraceae), endangered plant of Korea. Mitochondrial DNA B Resour. 5(3):3360–3362. doi:10.1080/23802359.2020.1821819.

Kolodner R, Tewari KK. 1979. Inverted repeats in chloroplast DNA from higher plants. Proc Natl Acad Sci U S A. 76(1):41–45. doi:10.1073/pnas.76.1.41.

Li J, Jiang M, Chen H, Yu J, Liu C. 2020. Intraspecific variations among the chloroplast genomes of *Artemisia scoparia* (Asteraceae) from Pakistan and China. Mitochondrial DNA Part B Resour. 5(3):3182–3184. doi:10.1080/23802359.2020.1810167.

Liu F, Movahedi A, Yang W, Xu D, Jiang C. 2023. The complete plastid genome and characteristics analysis of *Achillea millefolium*. Funct Integr Genomics. 23(2):192. doi: 10.1007/s10142-023-01121-0.

Liu S, Ni Y, Li J, Zhang X, Yang H, Chen H, Liu C. 2023. CPGView: a package for visualizing detailed chloroplast genome structures. Mol Ecol Resour. 23(3):694–704. doi:10.1111/1755-0998.13729.

Liu X, Zhou B, Yang H, Li Y, Yang Q, Lu Y, Gao Y. 2018. Sequencing and analysis of *Chrysanthemum carinatum* Schousb and *Kalimeris indica*. The complete chloroplast genomes reveal two inversions and rbcL as barcoding of the vegetable. Molecules. 23(6):1358. doi:10.3390/molecules23061358.

Liu Y, Huo N, Dong L, Wang Y, Zhang S, Young HA, Feng X, Gu YQ. 2013. Complete chloroplast genome sequences of Mongolia medicine *Artemisia frigida* and phylogenetic relationships with other plants. PLoS One. 8(2):e57533. doi:10.1371/journal.pone.0057533.

Manzo E, Tomasello S. 2024. The first complete chloroplast genome for the species-rich genus *Anthemis* (Asteraceae). mtDNA. Mitochondrial DNA B Resour. 9(9):1248–1253. doi:10.1080/23802359.2024.2405528.

Min J, Park J, Kim Y, Kwon W. 2019. The complete chloroplast genome of *Artemisia fukudo* Makino (Asteraceae): providing insight of intraspecies variations. Mitochondrial DNA Part B Resour. 4(1):1510–1512. doi:10.1080/23802359.2019.1601044.

Naito K. 2023. How to sequence and assemble plant genomes. Methods Mol Biol. 2632:57–77. doi:10.1007/978-1-0716-2996-3_5.

Peng J, Zhao Y, Li C, Xu Z. 2018. The complete chloroplast genome and phylogeny of *Artemisia selengensis* in Dongting Lake. Mitochondrial DNA B Resour. 3(2):907–908. doi:10.1080/23802359.2018.1501322.

Rang FJ, Kloosterman WP, de Ridder J. 2018. From squiggle to basepair: computational approaches for improving nanopore sequencing read accuracy. Genome Biol. 19(1):90. doi:10.1186/s13059-018-1462-9.

Saeidnia S, Gohari AR, Mokhber-Dezfuli N, Kiuchi F. 2011. A review on phytochemistry and medicinal properties of the genus *Achillea*. DARU. 19(3):173.

Scheunert A, Dorfner M, Lingl T, Oberprieler C. 2020. Can we use it? On the utility of de novo and reference-based assembly of Nanopore data for plant plastome sequencing. PLoS One. 15(3):e0226234. doi:10.1371/journal.pone.0226234.

Sheidai M, Azanei N, Attar F. 2009. New chromosome number and unreduced pollen formation in *Achillea* species (Asteraceae). Acta Biol Szeged. 53(1):39–43.

Shi L, Chen H, Jiang M, Wang L, Wu X, Huang L, Liu C. 2019. CPGAVAS2, an integrated plastome sequence annotator and analyzer. Nucleic Acids Res. 47(W1):W65–W73. doi:10.1093/nar/gkz345.

Shi X, Xie K. 2020. The complete chloroplast genome sequence of *Stilpnolepis centiflora* (Asteraceae), an endemic desert species in Northern China. Mitochondrial DNA B Resour. 5(3):3545–3546. doi:10.1080/23802359.2020.1829127.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313. doi:10.1093/bioinformatics/btu033.

Timme RE, Kuehl JV, Boore JL, Jansen RK. 2006. A comparison of the first two sequenced chloroplast genomes in asteraceae: lettuce and sunflower. LBNL: 59386. doi:10.2172/960402.

Vargas OM, Ortiz EM, Simpson BB. 2017. Conflicting phylogenomic signals reveal a pattern of reticulate evolution in a recent high-Andean diversification (Asteraceae: astereae: diplostephium). New Phytol. 214(4):1736–1750. doi:10.1111/nph.14530.

Veryser L, Taevernier L, Wynendaele E, Verheust Y, Dumoulin Ann, De Spiegeleer B. 2017. N-alkylamide profiling of *Achillea ptarmica* and *Achillea millefolium* extracts by liquid and gas chromatography-mass spectrometry. J Pharm Anal. 7(1):34–47. doi:10.1016/j.jpha.2016.09.005.

Won SY, Jung JA, Kim JS. 2018. The complete chloroplast genome of *Chrysanthemum boreale* (Asteraceae). Mitochondrial DNA B Resour. 3(2):549–550. doi:10.1080/23802359.2018.1468225.

Yu J, Xia M, Xu H, Han S, Zhang F. 2021. The complete chloroplast genome sequence of *Neopallasia pectinata* (Asteraceae). Mitochondrial DNA B Resour. 6(2):430–431. doi:10.1080/23802359.2020.1870893.

Zhang Y, Iaffaldano BJ, Zhuang X, Cardina J, Cornish K. 2017. Chloroplast genome resources and molecular markers differentiate rubber dandelion species from weedy relatives. BMC Plant Biol. 17(1):34. doi:10.1186/s12870-016-0967-1.

Zhao Y, Qu D, Ma Y. 2022. Characterization of the chloroplast genome of *Argyranthemum frutescens* and a comparison with other species in Anthemideae. Genes (Basel). 13(10):1720. doi:10.3390/genes13101720.

Zhou W, Armijos CE, Lee C, Lu R, Wang J, Ruhlman TA, Jansen RK, Jones AM, Jones CD. 2023. Plastid genome assembly using long-read data. Mol Ecol Resour. 23(6):1442–1457. doi:10.1111/1755-0998.13787.