GAYANA BOTANICA
Gayana Bot. (2025) vol. 82, No. 2, 185-191
SHORT COMMUNICATION
http://dx.doi.org/10.4067/s0717-6643202500020185

The complete plastome of the endemic and endangered species *Chorizanthe novoana* (Polygonaceae)

Reporte del plastoma completo de la especie endémica y en peligro *Chorizanthe novoana* (Polygonaceae)

Sebastián Torres-Sepúlveda[1,*] & Oscar Toro-Núñez[1]

[1]Laboratorio de Biosistemática y Taxonomía Vegetal, Departamento de Botánica, Facultad de Ciencias Naturales y Oceanográficas, Universidad de Concepción, Concepción, Chile.
*Corresponding author: sebtorres2018@udec.cl

ABSTRACT

This study presents the complete sequence of the plastid genome of the Chilean endemic and endangered *Chorizanthe novoana* (Polygonaceae). The total sequence length was 157,451 bp, with LSC, SSC and two IR regions of 87,239 bp, 18,076 bp, and 26,068 bp length, respectively. The GC content was 36.1% and 132 genes were identified: 90 are protein coding, 38 tRNA and 8 rRNA. 56 microsatellites were found, mostly mononucleotides. Phylogenetic analyses confirm the position of *Chorizanthe* in the subfamily Eriogonoideae.

Keywords: chloroplast genome, Eriogonoideae, Next Generation Sequencing.

RESUMEN

Este estudio presenta la secuencia completa del genoma cloroplastidial de *Chorizanthe novoana* (Polygonaceae), especie endémica y en peligro de Chile. La secuencia mide 157.451 pb, con regiones LSC, SSC y dos IR de 87.239 pb, 18.076 pb y 26.068 pb, respectivamente. El contenido GC fue 36,1% y se identificaron 132 genes: 90 codificadores de proteínas, 38 tRNA y 8 rRNA. Se detectaron 56 microsatélites, mayormente mononucleótidos. Los análisis filogenéticos confirman la inclusión de *Chorizanthe* en la subfamilia Eriogonoideae.

Palabras clave: Eriogonoideae, genoma cloroplastidial, Secuenciación de Siguiente Generación.

INTRODUCTION

The Polygonaceae is a globally distributed plant family comprising approximately 48 genera and 1,200 species of perennial and annual herbs, vines, shrubs, and trees (Brandbyge 1993; Freeman & Reveal 2005). While recent systematic analyses have clarified most parts of its classification — particularly recognizing the subfamilies Polygonoideae and Eriogonoideae — patterns of diversification at the genus and species levels still remain poorly understood.

A particularly noteworthy example is the amphitropical genus *Chorizanthe* R. Br. ex Benth., which has received limited attention since its most recent systematic study, primarily focused on North American representatives (Kempton 2012). Kempton's analysis, based on nuclear and chloroplast markers and encompassing 163 taxa within Eriogonoideae, demonstrated the non-monophyly of the subtribes Eriogonineae and Chorizanthineae. According to Reveal and Hardham (1989), approximately 50 species of Chorizanthe were recognized, from which 11 were known to occur exclusively in South America. However, more recent work — especially through the description of new

species in Chile (Teillier *et al.* 2019) — has increased the number of South American species to 19. Nevertheless, no studies have investigated the phylogenetic relationships or genetic characteristics of these southern taxa, limiting our understanding of their diversification processes and underlying mechanisms.

In recent years, complete plastome sequencing has become a powerful tool in plant systematics, owing to its uniparental inheritance, low substitution rates, and conserved gene order (Gitzendanner *et al.* 2017). Plastomes offer valuable genetic markers and are increasingly applied in phylogenetics, phylogenomics, population genetics, and related disciplines (Provan *et al.* 2001; Kim *et al.* 2018; Zhang *et al.* 2020). Advances in next-generation sequencing technologies (e.g., Illumina, PacBio, Nanopore) have further facilitated plastome sequencing and annotation, making them more accessible and efficient (Jin *et al.* 2020; Zhou *et al.* 2023; Liu *et al.* 2025).

In this context, generating new genomic resources — such as plastomes from recently described endemic species of *Chorizanthe* — is essential to bridging the knowledge gap in the systematics and conservation of both the genus and the Polygonaceae family. Here, we present the first report of the sequenced and annotated plastome of *Chorizanthe novoana* Teillier & Macaya, a critically endangered species endemic to Chile's Valparaíso Region (MMA 2024).

METHODOLOGY

Leaf material and DNA extraction

A specimen of *Chorizanthe novoana* was collected at Mirador Laguna Verde, Valparaíso Province, Chile (33°04'53.3" S; 71°39'58.4" W, 166 m.a.s.l.). A voucher specimen is deposited at the Herbarium of the Universidad de Concepción (CONC 198237). Total genomic DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Düsseldorf, Germany), following the manufacturer's protocol. DNA quality was assessed with GelRed™ staining (Biotium Inc., USA) under a UV transilluminator, and quantity was measured using the Qubit 2.0 Fluorometer (Invitrogen Inc., Massachusetts, USA).

Sequencing, assembly and annotation of the plastome

Sequencing was carried out by BGI Genomics (Shenzhen, Guangdong, China) using the BGISEQ-500 platform (Mak *et al.* 2017). A total of 3.5 million PE150 reads were used for plastome assembly with GetOrganelle v1.7.1 (Jin *et al.* 2020), using default settings with all reads included. The resulting contig was visualized and examined using Bandage v0.8.1 (Wick *et al.* 2015) to detect and remove potential paralogous sequences and selecting the contig with the highest coverage. Annotation was performed using GeSeq on the CHLOROBOX platform (Tillich *et al.* 2017), with *Ruprechtia albida* (GenBank accession OK661153) as the reference. The final sequence of *C. novoana*, and its annotations, were deposited in the Genbank Accession Number PX309297.1

Microsatellite analysis

Microsatellite detection was performed using Krait v1.5.1 (Du *et al.* 2018). The minimum repeat thresholds for SSR detection were set as follows: 10 for mononucleotides, 5 for dinucleotides, 4 for trinucleotides, and 3 for tetra-, penta-, and hexanucleotides (Zhao *et al.* 2019; Yang *et al.* 2021). The motif standardization level was set to Level 3, as recommended by the software authors.

Phylogenetic analysis

Phylogenetic relationships were inferred by comparing the *Chorizanthe novoana* plastome to plastomes from species within the subfamilies Polygonoideae and Eriogonoideae available in GenBank. Representatives from ten genera were selected for Polygonoideae, while only three genera from Eriogonoideae were included due to limited available data (Table 1). The plastome of *Plumbago auriculata* (Plumbaginaceae, GenBank accession MH286308) was used as an outgroup based on its sister relationship to Polygonaceae (Zhang *et al.* 2022).

All plastomes were aligned using MAFFT v7.526 (Katoh & Standley 2013). Coding sequences (CDS) were annotated and extracted using *Chorizanthe novoana* as the reference plastome in Geneious Prime v2025.1.2 (Kearse *et al.* 2012). To avoid redundancy, CDS present in both inverted repeat regions were removed (Cai *et al.* 2021). The remaining sequences were concatenated using SEGUL v0.22.1 (Handika & Esselstyn 2024).

Phylogenetic analysis was conducted with IQ-TREE2 v2.2.03 (Minh *et al.* 2020), employing optimal partitioning and substitution models inferred via ModelFinder (Kalyaanamoorthy *et al.* 2017) and PartitionFinder (Lanfear *et al.* 2012), with unlinked marginal rates. Branch support was evaluated with 1,000 ultrafast bootstrap replicates (UFBoot2) (Hoang *et al.* 2018) and the Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) (Guindon *et al.* 2010). Node support was considered significant if UFBoot2 > 0.95 and SH-aLRT > 0.85 (Hoang *et al.* 2018). The final tree was visualized using FigTree v1.4 (Rambaut 2012).

Table 1. Obtained species from GenBank and his respective accession number. / Especies obtenidas desde GenBank y su respectivo número de acceso.

Subfamily	Species	Genbank Accession number
Eriogonoideae	*Coccoloba uvifera*	OK661151
	Ruprechtia albida	OK661153
	Triplaris americana	OK661152
Polygonoideae	*Persicaria neofiliformis*	OK661145
	Fagopyrum cymosum	MZ702796
	Fallopia multiflora	OK661155
	Koenigia alpina	MZ573783
	Polygonum aviculare	OK661156
	Rumex japonicus	MK058527
	Bistorta macrophylla	OK661158
	Calligonum arborescens	MN202599
	Reynoutria japonica	OK661148
	Pteroxygonum denticulatum	OK661160

RESULTS

Plastome assembly and structure

The assembled plastome of *Chorizanthe novoana* was obtained without evidence of paralogous copies in the final assembly, being assembled with an average depth of 150X. The complete plastome measured 157,451 bp and exhibited the typical quadripartite structure, consisting of a large single-copy (LSC) region of 87,239 bp, a small single-copy (SSC) region of 18,076 bp, and a pair of inverted repeats (IRa and IRb), each 26,068 bp in length. The overall guanine-cytosine (GC) content was 36.1%. A total of 132 genes were identified, including 90 protein-coding genes, 38 transfer RNA (tRNA) genes, and 8 ribosomal RNA (rRNA) genes. Within the IR regions, 17 duplicated genes were detected, of which six were protein-coding genes, seven rRNAs genes and four tRNAs genes (Fig.1).

Microsatellite (SSR) analysis

A total of 56 microsatellites (SSRs) were detected within the plastome of *Chorizanthe novoana*. These included 34 mononucleotide repeats (60.71%), 7 dinucleotide repeats (12.5%), 3 trinucleotide repeats (5.36%), and 12 tetranucleotide repeats (21.43%). Two types of mononucleotide motifs were identified, with the polyA motif being the most frequent (33 repeats). Among the dinucleotide repeats, only the "AT" motif was observed, occurring seven times. Two trinucleotide motifs were detected, with "AAT" being the most common. The predominant tetranucleotide motifs were "AAAT" and "AATT", each occurring four times. The total SSR length was 637 bp, with an average repeat length of 11.38 bp. Most SSRs were located in intronic and intergenic regions; however, a few were identified within coding regions, including *rpo*C2, *atp*B, and *ndh*E.

Phylogenetic analysis

The best model selected for phylogenetic reconstruction was a system of eight partitions for the genes analyzed. The optimal model for phylogenetic reconstruction included eight partitions across the analyzed genes (Table 2). The maximum likelihood tree had a log-likelihood score of -292,790.0130 and showed strong branch support throughout (Fig. 2). *Chorizanthe novoana* was recovered within the subfamily Eriogonoideae, closely grouped with *Ruprechtia albida*, and forming a well-supported clade alongside *Coccoloba uvifera* and *Triplaris americana*. All other species in the analysis formed a sister clade corresponding to the subfamily Polygonoideae, which was subdivided into distinct lineages.

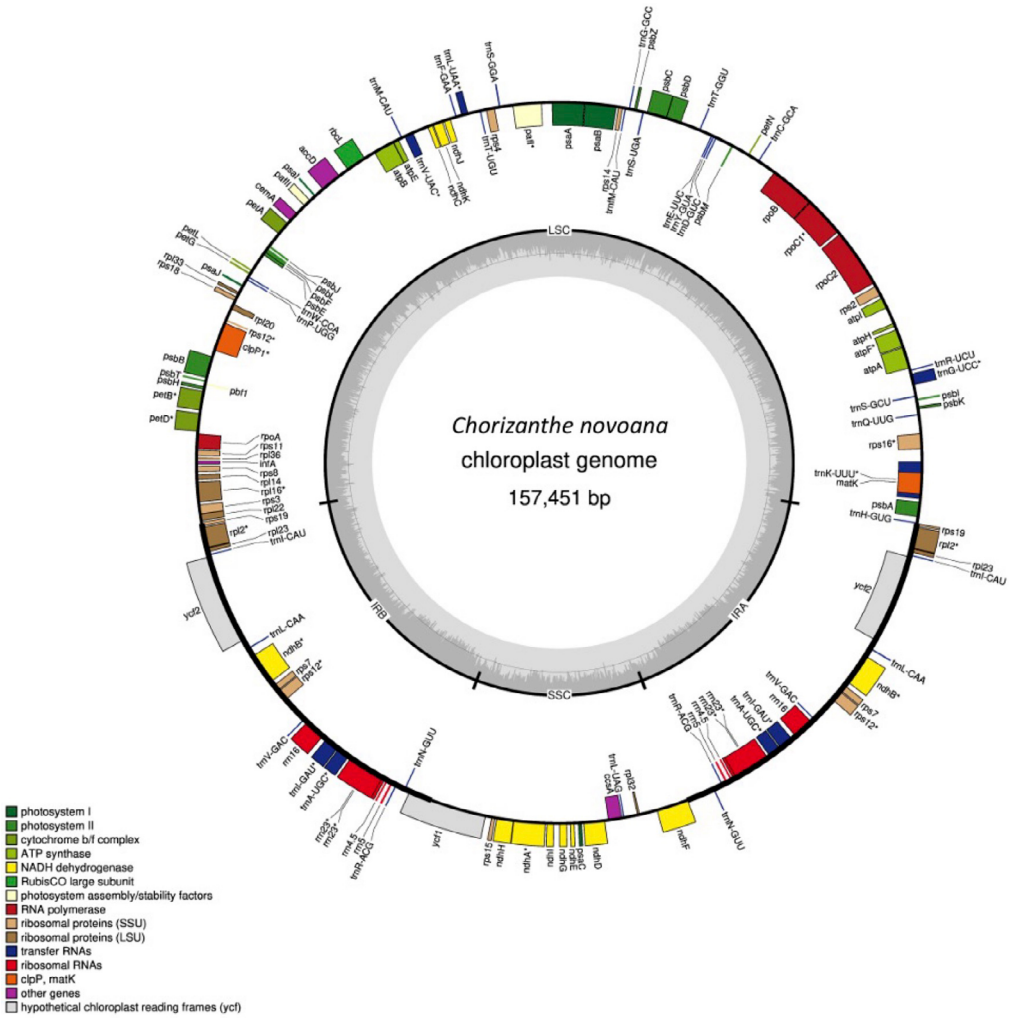

Figure 1. Complete chloroplast genome of *Chorizanthe novoana*, including protein-coding genes, tRNAs and rRNAs. / Genoma cloroplastidial completo de Chorizanthe novoana, incluyendo genes codificantes de proteinas, tRNAs y rRNAs.

Table 2. Partition model scheme implemented in ML tree inference analysis. / Esquema del modelo de partición implementado en el análisis de inferencia de árboles ML.

Partition	Model
accD+ccsA+matK+rpl22+rpl33+rps15	TVM+F+G4
atpA+atpB+atpE+atpF+infA+ndhE+ndhK+pafI+petA+petG+ petL+psaJ+psaI+psbK+psbT+psbZ+psbI+rbcL+rpl14+rpl36+ rpoB+rpoC1+rps2+rps4+rps11+rps14+rps16+rps18	GTR+F+I+I+R2
atpH+atpI+ndhC+ndhJ+pbf1+psaA+ psaB+psaC+psbA+psbE+psbF+psbJ+psbL	GTR+F+I+I+R2
cemA+clpP1+pafII+rpl20+rpoA+rpoC2+rps3+rps8	TVM+F+I+I+R2
ndhA+ndhD+ndhG+ndhH+ndhI+psbM	TVM+F+I+G4
ndhF+rpl32	TVM+F+R2
petN+psbB+psbC+psbD	GTR+F+I+G4
psbH	HKY+F

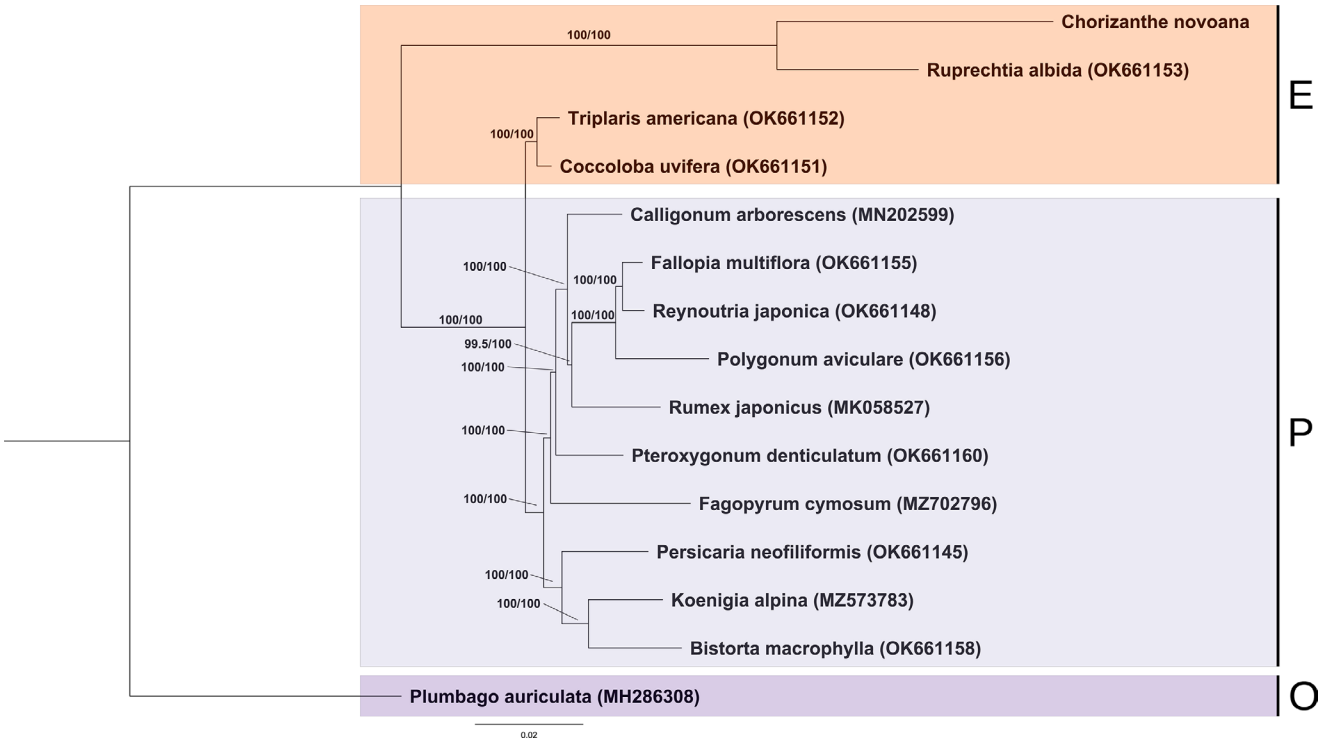

Figure 2. Maximum likelihood tree with SH-aLRT/ultrafast bootstrap support (100/100) for *C. novoana* and the other species with their respective accession numbers. “E”: Eriogonoideae; "P": Polygonoideae; “O”: outgroup. / Árbol de máxima verosimilitud con soporte SH-aLRT/bootstrap ultrarrápido (100/100) para *C. novoana* y las demás especies con sus respectivos números de acceso. "E": Eriogonoideae; "P": Polygonoideae; "O": grupo externo.

DISCUSSION

The structure and composition of the *Chorizanthe novoana* plastome suggests a conserved molecular organization, consistent with patterns observed in other members of the Polygonaceae family. The total sequence length (157,451 bp) falls within the lower range of the analyzed group, slightly exceeding the plastome size of *R. albida* (157,255 bp) but remaining smaller than those of *C. uvifera* (169,360 bp) and *T. americana* (171,340 bp). Similarly, the guanine-cytosine (GC) content (36.1%) is on the lower end for the subfamily Eriogonoideae (Zhang *et al.* 2022). The number of genes is consistent with those reported in other Polygonaceae species, such as *T. americana* (Zhang *et al.* 2022), suggesting overall structural conservation within the family.

Within the inverted repeat regions, exons 2 and 3 of the rps12 gene were detected, whereas exon 1 was located in the LSC region — a known feature of plastomes with the trans-spliced structure of this gene (Hildebrand *et al.* 1987). This configuration has also been documented in other members of Polygonaceae, such as the genus *Polygonatum (*Yao *et al.* 2025).

Consistent with other plastomes in the family, 56 simple sequence repeats (SSRs) were identified in *Chorizanthe novoana*, predominantly of the mononucleotide type. This count falls within the range reported for other Polygonaceae species, with recent studies indicating between 48 and 77 SSRs per plastome, ranging from mononucleotides (average of 33) to hexanucleotides (typically a single motif, e.g., in *Oxyria sinensis*) (Yang *et al.* 2021). Notably, while pentanucleotide motifs are commonly found in other Polygonaceae plastomes, none were detected in *C. novoana*. Conversely, the number of tetranucleotide repeats in *C. novoana* surpassed the previously reported maximum in the family, exceeding the nine repeats observed in *Rheum palmatum* (Yang *et al.* 2021). These features may preliminarily reflect distinctive aspects of the *C. novoana* plastome within Polygonaceae, although more extensive sampling among individuals of *C. novoana* and related species within the genus *Chorizanthe* is required to validate this observation.

The phylogenetic analysis confirms the placement of *Chorizanthe* within the subfamily Eriogonoideae, in agreement with previous studies (Sanchez *et al.* 2009; Kempton 2012). Interestingly, Eriogonoideae has been

subject to varying taxonomic interpretations depending on whether molecular or morphological data are used, a discrepancy that warrants further examination through expanded phylogenomic evidence. Our analysis was limited to three available plastomes from Eriogonoideae (*C. uvifera*, *R. albida*, and *T. americana*), which restricts the resolution of broader phylogenetic patterns. Nonetheless, this study contributes valuable genomic data that may support a more comprehensive phylogenomic approach to resolving diversification within the group.

In conclusion, this study presents the first report on the structure and composition of the plastome of *Chorizanthe novoana*, a critically endangered species endemic to Chile (MMA 2024). The plastome represents a valuable genetic resource for future research involving this species and others within Polygonaceae. Its inclusion in phylogenetic analyses will enhance our understanding of evolutionary relationships, diversification processes, conservation priorities, and potential taxonomic revisions based on molecular evidence.

ACKNOWLEDGMENTS

We thank the ANID Fondecyt Initiation N° 11220556 project for the financial support that made this research possible. We also thank Patricio Novoa and Sebastian Teillier for providing critical information to obtain the required plant material for analyses and Ignacio Aravena for his help with the English translation.

REFERENCES

Brandbyge, J. 1993. Polygonaceae. In: Kubitzki, K., Rohwer, J.G., Bittrich, V. (Eds.) Flowering Plants Dicotyledons, pp. 531-544. Springer Berlin Heidelberg.

Cai, X.-L., Landis, J. B., Wang, H.-X., Wang, J.-H., Zhu, Z.-X., Wang, H.-F. 2021. Plastome structure and phylogenetic relationships of Styracaceae (Ericales). BMC Ecology and Evolution 21(1): 103.

Du, L., Zhang, C., Liu, Q., Zhang, X., Yue, B. 2018. Krait: An ultrafast tool for genome-wide survey of microsatellites and primer design. Bioinformatics 34(4): 681-683.

Freeman, C.C., Reveal, J.L. 2005. Poylgonaceae Jussieu. Flora of North America 5: 216-601.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Systematic Biology 59(3): 307-321.

Handika, H., Esselstyn, J.A. 2024. SEGUL: Ultrafast, memory-efficient and mobile-friendly software for manipulating and summarizing phylogenomic datasets. Molecular Ecology Resources 24(7): e13964.

Hildebrand, M., Hallick, R.B., Passavant, C.W., Bourque, D.P. 1988. Trans-splicing in chloroplasts: The rps 12 loci of Nicotiana tabacum. Proceedings of the National Academy of Sciences 85(2): 372-376.

Hoang, D.T., Chernomor, O., Von Haeseler, A., Minh, B.Q., Vinh, L.S. 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution 35(2): 518-522.

Jin, J.-J., Yu, W.-B., Yang, J.-B., Song, Y., dePamphilis, C.W., Yi, T.-S., Li, D.-Z. 2020. GetOrganelle: A fast and versatile toolkit for accurate de novo assembly of organelle genomes. Genome Biology 21(1): 241.

Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., Von Haeseler, A., Jermiin, L.S. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods 14(6): 587-589.

Katoh, K., Standley, D.M. 2013. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Molecular Biology and Evolution 30(4): 772-780.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A. 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28(12): 1647-1649.

Kempton, E.A. 2012. Systematics of Eriogonoideae s. S. (Polygonaceae). Systematic Botany 37: 723-737.

Kim, H.T., Kim, J.S., Lee, Y.M., Mun, J., Kim, J. 2019. Molecular markers for phylogenetic applications derived from comparative plastome analysis of Prunus species. Journal of Systematics and Evolution 57(1): 15-22.

Lanfear, R., Calcott, B., Ho, S.Y.W., Guindon, S. 2012. PartitionFinder: Combined Selection of Partitioning Schemes and Substitution Models for Phylogenetic Analyses. Molecular Biology and Evolution 29(6): 1695-1701.

Liu, Q., Huang, H., Lin, J., Liu, F., Wang, X., Chen, R., Geng, X. 2025. Decoding the Chloroplast Genome of Bitterwood (*Picrasma quassioides*): Structure, Variability, and Evolutionary Relationships. Ecology and Evolution 15(4): e71245.

Mak, S.S.T., Gopalakrishnan, S., Carøe, C., Geng, C., Liu, S., Sinding, M.-H.S., Kuderna, L.F.K., Zhang, W., Fu, S., Vieira, F.G., Germonpré, M., Bocherens, H., Fedorov, S., Petersen, B., Sicheritz-Pontén, T., Marques-Bonet, T., Zhang, G., Jiang, H., Gilbert, M.T.P. 2017. Comparative performance of the BGISEQ-500 vs Illumina HiSeq2500 sequencing

platforms for palaeogenomic sequencing. GigaScience 6(8): gix049.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., Von Haeseler, A., Lanfear, R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution 37(5): 1530-1534.

MMA. 2024. Clasificación de especies. Ministerio del Medio Ambiente, Chile. https://clasificacionespecies.mma.gob.cl/ Accessed: May 22, 2025.

Gitzendanner, P.S., Soltis, T.-S., Yi, D.-Z., Li, D.E. 2018. In: Chaw, S-M., Jansen, R. (Eds.) Plastome Phylogenetics: 30 Years of Inferences Into Plant Evolution. Advances in Botanical Research, Vol. 85, pp. 293-313. Elsevier.

Provan, J., Powell, W., Hollingsworth, P.M. 2001. Chloroplast microsatellites: New tools for studies in plant ecology and evolution. Trends in Ecology & Evolution 16(3): 142-147.

Rambaut, A., Drummond, A.J., Xie, D., Baele, G., Suchard, M.A. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Systematic Biology 67(5): 5.

Sanchez, A., Schuster, T.M., Kron, K.A. 2009. A Large-Scale Phylogeny of Polygonaceae Based on Molecular Data. International Journal of Plant Sciences 170(8): 1044-1055.

Teillier, S., Macaya-Berti, J., Novoa, P. 2019. Revisión de Chorizanthe R. Br. Ex Benth. (Polygonaceae) en Chile. Gayana Botánica 76(1): 34-60.

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

Wick, R.R., Schultz, M.B., Zobel, J., Holt, K.E. 2015. Bandage: Interactive visualization of *de novo* genome assemblies. Bioinformatics 31(20): 3350-3352.

Yang, B., Li, L., Liu, J., Zhang, L. 2021. Plastome and phylogenetic relationship of the woody buckwheat *Fagopyrum tibeticum* in the Qinghai-Tibet Plateau. Plant Diversity 43(3): 198-205.

Yao, J., Zheng, Z., Xu, T., Wang, D., Pu, J., Zhang, Y., Zha, L. 2025. Chloroplast Genome Sequencing and Comparative Analysis of Six Medicinal Plants of Polygonatum. Ecology and Evolution 15(1): e70831.

Zhang, H., Zhang, X., Sun, Y., Landis, J.B., Li, L., Hu, G., Sun, J., Tiamiyu, B.B., Kuang, T., Deng, T., Sun, H., Wang, H. 2022. Plastome phylogenomics and biogeography of the subfam. Polygonoideae (Polygonaceae). Frontiers in Plant Science 13: 893201.

Zhang, H.-R., Wei, R., Xiang, Q.-P., Zhang, X.-C. 2020. Plastome-based phylogenomics resolves the placement of the sanguinolenta group in the spikemoss of lycophyte (Selaginellaceae). Molecular Phylogenetics and Evolution 147: 106788.

Zhao, Y., Yang, Z., Zhao, Y., Li, X., Zhao, Z., Zhao, G. 2019. Chloroplast Genome Structural Characteristics and Phylogenetic Relationships of Oleaceae. Chinese Bulletin of Botany 54(4): 441-454.

Zhou, W., Armijos, C.E., Lee, C., Lu, R., Wang, J., Ruhlman, T.A., Jansen, R.K., Jones, A.M., Jones, C.D. 2023. Plastid Genome Assembly Using Long-read data. Molecular Ecology Resources 23(6): 1442-1457.

Received: 26.05.2025
Accepted: 03.09.2025

Editor: Alfredo Saldaña