




Nuclear DNA-based phylogenetic analysis of *Neocinnamomum* species

Qishao Li^{1,2} , Linyi Yang³, Qunfei Yu^{4,5}, Wenbin Xu^{4,5}, Yaxuan Xin^{1,2}, Yu Song⁶ 
and Peiyao Xin^{1,2} 

Research Article

Cite this article: Li Q, Yang L, Yu Q, Xu W, Xin Y, Song Y, Xin P (2023). Nuclear DNA-based phylogenetic analysis of *Neocinnamomum* species. *Plant Genetic Resources: Characterization and Utilization* **21**, 323–330. <https://doi.org/10.1017/S1479262123000771>

Received: 30 May 2023

Revised: 23 September 2023

Accepted: 25 September 2023

First published online: 20 October 2023

Keywords:

genetic relationship; genome; *Neocinnamomum*; nuclear DNA sequence; phylogenetic analysis

Corresponding author:

Yu Song; Email: songyu@gxnu.edu.cn;

Peiyao Xin; Email: xpytyx@163.com

¹Southwest Research Center for Landscape Architecture Engineering, National Forestry and Grassland Administration, Southwest Forestry University, Kunming, China; ²Key Laboratory of Forest Resources Conservation and Utilization in the Southwest Mountains of China Ministry of Education, Southwest Forestry University, Kunming, China; ³Yunnan Forestry Technological College, Kunming, China; ⁴Center for Integrative Conservation & Key Laboratory of Tropical Plant Resources and Sustainable Use, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, Yunnan, China; ⁵University of Chinese Academy of Sciences, Beijing, China and ⁶Key Laboratory of Ecology of Rare and Endangered Species and Environmental Protection (Ministry of Education) and Guangxi Key Laboratory of Landscape Resources Conservation and Sustainable Utilization in Lijiang River Basin, Guangxi Normal University, Guilin, Guangxi, China

Abstract

Neocinnamomum plants are evergreen shrubs or small trees belonging to the Neocinnamomeae tribe of the Lauraceae family. Their seeds are rich in fatty acids, and their leaves are often used in traditional Chinese medicine. Presently, only a few studies have been performed on *Neocinnamomum* plants; therefore, the genome-based phylogeny among *Neocinnamomum* species has not been determined, which limits the germplasm innovation of this genus. In this study, by the Illumina (next-generation sequencing) and third-generation sequencing technologies, the whole genomes of seven *Neocinnamomum* species samples were sequenced, their nuclear DNA (nrDNA) sequences were assembled and characterized, and their phylogeny was reconstructed. The results revealed four hypervariable regions (i.e. transcribed spacer regions) in the nrDNA sequences, among which the highest degree of variation was observed in the external transcribed spacer (ETS) region localized behind the 26S gene. A total of 27 insertions/deletions and 184 single-nucleotide polymorphisms, both localized mainly in the ETS and internal transcribed spacer regions, were identified. Phylogenetic trees were constructed based on the nrDNA sequences using the maximum likelihood (ML) and Bayesian inference (BI) methods with *Caryodaphnopsis henryi* as the outgroup. The ML tree divided the seven *Neocinnamomum* species into four clades. Clade I consisted of *Neocinnamomum caudatum* var. *macrocarpum* and *Neocinnamomum caudatum*, clade II included *Neocinnamomum delavayi* and *Neocinnamomum mekongense*, clade III included *Neocinnamomum fargesii* and a branch species of *N. delavayi* and clade IV included *Neocinnamomum lecomtei*, constituting a monophyletic and basal group. The BI tree shared the same topological structure as the ML tree, and all the support values of the BI tree were one except for that of one *Neocinnamomum* species (0.98). The results of this study provide new evidence regarding the phylogenetic evolution of the *Neocinnamomum* plants.

Introduction

Neocinnamomum plants are evergreen shrubs or small trees belonging to the Neocinnamomeae tribe of the Lauraceae family. To date, five *Neocinnamomum* species have been recorded in the Chinese Flora (A Editorial Committee of Flora of China, 1982), namely *Neocinnamomum caudatum* (*N. caudatum*), *Neocinnamomum fargesii* (*N. fargesii*), *Neocinnamomum lecomtei* (*N. lecomtei*), *Neocinnamomum mekongense* (*N. mekongense*) and *Neocinnamomum delavayi* (*N. delavayi*). Two additional *Neocinnamomum* species, *Neocinnamomum atjehense* (*N. atjehense*) and *Neocinnamomum parvifolium* (revised to a branch species of *N. delavayi*) that are distributed outside China, are recorded in the World Flora. On the International Plant Names Index website, 14 *Neocinnamomum* species are listed, and seven species are left after removing synonyms, including *Neocinnamomum caudatum* var. *macrocarpum* (Xu *et al.*, 2017). *Neocinnamomum* plants are mainly distributed in the south-central, western and northwestern provinces of China, as well as in Nepal and Sikkim. They mainly grow as shrubs in sparse or dense forests or at forest margins, at an altitude of 1100–2300 m, along the banks of river valleys, ditch sides or on well-drained limestones (Zhang and Jia, 2013). *Neocinnamomum* plants are important oil trees and valuable medicinal plant resources in China, with multiple economic applications (Nanjing Medical University, 1977; Wang and Yang, 2013). Presently, only a few studies have been performed on the genomes of *Neocinnamomum* plants compared with other oil plants; moreover, the



degree of development and utilization of their wild resources is extremely low (Gan *et al.*, 2018).

Ribosomal DNA is parentally inherited, and phylogenetic analysis can comprehensively reveal the parental pedigree and system network evolution relationship (Kellogg and Bennetzen, 2004). Presently, the research and application of nuclear genomic sequences mainly focus on the internal transcribed spacer (ITS) sequences of ribosomal DNA (Liston *et al.*, 1996; Mayol and Rosselló, 2001; Wei *et al.*, 2003; Won and Renner, 2005; Kan *et al.*, 2007; Zheng *et al.*, 2008). The general length of ITS sequences is 565–700 base pairs (bp), whereas that of the ribosomal genome ranges from 6000 to 8000 bp and has more genetic loci and information. The plant nuclear DNA (nrDNA) sequences have highly tandem repeats; however, these repeated sequences are homozygous or near-homozygous in most species because of homogeneous evolution, especially ITS1 and ITS2. Owing to their fast evolution, relatively short sequences and many mutant sites, ITS1 and ITS2 of the nuclear genome have been used in systematic studies (Schaal and Learn, 1988; Baldwin *et al.*, 1999). ITS sequences are used in the studies of evolutionary genetics, including species identification, polyploid origin, hybridization and evolution. Artyukova *et al.* (2005) established the phylogenetic relationship between species in the Far East and other species in the family by analysing the ITS sequence of eight species of Araliaceae growing in the Far East of Russia (Artyukova *et al.*, 2005). Luo *et al.* (2005) analysed the ITS sequences of 51 species and one subgenus of *Aconitum* from eastern Asia, northern America and Europe, the results established the taxonomic status of the subgenus *Aconitum* of Ranunculaceae, and were also supported by morphological characteristics such as seeds and petals, which corrected the errors of traditional classification (Luo *et al.*, 2005). Sudheer Pamidimarri *et al.* (2009) used the ITS sequence to study the phylogenetic relationship and genetic differentiation of *Jatropha curcas* L. (Sudheer Pamidimarri *et al.*, 2009). A large number of studies have shown that ITS sequence analysis has irreplaceable advancement and reliability in the study of genetic relationship of medicinal plants, which will become the trend of future research.

Owing to the small genetic differences among species, accurate analysis of the relationship among species based on a small number of genetic markers is difficult. In recent years, in many studies, the use of the complete mitochondrial genome, chloroplast genome and nrDNA was preferred to explore the phylogenetic relationship among species (Hsiao *et al.*, 1994; Liu *et al.*, 2014; Dong *et al.*, 2020; Kang *et al.*, 2021; Xin *et al.*, 2023). Among the three types of genomes present in cells, the nuclear genome is the largest, most dominant and inherited from both parents (Li, 2006). Presently, the publicly available nuclear genomic data of the *Neocinnamomum* plants are relatively limited, which slows down the systematic genome-based research on the *Neocinnamomum* plants. In this study, we assembled nrDNA sequences of seven *Neocinnamomum* species using genome resequencing technologies, constructed high-resolution phylogenetic trees and studied the phylogenetic relationships among the *Neocinnamomum* species. The aims of this study were as follows: (1) constructing a phylogenetic framework with high support values; (2) exploring the phylogenetic differentiation process in the *Neocinnamomum* species; and (3) providing a scientific basis for the development and utilization of germplasm resources, genetic improvement and constructing germplasm resources of the *Neocinnamomum* plants.

Materials and methods

Plant materials

In this study, a total of 51 leaf samples from seven *Neocinnamomum* species were collected from the healthy and mature branches of at least three plants per species (details shown in online Supplementary Table S1). These samples were collected mainly in Yunnan, Sichuan and Guangxi from an altitude ranging from 1100 to 2300 m above sea level. From each plant, two to three leaves were collected and packed in Ziplock bags and dried with silica gel for later use for DNA isolation. These leaves were photographed with colour cards and rulers as references. The specimens were flattened with dry newspaper and cardboard and dried in an electric blast drying oven.

Sequencing

Samples were sent to the Kunming Institute of Botany, Chinese Academy of Sciences (Kunming, China) for nrDNA sequencing. Illumina (Illumina, Inc., San Diego, California, USA) (next-generation sequencing) and third-generation sequencing technologies were used, which enable the sequencing of millions of DNA bases parallelly, allowing an in-depth, detailed and comprehensive analysis of the sequenced genome and transcriptome. The core principle is sequencing-by-synthesizing, which determines the DNA base by capturing the marker's signal of the newly synthesized end (Liang *et al.*, 2020).

Genome assembly and annotation

The obtained sequencing data of the nrDNA sequences were assembled using the GetOrganelle software package (Jin *et al.*, 2020). The assembly workflow included (1) extracting all possible reads by extension based on the reference target genes or genomes in the database; (2) extracting the reads of the target genome; (3) removing non-target gene fragments and low-coverage contigs; and (4) final assembly of the genome and output results.

Bandage (Wick *et al.*, 2015) software was used to view, edit and export complete nrDNA sequences. For the samples without a complete structure, the log files were inspected to adjust the corresponding parameters and re-analyse the data; moreover, the output scaffolds.fasta file was examined and further filtered, with closely related species as reference sequences, to generate a complete structure.

The online software GeSeq (Tillich *et al.*, 2017) and Geneious (Kearse *et al.*, 2012) were used to annotate the assembled nrDNA sequences, which were then manually corrected using the open reading frames of the corresponding sequences. Geneious was used to annotate the nrDNA sequences of the seven *Neocinnamomum* species using *Eriobotrya fragrans* as the reference.

Genome characterization and mutation analysis

Geneious was used to calculate the size, total GC content, gene content and pseudogenes of each region of the nrDNA. The nrDNA sequences were aligned using an online software Multiple Alignment using Fast Fourier Transform (MAFFT) v7 (Katoh and Standley, 2013) and were manually checked by MEGA X (Kumar *et al.*, 2018). Insertions/deletions (Indels) and single-nucleotide polymorphisms (SNPs) were counted and localized using manual statistical methods. Additionally, nrDNA

hypervariable regions were analysed using mVASTA and DNAsp, and the sliding window method was performed in DnaSP v6.0 (Rozas *et al.*, 2017) to evaluate the nucleotide diversity (π) value of the nrDNA sequences.

Genome sequence alignment and phylogenetic analysis

The ribosomal gene matrix was first aligned by using MAFFT and then manually adjusted and corrected by Geneious to obtain a reliable comparison matrix for subsequent phylogenetic analysis. A maximum likelihood (ML) phylogenetic tree was constructed by IQ-TREE software, with the general time-reversible model and a fast bootstrap value (bootstrap iteration) of 1000. An additional Bayesian inference (BI) tree was constructed by the CIPRES online software (<http://www.phylo.org/>). FigTree was used to view, edit and obtain the support values and branch length of the phylogenetic trees.

Results

Geographical distribution of the *Neocinnamomum* plants

ArcGIS software was used to create a heat map of the geographical distribution of the *Neocinnamomum* species (Fig. 1), which were mainly distributed in Yunnan, Sichuan, Tibet, Guizhou, Hubei, Chongqing and Guangxi of China, as well as in Myanmar, Laos and Vietnam. In China, Guangxi was most abundant in *Neocinnamomum* species (six species), followed by Yunnan (five species), Tibet (three species), Sichuan (three species),

Guizhou (three species), Hubei (two species) and Chongqing (one species).

nrDNA sequence analysis

Genome size and gene content

Across the 51 *Neocinnamomum* samples, the nrDNA length ranged from 6716 to 6761 bp, and the GC content ranged from 56.3 to 56.9%. The lengths of the external transcribed spacer (ETS) region (including ETS1 and ETS2, namely 5' ETS and 3' ETS, respectively), 18S gene, 5.8S gene, 26S gene, ITS1 and ITS2 were 926–940, 1810, 159, 3384–3385, 228–264 and 203–208 bp, respectively, and the corresponding GC contents were 57.1–59.2, 50.4–50.5, 57.9, 57.9–58.3, 63.1–65.9 and 68.8–70.4%, respectively (online Supplementary Table S2).

nrDNA mutation analysis

A total of 27 indels were identified in all *Neocinnamomum* nrDNA sequences (Table 1), with the 26S gene having one, both the ETS regions having 12, ITS1 having 13, ITS2 having one and 18S and 5.8S genes having no indels. A total of 184 SNPs were detected in the nrDNA sequences (Fig. 2), including 51 transitions (TS) and 133 transversions (TV). The ratio of TS to TV was 1:0.38. A total of 20 SNPs were detected in the three RNA genes, with the 5.8S gene having none, the 18S gene having two and the 26S gene having 18 SNPs. Among the transcribed

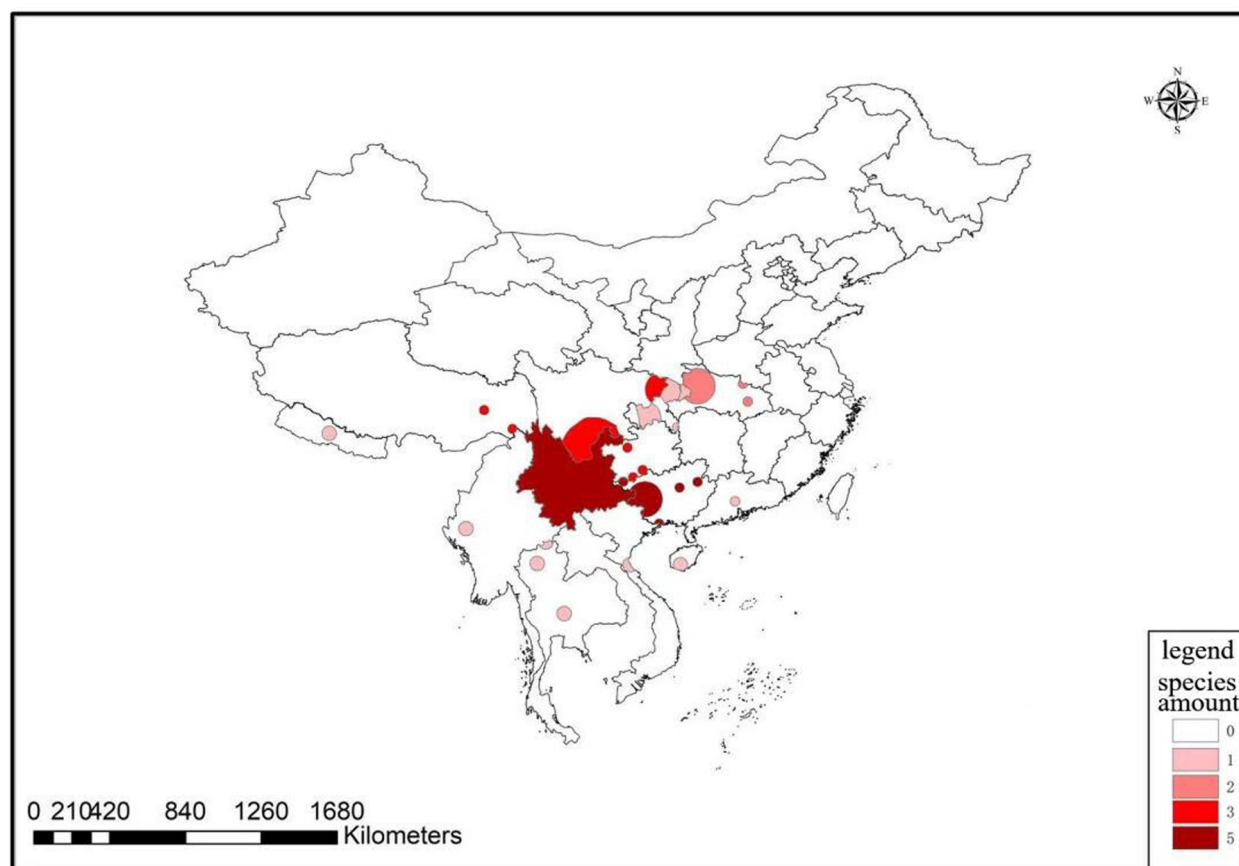


Figure 1. Geographical distribution of *Neocinnamomum* species.

Table 1. Pattern and number of indels in nrDNA sequences of *Neocinnamomum* plants

Number	Location	Nucleotide(s)	Size/bp	Type
1	26S	A/G	1	Insertion
2	ETS	G	1	Insertion
3	ETS	T	1	Insertion
4	ETS	T	1	Insertion
5	ETS	C	1	Insertion
6	ETS	TGC	3	Insertion
7	ETS	C	1	Insertion
8	ETS	GACC	4	Insertion
9	ETS	T	1	Insertion
10	ETS	C	1	Deletion
11	ETS	G	1	Insertion
12	ETS	TTTATTTGTATA	14	Deletion
13	ETS	A	1	Insertion
14	ITS1	A	1	Insertion
15	ITS1	CT	2	Insertion
16	ITS1	T	1	Insertion
17	ITS1	CTGTCGTCC	9	Deletion
18	ITS1	CCTGTCGTCCTCGCGGCACAT	23	Deletion
19	ITS1	GCG	3	Insertion
20	ITS1	A	1	Insertion
21	ITS1	A	1	Insertion
22	ITS1	A	1	Deletion
23	ITS1	C	1	Insertion
24	ITS1	T	1	Deletion
25	ITS1	AGT	3	Insertion
26	ITS1	AGGGGA	6	Deletion
27	ITS2	G	1	Insertion

spacer regions, 118 SNPs were detected in both the ETS regions combined, 20 in ITS2 and 26 in ITS1.

Analysis of hypervariable regions

The Pi values of the nrDNA sequences of the 51 *Neocinnamomum* samples were 0–0.09111 (Fig. 3), with an average Pi value being 0.020188. Four hypervariable regions, with a Pi value greater than 0.04, were found in the nrDNA sequences, and all were located in the transcribed spacer regions. Among them, ETS2 had the highest Pi value (0.09111), followed by ETS1 (0.06039), whereas the Pi values of ITS1 and ITS2 were relatively low, with a maximum value of 0.04292.

Phylogenetic analysis

The length of the nrDNA sequences of the 51 *Neocinnamomum* samples ranged from 6716 to 6761 bp and the alignment matrix was 6790 bp in length after manual adjustment. In this study, two phylogenetic trees (i.e. ML and BI trees) were reconstructed

based on the nrDNA sequences of the seven *Neocinnamomum* species (Fig. 4), with *Caryodaphnopsis henryi* as the outgroup. The ML tree divided the *Neocinnamomum* species into four clades, with clade I consisting of *N. macrocarpum* and *N. caudatum*, clade II with *N. delavayi* and *N. mekongense*, clade III with *N. fargesii* and a branch species of *N. delavayi* and clade IV with *N. lecomtei* constituting the most basal clade. All support values of the BI tree were one, except for one *Neocinnamomum* species, which had a support value of 0.98. The topological structure of the BI tree was consistent with that of the ML tree.

Discussion

In this study, the phylogenetic tree constructed by nrDNA divided the *Neocinnamomum* species into four clades, with clade I consisting of *N. macrocarpum* and *N. caudatum*, clade II with *N. delavayi* and *N. mekongense*, clade III with *N. fargesii* and a branch species of *N. delavayi* and clade IV with *N. lecomtei* constituting the most basal clade. Our result contrasts with an earlier phylogenetic study based on the complete chloroplast genomes of

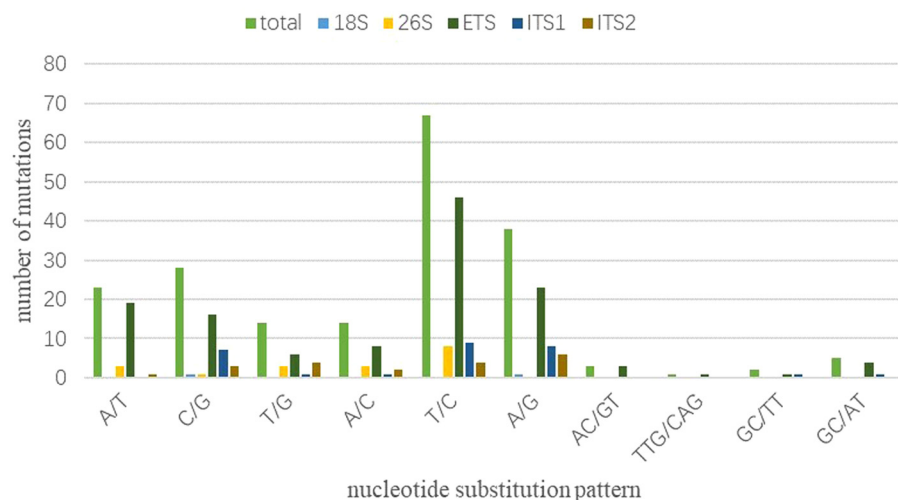


Figure 2. Nucleotide substitution patterns in the nuclear DNA sequences of the *Neocinnamomum* plants.

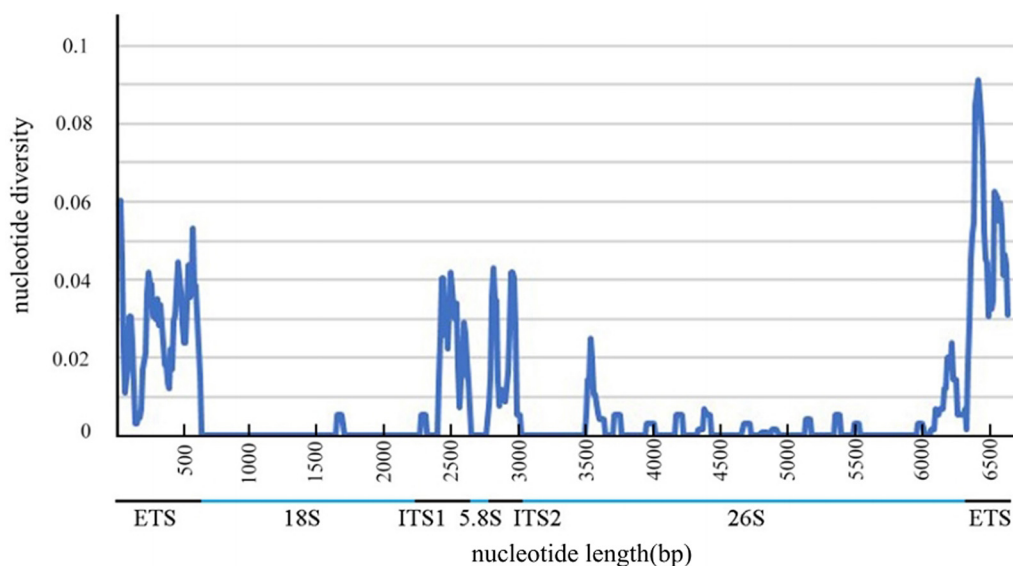


Figure 3. Nucleotide diversity in hypervariable regions of the nuclear DNA sequences of the *Neocinnamomum* plants.

N. delavayi, *N. mekongense* and *N. lecomtei* using the ML method and revealed that *N. delavayi* was a sister species to *N. mekongense* and *N. lecomtei*, having high support values (Ren *et al.*, 2019). One explanation for this is that in most higher plants, the mitochondrial genome and chloroplast genome belong to maternal inheritance, while the nrDNA sequence belongs to parental inheritance, which can more truly reflect the genetic relationship between species (Zong, 2019). In the division of interspecific relationship, ribosomal genome and ITS sequence are more suitable for the division of interspecific relationship than chloroplast genome and mitochondrial genome. Qu (2022) constructed three phylogenetic trees of 44 *Carvodayphnopsis* using nrDNA, chloroplast and mitochondrial genomes. There were serious conflicts among the three phylogenetic trees. Compared with the ribosomal gene tree, chloroplast and mitochondrial genomes performed poorly in the division of interspecific relationships. The results not only verify the effectiveness of nuclear genome sequences in the classification of interspecific relationships, but also show that chloroplast and mitochondrial genomes have limitations in the definition of subgenus relationships (Qu, 2022).

Hybridization has been reported in plant lineages and is considered to be a key mechanism of plant evolution and far more common than in of animals. Hybridization can result in gene tree conflicts, especially between single-parent inheritance chloroplast or mitochondrial fragments and biparental inheritance nuclear gene fragments (Soltis and Kuzoff, 1995). Sequences from different genomes of hybrid species often reflect different genetic systems (e.g. mitochondrial genes from fathers, plastid genes from maternal lines and nuclear genes from two parents), which may lead to inconsistencies between these different data sources, resulting in differences between the nrDNA phylogenetic tree and the phylogenetic tree constructed by chloroplast and mitochondrial genomes. For example, a serious conflict between nuclear and plastid gene trees was demonstrated in the construction of *Cymbidium* phylogeny using ITS and chloroplast genes (Zhang *et al.*, 2021).

Neocinnamomum has been confirmed to be monophyletic by many scholars, but there are still many controversies in the systematic study of *Neocinnamomum*. Li *et al.* (2016) performed a phylogenetic study on *Caryodaphnopsis*. Among the four *Neocinnamomum* species, *N. delavayi* and *N. mekongense* were closely related sister

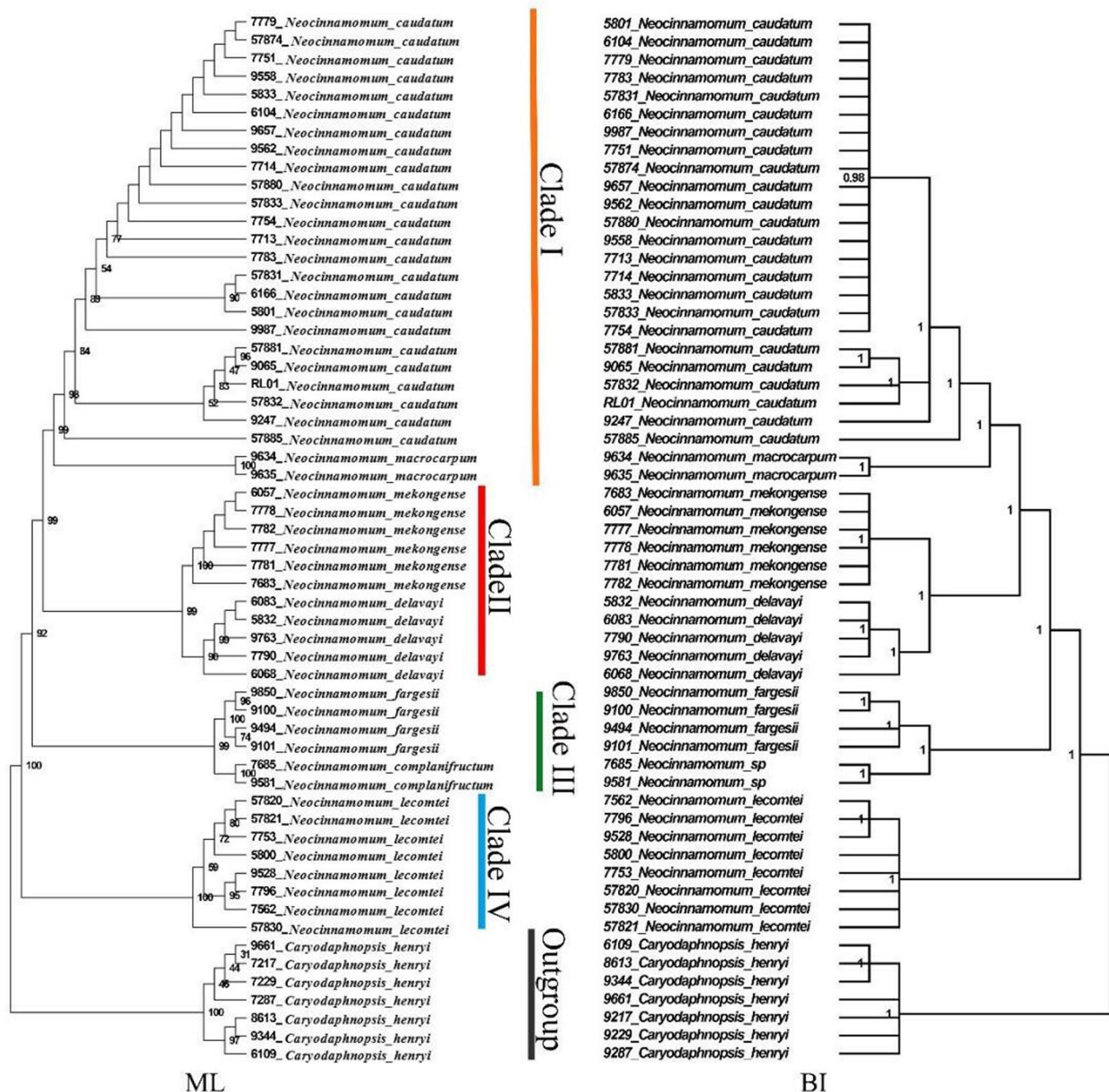


Figure 4. Nucleotide polymorphism in hypervariable regions of the nuclear DNA sequences of the *Neocinnamomum* plants.

groups, and then they were clustered with *N. fargesii* and *N. lecomtei*, respectively (Li *et al.*, 2016), which was consistent with the results of this study. Wang *et al.* (2010) suggested *N. delavayi* and *N. mekongense*, and *N. lecomtei* and *N. caudatum* as sister groups, respectively, whereas *N. fargesii* was determined as a basal and monophyletic group. Our findings were not consistent with the above results (Wang *et al.*, 2010). The reason may be that Wang *et al.* (2010) only used the chloroplast intergenic spacer and the three fragment sequences of nrDNA to construct a phylogenetic tree, so its accuracy is lower than that of the phylogenetic tree constructed by the chloroplast genome and the long sequence of nrDNA.

Conclusions

In this study, we sequenced the genomes of 51 individual leaf samples of seven *Neocinnamomum* species by next-generation

and third-generation sequencing technologies with a sequencing depth of 30×, and about 6 Gb raw data and 2 Gb clean data were obtained for each individual. All 51 nrDNA sequences were assembled using GetOrganelle software, and nrDNA mutations were analysed. Four hypervariable regions (i.e. transcribed spacer regions) were found in the nrDNA sequences, with ETS (located downstream of the 26S gene) having the highest degree of variation. A total of 27 indels and 184 SNPs were detected in the nrDNA sequences, both of which mainly occurred in the ETS and ITS regions. Finally, the ML and BI phylogenetic trees were constructed based on the nrDNA sequences. The ML and BI trees constructed based on the ribosomal genes divided the *Neocinnamomum* species into four clades. Clade I included *N. macrocarpum* and *N. caudatum*, clade II included *N. delavayi* and *N. mekongense*, clade III included *N. fargesii* and a branch species of *N. delavayi* and clade IV included *N. lecomtei*

constituting the most basal clade. In the BI tree, all the support values were one except for that of one *Neocinnamomum* species (PP = 0.98), and the topological structure of the BI tree was consistent with that of the ML tree.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1479262123000771>.

Data. The plastid marker matK sequence data of the 163 Lauraceae species will be submitted to LCGDB (<https://lcgdb.wordpress.com>) through the revision process.

Acknowledgements. The authors thank the financial support from the Yunnan Province Science and Technology Talents and Platform Project.

Author contributions. Sampling, Q. Y. and W. X.; data analysis, L. Y., Y. X. and Q. L.; thesis writing, Q. L.; conceived and designed the experiments, Y. S. and P. X.; project fund support provider, P. X. All authors have read and agreed to the published version of the manuscript.

Funding statement. This research was funded by the Yunnan Province Science and Technology Talents and Platform Project (No. 202205AF150022).

Competing interest. None.

References

- A Editorial Committee of Flora of China, Chinese Academy of Sciences** (1982) *Flora of China (Volume 31)*. Beijing, China: Beijing Science Press, pp. 115–120. ISBN 978-7-03-027166-2.
- Artyukova EV, Gontcharov AA, Kozyrenko MM, Reunova GD and Zhuravlev YN** (2005) Phylogenetic relationships of the far eastern Araliaceae inferred from ITS sequences of nuclear rDNA. *Russian Journal of Genetics* **41**, 649–658.
- Baldwin BG, Sanderson MJ, Porter JM, Wojciechowski MF and Donoghue C** (1999) The ITS region of nuclear ribosomal DNA: a valuable source of evidence on angiosperm phylogeny. *Annals of the Missouri Botanical Garden* **82**, 247–277.
- Dong S, Chen L, Liu Y, Wang Y, Zhang S, Yang L, Lang X and Zhang S** (2020) The draft mitochondrial genome of *Magnolia biondii* and mitochondrial phylogenomics of angiosperms. *PLoS ONE* **15**, e0231020.
- Gan Y, Song Y, Chen Y, Liu H, Yang D, Xu Q and Zheng Z** (2018) Transcriptome analysis reveals a composite molecular map linked to unique seed oil profile of *Neocinnamomum caudatum* (Nees) Merr. *BMC Plant Biology* **18**, 303.
- Hsiao C, Chatterton NJ and Asay KH** (1994) Phylogenetic relationships of 10 grass species: an assessment of phylogenetic utility of the internal transcribed spacer region in nuclear ribosomal DNA in monocots. *Genome* **37**, 112–120.
- Jin JJ, Yu WB, Yang J, Song Y, DePamphilis CW, Yi TS and Li DZ** (2020) GetOrganelle: a fast and versatile toolkit for accurate de novo assembly of organelle genomes. *Genome Biology* **21**, 241.
- Kan XZ, Wang SS, Ding X and Wang XQ** (2007) Structural evolution of nrDNA ITS in Pinaceae and its phylogenetic implications. *Molecular Phylogenetics and Evolution* **44**, 765–777.
- Kang DH, Ong HG, Lee JH, Jung EK, Kyaw NO, Fan Q and Kim YD** (2021) A new broad-leaved species of loquat from eastern Myanmar and its phylogenetic affinity in the genus *Eriobotrya* (Rosaceae). *Phytotaxa* **482**, 279–290.
- Katoh K and Standley DM** (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**, 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 and Drummond A** (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649.
- Kellogg EA and Bennetzen JL** (2004) The evolution of nuclear genome structure in seed plants. *American Journal of Botany* **91**, 1709–1725.
- Kumar S, Stecher G, Li M, Knyaz C and Tamura K** (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547–1549.
- Li ZZ** (2006) *Molecular Phylogeny of the Genus Actinidia Based on Nuclear DNA Genetic Markers and Cytoplasm DNA Sequence Analysis* (D.Agr thesis). University of Chinese Academy of Sciences, Beijing, China.
- Li L, Madriñán S and Li J** (2016) Phylogeny and biogeography of *Caryodaphnopsis* (Lauraceae) inferred from low-copy nuclear gene and ITS sequences. *Taxon* **65**, 433–443.
- Liang XH, Liu TT and Zhang Y** (2020) Application of next-generation sequencing based bulked segregant analysis in gene mapping of woody plants. *World Forestry Research* **33**, 54–61.
- Liston A, Robinson WA, Oliphant JM and Alvarez-Buylla ER** (1996) Length variation in the nuclear ribosomal DNA internal transcribed spacer region of non-flowering seed plants. *Systematic Botany* **21**, 109–120.
- Liu Y, Cox CJ, Wang W and Goffinet B** (2014) Mitochondrial phylogenomics of early land plants: mitigating the effects of saturation, compositional heterogeneity, and codon-usage bias. *Systematic Biology* **63**, 862–878.
- Luo Y, Zhang FM and Yang QE** (2005) Phylogeny of *Aconitum* subgenus *Aconitum* (Ranunculaceae) inferred from ITS sequences. *Plant Systematics and Evolution* **252**, 11–25.
- Mayol M and Roselló JA** (2001) Why nuclear ribosomal DNA spacers (ITS) tell different stories in *Quercus*. *Molecular Biology and Evolution* **19**, 167–176.
- Nanjing Medical University** (1977) *Dictionary of Traditional Chinese Pharmacy*. Shanghai, China: Shanghai People's Publishing House, pp. 124–129. ISBN 978-7-31-318819-9.
- Qu YY** (2022) *Phylogenetic Relationship Reconstruction and Morphological Traits Studies of Asian Caryodaphnopsis (Lauraceae)* (M.Agr thesis). Southwest Forestry University, Yunnan, China.
- Ren SL, Song Y, Zhao ML and Xu WB** (2019) The plastid genome sequence of *Neocinnamomum delavayi* (Lec.) Liou. *Mitochondrial DNA Part B: Resources* **4**, 3711–3712.
- Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P and Ramos-Onsins SE** (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Molecular Biology and Evolution* **34**, 3299–3302.
- Schaal BA and Learn GH** (1988) Ribosomal DNA variation within and among plant populations. *Annals of the Missouri Botanical Garden* **4**, 1207–1216.
- Soltis DE and Kuzoff RK** (1995) Discordance between nuclear and chloroplast phylogenies in the Heuchera group (Saxifragaceae). *Evolution* **49**, 727–742.
- Sudheer Pamidimarri DVN, Chattopadhyay B and Reddy MP** (2009) Genetic divergence and phylogenetic analysis of genus *Jatropha* based on nuclear ribosomal DNA ITS sequence. *Molecular Biology Reports* **36**, 1929–1935.
- Tillich M, Lehwark P, Pellizzer T, Ulbricht-Jones ES, Fischer A, Bock R and Greiner S** (2017) GeSeq – versatile and accurate annotation of organelle genomes. *Nucleic Acids Research* **45**, W6–W11.
- Wang Z and Yang J** (2013) Comprehensive evaluation of woody oil-bearing plants in Yunnan as sources for biodiesel. *Plant Diversity* **35**, 630–640.
- Wang ZH, Li J, Conran JG and Li HW** (2010) Phylogeny of the southeast Asian endemic genus *Neocinnamomum* H. Liu (Lauraceae). *Plant Systematics and Evolution* **290**, 173–184.
- Wei XX, Wang XQ and Hong DY** (2003) Marked intragenomic heterogeneity and geographical differentiation of nrDNA ITS in *Larix potaninii* (Pinaceae). *Journal of Molecular Evolution* **57**, 623–635.
- Wick RR, Schultz MB, Zobel J and Holt KE** (2015) Bandage: interactive visualisation of de novo genome assemblies. *Bioinformatics* **31**, 3350–3352.
- Won H and Renner SS** (2005) The internal transcribed spacer of nuclear ribosomal DNA in the gymnosperm Gnetum. *Molecular Phylogenetics and Evolution* **36**, 581–597.
- Xin YX, Yu WB, Eiadthong W, Cao ZY, Li QS, Yang ZX, Zhao WZ and Xin PY** (2023) Comparative analyses of 18 complete chloroplast genomes from eleven *Mangifera* species (Anacardiaceae): sequence characteristics and phylogenomics. *Horticulturae* **9**, 86.
- Xu WB, Xia BS and Zhang SJ** (2017) A new variety of *Neocinnamomum* H. Liou from Guangxi, China. *Guihaia* **37**, 855–858.
- Zhang SR and Jia YX** (2013) *Study on Plant Germplasm Resources for Remediation of Heavy Metal Contaminated Soil – A Case Study of*

- Western Sichuan Mining Area*. Beijing, China: Beijing Science Press. ISBN 978-7-03-038269-6.
- Zhang GQ, Chen GZ, Chen LJ, Zhai JW, Huang J, Wu XY, Li MH, Peng DH, Rao WH, Liu ZJ and Lan SR** (2021) Phylogenetic incongruence in *Cymbidium* orchids. *Plant Diversity* **43**, 452–461. eCollection 2021 Dec.
- Zheng XY, Cai DY, Yao LH and Teng Y** (2008) Non-concerted ITS evolution, early origin and phylogenetic utility of ITS pseudogenes in *Pyrus*. *Molecular Phylogenetics and Evolution* **48**, 892–903.
- Zong D** (2019) *Phylogeny Utility of Nuclear and Chloroplast Genomes of Populus* (M.Agr thesis). Southwest Forestry University, Yunnan, China. doi: [10.27416/d.cnki.gxnlc.2019.000006](https://doi.org/10.27416/d.cnki.gxnlc.2019.000006).