Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Characterization of the Mitochondrial Genome and Phylogenetic Analysis of Formica candida

Anqi Wang
Southwest Forestry University
Colin Perrot
L'Institut Agro Dijon
Bin Yang
Southwest Forestry University
Tao Yuan
Southwest Forestry University
Zhenghui Xu
Southwest Forestry University
Zhongping Xiong
xzping@swfu.edu.cn

Southwest Forestry University

Research Article

Keywords: Formica candida, mitochondrial genome, protein-coding genes, phylogenetic tree

Posted Date: June 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9877814/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Ants (Hymenoptera: Formicidae) represent one of the largest and most widely distributed families of terrestrial insects, holding significant value in ecological research. The mitochondrial genome has been widely used as a reliable genetic marker for species identification and phylogenetic analysis. *Formica candida* is one of the ant species known to inhabit the highest altitudes. In this study, we conducted the first systematic research on *Formica candida* in China. By extracting total DNA and performing shallow genome sequencing, assembly, and annotation, we obtained the complete mitochondrial genome sequence of *Formica* candida, which is 17,558 bp in length and has been deposited in GenBank under accession number OR251196. Simultaneously, a total of 31 complete and well-annotated mitochondrial genomes of Formicidae were retrieved from the NCBI database. Phylogenetic trees were constructed using Maximum Likelihood (ML) and Bayesian Inference (BI) methods. The results showed that the BI tree and ML tree based on the 13 mitochondrial protein-coding genes (PCGs) exhibited congruent topologies, consistently revealing the phylogenetic relationships within the subfamily Formicinae and the genus *Formica*. The major clades, including *Formica* and *Lasius* from China and adjacent regions, were clearly delineated. Our findings enrich the basic biological data for *Formica candida* and provide a valuable reference for the phylogeny of Formicidae and related research.

Introduction

Ants are the common name for insects belonging to the order Hymenoptera, family Formicidae, within the class Insecta of the phylum Arthropoda. They are the most widely distributed and numerous social insects on Earth. To date, 16 subfamilies, 342 genera, and 14,263 species of ants have been recorded worldwide (Bolton n.d.), In China, 1,333 species across 10 subfamilies and 109 genera have been recorded (Hölldobler and Wilson 1990; AntWiki n.d.). Ants play a significant role as ecological indicators, performing functions such as improving the physicochemical properties of soil, influencing soil microbial biomass, acting as pollinators, and dispersing seeds within ecosystems (Hölldobler and Wilson 1990).

Mitochondrial genomes are characterized by rapid evolutionary rates, low recombination frequencies, and strict maternal inheritance, and have been widely applied in research on insect taxonomic identification, population genetic structure analysis, phylogenetic reconstruction, and the analysis of rapidly radiating lineages (Cameron 2025). In recent years, with the rapid development of high-throughput sequencing technologies, mitochondrial genome research on various insect groups has continued to deepen (Li et al. 2025). For example, Zhang et al. (2022) first reported the complete mitochondrial genome of the Chinese red ant (*Formica sinae*) (Zhang et al. 2022). Through comparative genomic analysis, they found that no gene rearrangements had occurred within the genus *Formica* and further confirmed the monophyly of the various subfamilies within the subfamily Formicinae (Xiong et al. 2023). Xiong et al. (2023) identified novel tRNA gene rearrangements in poneroid ants and clarified the phylogenetic position of the subfamily Paraponerinae (Vidal et al. 2023); Vidal et al. (2023) reported the first complete mitochondrial genome of the genus Odontomachus (Chen et al. 2016). These studies

collectively demonstrate that mitochondrial genomes hold significant value for application in ant systematics research.

Formica candida belongs to the subfamily Formicinae of the family Formicidae within the order Hymenoptera and is one of the ant species known to inhabit the highest elevations (Guénard and Dunn 2012). Although previous studies have analyzed its morphological identification characteristics (Huang and Chen 2018; Zhigulskaya et al. 2022), its molecular characteristics and phylogenetic position remain unclear. This study investigates the molecular biological characteristics and phylogenetic position of *Formica candida* by analyzing the structure of its mitochondrial genome and constructing a phylogenetic tree of the Formicidae family, thereby providing a reference for the systematic classification and conservation and utilization of this species.

Materials and methods

1.1 Sample collection

All specimens of *Formica candida* used in this study are housed in the Southwest Forestry University Herbarium. The relevant specimens were collected by the university's ant research team in Huoge Village, Kongse Township, Daofu County, Sichuan Province, at coordinates N31°07'22.25", E100°57'39.70", at an elevation of 3,069 meters. Immediately after collection, the specimens were fixed and preserved in 75% ethanol on-site. Upon return to the laboratory, they were identified to species based on morphological characteristics, prepared as dried specimens, and uniformly stored in the Southwest Forestry University Herbarium.

1.2 DNA extraction from *Formica candida*

Approximately 25 mg of *Formica candida* tissue was homogenized in liquid nitrogen prior to genomic DNA extraction using a spin-column-based method. The powdered tissue was lysed in 180 µL of Buffer ACL supplemented with 20 µL of Proteinase K and incubated at 56°C until complete digestion. Because insect tissues generally require prolonged digestion, the incubation was maintained for 6–8 h. When necessary, RNase A treatment was performed after lysis to remove residual RNA.

Following digestion, 200 µL of Buffer CL was added to the lysate and mixed thoroughly by inversion. If precipitates were observed, the lysate was heated at 70°C for 10 min to improve dissolution. Subsequently, 200 µL of absolute ethanol was added to promote DNA binding to the silica membrane. The resulting mixture, including any translucent fibrous precipitate, was transferred to a spin column and left to stand for 2 min before centrifugation at 10,000 rpm for 1 min. The flow-through was discarded, and the column was washed sequentially with 500 µL of ethanol-containing CW1 Solution and 500 µL of ethanol-containing CW2 Solution, with each washing step followed by centrifugation for 30 s.

To remove residual wash buffer and ethanol, the column was centrifuged at 12,000 rpm for 2 min and then air-dried at room temperature for several minutes. Genomic DNA was eluted by adding 50–200 µL

of CE Buffer to the membrane, followed by a 3 min incubation and centrifugation at 12,000 rpm for 2 min. When a higher DNA yield was required, the elution step was repeated. The purified genomic DNA was subsequently used for library construction and mitochondrial genome sequencing, while remaining DNA was stored at − 20°C until further use.

This extraction procedure was based on a modified SDS lysis system and did not require phenol/chloroform extraction. The quality of the extracted DNA was assessed spectrophotometrically, with OD260/280 values generally ranging from

1.7 to 1.9. Particular attention was paid to complete tissue lysis, prior supplementation of CW1 and C

W2 Solutions with absolute ethanol according to the manufacturer's instructions, and sufficient air-drying of the spin column before elution, as residual ethanol may inhibit downstream molecular reactions (Vidal et al. 2023; Liu et al. 2024).

1.3 Assembly and annotation of mitochondrial genes

The clean data were assembled and functionally annotated using MitoZ v2.4-alpha (Meng et al. 2019), yielding the complete mitochondrial genome sequence of *Formica candida*. The circularized sequence was uploaded to MITOS2 for re-annotation. The annotation results from both software tools were combined and cross-validated with GenBank BLASTN searches, followed by manual verification and curation (Bernt et al. 2013). The finalized sequence and annotation were subsequently submitted to NCBI GenBank to obtain an accession number (National Center for Biotechnology Information n.d.). Transfer RNA (tRNA) genes and their secondary structures were predicted using MITOS2 and RNAfold. The mitochondrial genome circular map was generated with Proksee v.6.0.2 (Grant et al. 2023). Base composition statistics were computed using PhyloSuite v2, through which the amino acid usage frequency and relative synonymous codon usage (RSCU) of the 13 protein-coding genes were analyzed (Zhao et al. 2025), and the AT-skew = (A − T)/(A + T) and GC-skew = (G − C)/(G + C) were calculated (Perna and Kocher 1995).

1.4 Construction methods of the phylogenetic trees

For the first tree, on the Formicinae subfamily, Mitochondrial genome sequences of Formicinae species were retrieved and downloaded from the NCBI database (National Center for Biotechnology Information n.d.), and the resulting data were compiled and organized. Amplification and alignment were performed on 23 Formicidae species, including *Formica candida*, and phylogenetic trees were constructed using both maximum likelihood (ML) and Bayesian (BI) methods. The selection criteria for the first tree were as follows: only taxonomic units distributed in China and neighboring regions were retained; a secondary screening based on sequence quality was conducted to exclude entries with incomplete genomes, sequence lengths shorter than 1,500 bp, or low reliability. Ultimately, *Stigmatomma silvestrii* was selected as an outgroup for the phylogenetic analysis.

For the second phylogenetic tree, which focuses on the genus *Formica* and likewise comprises both an ML and a BI tree, the selection criteria were similar to those applied for the first tree but were slightly relaxed: all genomes of sufficient length likely to contain the 13 PCGs (Protein Coding genes) were retained, and certain genomes designated as "Unverified" or "TPA_asm" were also included. *Lasius spathepus* was selected as the outgroup.

The processing of the data retrieved from NCBI was carried out using the Galaxy platform, employing in particular the MITOS2 tool (The Galaxy Community 2024; National Center for Biotechnology Information n.d.). Both trees were constructed using PhyloSuite v2 (Zhao et al. 2025) with the following settings: for the ML trees, 5,000 bootstrap replicates were performed under the GTR + GAMMA model, and for the BI trees, 5,000,000 generations were run, also under the GTR model. The resulting trees were subsequently visualized and exported using iTOL (Letunic and Bork 2024). Finally, the trees were formatted and finalized using Canva (Canva n.d.).

See Table 1 for details on all taxa and mitochondrial genome sequence information used in this study.

Table 1
Taxa and mitochondrial genome sequences used in the phylogenetic trees analyses

Serial number	Species	Length(bp)	GenBank Accession number
1	*Formica fusca*	16,673 bp	NC_026132.1
2	*Formica lemani*	15,970 bp	PV453803.1
3	*Formica rufa*	21,861 bp	OZ194008.1
4	*Formica candida*	17,558 bp	OR251196.1
5	*Formica sinae*	17,432 bp	NC_060873.1
6	*Formica glauca*	16,492 bp	NC_069645.1
7	*Formica aserva*	19,135 bp	CM073401.1
8	*Formica neogagates*	17,431 bp	BK012298
9	*Formica cunicularia*	18,081 bp	ON620085
10	*Formica sp. DM658*	14,348 bp	MT941916
11	*Formica rufibarbis*	12,899 bp	MT862421
12	*Formica moki*	16,932 bp	BK012606
13	*Formica podzolica*	14,190 bp	BK012706
14	*Formica selysi*	16,752 bp	NC_026711
15	*Cataglyphis aenescens*	17,197 bp	NC_060872.1
16	*Anoplolepis gracilipes*	16,943 bp	NC_039576.1
17	*Acropyga sauteri*	16,976 bp	NC_046398.1
18	*Lepisiota frauenfeldi*	17,090 bp	NC_061556.1
19	*Nylanderia flavipes*	16,687 bp	NC_049861.1
20	*Lasius flavus*	21,365 bp	OZ287120.1
21	*Lasius spathepus*	18,951 bp	NC_053901.1
22	*Lasius fuliginosus*	18,749 bp	OX424457.1
23	*Lasius platythorax*	20,726 bp	OZ020550.1
24	*Lasius brunneus*	18,027 bp	OZ179867.1
25	*Colobopsis nipponica*	17,431 bp	NC_053900.1
26	*Polyrhachis lamellidens*	16,335 bp	AP036026.1
27	*Polyrhachis dives*	17,117 bp	PV402686.1

Serial number	Species	Length(bp)	GenBank Accession number
28	*Camponotus friedae*	18,808 bp	CM078454.1
29	*Camponotus atrox*	16,540 bp	KT159775.1
30	*Camponotus japonicus*	16,422 bp	NC_061037.1
31	*Stigmatomma silvestrii*	18,861 bp	MT215092.1

Results and analysis

2.1 Characteristics of the mitochondrial genome of *Formica candida*

The mitochondrial genome of *Formica candida* is a typical double-stranded closed circular molecule with a total length of 17,558 bp. It contains 37 genes and one control region. There are two overlapping regions between genes outside the mitochondrial control region: the tRNA-Lys gene overlap region is 1 bp long, and the atp8 gene overlap region is 7 bp long. The total length of the intergenic regions outside the mitochondrial control region is 2,418 bp, distributed among 28 genes, with lengths ranging from 2 to 1,528 bp. Among these, the longest intergenic region is located between the 12S rRNA and tRNA-Met genes (Fig. 1).

2.2 Analysis of Protein Coding Genes (PCG)

Among all protein-coding genes (PCGs) of *Formica candida*, the majority utilize ATG as the start codon, with the exception of nad2, which employs TTG. The predominant stop codon is TAA, while certain genes use a truncated single T or TAG. Collectively, all PCGs encode a total of 3,705 amino acids. The five most frequently encoded amino acids are Leu (559 residues), Ile (488 residues), Phe (403 residues), Ser (361 residues), and Met (333 residues), accounting for 15.09%, 13.17%, 10.88%, 9.74%, and 8.99% of the total amino acid composition, respectively. The most commonly used codons are UUA (Leu), AUU (Ile), UUU (Phe), AUA (Met), and AAU (Asn), indicating a high AT content in the PCG sequences (Fig. 2).

2.3 Analysis of Mitochondrial tRNA Genes of *Formica candida*

The secondary structures of the 22 tRNA genes from Formica candida were predicted using the online tool RNAfold (University of Vienna and Institute for Theoretical Chemistry n.d.), and the results are shown in Fig. 3. The sequence lengths of the 22 tRNA genes in the mitochondrial genome of Formica candida range from 52 to 76 bp. The tRNA-Ala gene is the shortest at 52 bp, while the tRNA-Asp gene is the longest at 76 bp. Among these, 14 tRNA genes are located on the heavy strand, while 8 tRNA genes and 2 rRNA genes are located on the light strand.

2.4 Control regions and repetitive regions of the mitochondrial genome

The control region of the mitochondrial genome of *Formica candida* is located between the rrn12 and trnM genes, with a total length of 1,528 bp. This region is predominantly non-coding and is one of the most highly variable regions in the mitochondrial genome. Two distinct types of repeat units were identified within the control region, exhibiting a certain degree of structural regularity (Fig. 4): Repeat unit 1 is located in the proximal region of the control region and consists of two repeat fragments, distributed in the 33–199 bp and 200–366 bp intervals, respectively; Repeat unit 2 is located in the posterior region of the control region and similarly comprises two repeat fragments, situated in the 1080–1170 bp and 1235–1325 bp intervals, respectively.

2.5 Phylogenetic relationship of Formica candida

Figure 5 shows the phylogenetic relationships within the subfamily Formicinae inferred from the concatenated dataset of 13 mitochondrial protein-coding genes (PCGs). The Bayesian inference (BI) tree is presented on the left, whereas the maximum likelihood (ML) tree is shown on the right. Overall, the two analytical approaches recovered highly congruent topologies, indicating a broadly consistent phylogenetic signal. Using *Stigmatomma silvestrii* as the outgroup, the sampled genera were generally resolved as distinct and well-supported monophyletic clades. Within the *Formica* clade, *Formica candida*, *Formica rufa*, and *Formica glauca* formed a strongly supported monophyletic group in both the ML and BI analyses. The remaining major clades, corresponding to the genera *Lasius*, *Camponotus*, and *Polyrhachis*, also showed high levels of support, with bootstrap values and Bayesian posterior probabilities approaching 100. Nevertheless, some internal nodes within *Lasius* received comparatively lower support, with a bootstrap value of 68. In addition, deeper relationships among *Formica*, *Lasius*, *Nylanderia*, *Anoplolepis*, *Lepisiota*, *Acropyga*, and *Cataglyphis* were supported by relatively moderate values, notably 65 in the ML analysis. These results suggest that the deeper internodal relationships should be interpreted with caution.

Figure 6 presents the phylogenetic relationships within the genus *Formica* based on ML and BI analyses, with *Lasius spathepus* used as the outgroup. The results confirm that *Formica candida* is nested within the genus *Formica* and is closely associated with a clade including *Formica rufa*, *Formica aserva*, *Formica podzolica*, *Formica lemani*, *Formica selysi*, *Formica fusca*, and *Formica moki*. Overall, two main clades were recovered within *Formica*. The first comprised *Formica cunicularia*, *Formica sinae*, *Formica glauca*, *Formica* sp. DM658, and *Formica rufibarbis*, whereas the second included *Formica candida* and the closely related taxa mentioned above. However, several nodes in the ML tree showed limited support. Within the clade containing *Formica candida*, bootstrap values of 57, 74, and 49 were observed, while the support value separating the two major *Formica* clades was only 32. These relatively low support values indicate that the corresponding relationships should be interpreted cautiously. This is particularly important because some sequences included in the second analysis were of slightly lower quality;

although they were incorporated to increase taxon sampling and broaden the comparative framework, they may have contributed to the reduced nodal support observed in parts of the tree.

Discussion

The mitochondrial genome of Formica candida is 17,558 bp in size and contains 37 genes plus one control region, which is largely consistent with other Formicidae species (Li et al. 2023). A total of 28 intergenic spacers are detected (total length 2,418 bp), among which the space between 12S-rRNA and tRNA-Met reaches 1,528 bp, a length significantly greater than that observed in other ant species. This may result from historical gene rearrangements or the insertion of repetitive sequences (Zhang et al. 2022); similar phenomena have been reported in ponerine ants. The overlap regions involving tRNA-Lys and atp8 (1 bp and 7 bp, respectively) are consistent with those found in other Formica species, suggesting that these overlapping features are relatively conserved (Sun et al. 2009).

Among the protein-coding genes (PCGs), all initiate with ATG except for *nad2*, which uses TTG as the start codon. The predominant stop codon is TAA, while a few genes employ a truncated single T or TAG, a pattern consistent with other *Formica* species (Huang and Chen 2018). The high AT content of the PCG sequences (approximately 81%–84%) accounts for the observed codon usage bias (e.g., UUA, AUU), and all PCGs are under purifying selection, reflecting their functional constraints in oxidative phosphorylation (Lorenz et al. 2011). The non-canonical start codon of *nad2* may be associated with the regulation of translation efficiency (Lorenz et al. 2011).

The secondary structures of the 22 tRNA genes of *Formica candida* were predicted using RNAfold, and all tRNAs are capable of folding into the typical cloverleaf structure (Lorenz et al. 2011). Their lengths range from 52 to 76 bp; 14 tRNAs are located on the heavy strand, while 8 tRNAs and both rRNAs reside on the light strand, a configuration consistent with other *Formica* species (Lin and Song 2023). The comparative analysis of 12 *Formica* species conducted by Liu et al. (2024) also confirms that tRNA secondary structure is highly conserved within the genus *Formica* (Lin and Song 2023). However, certain ponerine species (e.g., *Pachycondyla*) exhibit the loss of the DHU or T arm, whereas the tRNA structures of *F. candida* are complete, suggesting that the subfamily Formicinae displays greater structural conservation in tRNAs than Ponerinae, a difference that may be attributable to divergent selective pressures between the two subfamilies.

The control region (1,528 bp), situated between rrn12 and trnM, represents one of the most variable regions of the mitochondrial genome (Cameron 2025), with length variation primarily driven by the insertion or deletion of repetitive sequences (Liu et al. 2016). The control region of *F. candida* is relatively long compared to other Formicidae species. It is hypothesized that this elongated control region and its internal repeat units constitute a molecular manifestation of adaptive structural diversity in the mitochondrial genome.

The Bayesian inference (BI) and Maximum Likelihood (ML) trees reconstructed from the 13 mitochondrial protein-coding genes (PCGs) exhibit highly congruent topologies, indicating that this

dataset reliably recovers the phylogenetic relationships within the subfamily Formicinae and the genus *Formica* (Borowiec et al. 2019). Major lineages—including the genera *Lasius*, *Camponotus*, *Polyrhachis*, and *Formica*—all form clearly delineated clades, with high support values at most nodes, demonstrating that the 13 mitochondrial PCGs can effectively discriminate the principal formicine lineages present in the region and therefore hold both systematic and regional biogeographical significance. *Formica* is unequivocally nested within the formicine clade, consistent with traditional taxonomic placement, confirming that the genus *Formica* shares well-defined phylogenetic affinities with other formicine genera (Ward 2014). A finer-scale phylogenetic analysis focusing on the internal relationships of *Formica* clarifies the phylogenetic position of *Formica candida* within the genus (Zhang et al. 2022). The results indicate that *F. candida* is closely related to a clade comprising *Formica rufa*, *Formica aserva*, *Formica podzolica*, *Formica lemani*, *Formica selysi*, and *Formica fusca*, pointing to a close matrilineal relationship with certain wood ant taxa or closely allied lineages (Chen et al. 2025). Although the complete mitochondrial genome provides a robust marker for establishing the generic placement of *F. candida* (Zhang et al. 2025), several internal nodes within the *Formica* tree receive only moderate bootstrap support, underscoring the need to integrate whole-genome sequencing and denser taxon sampling to resolve the deeper phylogenetic relationships within this ecologically important group.

The present study represents the first systematic mitogenomic investigation of *Formica candida* in China, yielding its complete mitochondrial genome (GenBank accession number OR251196). The structural features of the genome—including gene content, the distribution of overlap and intergenic regions, codon usage bias, tRNA secondary structures, and repetitive sequence units within the control region—are thoroughly characterized. Furthermore, Bayesian and Maximum Likelihood trees were reconstructed based on the 13 PCGs to assess the phylogenetic position of *F. candida* within the subfamily Formicinae and the genus *Formica*.

Declarations

Competing interests

The authors declare that they have no competing interests.

LLM using

Deepseek was used in some sections to correct spelling and stylistic errors and to make this document more enjoyable to read.

Funding

This study was supported by the National Natural Science Foundation Project of China (31860615:31860166:32060422)

Author Contribution

Wang Anqi and Colin Perrot were responsible for the writing of the manuscript. Yang Bin and Yuan Tao were responsible for the experimental work. Xu Zhenghui and Xiong Zhongping were responsible for the conception of the study, provision of experimental instruments, and overall supervision. All authors read and approved the final manuscript.

Data Availability

The sequenced mitochondrial DNA of Formica candida were deposited into the NCBI database under accession number : OR251196.1, and can be viewed at the following address : https://www.ncbi.nlm.nih.gov/nuccore/OR251196.1.

References

1. AntWiki (n.d.) AntWiki. In: AntWiki (ed) https://www.antwiki.org/wiki. Accessed 13 May 2026
2. Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsch G, Pütz J, Middendorf M, Stadler PF (2013) MITOS: improved de novo metazoan mitochondrial genome annotation. Mol Phylogenet Evol 69(2):313–319. https://doi.org/10.1016/j.ympev.2012.08.023
3. Bolton B (n.d.), AntCat (eds) https://www.antcat.org/. Accessed 13 May 2026
4. Borowiec ML, Rabeling C, Brady SG, Fisher BL, Schultz TR, Ward PS (2019) Compositional heterogeneity and outgroup choice influence the internal phylogeny of the ants. Mol Phylogenet Evol 134:111–121. https://doi.org/10.1016/j.ympev.2019.01.024
5. Cameron SL (2025) Insect Mitochondrial Genomics: A Decade of Progress. Annu Rev Entomol 70(1):83–101. https://doi.org/10.1146/annurev-ento-013024-015553
6. Canva (n.d.) Canva. In: Canva. https://www.canva.com/. Accessed 14 May 2026
7. Chen Q, Deng M, Dai X, Wang W (2025) Phylogenomic data exploration with increased sampling provides new insights into the higher-level relationships of butterflies and moths (Lepidoptera). J Syst Evol. https://doi.org/10.1111/jse.13180
8. Chen Z, Zhou S, Huang J (2016) Seven species new to science and one newly recorded species of the ant genus Myrmica Latreille, 1804 from China, with proposal of a new synonym (Hymenoptera, Formicidae). ZooKeys (551):85–128. https://doi.org/10.3897/zookeys.551.6005
9. Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C, Graham M, Van Domselaar G, Stothard P (2023) Proksee: in-depth characterization and visualization of bacterial genomes. Nucleic Acids Res 51(W1):W484–W492. https://doi.org/10.1093/nar/gkad326
10. Guénard B, Dunn RR (2012) A checklist of the ants of China. Zootaxa 3558(1). https://doi.org/10.11646/zootaxa.3558.1.1
11. Hölldobler B, Wilson EO (1990) The Ants, Harvard

12. Huang H, Chen Y (2018) [Morphometric analysis of Formica candida and Formica picea (Hymenoptera: Formicidae)]. J Guangxi Norm Univ 33(2):96–102. https://doi.org/10.16088/j.issn.1001-6600.2015.02.015
13. Letunic I, Bork P (2024) Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res 52(W1):W78–W82. https://doi.org/10.1093/nar/gkae268
14. Li F, Wang X, Zhou X (2025) The Genomics Revolution Drives a New Era in Entomology. Annu Rev Entomol 70(Volume 70, 2025):379–400. https://doi.org/10.1146/annurev-ento-013024-013420
15. Li T, Xu Z-H, Liu X, Xiong Z-P, Zhang X-M (2023) Characterization and phylogenetic implication of complete mitochondrial genome of the Cataglyphis aenescens (Hymenoptera: Formicidae): Genomic comparisons in Formicinae. Entomol Res 53(2):45–54. https://doi.org/10.1111/1748-5967.12631
16. Lin X, Song N (2023) The First Complete Mitochondrial Genome of the Genus Pachycondyla (Formicidae, Ponerinae) and Insights into the Phylogeny of Ants. Genes 14(8):1528. https://doi.org/10.3390/genes14081528
17. Liu M, Hu S-Y, Li M, Sun H, Yuan M-L (2024) Comparative mitogenomic analysis provides evolutionary insights into Formica (Hymenoptera: Formicidae). PLoS ONE 19(6):e0302371. https://doi.org/10.1371/journal.pone.0302371
18. Liu N, Duan X-Y, Qian Z-Q, Wang X-Y, Li X-L, Ding M-Y (2016) Characterization of the complete mitochondrial genome of the myrmicine ant Vollenhovia emeryi (Insecta: Hymenoptera: Formicidae). Conserv Genet Resour 8(3):211–214. https://doi.org/10.1007/s12686-016-0535-x
19. Lorenz R, Bernhart SH, Höner zu Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL (2011) ViennaRNA Package 2.0. Algorithms Mol Biol 6(1):26. https://doi.org/10.1186/1748-7188-6-26
20. Meng G, Li Y, Yang C, Liu S (2019) MitoZ: a toolkit for animal mitochondrial genome assembly, annotation and visualization. Nucleic Acids Res 47(11):e63. https://doi.org/10.1093/nar/gkz173
21. National Center for Biotechnology Information (n d.) National Center for Biotechnology Information. In: NCBI. https://www.ncbi.nlm.nih.gov/. Accessed 14 May 2026
22. Perna NT, Kocher TD (1995) Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol 41(3):353–358. https://doi.org/10.1007/BF00186547
23. Sun Z, Wan D-G, Murphy RW, Ma L, Zhang X-S, Huang D-W (2009) Comparison of base composition and codon usage in insect mitochondrial genomes. Genes Genomics 31(1):65–71. https://doi.org/10.1007/BF03191139
24. The Galaxy Community (2024) The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. Nucleic Acids Res 52(W1):W83–W94. https://doi.org/10.1093/nar/gkae410
25. University of Vienna Institute for Theoretical Chemistry (n.d.) RNAfold web server. http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi. Accessed 13 May 2026

26. Vidal MPV-BT, Rodrigues-Oliveira IH, Silva IB, Rocha RR, Pasa R, Kavalco KF, Menegidio FB, Garcia C (2023) Big jaw, small genome: first description of the mitochondrial genome of Odontomachus (Formicidae, Ponerinae): evolutionary implications for Ponerinae ants. Mitochondrial DNA Part A 34(1–8):29–40. https://doi.org/10.1080/24701394.2024.2343477
27. Ward PS (2014) The Phylogeny and Evolution of Ants. Annu Rev Ecol Evol Syst 45(1):23–43. https://doi.org/10.1146/annurev-ecolsys-120213-091824
28. Xiong Z, He D, Guang X, Li Q (2023) Novel tRNA Gene Rearrangements in the Mitochondrial Genomes of Poneroid Ants and Phylogenetic Implication of Paraponerinae (Hymenoptera: Formicidae). Life 13(10):2068. https://doi.org/10.3390/life13102068
29. Zhang X-M, Li T, Liu X, Xu Z-H (2022) Characterization and Phylogenetic Implication of Complete Mitochondrial Genome of the Medicinal Ant Formica sinae (Hymenoptera: Formicidae): Genomic Comparisons in Formicidae. J Med Entomol 59(6):1971–1979. https://doi.org/10.1093/jme/tjac128
30. Zhang YM, Bossert S, Spasojevic T (2025) Evolving perspectives in Hymenoptera systematics: Bridging fossils and genomes across time. Syst Entomol 50(1):1–31. https://doi.org/10.1111/syen.12645
31. Zhao D, Ye T, Gao F, Jakovlić I, La Q, Tong Y, Liu X, Song R, Liu F, Lian Z, Zou H, Li W-X, Wang G-T, Zeng B, Zhang D (2025) PhyloSuite v2: The development of an all-in-one, efficient and visualization-oriented suite for molecular dating analysis and other advanced features. iMeta 4(6):e70095. https://doi.org/10.1002/imt2.70095
32. Zhigulskaya ZA, Shekhovtsov SV, Poluboyarova TV, Berman DI (2022) Formica picea and F. candida (Hymenoptera: Formicidae): Synonyms or Two Species? Diversity 14(8):613. https://doi.org/10.3390/d14080613

Figures

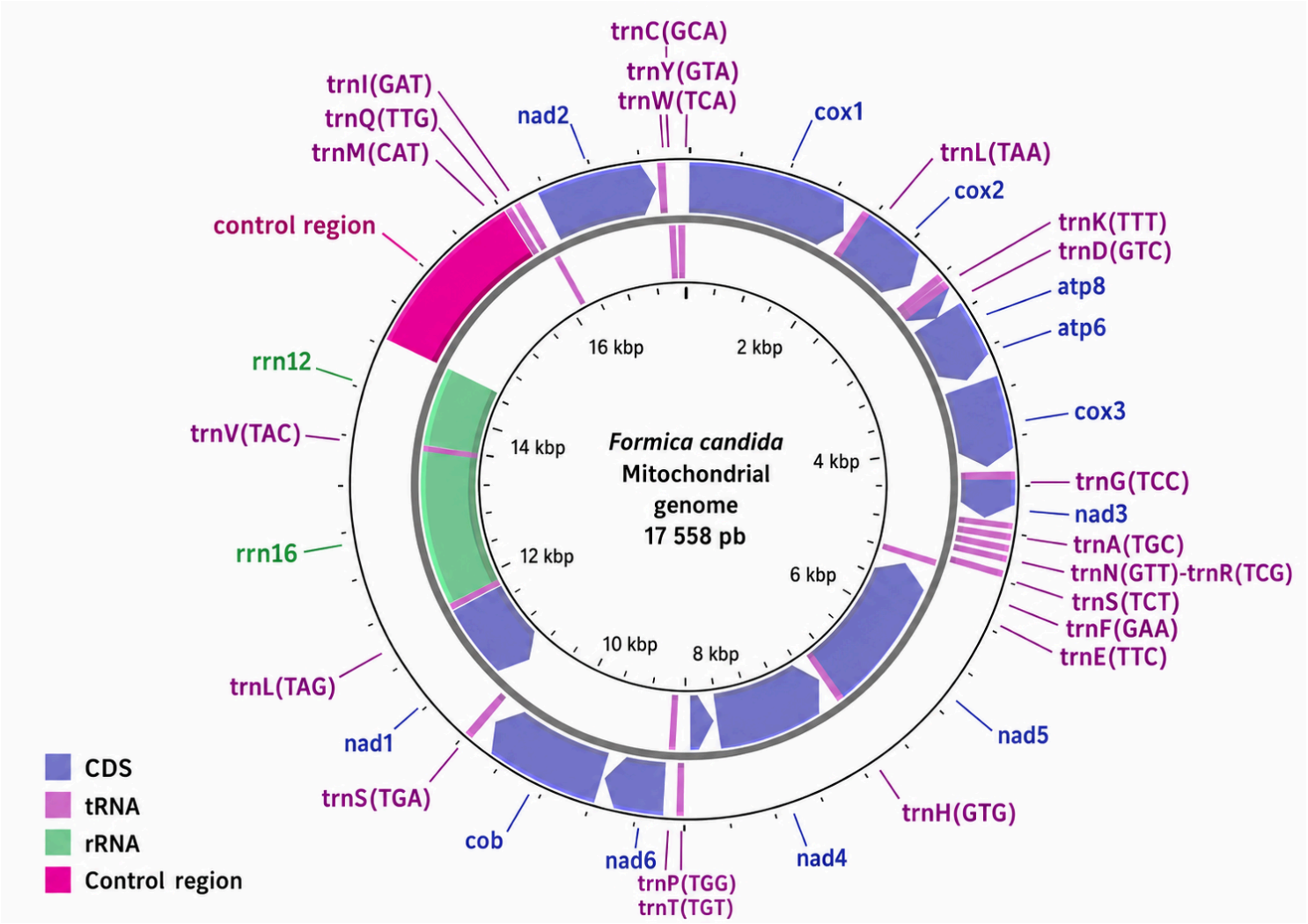


Figure 1

Circular structure of mitochondrial genome of *Formica candida*

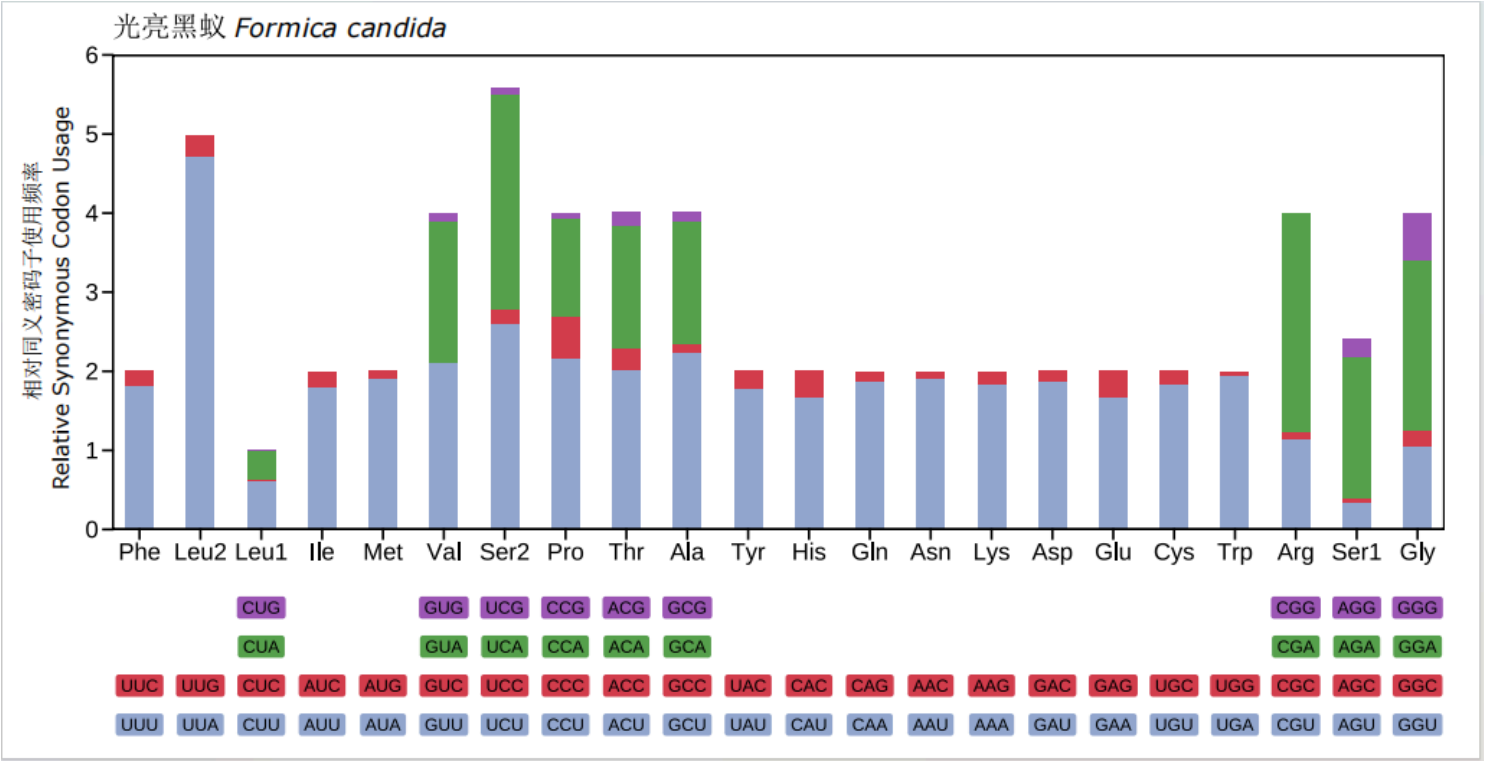


Figure 2

Relative synonymous codon usage of the mitochondrial genome of *Formica candida*

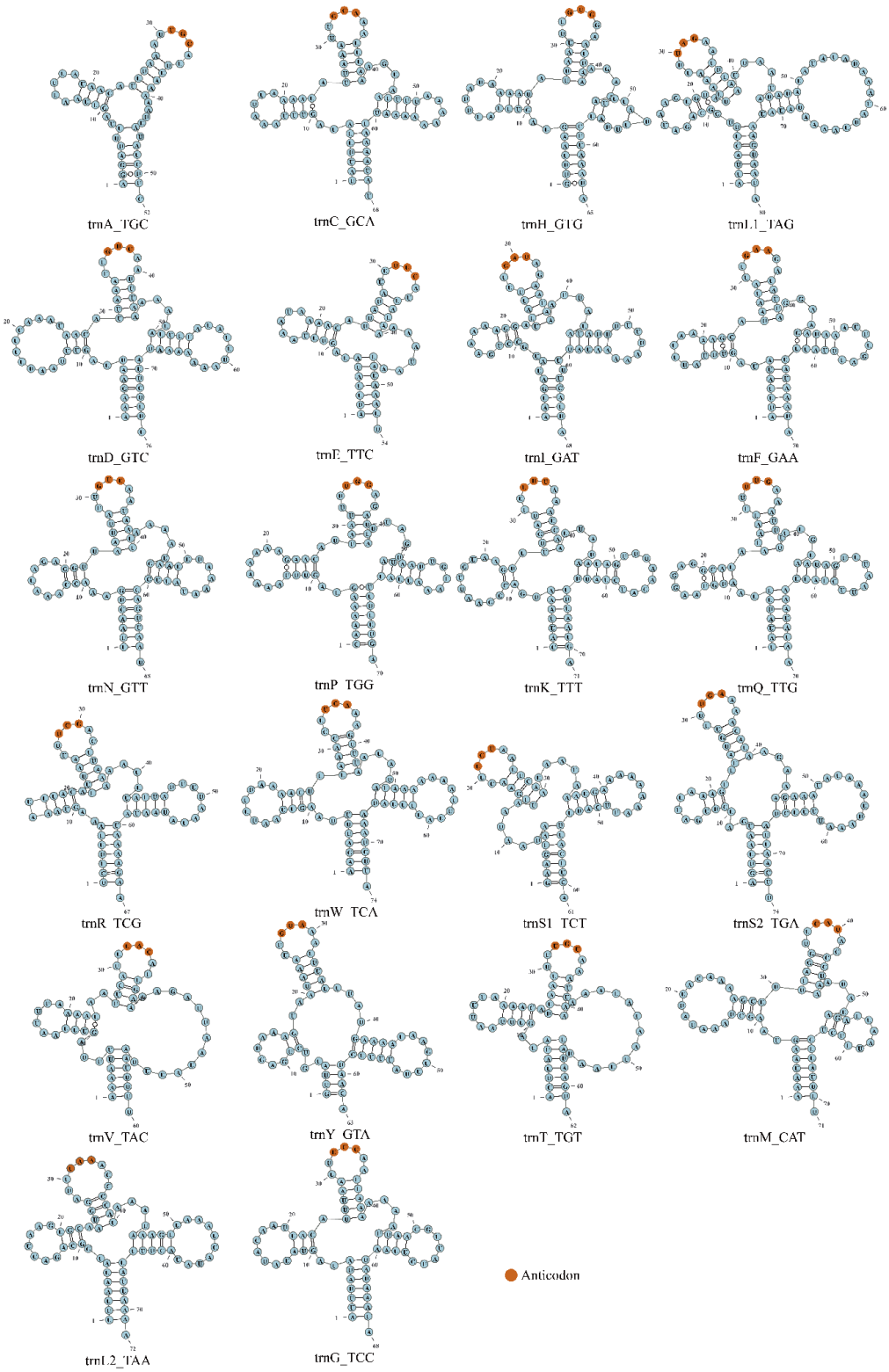


Figure 3

Predicted cloverleaf secondary structures for tRNAs genes of *Formica candida*

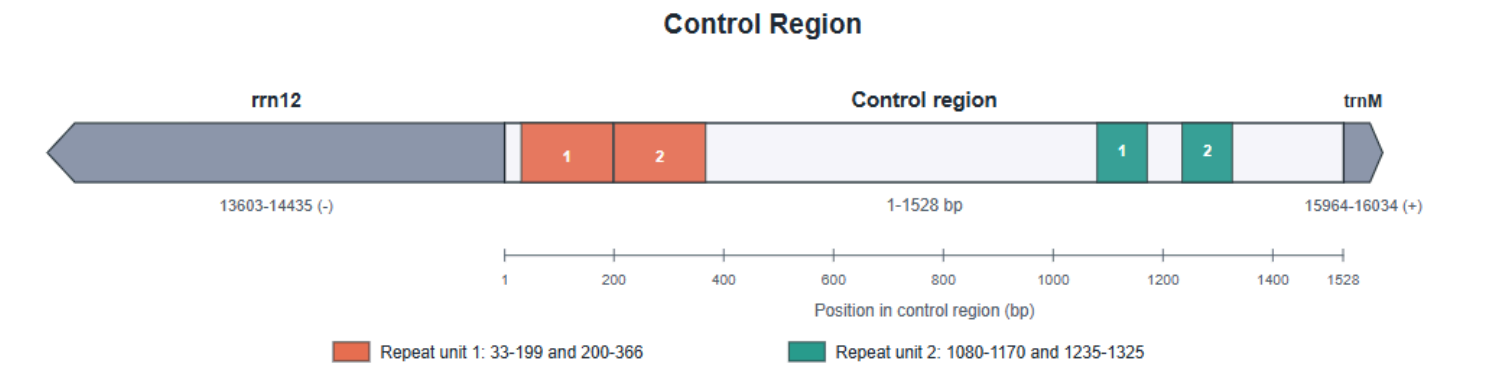


Figure 4

Structure of the mitochondrial control region and repeat regions of *Formica candida*

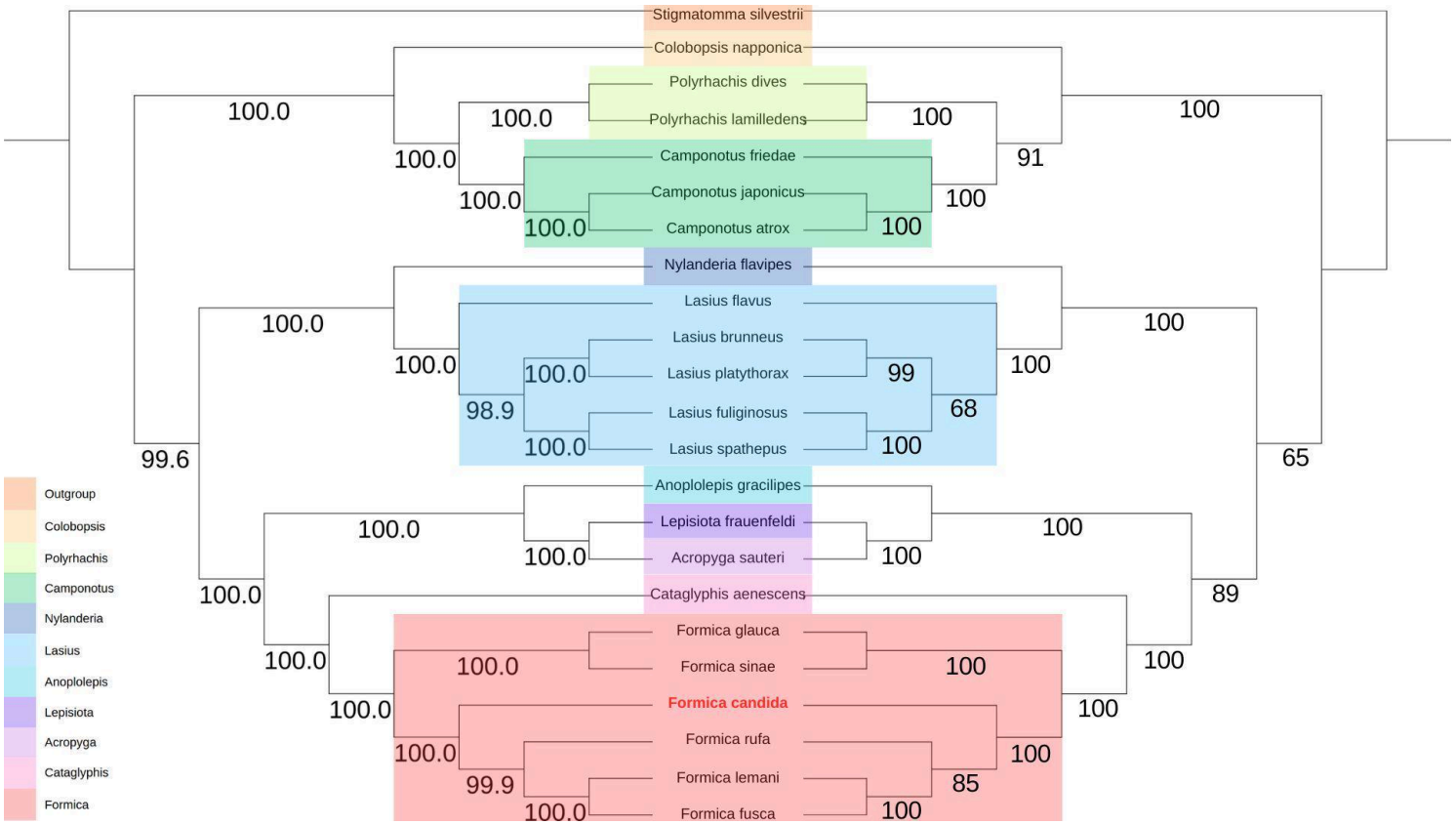


Figure 5

Bayesian inference tree (left) and maximum-likelihood tree (right) of the Formicinae subfamily based on the concatenated nucleotide sequences of 13 mitochondrial protein-coding genes

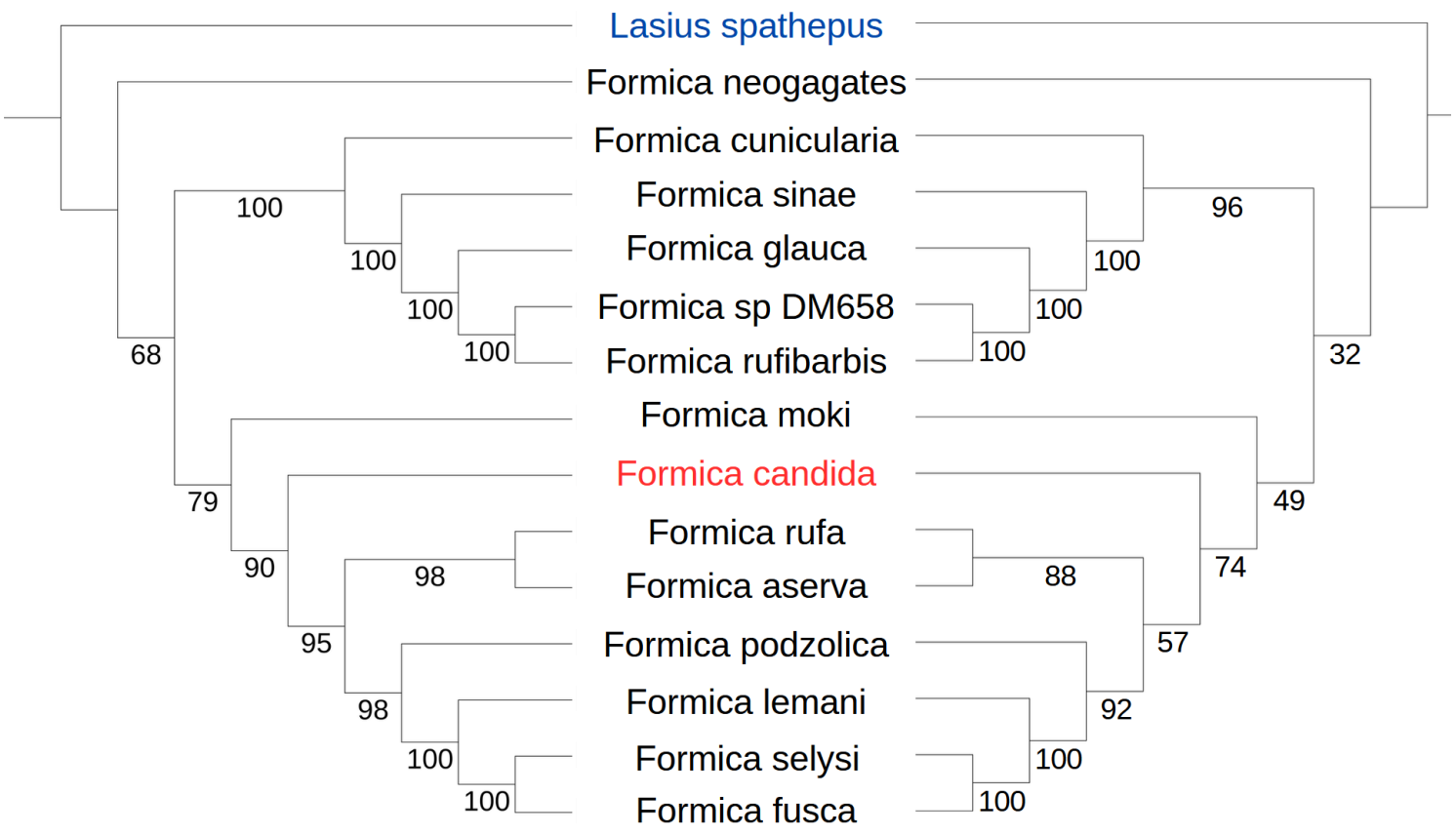


Figure 6

Bayesian inference tree (left) and maximum-likelihood tree (right) of the *Formica* genus based on the concatenated nucleotide sequences of 13 mitochondrial protein-coding genes