Cumhuriyet Sci. J., 43(4) (2022) 577-583
DOI: https://doi.org/10.17776/csj.1195087

Cumhuriyet Science Journal

| csj.cumhuriyet.edu.tr | Founded: 2002 ISSN: 2587-2680 e-ISSN: 2587-246X Publisher: Sivas Cumhuriyet University

The Complete Mitogenome of Redheaded Pine Sawfly, *Neodiprion lecontei* (Hymenoptera: Diprionidae): Duplication of *trnR* Gene and Rearrangement in the ARNS1EF Gene Cluster

Ertan Mahir Korkmaz [1,a,*]

[1] *Department of Molecular Biology and Genetics, Faculty of Science, Sivas Cumhuriyet University, Sivas, 58140, Türkiye*

**Corresponding author*

Research Article

History
Received: 26/10/2022
Accepted: 20/12/2022

ABSTRACT

Neodiprion is a genus belonging to the small sawfly family Diprionidae, feeding the plant family Pinaceae entirely. Here, the complete mitogenome of the redheaded pine sawfly *Neodiprion lecontei* (Hymenoptera: Diprionidae) was assembled, annotated as third party annotation from the raw genome dataset of *N. lecontei* and comparatively characterised. The length of *N. lecontei* mitogenome was 16,067 bp in size, with an AT content of 81.32%. The initiation codons of protein coding genes (PCGs) are ATN (except for *nad6* (TTA-Phe), while termination codons are TAA or T–. tRNA genes favoured usual anticodons except for *trnS1* which preferred an unusual anticodon GCU. Compared with the *Neodiprion sertifer* mitogenome, the ARNS1EF gene cluster was rearranged as RAS1RNEF and *trnR* gene has a duplicated copy, revealing a new event not formerly reported in Symphyta. The phylogeny confirms the position of *N. lecontei* within the family of Diprionidae and supports the monophyly of included genera and families in Tenthredinoidea.

Keywords: Mitochondrial genome, Rearrangement, Sawflies; Tenthredinoidea, Symphyta

[a] *ekorkmaz@cumhuriyet.edu.tr* *https://orcid.org/0000-0003-0699-1354*

Introduction

Mitochondria have a central role in the production of metabolic energy in nearly all living eukaryotic organisms [1]. In addition to its vital functional importance involved in maintaining an accurate energy balance and cellular lifecycle, this organelle has also been extensively used in terms of mitogenome information to investigate genome features and to infer evolutionary relationships from populations to species or higher level of taxa [2–4]. A typical insect mitogenome consists of 14–25 kb with a quite conserved gene content, containing 22 transfer RNAs (tRNAs), 13 protein coding genes (PCGs), two ribosomal RNAs (rRNAs) and one large control region (A + T-rich region) [2,5]. In the last decade, revolutionary advances in next-generation sequencing technology and bioinformatics have also increased the number of insect mitogenomes. In the last release of organelle section of database of NCBI (September 2022) using the "Insecta, mitochondrion" as keywords and filtering the sequence length >10,000 bp, there are complete or nearly complete mitogenomes of more than 9,000 insect species. These contain the mitogenomes from only 978 hymenopteran species, one of the "big four (Coleoptera, Hymenoptera, Diptera and Lepidoptera)" of insect orders including over 150,000 species with remarkable different life strategies [6,7].

The suborder Symphyta (also known as sawflies) is the paraphyletic lineage of Hymenoptera with eight extant phytophagous superfamilies and more than 8900 extant described species [8]. The great majority of this suborder are considered as pest in agriculture and forestry, largely due to their plant-eating lifestyles during larval stage, however, to date, complete or nearly complete mitogenomes of 88 symphytan species have been reported (NCBI, September 2022), with only approximately 9 % of the sequenced hymenopteran mitogenomes. Due to the limited available mitogenome data of sawflies, gene rearrangements are considered to be relatively conserved, but substitution rate is high [9–11], indicating the necessity of more representative mitogenome from Symphyta to infer mitogenome architecture and features.

Here, the complete mitogenome of the redheaded pine sawfly *Neodiprion lecontei* (Fitch, 1859) (Hymenoptera: Diprionidae) was assembled and annotated for the first time. This pest species feeds multiple pine (*Pinus*) tree species throughout its native range in North America [12]. So far, only two mitogenomes from *Neodiprion sertifer* and *N. fabricii*, have been reported for Diprionidae [3]. The mitogenome of *N. lecontei* was also compared with the previously reported mitogenomes of *Neodiprion* for investigating of the mitogenome architectures and features of the Diprionidae.

Materials and Methods

Mitogenome Assembly, Annotation and Analyses

The raw sequencing data of *N. lecontei* was downloaded from the NCBI Sequence Read Archive (SRA) database under the SRA accession numbers of SRR1955932, SRR1956520 and SRR1956730. Quality control steps were performed to get clean reads from the raw sequencing datasets. The adapter sequences, low quality and possible contaminated reads were removed from raw reads by using Fastp v0.20.0 [13] and Lighter v1.0.7 [14]. The obtained clean reads from three datasets were merged into a single fastq file and then the reads were assembled into contigs using both a reference assembly using the mitogenome of *N. sertifer* (MK994526, [3]) from the same genus under the 'iterate up to five times' and 'medium–low sensitivity' parameters in Geneious R9 [15] and *de novo* assembly using SPAdes v3.15.3 [16] in DOE Systems Biology Knowledgebase (KBase) platform [17]. The obtained *de novo* contigs were then mapped with the mitogenome produced under the first approach. The sets of selected assemblies generated by these approaches were finally aligned, compared as manual and gathered into a single contig.

The identification of tRNA genes were performed based on their accepted secondary structure and anticodon sequence by MITOS web server [18] with the invertebrate genetic code option and ARWEN v1.2 [19] with the default search options and the genetic code as mito/chloroplast. The boundaries and locations of rRNA genes and PCGs were manually designated comparing with the known symphytan homologous gene sequences using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html) and BLAST searches in GenBank. The boundaries of the rRNA genes were predicted from location of the adjacent tRNA genes or comparison with homologous symphytan rRNA genes. Overlapping regions and intergenic spacers between genes were estimated manually. The complete mitogenome was graphically mapped by Geneious R9 [15]. Finally, the mitogenome sequence of the redheaded pine sawfly *N. lecontei* was submitted to GenBank under the accession number of TPA: BK062819. Basic statistics on this third party annotated mitogenome nucleotide composition were calculated by MEGA v6.0 [20]. The formulae: AT-skew=[A−T] / [A+T] and GC-skew=[G−C] / [G+C] [21] were used to calculate the base compositional differences between different strands, degenerated codon positions and the genes coded on the alternative strands. The relative synonymous codon usage (RSCU) was also computed for all protein-coding genes by MEGA v6.0.

Table 1. List of sawfly mitogenomes used in phylogenetic analyses.

Family	Species	Accession Number	Family	Species	Accession Number
Diprionidae	*Neodiprion lecontei*		Tenthredinidae	*Allantus luctifer*	KJ713152
	Neodiprion sertifer	MK994526		*Allantus togatus*	MW464859
	Nesodiprion biremis	ON964465		*Analcellicampa xanthosoma*	MH992752
	Nesodiprion japonicus	ON964464		*Asiemphytus rufocephalus*	KR703582
Argidae	*Arge aurora*	MN913350		*Birmella discoidalisa*	MF197548
	Arge bella	MF287761		*Colochela zhongi*	MT702984
	Arge similis	MG923484		*Conaspidia wangi*	MW415019
Athaliidae	*Athalia birmanica*	ON840085		*Eutomostethus vegetus*	MT663219
	Athalia japonica	ON964466		*Hemathlophorus* sp.	MW632125
	Athalia proxima	MN527306		*Macrophya dolichogaster*	MW544890
Cimbicidae	*Cimbex luteus*	MW136447		*Metallus mai*	MW255941
	Corynis lateralis	KY063728		*Monocellicampa pruni*	JX566509
	Labriocimbex sinicus	MH136623		*Moricella rufonota*	MW487926
	Leptocimbex clavicornis	MT478109		*Neostromboceros nipponicus*	MW632127
	Leptocimbex praiaformis	MT478110		*Sinopoppia nigroflagella*	MW487927
	Leptocimbex yanniae	MT478111		*Siobla xizangensis*	MN562486
	Praia tianmunica	MT665975		*Strongylogaster xanthocera*	MW324676
	Trichiosoma anthracinum	KT921411		*Taxoblenus sinicus*	MW632126
	Trichiosoma vitellina	MN853777		*Taxonus zhangi*	MZ461490
Pergidae	*Perga condei*	AY787816		*Tenthredo tienmushana*	KR703581
Xyelidae	*Macroxyela ferruginea*	MK270536		*Xenapatidea procincta*	MW487928
	Xyela sp.	MG923517			

Phylogenetic analyses were carried out using the dataset of 41 species from Tenthredinoidea, representing six families and of two species from the family Xyelidae as outgroup (Table 1). Nucleotide sequences of each PCG were aligned individually under codon-based alignments using ClustalW as implemented in MEGA v6.0. The alignment of RNA genes was performed using MAFFT algorithm [22] as implemented in Geneious R9. The obtained alignments were concatenated with SequenceMatrix v1.7.8 [23]. The best-fitting partitioning scheme and model of each partition were selected by PartitionFinder v1.1.1 [24] using Bayesian information criterion (BIC) and the "greedy" algorithm based on branch lengths estimated as ''unlinked''. The data blocks were stated by codons and genes to create an input configuration partition file with 63 (with all codon positions) and 50 (without 3rd codon positions). All subsequent phylogenetic analyses were performed using the best partitioning schemes and related models (Table 2). The genetic saturation

levels of genes and different codon positions were measured by correlation test implemented in R core packages [25] comparing the distances measured by applying the best-fitting model evolution GTR + G + I proposed by PartitionFinder v1.1.1 with the uncorrected p-distances. The distance values were estimated with PAUP v4.0b10 [26]. The phylogenetic analyses were performed with the dataset of nine PCGs with the first two codon positions plus two rRNAs and 22 tRNAs (9P12RNA). In the preference of this dataset, the result of the test of substitution saturation was considered, which exhibited lower degrees of correlation between 3rd codon positions of all PCGs and the all codon positions of *atp8, nad4l, nad6 and nad2.* Maximum Likelihood (ML) analyses were performed in IQ-TREE v2.1.4 [27] using default parameters under the proposed substitution model (GTR + G + I) with 1000 bootstrap replicates using the fast bootstrapping option implemented in IQTree. Bayesian Inference (BI) analyses were carried out in MrBayes v3.2.2 [28] under the unlinked branch lengths of each partition scheme with two independent runs of five million generations, sampling every 5000 generations. Each run was assessed for stationary using Tracer v1.6 [29]. After the assessment, the first 20% of trees were excluded as burn-in. The generated majority-rule consensus tree (BI tree) from the remaining trees was visualised by FigTree v1.4.2 [30].

Table 2. The best partition scheme selected by PartitionFinder for each dataset used in phylogenetic analyses.

Subsets	Partition scheme	Model
1	*atp6* 1st + *cox1* 1st + *cox2* 1st + *cox3* 1st + *cytb* 1st + *trnK* + *trnM* + *trnS2*	GTR + G + I
2	*atp6* 2nd + *cox1* 2nd + *cox2* 2nd + *cox3* 2nd + *cytb* 2nd + *nad1* 2nd + *nad2* 2nd + *nad3* 2nd + *nad4* 2nd + *nad4l* 2nd + *nad5* 2nd + *nad6* 2nd	GTR + G + I
3	*atp6* 3rd + *atp8* 3rd + *cox1* 3rd + *cox2* 3rd + *cox3* 3rd + *cytb* 3rd + *nad2* 3rd + *nad3* 3rd + *nad6* 3rd	GTR + G + I
4	*nad1* 1st + *nad4* 1st + *nad4l* 1st + *nad5* 1st + *rrnL* + *rrnS* + *trnF* + *trnH* + *trnL1* + *trnQ* + *trnV*	GTR + G + I
5	*nad1* 3rd + *nad4* 3rd + *nad4l* 3rd + *nad5* 3rd	GTR + G + I
6	*atp8* 1st + *atp8* 2nd + *nad2* 1st + *nad3* 1st + *nad6* 1st + *trnA* + *trnC* + *trnD* + *trnE* + *trnG* + *trnI* + *trnL2* + *trnN* + *trnP* + *trnR* + *trnS1* + *trnT* + *trnW* + *trnY*	GTR + G + I

Results and Discussion

Genome architecture and nucleotide composition

The complete mitogenome of the redheaded pine sawfly *N. lecontei* was characterised and comparatively analysed with the mitogenome of *N. sertifer* [3] (Fig. 1 and Table 3). The complete mitogenome of *N. lecontei* was 16,067 bp in length, including 13 PCGs, 23 tRNAs, two rRNAs and A+T rich region (Fig. 1 and Table 3). Fourteen genes are located on the minority N strand, while 24 are encoded by the majority J strand (Table 3). Mitogenome architecture closely matched all previously reported symphytan mitogenomes [10,11,31] and was nearly consistent with that of the inferred insect ancestral mitogenome. The orientation and position of the predicted genes of *N. lecontei* mitogenome were almost identical with *N. sertifer* [3], except for *trnR* gene duplication and rearrangement of ARNS1EF gene cluster (Fig. 1).

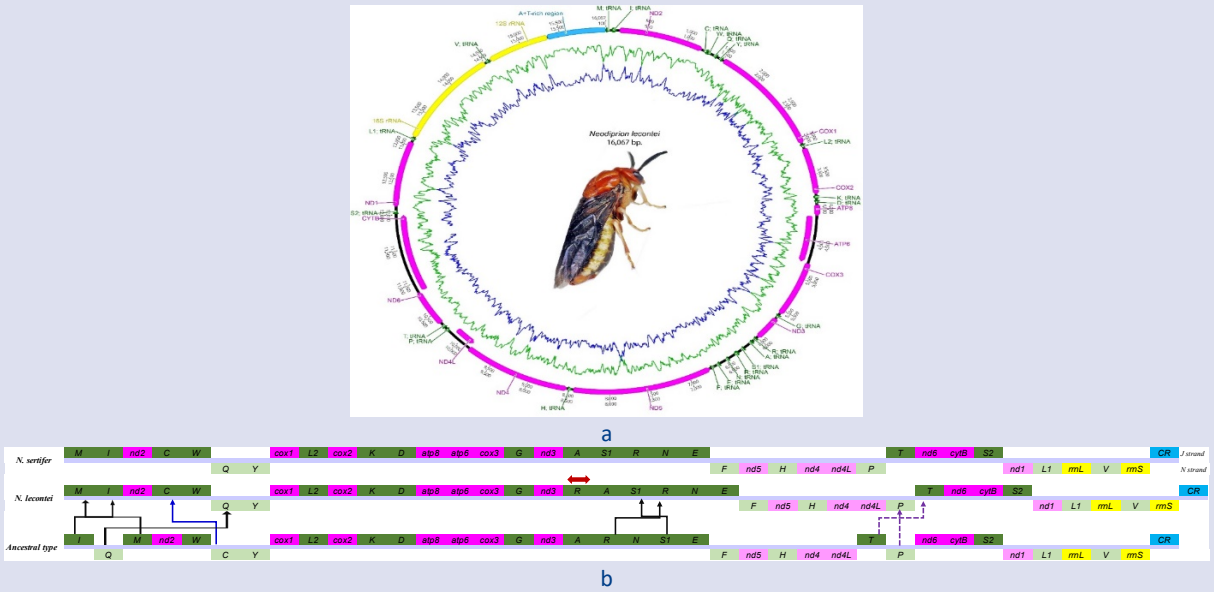

Fig.1. a. Circular map of the mitogenome of *Neodiprion lecontei*. Genes encoded on the J and N strands are marked with right and left arrows, respectively. rRNA genes, PCGs and control region are shown as yellow, pink, and cyan, respectively. tRNA genes are labelled by the single letter amino acid code and marked in green. The skew values of AT% and GC % are displayed with line blue and green respectively. Photo of the species was taken by Ott (2010). b. Mitogenome architecture of *Neodiprion lecontei* referenced with the ancestral type and *Neodiprion sertifer* mitogenome. Same colours were preferred in PCGs, rRNA and tRNA genes, and AT-rich region. Gene rearrangements are specified with arrows (black show transposition; blue shows remote inversion; purple show shuffling and red show duplication).

The duplicated copy of this tRNA was inserted upstream of trnA (arranged as RAS1RNEF, Fig. 1), representing a new pattern for Symphyta. The presence of long intergenic regions among the rearranged and/or duplicated genes might be explained by tandem duplication and random loss (TDRL) mechanism proposed as most widely accepted mechanism for gene rearrangements in insect mitogenomes [2,3,31]. The total length of the intergenic regions was 443 bp in 23 locations with a size ranging from 1 to 74 bp (Table 3). These were mainly found in the ARNS1EF gene cluster, with a 55.76% (247 bp) of total length of the intergenic regions. In spite of the commonly observed pattern in hymenopteran mitogenomes [10,32], only three overlapping regions were detected between *atp8* and *atp6* (7 bp), *nd4* and *nd4L* (1 bp), and *nd6* and *cytB* (1 bp) (Table 3).

Table 3. Mitogenome summary of Neodiprion lecontei (16,067 bp)

Gene	Strand	Size (bp)	Start codon	Stop codon	Anticodon	IGN
trnM	J	69			CAT	0
trnI	J	70			GAT	28
nd2	J	1050	ATG	TAA		6
trnC	N	64			GCA	0
trnW	J	66			TCA	2
trnQ	N	69			TTG	15
trnY	N	66			GTA	9
cox1	J	1540	ATA	T--		0
trnL2	J	66			TAA	1
cox2	J	675	ATG	TAA		2
trnK	J	72			CTT	0
trnD	J	67			GTC	0
atp8	J	162	ATT	TAA		-7
atp6	J	681	ATG	TAA		3
cox3	J	804	ATG	TAA		3
trnG	J	69			TCC	0
nd3	J	351	ATT	TAA		6
trnR	J	69			TCG	1
trnA	J	67			TGC	74
trnS1	J	61			TCT	30
trnR	J	69			TCG	73
trnN	J	65			GTT	64
trnE	J	65			TTC	3
trnF	N	69			GAA	2
nd5	N	1728	ATT	TAA		0
trnH	N	66			GTG	4
nd4	N	1341	ATG	TAA		-1
nd4L	N	291	ATT	TAA		31
trnP	N	67			TGG	20
trnT	J	63			TGT	26
nd6	J	513	TTA	TAA		-1
cytB	J	1134	ATG	TAA		12
trnS2	J	69			TGA	28
nd1	N	951	ATT	TAA		0
trnL1	N	67			TAG	0
rrnL	N	1374				0
trnV	N	67			TAC	0
rrnS	N	780				0
A + T-rich region		716				

As similar to the reported hymenopteran mitogenomes [11,31,33–35], a bias to A and T nucleotides was observed in the mitogenome of N. lecontei, with an average 81.32% AT content (Table 4). Similar to the mitogenome of N. sertifer [3], the AT content was high in tRNA genes (84.11%), while low in PCGs (79.48%). AT bias was also strong in N strand (82.00% AT content on average) than in J strand (77.91%), with G strand being richer in G (11.51%) than C (6.50%), and the J strand showing an opposite skew for C (13.37%) and G (9.71%) (Table 4). The AT content of 3rd codon positon (91.04%) was higher than those of the 1st (75.62%) and 2nd codon positions (71.79%), indicating the robust effect of mutational pressure and the reduced effect of selection acting on the third codon position. The presence of T bias in the second codon position (50.05% T content) might be explained by the pressure of mutation on mitochondrial proteins in favour of hydrophobic amino acids with codons having a T at the second codon positon as phenylalanine, leucine, isoleucine, and methionine [3,10]. A positive AT skew (0.066) and a negative GC skew (−0.182) were counted in the whole mitogenome (Table 4), which is widely reported pattern in the mitogenome of sawflies. However, a deviation out of strand asymmetries is observed in the PCGs: T- and G-skewed (-0.133 and 0.014). The PCGs encoded on the J strand are T- (−0.049) and C-skewed (−0.120), while those of the N strand are T- (−0.260) and G-skewed (0.278) (Table 4). This deviation is most probably related with the effect of the mutational pressures on the mitogenome such as exposure to more DNA damage during replication [37].

Protein Coding Genes

In comparison with the *N. sertifer* mitogenome, the lengths of PCGs were same, except for *nad2*, *cox1*, *cox3* and *nad4L* (Table 3). The *cox1* gene of *N. lecontei* is six codons shorter than that of *N. sertifer*. As widely reported for animal mitogenomes [37], the initiation codons were found as isoleucine (ATY) or methionine (ATR), except for *nad6* (TTA-Phe) (Table 3). The termination codon was found as TAA, except for cox1 which has truncated termination codon (T-) and its product was probably completed after posttranscriptional polyadenylation of mature mRNA [38]. The RSCU values of *N. lecontei* and *N. sertifer* exhibited a similarity for codon usage bias and pictured an important relation between codon preference and nucleotide composition (Fig. 2). Similar to that of other known symphytan mitogenomes, AUU-Ile, UUA-Leu, AUA-Met and UUU-Phe are the most commonly used codons, consisting 40.08% of total content [3,11,34]. The codons with RSCU greater than 2.00 had T or A in the third codon position (Fig. 2).

Table 4. Nucleotide composition of the *Neodiprion lecontei* and *N. sertifer* mitogenomes

Feature	Species	T%	C%	A%	G%	A+T%	AT-skew	GC-skew
Whole genome								
	N. lecontei	37.99	11.04	43.33	7.64	81.32	0.066	-0.182
	N. sertifer	38.30	11.03	43.14	7.53	81.44	0.059	-0.189
Protein-coding genes								
	N. lecontei	45.01	10.11	34.47	10.40	79.48	-0.133	0.014
	N. sertifer	44.57	10.57	34.45	10.42	79.02	-0.128	-0.007
First codon position								
	N. lecontei	37.96	9.60	37.66	14.78	75.62	-0.004	0.212
	N. sertifer	37.62	10.01	37.81	14.57	75.43	0.003	0.186
Second codon position								
	N. lecontei	50.05	16.18	21.74	12.03	71.79	-0.394	-0.147
	N. sertifer	50.31	16.12	21.70	11.88	72.01	-0.397	-0.151
Third codon position								
	N. lecontei	47.03	4.57	44.01	4.39	91.04	-0.033	-0.020
	N. sertifer	45.77	5.58	43.85	4.80	89.62	-0.021	-0.075
Protein-coding genes-J								
	N. lecontei	40.88	12.37	37.03	9.71	77.91	-0.049	-0.120
	N. sertifer	40.88	12.71	36.81	9.59	77.69	-0.052	-0.140
First codon position								
	N. lecontei	33.33	11.68	40.58	14.41	73.91	0.098	0.105
	N. sertifer	32.93	12.31	40.55	14.21	73.48	0.104	0.072
Second codon position								
	N. lecontei	47.37	18.45	22.41	11.77	69.78	-0.358	-0.221
	N. sertifer	47.72	18.25	22.54	11.49	70.26	-0.358	-0.227
Third codon position								
	N. lecontei	41.95	6.99	48.11	2.95	90.06	0.068	-0.406
	N. sertifer	42.00	7.59	47.33	3.08	89.33	0.060	-0.423
Protein-coding genes-N								
	N. lecontei	51.64	6.50	30.36	11.51	82.00	-0.260	0.278
	N. sertifer	50.46	7.13	30.67	11.74	81.13	-0.244	0.244
First codon position								
	N. lecontei	45.37	6.26	32.99	15.38	78.36	-0.158	0.421
	N. sertifer	45.14	6.32	33.40	15.14	78.54	-0.149	0.411
Second codon position								
	N. lecontei	54.35	12.53	20.67	12.46	75.02	-0.449	-0.003
	N. sertifer	54.44	12.71	20.35	12.50	74.79	-0.456	-0.008
Third codon position								
	N. lecontei	55.18	0.70	37.44	6.68	92.62	-0.192	0.810
	N. sertifer	51.81	2.36	38.26	7.57	90.07	-0.150	0.525
tRNA genes								
	N. lecontei	41.63	6.87	42.48	9.01	84.11	0.010	0.135
	N. sertifer	41.56	6.87	42.65	8.91	84.21	0.013	0.129
rRNA genes								
	N. lecontei	44.96	5.62	39.11	10.31	84.07	-0.070	0.294
	N. sertifer	45.07	5.58	38.89	10.46	83.96	-0.074	0.304

Transfer RNA and Ribosomal RNA Genes

The predicted tRNAs of *N. lecontei* mitogenome were almost same position and orientation with *N. sertifer* (Fig. 1). Their length ranged in size from 63 bp (*trnT*) to 72 bp (*trnK*) (Table 3), with 84.11% of AT content. These tRNAs also folded into a conserved clover-leaf structure, except for *trnS1* with a dihydrouridine (DHU) arm as expected. The anticodons of the tRNAs were identical with the reported symphytan mitogenomes, excepting *trnS1* which prefers GCU as an anticodon (Table 3). The *rrnL* gene was between *trnL1* and *trnV* (Fig. 1) and its length was 1374 bp, with an 84.3% AT content (Tables 3, 4). This was similar to homolog genes of *N. sertifer* and other reported hymenopteran species [3,10,11]. The conservation level of the nucleotide positions is relatively high with an average of 72.01%. The length of the *rrnS* gene was 780 bp with an 83.4% AT content (Tables 3, 4).

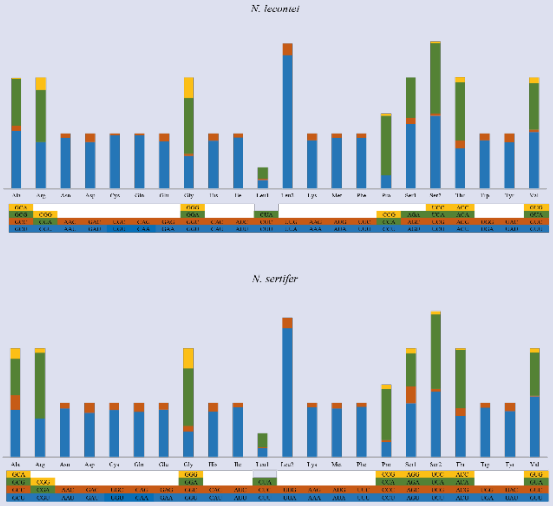

Figure 2. Relative synonymous codon usage (RSCU) of the pine sawfly mitogenomes. Codon families are provided on the x axis. The stop codons are not given.

Phylogeny of Tenthredinoidea

The same and well supported phylogenetic tree topologies were recovered in both analyses (Fig. 3). The recovered trees confirmed the taxonomic position of *N. lecontei* within Diprionidae and supported a relationship of (Pergidae + Argidae) + (Athalidae + ((Diprionidae + Cimbicidae) + Tenthredinidae)) in the Tenthredinoidae and this is reliable with the most of the reported phylogenies [33,39–40]. The monophyly of included genera and families were also highly supported (Fig. 3). These results highlight that the mitogenome dataset verifies useful in the built of the phylogeny of Tentredinoidae as well as of Symphyta.

Conclusion

The annotation and characterisation of the complete mitogenome of redheaded pine sawfly *N. lecontei* and the phylogenetic replacements of tenthridinoid families allow us to designate several conclusions: (i) the mitogenome architecture and orientation are mostly reliable with the reported symphytan mitogenomes; (ii) *trnR* gene duplication and rearrangement of ARNS1EF gene cluster seem to be unique to this species; (iii) the phylogenetic analyses confirm the position of *N. lecontei* and also support the monophyly of the tenthredinoid families.

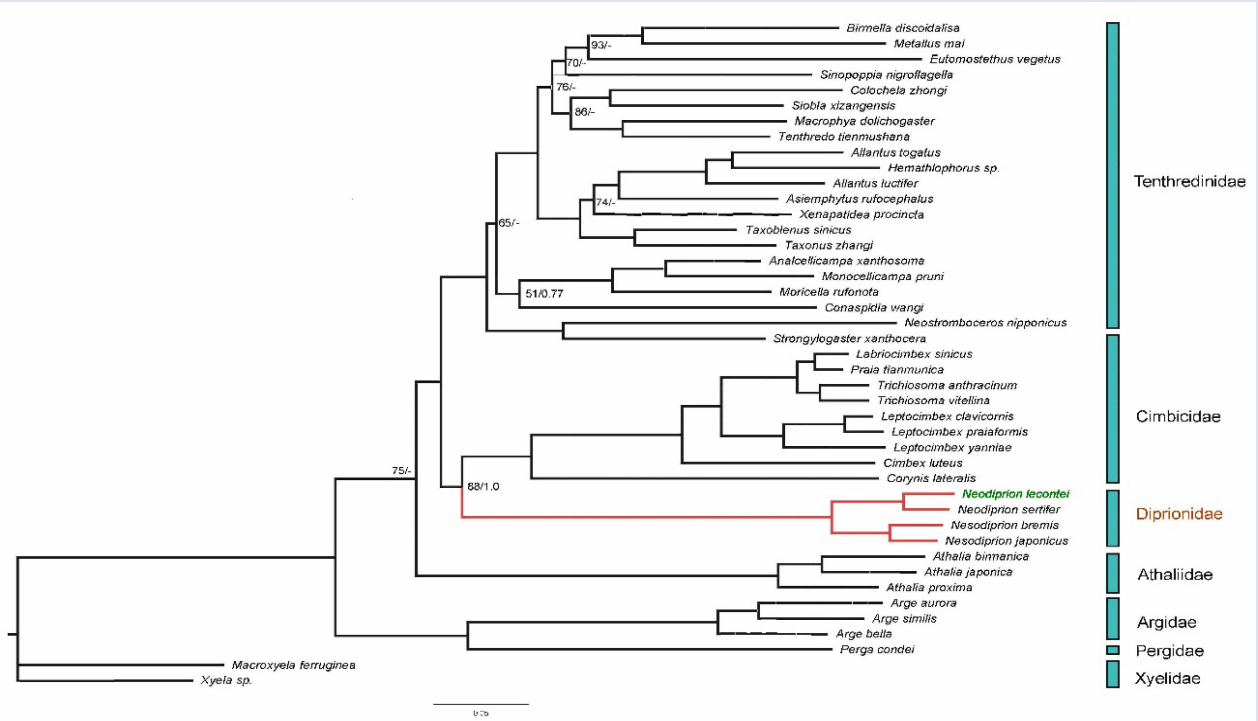

Fig. 3. Phylogenetic tree of the superfamily of Tentredinoidea. The trees were constructed applying BI and ML methods under a concatenated 9P12RNA dataset (9 PCGs, two rRNAs and 22 tRNAs); both analyses generated the same tree topology. The outgroups were selected from the members of Xyelidae. Support values ≥ 95% in ML and 0.95 in BI were shown.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. I thank Dr. Özgül DOĞAN for her help in the analysis stages. The members of Evolutionary Bioinformatics Research Group (EBRG) at Sivas Cumhuriyet University are thanked for their contributions in visualising and drawing some of the figures.

Conflicts of interest

The author declares no conflicts of interest.

References

[1] Carlucci A., Lignitto L., Feliciello A., Control of mitochondria dynamics and oxidative metabolism by cAMP, AKAPs and the proteasome, *Trends Cell Biol.*, 18 (2008) 604–613.

[2] Cameron S.L., Insect mitochondrial genomics: implications for evolution and phylogeny, *Annu. Rev. Entomol.*, 59 (2014) 95–117.

[3] Aydemir, M. N., Korkmaz, E. M., Comparative mitogenomics of Hymenoptera reveals evolutionary differences in structure and composition, International journal of biological macromolecules., 144 (2020) 460-472.

[4] Ballard J.W.O., Pichaud N., Mitochondrial DNA: more than an evolutionary bystander, *Funct. Ecol.,* 28 (2014) 218–231.

[5] Boore J.L., Animal mitochondrial genomes, *Nucleic Acids Res.*, 27 (1999) 1767–80.

[6] Grimaldi D., Engel M.S., Evolution of the insects., Cambridge University Press, New York, (2005).

[7] Klopfstein S., Vilhelmsen L., Heraty J.M., Sharkey M., Ronquist F., The hymenopteran tree of life: Evidence from

protein-coding genes and objectively aligned ribosomal data, *PLoS One*., 8 (2013).

[8] Taeger A., Liston A.D., Prous M., Groll E.K., Gehroldt T., Blank S.M. ECatSym – Electronic World Catalog of Symphyta (Insecta, Hymenoptera). Program version 5.0 (19 Dec 2018), data version 40 (23 Sep 2018). – Senckenberg Deutsches Entomologisches Institut (SDEI), Müncheberg. https://sdei.de/ecatsym/ Access: 15 Oct 2022

[9] Cheng Y., Yan Y., Wei M., Niu G., Characterization of mitochondrial genomes of three new species: *Leptocimbex praiaformis*, *L. clavicornis* , and *L. yanniae* (Hymenoptera: Cimbicidae), *Entomol. Res.,* 51 (2021) 287–304.

[10] Doğan Ö., Korkmaz E.M., Nearly complete mitogenome of hairy sawfly, *Corynis lateralis* (Brullé, 1832) (Hymenoptera: Cimbicidae): Rearrangements in the IQM and ARNS1EF gene clusters, *Genetica*., 145 (2017) 341–350.

[11] Song S.N., Tang P., Wei S.J., Chen X.X., Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans, *Sci. Rep*., 6 (2016).

[12] Bagley R.K., Sousa V.C., Niemiller M.L., Linnen C.R., History, geography and host use shape genomewide patterns of genetic variation in the redheaded pine sawfly (*Neodiprion lecontei*), *Mol. Ecol*., 26 (2017) 1022–1044.

[13] Chen, S., Zhou, Y., Chen, Y., Gu, J., fastp: an ultra-fast all-in-one FASTQ preprocessor, *Bioinformatics.*, *34*(17) (2018) i884-i890.

[14] Song, L., Florea, L., Langmead, B., Lighter: fast and memory-efficient sequencing error correction without counting, *Genome biology.*, 15(11) (2014) 1-13.

[15] 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., Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data, *Bioinformatics*., 28 (2012) 1647–9.

[16] Prjibelski, A., Antipov, D., Meleshko, D., Lapidus, A., Korobeynikov, A., Using SPAdes de novo assembler, *Current protocols in bioinformatics.*, 70(1) (2020) e102.

[17] Arkin AP., Cottingham RW., Henry CS., Harris NL., Stevens RL., Maslov S., Dehal P., Ware D., Perez F., Canon S., KBase: the United States department of energy systems biology knowledgebase, *Nat Biotechnol*., 36(7) (2018) 566–9.

[18] Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, F., Middendorf. M., Stadler, P. F., MITOS: improved de novo metazoan mitochondrial genome annotation, *Molecular phylogenetics and evolution.*, *69*(2) (2013) 313-319.

[19] Wyman, S. K., Jansen, R. K., Boore, J. L., Automatic annotation of organellar genomes with DOGMA, *Bioinformatics.*, *20*(17) (2004) 3252-3255.

[20] Tamura K., Stecher G., Peterson D., Filipski A., Kumar S., MEGA6: Molecular evolutionary genetics analysis version 6.0, *Mol. Biol. Evol.*, 30 (2013) 2725–2729.

[21] Perna N.T., Kocher T.D., Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes, *J. Mol. Evol.*, 41 (1995) 353–358.

[22] Katoh K., Standley D.M., MAFFT multiple sequence alignment software version 7: Improvements in performance and usability, *Mol. Biol. Evol.,* 30 (2013) 772–780.

[23] Vaidya G., Lohman D.J., Meier R., SequenceMatrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information, *Cladistics.*, 27 (2011) 171–180.

[24] Lanfear R., Calcott B., Ho S.Y.W., Guindon S., PartitionFinder: Combined selection of partitioning schemes and substitution models for phylogenetic analyses, *Mol. Biol. Evol.*, 29 (2012) 1695–1701.

[25] R Core Team, R: A language and environment for statistical computing, Vienna, Austria, 2020.

[26] Swofford D.L., PAUP. Phylogenetic analysis using parsimony (and other methods), version 4, Sinauer Assoc., Sunderland, Massachusetts. (2002) 294–307.

[27] Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., von Haeseler A., Lanfear R., IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era, *Mol. Biol. Evol*., 37 (2020) 1530–1534.

[28] Ronquist F., Teslenko M., Van Der Mark P., Ayres D.L., Darling A., Höhna S., Larget B., Liu L., Suchard M.A., Huelsenbeck J.P., Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space, *Syst. Biol*., 61 (2012) 539–542.

[29] Rambaut A., Suchard M.A., Xie D., Drummond A.J., Tracer v1.6., (2014).

[30] Rambaut A., FigTree v1.4.2, a graphical viewer of phylogenetic trees, Available from left angle bracket. http://tree.bio.ed.ac.uk/software/figtree/ (2014).

[31] Dowton M., Cameron S.L., Dowavic J.I., Austin A.D., Whiting M.F., Characterization of 67 mitochondrial tRNA gene rearrangements in the Hymenoptera suggests that mitochondrial tRNA gene position is selectively neutral, *Mol. Biol. Evol.*, 26 (2009) 1607–17.

[32] Castro L.R., Dowton M., The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: Symphyta: Pergidae), *Mol. Phylogenet. Evol*., 34 (2005) 469–79.

[33] Niu G., Jiang S., Doğan Ö., Korkmaz E.M., Budak M., Wu D., Wei M., Mitochondrial phylogenomics of Tenthredinidae (Hymenoptera:Tenthredinoidea) supports the monophyly of Megabelesesinae as a subfamily, *Insects*., 12 (2021) 495.

[34] Wu R., Wei M., Liu M., Niu G., Advancement in sequencing the mitochondrial genome of *Birmella discoidalisa* Wei, 1994 (Hymenoptera: Tenthredinidae) and the phylogenetic classification of *Fenusini*, *Mitochondrial DNA Part B*., 4 (2019) 4100–4101.

[35] Wei S.J., Shi M., He J., Sharkey M., Chen X., The complete mitochondrial genome of *Diadegma semiclausum* (Hymenoptera: Ichneumonidae) indicates extensive independent evolutionary events, *Genome*., 52 (2009) 308–319.

[36] Wei S.J., Shi M., Sharkey M.J., van Achterberg C., Chen X., Comparative mitogenomics of Braconidae (Insecta: Hymenoptera) and the phylogenetic utility of mitochondrial genomes with special reference to Holometabolous insects, *BMC Genomics*., 11 (2010) 371.

[37] Crozier R.H., Crozier Y.C., The mitochondrial genome of the honeybee *Apis mellifera*: Complete sequence and genome organization, *Genetics.,* 133 (1993) 97–117.

[38] Ojala D., Montoya J., Attardi G., tRNA punctuation model of RNA processing in human mitochondria, *Nature,* 290 (1981) 470–474.

[39] Malm T., Nyman T., Phylogeny of the symphytan grade of Hymenoptera: New pieces into the old jigsaw(fly) puzzle, *Cladistics*., 31 (2015) 1–17.

[40] Niu G., Budak M., Korkmaz E.M., Doğan Ö., Nel A., Wan S., Cai C., Jouault C., Li M., Wei M., Phylogenomic analyses of the Tenthredinoidea support the familial rank of Athaliidae (Insecta, Tenthredinoidea), *Insects*., 13 (2022) 858