Complete mitochondrial genome sequence of *Labriocimbex sinica*, a new genus and new species of Cimbicidae (Hymenoptera) from China (#34672)

First submission

Editor guidance

Please submit by **2 Mar 2019** for the benefit of the authors (and your $200 publishing discount).

Structure and Criteria
Please read the 'Structure and Criteria' page for general guidance.

Custom checks
Make sure you include the custom checks shown below, in your review.

Author notes
Have you read the author notes on the guidance page?

Raw data check
Review the raw data. Download from the location described by the author.

Image check
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files
Download and review all files from the materials page.

11 Figure file(s)
4 Table file(s)
2 Other file(s)

Custom checks

DNA data checks

- Have you checked the authors data deposition statement?
- Can you access the deposited data?
- Has the data been deposited correctly?
- Is the deposition information noted in the manuscript?

New species checks

- Have you checked our new species policies?
- Do you agree that it is a new species?
- Is it correctly described e.g. meets ICZN standard?

Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- Clear, unambiguous, professional English language used throughout.
- Intro & background to show context. Literature well referenced & relevant.
- Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- Figures are relevant, high quality, well labelled & described.
- Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- Original primary research within Scope of the journal.
- Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- Rigorous investigation performed to a high technical & ethical standard.
- Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- Data is robust, statistically sound, & controlled.
- Speculation is welcome, but should be identified as such.
- Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Complete mitochondrial genome sequence of *Labriocimbex sinica*, a new genus and new species of Cimbicidae (Hymenoptera) from China

Yuchen Yan [1], **Yaoyao Zhang** [1], **Qianying Ren** [1], **Shiyu Du** [1], **Bocheng Lan** [1], **Gengyun Niu** [Corresp., 2], **Meicai Wei** [Corresp. 2]

[1] Lab of Insect Systematics and Evolutionary Biology, Key Laboratory of Cultivation and Protection for Non-Wood Forest Trees, Central South University of Forestry and Technology, Changsha, Hunan, China
[2] Jiangxi Normal University, Nanchang, Jiangxi, China

Corresponding Authors: Gengyun Niu, Meicai Wei
Email address: niug@jxnu.edu.cn, weim@jxnu.edu.cn

Labriocimbex sinica Yan & Wei gen. et sp. nov. of Cimbicidae is described. Its mitochondrial genome is also reported here. The new genus is closely similar to *Pseudoclavellaria* Schultz and *Trichiosoma* Leach. A key to the genera of Trichiosomini is provided. To identify the systematic placement of Cimbicidae, the mitochondrial genome of *L. sinica* was assembled using high-throughput sequencing data. The complete mitochondrial genome of *L. sinica* was obtained with a length of 15405 bp (GenBank: MH136623; SRA: SRR8270383) and a typical set of 37 genes (22 transfer RNAs [tRNAs], 13 protein-coding genes [PCGs], and two rRNAs). The results demonstrated that all PCGs were initiated by ATN codons, and ended with TAA or T stop codons. The study revealed that all tRNA genes of this species had a typical clover-leaf secondary structure, except *trnS1*. Remarkably, the secondary structure of the *rrnS* and *rrnL* of *L. sinica* was much different from that of *Corynis lateralis* (*C. lateralis*). Phylogenetic analyses verified the monophyly and positions of three Cimbicidae species within the superfamily Tenthredinoidea and demonstrated a (Tenthredinidae + Cimbicidae) + (Argidae + Pergidae) relationship in Symphyta with strong nodal supports. Furthermore, we found that a phylogenetic tree based on two methods showed that *L. sinica* is a sister group of *Trichiosoma anthracinum* (*T. anthracinum*) with high support values.

Complete mitochondrial genome sequence of *Labriocimbex sinica*, a new genus and new species of Cimbicidae (Hymenoptera) from China

Yuchen YAN[1], Yaoyao ZHANG[1], Qianying REN[1], Shiyu Du[1], Bocheng LAN[1], Nengyun NIU[2], Meicai WEI[2],

[1]Lab of Insect Systematics and Evolutionary Biology; Key Laboratory of Cultivation and Protection for Non-Wood Forest Trees, Central South University of Forestry and Technology, Changsha, 410018, China.
[2] Jiangxi Normal University, Nanchang, Jiangxi, China

Corresponding Author:
Meicai WEI[2]
99 Ziyang Road, Nanchang, Jiangxi 330022, China
Email address: Email address: weim@jxnu.edu.cn

Nengyun NIU[2]
99 Ziyang Road, Nanchang, Jiangxi 330022, China
Email address: niug@jxnu.edu.cn

Abstract:

Labriocimbex sinica Yan & Wei gen. et sp. nov. of Cimbicidae is described. Its mitochondrial genome is also reported ~~here~~. The new genus is ~~closely~~ similar to *Pseudoclavellaria* Schultz and *Trichiosoma* Leach. A key to the genera of Trichiosomini is provided. To identify the ~~systematic~~ placement of Cimbicidae, the mitochondrial genome of *L. sinica* was assembled using high-throughput sequencing data. The complete mitochondrial genome of *L. sinica* was obtained with a length of 15405 bp (GenBank: MH136623; SRA: SRR8270383) and a typical set of 37 genes (22 transfer RNAs [tRNAs], 13 protein-coding genes [PCGs], and two rRNAs). The results demonstrated that all PCGs were initiated by ATN codons, and ended with TAA or T stop codons. The study ~~revealed~~ that all tRNA genes of this species ~~had~~ a typical clover-leaf secondary structure, except *trnS1*. Remarkably, the secondary structure of the *rrnS* and *rrnL* of *L. sinica* was much different from that of *Corynis lateralis* ~~(C. lateralis)~~. Phylogenetic analyses verified the monophyly and positions of three Cimbicidae species within the superfamily Tenthredinoidea and demonstrated a (Tenthredinidae + Cimbicidae) + (Argidae + Pergidae) relationship ~~in Symphyta~~ with strong nodal supports. Furthermore, we found that a phylogenetic tree based on two methods ~~showed~~ that *L. sinica* is ~~a~~ sister group of *Trichiosoma anthracinum* (*T. anthracinum*) with high support values.

INTRODUCTION

Hymenoptera is one of the large insect order including more than 153,000 species which possess very diverse life strategies (Peters *et al*., 2017). Currently, complete or nearly complete mitochondrial genomes have been reported for 199 hymenopteran species (NCBI, January 2019). The Cimbicidae ~~represents~~ a small family of Tenthredinoidea of the phytophagous suborder Symphyta (Hymenoptera), with about 197 valid species and 26 genera around the world. Within China, about 63 species and 13 genera have already been recorded (Taeger *et al*., 2010; Yan and Wei, 2010; Blank *et al*., 2012; Yan and Wei, 2013; Yan *et al*., 2014; Yan and Wei, 2016; Yan *et al*., 2018). The monophyly of Tenthredinoidea is supported by both morphological (Wei and Nie, 1997) and molecular data (Malm and Nyman, 2015) as well as both combined (Ronquist *et al*., 2012; Sharkey *et al*. 2012; Klopfstein *et al*., 2013). However, the relationships among core tenthredinoids ~~were~~ unclear. Cimbicidae was inferred as the sister to Argidae + Pergidae proposed by morphological analyses (Wei and Nie, 1997; Vilhelmsen, 2001; 2015; 2018). The disaccord with several recent studies may ~~suffer from~~ the limited dataset, by molecular or combined analyses, which have placed Cimbicidae as sister to Diprionidae (Schulmeister, 2003; Schmidt and Walter, 2014; Isaka and Sato, 2015; Malm and Nyman, 2015) or a clade including Diprionidae ~~form~~ a monophylum as sister to the remaining tenthredinoids (Heraty *et al.*, 2011; Ronquist *et al.*, 2012; Klopfstein *et al.*, 2013).

The monophyly of Cimbicidae has never been contested (Vilhelmsen, ~~2018~~). Adult Cimbicidae ~~were~~ primarily characterized by their clubbed antennae, one or more of the apical antennomeres being expanded. They ~~were~~ small (6mm) to very large insects (3cm), making them the largest true sawflies known (Vilhelmsen, ~~2018).~~ Some of the species ~~were~~ economically important pests

causing serious defoliation of woody plants such as elm, willow, honeysuckle and snowberry (Gauld and Bolton 1988). Malaise (1934) established the classification system of Cimbicidae: subfamily, tribe, subtribe and genus. Benson (1938) carried out a comprehensive study of sawfly, especially the family members of Cimbicidae, which was further determined by the classification status of Cimbicidae. It included four subfamilies: Abiinae, Cimbicinae, Pachylostictinae and Coryninae. The Cimbicinae is the most diverse subfamily, encompassing Cimbicini and Trichiosomini (Abe and Smith, 1991), that was best-supported monophyletic clade in the recent analysis of Cimbicidae(Vilhelmsen, 2018), but it is not supported as a monophyly in Cladistics analyses with sufficient representation of cimbicid taxa of China (Deng, 2000). Deng (2000) proposed Trichiosomini tribe included *Pseudoclavellaria*, *Leptocimbex* and *Trichiosoma*. The new genus is closely similar to *Pseudoclavellaria* Schultz and *Trichiosoma* Leach by some morphological characters. Vilhelmsen (2018) sustained *Labriocimbex* placed as a grade basal to Cimbicinae, it was usually placed as the sister to *Cimbex* + *Odontocimbex* in the all strict consensus trees.

So far, mitochondrial genome of two species, *T. anthracinum* (GenBank accession KT921411) and *C. lateralis* (GenBank accession KY063728) have been reported for the family (Song *et al.* 2016; Doğan and Korkmaz, 2017). Here, we reported one complete mitochondrial genome of *Labriocimbex*. We also compared it with the previously reported mitochondrial genome of *T. anthracinum* and *C. lateralis* for the better understanding of the mitochondrial genome characteristics of the Cimbicidae. Finally, we have performed phylogenetic analyses to verify the phylogenetic position of *Labriocimbex* using a mitochondrial genome dataset of 36 species of Hymenoptera (Symphyta species 34 and Apocrita species two) and four non-hymenopteran outgroups (Table 1).

MATERIALS & METHODS

Description of new species

Specimens were examined with a Leica S8APO dissection microscope. Adult images were taken with a Nikon D700 digital camera and a series of images edited using Helicon Focus (HeliconSoft), while detailed images were taken with Leica Z16 APO/DFC550. We used Adobe Photoshop CS 6.0 for further image processing. The terminology of sawfly genitalia follows Ross (1945), and that of general morphology follows Viitasaari (2002). For a few terms (e.g. middle fovea and lateral fovea), we followed Takeuchi (1952). Abbreviations used were: OOL = distance between the eye and outer edge of lateral ocelli; POL = distance between the mesal edges of the lateral ocelli; OCL = distance between a lateral ocellus and the occipital carina or hind margin of the head.

The holotype and all paratypes of the new species were deposited in the Insect Collection of Central South University of Forestry and Technology, Changsha, Hunan, China (CSCS). Part paratypes were from Lishui Academy of Forestry (LSAF).

The electronic version of this article in Portable Document Format (PDF) will represent a published work according to the International Commission on Zoological Nomenclature (ICZN), and hence the new names contained in the electronic version are effectively published under that

~~Code from the electronic edition alone.~~ This published work and the nomenclatural acts it contains have been registered in ZooBank, the online registration system for the ICZN. The ZooBank LSIDs (Life Science Identifiers) ~~can be resolved~~ and the associated information viewed through any standard web browser by appending the LSID to the prefix http://zoobank.org/. The LSID for this publication is: urn: lsid: zoobank.org: pub: EE7F5193-78B2-42CE-87C1-B3FE947CB70F. The online version of this work is archived and available from the following digital repositories: PeerJ, PubMed Central and CLOCKSS.

DNA library construction and sequencing

Total DNA was extracted from *L. sinica* using an E.Z.N.A.® Tissue DNA Kit (Omega, Norcross, GA) and was stored at −20°C, in accordance with the manufacturer's instructions. Sequencing libraries with approximately 250-bp insertions were constructed using a NEXT flex™ Rapid DNA-Seq Kit (Illumina, San Diego, CA) in accordance with the manufacturer's protocol. Each library was sequenced using an Illumina Hiseq 4000 to generate 150-bp paired end reads at BGI-Shenzhen, China. The sequencing reads have been deposited in NCBI SRA database under accession number: PRJNA507477.

Mitochondrial genome assembly

Next generation sequencing and bioinformatic analyses were performed by Shanghai Majorbio Bio-pharm Technology Co., Ltd. Reconstruction of the mitochondrial genome from Illumina reads were carried out using three different approaches to ensure the accuracy of the assemblies: SOAPdenovo v2.0 (Luo *et al.*, 2012), MITObim v1.8 (Hahn *et al.*, 2013) and NOVOPlasty v2.7.1 (Dierckxsens *et al.*, 2017). The assembled mitochondrial fragments were identified using BlastX and *T. anthracinum* (NC029733) mitochondrial genes as queries. Prediction and annotation of protein-coding, tRNA and rRNA genes were performed using DOGMA (http://dogma.ccbb.utexas.edu/) or MITOS (http://mitos.bioinf.uni-leipzig.de/index.py) with annotation from a reference mitochondrial genome. Queries were then corrected manually.

Mitochondrial genome annotation and secondary structure prediction

All RNA genes were identified by employing the online MITOS tool (http://mitos.bioinf.uni-leipzig.de/index.py) (Bernt *et al.*, 2013) with the invertebrate mitochondrial genetic code. The initiation and termination codons of PCGs were determined using Geneious v11.0.3 (http://www.geneious.com) with reference sequences from other symphytan species with subsequent manual adjustment. The A + T content of nucleotide sequences and relative synonymous codon usage (RSCU) were calculated using MEGA v7.0 (Kumar *et al.*, 2016). Strand asymmetry was calculated using the formulae (Perna and Kocher, 1995): GC–skew = (G – C) / (G + C) and AT–skew = (A – T) / (A + T), for the strand encoding the majority of the PCGs.

The secondary structures of the *rrnS* and *rrnL* were partitioned into four areas and six areas, respectively. The secondary structures of rRNAs were inferred using alignment to models predicted for *T. anthracinum*. First, the primary sequence and the secondary structure of this species were aligned in MARNA (Siebert and Backofen, 2005) to identify a consensus sequence

as well as a consensus structure in the output files. Secondly, the secondary structures of the *rrnS* and *rrnL* in *L. sinica* were predicted by specific structure models in SSU-ALIGN (Nawrocki, 2009). Finally, the structures were artificially transformed to their relative secondary structure with micro changes.

The predicted secondary structures of RNAs were drawn using VARNA v3-93 (Darty *et al.*, 2009) and RNAviz 2.0.3 (De Rijk *et al*., 2003). Helix numbering was performed as with the *Apis mellifera* rRNA secondary structure (Gillespie *et al*., 2006) including minor modifications.

Phylogenetic analysis

We used the Maximum Likelihood (ML) and Bayesian Inference (BI) methods to construct phylogenetic trees of selected species, using 13 PCGs and two rRNAs (Table 1). The mitochondrial genome sequences of selected species were downloaded from GenBank. A total of 13 PCGs were aligned by MUSCLE in MEGA v7.0 individually, and two rRNAs were aligned by MAFFT (https://www.ebi.ac.uk/Tools/msa/mafft/) (Katoh and Standley, 2013). Then, the aligned nucleotide sequences of PCGs and rRNAs were concatenated using Sequence Matrix v1.7.8 (Vaidya *et al.*, 2011) and partitioned into several data blocks.

The partitioned data block file was used to infer both partition schemes and substitution models in Partition Finder v1.1.1 (Lanfear *et al*., 2012), with "unlinked" branch lengths under the "greedy" search algorithm. The standard partitioning schemes "bic" and "aicc" were selected for BI and ML analyses, respectively. BI analyses were conducted with the GTR+I+G model and HKY+G model using MrBayes v3.2.2 (Ronquist *et al*., 2012). Four simultaneous Markov chains (three cold, one heated) were run for five million generations in two independent runs, with sampling every 100 generations and 25% of the first generations were discarded as burn-in.

The GTR+I+G, GTR+G and HKY+G for 13 PCGs were chosen. The best-fit model of nucleotide substitution and phylogenetic construction based on ML was created using the IQ-TREE web server (http://iqtree.cibiv.univie.ac.at/). The previous data block file was used as well as the original parameters. In addition, 0.1 was employed as the disturbance intensity and 1000 as the IQ-TREE stopping rule. All related files have been uploaded to Figshare were (https://figshare.com/s/b7e5b401b4881328c3b1).

RESULTS AND DISCUSSION

Description

***Labriocimbex* Yan & Wei, gen. nov.**

urn: lsid: zoobank. org: act: 29EB6C0E-881D-46E2-AEF0-3BDF5992EC37

Type species: *Labriocimbex sinica* Yan & Wei, sp. nov.

Description. Body middle to large-sized; black, without metallic luster and macula (Figure 1); head and thorax with dense and long yellowish brown hairs; clypeus distinctly broader than space between lower margin of eyes, ~~anterior~~ margin with arcuate notch, upper furrow deep

(Figure 2A); base of labrum much broader than apex, lateral margin of labrum distinctly narrowed upward at base edge (Figure 2A); mandibles ~~long and strong~~, with 2 symmetrical inner teeth, basal one truncate at apex (Figures 2F, 2G); maxillary palp with 6 palpomeres, apex 1–2 combined distinctly shorter than palpomere 4; labial palp with 4 palpomeres, short ~~and small~~ (Figure 2H); malar space 2.3 times longer than diameter of lateral ocellus, about as long as scape and pedicel combined; eyes moderately large, inner margins ~~of eyes~~ ~~sub-parallel~~, distance between eyes slightly longer than longest axis of ~~an~~ eye (Figures 2A, 2B); lateral of head distinctly dilated behind eyes in lateral view (Figure 2C); genal carina distinct cuspidally at lower half in male; postocellar area with ~~middle furrow~~ and lateral furrows distinct, frontal carina indistinct (Figure 2B). Antenna slightly longer than breadth of head ~~together~~, club of antenna strongly enlarged with obscure annular suture, with 5 antennomeres before club, antennomere 3 slender and distinctly longer than antennomeres 4 and 5 combined (Figure 2U). Propleuron and sternum merged, median suture of ~~prescutum~~ shallow, notaulus distinct; mesoscutellum slightly flat, anterior margin sub-truncated, posterior margin roundly triangular-~~like~~(Figure 2N); cenchri ~~keep away~~ from each other, distance between outer margin of cenchri longer than breadth of ~~a~~ mesoscutellum(Figure 2N). Coxa and femur of leg with long hairs; ventral side of middle and hind femur without tooth near apex (Figure 2I), hind coxa close to each other; inner spur of hind tibia as long as apical breadth of tibia, apex blunt and membranous (Figure 2M), about 0.4 times length of metabasitarsus; metabasitarsus slightly shorter than tarsomeres 2 and 3 combined, base narrower than apex(Figure 2D); 1st and 2nd tarsal pulvilli long, nearly ~~contacting to each other~~(Figure 2D); claw simple ~~and not bifurcate~~ (Figure 2J). Fore wing with crossvein 2r, base of vein Rs absent (Figure 1A); anal cell strongly narrowed in basal 1/3 with a short anal crossvein and apical anal cell about 2 times the length of basal anal cell; vein 2r-m and 2m-cu almost interstitial, pterostigma long and narrow; cell Rs and M closed in hind wing, apex of anal cell quadrate, petiole of anal cell longer than length of vein cu-a, jugum region only with 1 longitudinal vein, without crossvein. Sternites and basal abdominal terga with long hairs, posterior margin of abdominal tergum 1 roundly incised, without middle carina and lateral carina (Figure 2N). Genital plate of female developed with ~~incision wide~~ and arcuate ~~at~~ middle (Figure 2L); apical ovipositor sheath short and oblique in lateral view (Figure 2P); apex of lancet and lance curved upwards (Figure 2O, 2Q), each annulus with 1 pore, serrulae sub-truncate at apex, lateral teeth sharp (Figures 2R). Each sternite of male incised ~~at~~ middle and roundish, both sides roundish; penis valve broad, with apical lobe bulge, ventral hook small, lateral ridge distinct (Figure 2S); harpe small, slightly longer than broad (Figure 2T).

Etymology. The generic name is composed of "*labrio-*" and "*-cimbex*", emphasizing the shape of labrum differs from other genera of the family. Gender masculine.

Distribution. China.

Host plant: *Cerasus pseudocerasus* of Rosaceae.

Remarks. This new genus is similar to *Pseudoclavellaria* Schultz and *Trichiosoma* Leach. It differs from *Pseudoclavellaria* by the clypeus and labrum black; base of labrum much wider than its apex and about half length of the clypeus; antenna with 5 antennomeres before the club;

the anal cell strongly narrowed in basal 1/3 with a punctiform crossvein and apical anal cell about 2 times the length of basal anal cell, the vein 2r-m and 2m-cu in fore wing almost interstitial. It differs from *Trichiosoma* by the ventral side of hind femur without a subapical tooth; the 1st and 2nd tarsal pulvilli in male very long and nearly contacting to each other, and the different pattern of the male genitalia.

Key to genera of the tribe Trichiosomini

1 Ventral side of middle and hind femur with 1 distinct tooth near apex……*Trichiosoma* Leach
- Ventral side of femur without tooth……..……2
2 Head and thorax without dense and long hairs; club of antenna segmented; abdominal tergum 1 with distinct lateral carina; tarsal pulvilli short and small, distance between basal 2 pulvilli not shorter than length of a pulvillus; body usually slender ... *Leptocimbex* Semenov
- Head and thorax with dense and long hairs, club of antenna not segmented; abdominal tergum 1 without lateral carina; basal 2 tarsal pulvilli long and almost touched to each other; body stout.. ……3
3 labrum broad at base, distinctly narrowed toward apex; antenna with 5 antennomeres before club; inner spur of hind tibia as long as apical breadth of tibia; abdominal terga with long hairs; forewing with length of apical anal cell about 2 times basal anal cell...*Labriocimbex* Yan and Wei, gen. nov.
- Labrum narrow at base and distinctly broadened at apex, apical breadth of labrum slightly narrower than clypeus; antenna with 4 antennomeres before club; inner spur of hind tibia shorter than apical breadth of tibia, abdominal terga without long hairs; forewing with length of apical anal cell about as long as basal anal cell……..……*Pseudoclavellaria* Schultz

Labriocimbex sinica Yan & Wei sp. nov. **(Figures 1-2)**
urn: lsid:zoobank.org:act:E1454ED2-5321-4D39-97C2-EC8957D034C1

Female. (Holotype) Body length 21mm (Figure 1A). Black; apical 1/2 of mandible dark brown (Figures 2F, 2G); inner and ventral side of club of antenna largely brown, outer side dark brown (Figure 2U); cenchri pale yellowish brown; posterior half of mesepimeron, metapleuron largely, metanotum except for a small macula behind cenchri and most of metapostnotum, median triangular macula and narrow posterior margin of abdominal tergum 1(Figure 2N), yellow brown; apex of each tibia, tarsus and claw reddish brown, tarsal pulvillus grayish white (Figure 2D). Wings brownish hyaline, stigma black, basal 3/5 of vein C in fore wing, basal 3/7 of vein Sc+R and entire vein M+Cu pale yellow, vein A pale brown, other veins largely black; vein J and basal parts of all other veins in hind wing pale yellow (Figure 1A). Hairs on face and gena black at base and yellowish white at apex (Figure 2A); hairs on vertex of head and mesonotum black; hairs on pronotum and scutellum yellowish white largely except for black basal 0.2; hairs on mesopleuron, coxa and femora yellowish brown largely with less than basal 0.3 black; inner hairs of fore tibia reddish yellow; abdominal terga 1–2 and posterior margins of terga 3–4 with

yellowish white hairs; hairs on ventral side of terga and sternites 1–3 black in basal 0.4 and reddish yellow in apical 0.6.

Body densely microsculptured, matt; lower margin of orbit, small fovea lateral to lateral ocellus, apical half of mandible, narrow lateral side of mesoscutal lateral lobe, ventral part of trochanters and of femora distinctly shiny, ventral half of mesepisternum feebly microsculptured mixed with some minute punctures, shiny; venter of abdomen feebly shiny.

Apex of labrum thickened with middle notch (Figure 2A); median fovea round and deep, lateral foveae obscure (Figure 2B); middle of frons concave, lateral furrow of frons shallow; postocellar furrow distinct, interocellar furrow long and deep, postocellar area quadrate, middle furrow indistinct, lateral furrows shallow and weak, weakly divergent backwards (Figure 2B). Long hairs on gena clearly shorter than 1/3 head width in dorsal view. Club of antenna as long as length of antennomeres 4 and 5 combined, with obscure annular suture (Figure 2U).

Mesopleuron without middle ridge (Figure 2E); cenchrus 1.1 times broader than long, reniform (Figure 2N). Coxae and femora with dense hairs longer than breadth of femora (Figure 2I); inner hairs of fore tibia dense and short. Vein 2r in fore wing joining cell 2Rs at basal 0.4; cu-a joining cell 1M close to vein 1M. Abdominal terga 1–2 and posterior margin of terga 3–4 with dense and long hairs, other terga with sparse and short hairs. Sternites and ventral side of abdominal terga with spare and long hairs. Ovipositor sheath 0.8 times as long as metatarsomere 1 and 2 combined, apical margin roundish in lateral view (Figure 2P), acute at apex in dorsal view (Figure 2K). Lancet with 45 serrulae (Figure 2O), middle serrulae as Figure 2R, annular spine bands narrow, membranous area between serrulae roundly protruding, middle serrulae subtruncate at apex, with 5–6 proximal and 4–5 distal subbasal teeth (Figure 2R).

Male: Body length 21–22 mm (Figure 1B); body color and structure similar to female except for following parts: labrum broad and large; anterior margin of clypeus arc-shape, without incision; metathorax and abdominal tergum 1 entire black; subgenital plate slightly broader than long, apical margin round; apex of each sternite with clear middle incision, both sides roundly arcuate. Penis valve shown in Figure 2S, gonoforcep shown in Figure 2T.

Holotype. Female (CSCS13010_Lab001). China: Hunan Province, Wugang County, Mt. Yun, Yunfengge alt. 1380 m, 26°38.630' N, 110°37.299' E, April 13, 2013, Zejian Li leg..

Paratypes: 17 Females, 15 Males (CSCS13010_Lab002–033). Collecting information as the holotype. 18 Females, 10 Males (CSCS13015_Lab034–061), locality and collector as the holotype, April 15, 2013. 45 Females, 17 Males (CSCS13014_Lab062–123), locality and collecting time as the holotype, Liwei Qi, Biao Chu leg. 36 Females, 51 Males (CSCS11009_Lab124–210), China: Hunan Province, Wugang County, Mt. Yun, Shengli Temple, alt. 1145 m, 26°38.859' N, 110°37.026' E, April 18-22, 2011, Zejian Li, Liwei Qi leg. 17 Females, 22 Males (CSCS05001_Lab211–249), China: Hunan Province, Wugang County, Mt. Yun, alt. 800-1100 m, April 24-26, 2005, Meicai Wei, Shaobing Zhang, Wei Xiao leg. One Male (CSCS1999001_Lab250), China: Hunan Province, Wugang County, Mt. Yun, alt. 1300 m, April 3, 1999, Wei Xiao leg. Two Females, six Males (LSAF18029_Lab251–258), China: Zhejiang Province, Lin'an City, Mt. Tianmu, alt.1506, 30.349°N, 119.424°E, April 19, 2018,

Zejian Li, Mengmeng Liu leg. One Females (LSAF17053_Lab259), locality and collector as the former, April 16, 2017. One Females, 26 males (LSAF17054_Lab259–285), locality as the former, April 17, 2017, Tingting Ji leg. Four Females, two Males (CSCS18006_Lab286–291), China: Hunan Province, Wugang County, Mt. Yun, alt. 1124 m, 26°38.059' N, 110°37.017' E, April 03, 2018, Meicai Wei, Gengyun Niu, Hannan Wang leg. Seven Females, one Males (CSCS18007_Lab292–299), locality as the former, alt. 1129 m, 26°39.003' N, 110°37.027' E, April 04, 2018, Meicai Wei, Hannan Wang leg.

Distribution. China (Hunan, Zhejiang)

Etymology: The specific name of the new species refers to the distribution area, China.

Remarks. This new species is similar to *L. zaraeoides* (Malaise, 1939) comb. nov., but differs from the latter in the following characters: clypeal notch deep, depth about 1/2 length of clypeus; between clypeus and supraclypeal area with distinct upper furrow; long hairs on gena 3.5 times longer than diameter of lateral ocellus, longer than transversal radius of an eye; long hairs on mesopleuron about 4.5 times longer than diameter of lateral ocellus; abdominal tergum 1 largely black.

***Labriocimbex zaraeoides* (Malaise, 1939) comb. nov**.

Trichiosoma zaraeoides Malaise, 1939: 16–17.

Distribution. Northern Burma

Remarks. This species is similar to *L. sinica* Yan & Wei sp. nov., but differs from the latter in the following characters: clypeal notch shallow, depth about 1/4 length of clypeus; between clypeus and supraclypeal area flat, upper furrow of clypeus absent; long hairs on gena 2.5 times diameter of lateral ocellus, shorter than transversal radius of an eye; long hairs on mesopleuron about 3.5 times longer than diameter of lateral ocellus; abdominal tergum1 largely yellow brown.

General features of the *L. sinica* mitochondrial genome

We sequenced the complete mitochondrial genome of *L. sinica* (GenBank accession no. MH136623), a typical set of 37 genes, including 13 PCGs, 22 tRNAs and two rRNAs. Most of the genes were located on the J strand except for four PCGs (*ND1*, *ND4*, *ND4L* and *ND5*), two rRNAs and seven tRNAs (Table 2).

A total of 14 pairs of genes were directly adjacent, without overlapping or intergenic nucleotides. The total length of the intergenic regions was 268 bp in 18 locations with a size ranging from 1 to 50 bp (Table 2). The longest was located between *trnH* and *ND4*, while the second longest was 45 bp located between *rrnS* and *trnM*. In comparison with the mitochondrial genome of *T. anthracinum* and *C. lateralis* (Song *et al*. 2016; Doğan and Korkmaz, 2017), there were differences in the length of intergenic spacers and locations. The longest (414 bp) was located at the start of the mitochondrial genome before *trnY* in *T. anthracinum*. The longest length of the intergenic spacers was 345 bp located between the *ND6* and *CYTB* genes in *C. lateralis*. We found that homologous searches on the longest intergenic region of *L. sinica* revealed ~~no significant similarity to~~ any identified Symphyta sequence.

There were in total 32 overlapping nucleotides between neighboring genes in six locations, and the range of length of the overlapping sequence is from 3 to 14 bp: *trnM* and *trnQ*, *ATP8* and *ATP6*, *ND4* and *ND4L*, *trnN* and *trnS2*, and *ATP6* and *COIII*; and the longest was 14 bp between *ATP6* and *COIII* (Table 2). The common motifs such as: ATGATAA between *ATP8* and *ATP6*, and ATGTTAA between *ND4* and *ND4L*, which also exist in *T. anthracinum*, and are not found in *C. lateralis*, are common features of many other insect mitochondrial genomes (Song *et al*., 2016; Doğan and Korkmaz, 2017).

Protein-coding genes and codon usage

The mitochondrial genome of *L. sinica* contains 13 PCGs, and its length is 12456 bp, accounting for 80.86% of the total length (Table 3). All PCGs were initiated by ATN codons: two genes (*CYTB* and *ND6*) used ATA start codons, six genes (*COI*, *ND3*, *ND5*, *ND4*, *ND4L* and *ND1*) started with ATT, four genes (*ND2*, *COII*, *COIII* and *ATP6*) were initiated with ATG, and one gene (*ATP8*) started with ATC. The stop codons were generally TAA; *ND2*, *COI*, *COII*, *ATP8*, *ATP6*, *COIII*, *ND3*, *ND4*, *ND4L*, *ND6*, *CYTB*, and *ND1* ended with TAA, and only *ND5* ended with T (Table 2).

The codon usage of *L. sinica* also shows a significant bias towards A/T, Leu, Ile, Phe and Met, which were the most frequently used amino acids. TTA-Leu showed the highest RSCU of 5.04 (Table 4). Comparisons of the RSCU with those of *C. lateralis* and *T. anthracinum* showed a similar pattern for codon usage bias and reflected a significant correlation between codon preference and nucleotide composition, that is similar to other symphytan species (Dowton *et al*., 2009; Wei *et al*., 2010; 2015; Korkmaz *et al*., 2015, 2016, 2017; Song *et al*., 2015, 2016; Niu *et al*., 2018; Du *et al*., 2018; Ma *et al*., 2019; Tang *et al*., 2019). Codons rich in C and G, CGC-Arg was absent, CGG-Arg, GGC-Gly, AGC-Ser and ACG-Thr, were rarely used, which is similar to both cimbicid mitochondrial genomes (Table 4). The ratio can be calculated by rate of G + C rich codons (Pro, Ala, Arg, and Gly and A + T rich codons (Phe, Ile, Met, Tyr, Asn, and Lys), and it is 0.28 in *L. sinica*, which is similar to those of other symphytan species (0.28–0.31) (Korkmaz *et al*., 2015). The translation, initiation, and termination signals as well as the codon usage of the *L. sinica* mitochondrial genome do not display any unusual characteristics.

Gene rearrangement and nucleotide composition

Compared with the putative ancestral mitochondrial genome of insects, we detected several rearrangement events are observed in three tRNA gene clusters in *L. sinica* (Figure 3), The first rearrangement event is found the clusters of *trnI-trnQ-trnM*, *trnM* and *trnQ* was found to swap positions, in addition, *trnM-trnQ* was translocated from the *trnI-trnQ-trnM* cluster to a downstream position from *rrnS*; which have not been reported from symphytan mitogenomes to date (Du *et al*., 2018; Niu *et al*., 2018; Ma *et al*., 2019; Tang *et al.*, 2019). The second event is corresponding to the remote inversion of *trnY* and the translocation of *trnC* from a location between *trnW* and *cox1* to upstream of *trnI*, which has great similarity to the gene order and rearrangement events observed in *T. anthracinum*. In comparisons of the mitochondrial genomes of all sequenced species in Symphyta (Du *et al*., 2018; Niu *et al*., 2018; Ma *et al*., 2019; Tang *et al.*, 2019), the rearrangement of cluster of *trnW-trnC-trnY* of these two species are different from

those in other Symphyta species. Its (*trnW-trnC-trnY*) position is identical with the putative ancestral pattern in *C. lateralis.* The arrangement of cluster of *trnW-trnC-trnY* appears to be mostly conserved in almost all known symphytan mitogenomes, except for representative cimbicid species. The last event is only found in the TP cluster of *L. sinica*, and here *trnT* is inverted. The region from *COI* to *rrnS* of all sequenced species in Cimbicidae is conserved. The mitochondrial genome of the Symphyta species appears to be more conserved than that of the Apocrita (Song *et al*., 2016; Wei *et al*., 2014).

Similar to previously reported symphytan mitochondrial genomes (Ma *et al*., 2019; Doğan and Korkmaz, 2017; Song *et al*., 2016), the nucleotide compositions of *L. sinica* (43.5% A, 37.7% T, 7.7% G and 11.1% C) were biased towards A and T, with an average 81.2% A+T content; a stronger AT bias was found in the N strand (81.4% A+T content) than in the J strand (78.7%) (Table 3).

Further analysis of the PCGs indicated that the third codon position demonstrates the highest A + T content (93.5%), in agreement with symphytan mitochondrial genomes (Ma *et al*., 2019; Doğan and Korkmaz, 2017; Song *et al*., 2016). The gene with the highest A + T content was *ATP8* with 88.3% (Table 3). Here we observed that the AT-skew was slightly positive (0.0714), and the GC-skew was negative (−0.1809) when considering the whole genome (Table 3). This indicates that the occurrence of A is higher than that of T, and the occurrence of C is higher than that of G, which is a general phenomenon observed in all reported symphytan mitochondrial genomes, except for those of *Tremex columba* and *Xiphydria* sp. (Ma *et al*., 2019; Doğan and Korkmaz, 2017; Song *et al*., 2016; Wei *et al*., 2015; Castro and Dowton, 2005; Dowton *et al*., 2009). However, a deviation was found in the PCGs of *L. sinica*, in terms of AT-skew (−0.1389) and GC-skew (0.0348), which also occurred in both *C. lateralis* and *T. anthracinum*. This deviation can exert influences on the selection forces acting on the PCG codon positions, in accordance with study by Korkmaz (2015).

Transfer RNA genes

The mitochondrial genome of *L. sinica* contains 22 tRNAs, and 15 tRNAs were encoded by the J strand, while the remaining tRNAs were encoded by the opposite N-strand. The secondary structures of the tRNAs were predicted using Mitos. The result indicates that all tRNAs folded into a common clover-leaf structure, except *AGN*, where the dihydrouridine (DHU) arm was missing (Figure 4). The size of the tRNAs ranged from 64 bp (*trnG*) to 71 bp (*trnC, trnK*), and this usually depends on the length of the variable loop, TΨC loop and D-loops (Clary and Wolstenholme, 1985). The DHU arm was 3–4 bp, the AC arm was 4–5 bp, and the TΨC arm varied from 4–5 bp, while the amino acid acceptor (AA) stem and anticodon (AC) loops were conserved at 7 bp in all of the tRNA genes. However, the TΨC loops were less consistent, ranging from 1–9 bp, the D-loops ranged from 3–7 bp, and the variable loops ranged from 3–5 bp (Figure 4).

Mismatches occur in the mitochondrial tRNA gene, and mainly occur in the DHU arm, AA arm and AC arm, and sometimes occur in the TΨC arm. A total of 16 unmatched base pairs were scattered among the following tRNA genes, including 12 G-U mismatched pairs occurring in *trnA*, *trnD*, *trnQ*, *trnG*, *trnH*, *trnL1*, *trnP*, *trnF*, and *trnY*, and four U-U mismatches occurring in

trnR, *trnT* and *trnL1*. The number of mismatches were 24 (12 G–U pairs, five U–U pairs, three A–A pairs, two A–C pairs, one A–G pair and one C–U pair) in *C. lateralis* (Figure 5, adapted from Doğan and Korkmaz, 2017), and 18 (15 G–U pairs, two U–U pairs and 1 A–C pair) in *T. anthracinum* (Figure 4; adapted from Song *et al*., 2016), which is typical for Hymenoptera (Ma *et al*., 2019; Doğan and Korkmaz, 2017; Song *et al*., 2016; Castro and Dowton, 2005; Dowton *et al*., 2009b). The phenomenon of aberrant mismatches, loops, or extremely short arms for tRNA has been shown to be common in metazoan mitochondrial genomes (Wolstenholme, 1992). In addition, there were some tRNA structural differences between *L. sinica* and *T. anthracinum*. As shown in Figure 4, in *T. anthracinum*, *trnQ* and *trnM* demonstrated some significant structural differences mainly occurring in the stems and loops of DHU and TΨC (such as: *trnE*, *trnH*, *trnW*, and *trnY*). The identified anticodons were almost identical to those of the cimbicid species, with the exception of the anticodon of *trnS1* (*Agn*), which is UCU in *L. sinica* and *T. anthracinum*, and GUA in *C. lateralis*, as well as this is true of all previously reported of Symphyta (Ma *et al*., 2019; Doğan and Korkmaz, 2017; Song *et al*., 2016; Castro and Dowton, 2005; Dowton *et al*., 2009b).

Ribosomal RNA genes

The rRNA gene of *L. sinica rrnL* was 1341 bp in length with 84.2% A+T content, while *rrnS* was 791 bp in length with 84.1% A+T content (Table 3). This was in a comparable range to homologous genes in *T. anthracinum* (1351 bp; 800 bp) and *C. lateralis* (1351 bp; 493 bp *rrnS* partial gene), and also identical to all reported hymenopteran species (Gillespie, 2006; Wei *et al*., 2010, Doğan and Korkmaz, 2017; Song *et al*., 2016; Korkmaz *et al*., 2015). These genes accounted for mitochondrial genes essential for the translation of messenger RNA into mitochondrial RNA. Both genes were encoded on the N-strand (Table 2).

Similar to the known symphytan mitochondrial genomes, the *rrnL* gene is positioned between *trnL1* and *trnV* in three species of Cimbicidae (Figure 3). The predicted structure of *rrnL* in *L. sinica* is consistent with the observed pattern in *C. lateralis* and *T. anthracinum*, whereby 45 helices belonging to five domains were identified in those species (Figure 6, 7). Domain III is absent as in other arthropods (Korkmaz *et al*., 2015), and domain II is variable in base composition, forming a long stem with a big loop structure in the area II terminal. Domains IV and V are more conserved within the Tenthredinidae than domains I, II and VI. Eight helices (H563, H579, H777, H822, H2023, H2043, H2455 and H2547) of *rrnL* are highly conserved. The H183, H991, H1057, H1196 and H2077 helices display helical length and loop size/structure variability within three cimbicid *rrnL* genes (Figures 6, 7).

The *rrnS* secondary structure of *L. sinica* is between *trnV* and an AT-rich region, and contains four domains and 26 helices (Figure 8). Compared with *T. anthracinum,* it is significantly different in terms of base composition in domain II (Figure 8). Specifically, H47 is variable among the different hymenopteran species, having a large loop. The loop size is variable and determined by overall *rrnS* length, except for in the cephid species (Gillespie, 2006; Wei *et al*., 2010, Doğan and Korkmaz, 2017; Song *et al*., 2016; Korkmaz *et al*., 2015). The structures of

domains I and II of *C. lateralis* are missing, so they cannot be compared with those of *L. sinica*, but the structures are similar in domains III and IV (Figure 9). In *rrnS*, domain III and domain VI were more conserved within Tenthredinidae than domains I and II (Figures 8, 9). The anticodons of all predicted tRNAs are identical with those of other symphytan mitochondrial genomes.

Phylogenetic relationships

To investigate the phylogenetic relationship of *L. sinica* within the Symphyta, we used 36 species of Hymenoptera (34 Symphyta, two Apocrita) and four non-hymenopteran outgroups (Mecoptera, Diptera, Megaloptera, Coleoptera) for which mitochondrial genomes had previously been sequenced, or for which the mitochondrial genomes newly sequenced in the current study were available (Table 1). Phylogenetic relationships within the suborder Symphyta were reconstructed using both BI and ML analyses (Figure 9 and 10, respectively). The 34 Symphyta species can be divided into 11 families: there were 13 species belonging to Cephidae (Dowton *et al.*, 2009; Korkmaz *et al.*, 2015, 2016, 2017, 2018), six species belonging to Tenthredinidae (Wei *et al.*, 2014, 2015; Song *et al.*, 2015, 2016; Niu *et al.*, in press), three species belonging to Cimbicidae (Song *et al.*, 2016; Doğan and Korkmaz, 2017), three species belonging to Megalodontesidae (Niu *et al.*, 2018; Tang *et al.*, 2019), two species belonging to Argidae (Du *et al.*, 2018; Tang *et al.*, 2019), two species belonging to Pamphiliidae (Niu, *et al.*, 2018; Ma *et al.*, 2019), one species belonging to Pergidae (Castro and Dowton, 2005), one species belonging to Orussidae (Dowton *et al.*, 2009), one species belonging to Xyelidae (Ma *et al.*, 2019), one species belonging to Siricidae (Ma *et al.*, 2019), and one species belonging to Xiphydriidae (Ma et al., 2019).

The topologies of the two phylogenetic trees were almost identical, especially the clade consisting of (Tenthredinidae + Cimbicidae) + (Argidae + Pergidae), which was very stable with the highest nodal supports. The recovered trees supported a relationship consisting of Xyelidae + (Tenthredinoidea + ((Megalodontesidae + Pamphiliidae) + (Xiphydriidae + (Cephidae + (Orussidae + (Apocrita + Siricidae)))))) in the Hymenoptera. This relationship is also supported by both molecular (Malm and Nyman 2014; Peters *et al.*, 2017) and morphological studies (Schulmeister *et al.*, 2002, Vilhelmsen, 2001; 2015).

Both trees indicate that *L. sinica* is grouped with *T. anthracinum* and *C. lateralis* to form a sister group, that was shown to be the monophyly of Cimbicidae. We suggest that *Labriocimbex* belongs to the tribe of Trichiosomini as well as *Trichiosoma.* Additionally, we demonstrated that mitochondrial genome sequences can be used to solve phylogenetic relationships at different taxonomic levels within Symphyta.

CONCLUSIONS

Labriocimbex gen. nov. was regarded as a connecting link between the genera *Trichiosoma* Leach and *Pseudoclavellaria* Schultz. Most of the characteristics of the new genus suggest to place it in the tribe Trichiosomini. The most important of these characteristics include: clypeus black and broadly emarginated, the labrum triangular and tapering toward apex, the apical anal cell about 2 times as long as basal anal cell, the antenna with an apical club unsegmented and with 5 antennomeres before the club, the hind femora close to each other and without ventral

dent, the very large tarsal pulvilli, the long and dense hairs covering head, thorax, base of abdomen and legs. These characteristics separate this new genus from *Trichiosoma* and *Pseudoclavellaria*, and help to distinguish this new genus and new species.

The complete mitochondrial genome of *L. sinica* was obtained and was found to have a length of 15405 bp and a typical set of 37 genes (22 tRNAs, 13 PCGs, two rRNAs). The secondary structures of the 22 tRNAs and two rRNAs resemble those of Symphyta. In comparison with the structures of *T. anthracinum* and *C. lateralis*, some helices were highly variable in *rrnL* and *rrnS* in *C. lateralis*, and two tRNAs (*trnQ* and *trnM*) were missing in *T. anthracinum*. Phylogenetic reconstruction based on mitochondrial genomes (13 PCGs and two rRNAs) showed similarly high levels of support (100%) in both BI and ML analyses that the family Cimbicidae is a sister group of ((*L. sinica* + *T. anthracinum*) + *C. lateralis*).

The same results were obtained using two different analytical methods, and our findings agree with traditional morphological classification and recent molecular studies. The tree topologies confirm the newly sequenced taxonomic positions of the Cimbicidae species within the superfamily Tenthredinoidea, and reveal a relationship of (Tenthredinidae + Cimbicidae) + (Argidae + Pergidae) in Symphyta with strong nodal supports.

ACKNOWLEDGMENTS

Thank Dr. Zejian Li for the collection of specimens from Lishui Academy of Forestry, Zhejiang province, China. The members of Lab of Insect Systematics and Evolutionary Biology (LISEB) from Central South University of Forestry and Technology are thanked for their contributions in laboratory work. We thank all the reviewers for their comments.

REFERENCES

Abe M, Smith DR. 1991. The genus-group names of Symphyta (Hymenoptera) and their type species. *Esakia* 31: 1–115.

Beard CB, Hamm DM, Collins FH. 1993. The mitochondrial genome of the mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons with mitochondrial sequences of other insects. *Insect Molecular Biology* 2(2):103–124.

Benson RB. 1938. On the classification of Sawflies (Hymenoptera, Symphyta). *Transactions of the Entomological Society of London* 87: 353–384.

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. *Molecular Phylogenetics and Evolution* 69: 313–319.

Blank SM, Groll EK. Liston A, Prous M, Taeger A. 2012. ECatSym - Electronic World Catalog of Symphyta (Insecta, Hymenoptera). Program Version 4.0 Beta, Data Version 39. Digital Entomological Information, Müncheberg. URL http://sdei.senckenberg.de/ecatsym.

Castro LR, Dowton M. 2005. The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: Symphyta: Pergidae). *Molecular Phylogenetics and Evolution* 34(3): 469–479.

570 Dierckxsens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle
571 genomes from whole genome data. *Nucleic Acids Research* 45:18.
572 De Rijk P, Wuyts J, De Wachter R. 2003. RnaViz 2: an improved representation of RNA
573 secondary structure. *Bioinformatics* (Oxford, England) 19: 299–300.
574 Deng TJ. 2000. A systematic study of Cimbicidae from China (Hymenoptera: Symphyta) (In
575 Chinese, abstract in English). Central south forest university. Thesis for Master degree.
576 Zhuzhou, Hunan province, China.
577 Darty K, Denise A, Ponty Y. 2009. VARNA: Interactive drawing and editing of the RNA
578 secondary structure. *Bioinformatics* 25: 1974–1975.
579 Doğan Ö, Korkmaz EM. 2017. Nearly complete mitogenome of hairy sawfly, *Corynis lateralis*
580 (Brullé, 1832) (Hymenoptera: Cimbicidae): rearrangements in the IQM and ARNS1EF gene
581 clusters. *Genetica* 145: 341–350.
582 Dowton M, Cameron SL, Austin AD, Whiting MF. 2009. Phylogenetic approaches for the
583 analysis of mitochondrial genome sequence data in the Hymenoptera – a lineage with both
584 rapidly and slowly evolving mitochondrial genomes. *Molecular Phylogenetics and Evolution*
585 52: 512–519.
586 Du S, Niu G, Nyman T, Wei M. 2018. Characterization of the mitochondrial genome of *Arge*
587 *bella* Wei & Du sp. nov. (Hymenoptera: Argidae). *PeerJ* 6: e6131.
588 Foster PG, Jermiin LS, Hickey DA. 1997. Nucleotide Composition Bias Affects Amino Acid
589 Content in Proteins Coded by Animal Mitochondria. *Journal of Molecular Evolution* 44: 282–
590 288.
591 Gillespie JJ, Johnston JS, Cannone JJ, Gutell RR. 2006. Characteristics of the nuclear (18S, 5.8S,
592 28S and 5S) and mitochondrial (12S and 16S) rRNA genes of *Apis mellifera* (Insecta:
593 Hymenoptera): structure, organization, and retrotransposable elements. *Insect molecular*
594 *biology* 15: 657–686.
595 Gauld I, Bolton B.1988. The Hymenoptera. British Museum (Natural History). Oxford
596 University Press, Oxford.
597 Hahn C, Bachmann L, Chevreux B. 2013. Reconstructing mitochondrial genomes directly from
598 genomic next-generation sequencing reads – a baiting and iterative mapping approach. *Nucleic*
599 *Acids Research* 41:e129–e129.
600 Heraty J, Ronquist F, Carpenter JM. Hawks D, Schulmeister S, Dowling AP, Murray D, Munro
601 J, Wheeler WC, Schiff N, Sharkey M. 2011. Evolution of the hymenopteran megaradiation.
602 *Molecular Phylogenetics and Evolution* 60: 73–88.
603 Hyde J, Cooper SJB, Munguia P, Humphreys WF, Austin AD. 2018. The first complete
604 mitochondrial genomes of subterranean dytiscid diving beetles (*Limbodessus and Paroster*)
605 from calcrete aquifers of Western Australia. *Australian Journal of Zoology* 65:283–291.
606 Isaka Y, Sato, T. 2015. Was species diversification in Tenthredinoidea (Hymenoptera:
607 Symphyta) related to the origin and diversification of angiosperms? *The Canadian*
608 *Entomologist* 147: 443–458.

609 Korkmaz EM, Aydemir HB, Temel B, Budak M, Başıbüyük HH. 2017. Mitogenome evolution
610 in Cephini (Hymenoptera: Cephidae): Evidence for parallel adaptive evolution. *Biochemical*
611 *Systematics and Ecology* 71: 137–146.
612 Korkmaz EM, Budak M, Ördek MN, Başibüyük HH. 2016. The complete mitogenomes of
613 Calameuta filiformis (Eversmann, 1847) and Calameuta idolon (Rossi, 1794) (Hymenoptera:
614 Cephidae): The remarkable features of the elongated A+T rich region in Cephini. *Gene* 576:
615 404–411.
616 Korkmaz EM, Doğan Ö, Budak M, Başibüyük HH. 2015. Two nearly complete mitogenomes of
617 wheat stem borers, *Cephus pygmeus* (L.) and *Cephus sareptanus* Dovnar-Zapolskij
618 (Hymenoptera: Cephidae): An unusual elongation of *rrnS* gene. *Gene* 558: 254–264.
619 Korkmaz EM, Doğan Ö, Durel BS, Temel Altun B, Budak M, Başıbüyük HH. 2018.
620 Mitogenome organization and evolutionary history of the subfamily Cephinae (Hymenoptera:
621 Cephidae). *Systematic Entomology* 43(3):606–618.
622 Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis
623 version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870.
624 Klopfstein S, Vilhelmsen L, Heraty JM, Sharkey M, Ronquist F. 2013. The Hymenopteran Tree
625 of Life: Evidence from Protein-Coding Genes and Objectively Aligned Ribosomal Data).
626 *PLoS ONE* 8(8):e69344.
627 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7:
628 improvements in performance and usability. *Molecular Biology and Evolution* 30: 772–80.
629 Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Qi P, Liu Y. 2012. SOAPdenovo2:
630 an empirically improved memory-efficient short-read de novo assembler. *Giga science* 1: 18.
631 Lanfear R, Calcott B, Ho SYW, Guindon S. 2012. Partition Finder: Combined selection of
632 partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology*
633 *and Evolution* 29:1695–1701.
634 Malm T, Nyman T. 2015. Phylogeny of the symphytan grade of Hymenoptera: new pieces into
635 the old jigsaw (fly) puzzle. *Cladistics* 31: 1–17.
636 Malaise R.1934. Schwedisch-chinesische wissenschaftliche Expeditionnach den nordwestlichen
637 Provinzen Chinas unter Leitung vonDr. Sven Hedin und Prof. Sü Ping-Chang. Insekten
638 gesammelt vomschwedischen Arzt der Expedition Dr. David Hummel 1927–1930.
639 Malaise R. 1939. The Genus *Leptocimbex* Sem., and some other Cimbicidae. *Entomologisk*
640 *Tidskrift*, 60: 1–28.
641 Misof B. Fleck G. 2003. Comparative analysis of mt LSU rRNA secondary structures of
642 Odonates: Structural variability and phylogenetic signal. *Insect Molecular Biology* 12(6):535–
643 547.
644 Ma Y, Zheng BY, Zhu JC, Achterberg CV, Tang P, Chen XX. 2019. The first two mitochondrial
645 genomes of wood wasps (Hymenoptera: Symphyta): Novel gene rearrangements and higher-
646 level phylogeny of the basal hymenopterans. *International Journal of Biological*
647 *Macromolecules* 123: 1189–1196.

Nawrocki EP. 2009. Structural RNA homology search and alignment using covariance models. *Washington University in St. Louis.*

Niu G, Korkmaz EM, Doğan Ö, Zhang YY, Aydemir MN, Budak M, Du SY, Başıbüyük HH, Wei MC. 2018. The first mitogenomes of the superfamily Pamphilioidea (Hymenoptera: Symphyta): Mitogenome architecture and phylogenetic inference. *International Journal of Biological Macromolecules* 124: 185–199.

Perna NT, Kocher TD. 1995. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. *Journal of Molecular Evolution* 41: 353–358.

Peng Y, Chen B, LI TJ. 2017. Sequencing and analysis of the complete mitochondrial genome of *Parapolybia crocea* (Hymenoptera:Vespidae). *Acta Entomologica Sinica*60(4): 464–474.

Peters RS, Krogmann L, Mayer C, Donath A, Gunkel S, Meusemann K, Kozlov A, Podsiadlowski L, Petersen M, Lanfear R, Diez PA, Heraty J, Kjer KM, Klopfstein S, Meier R, Polidori C. Schmitt T. Liu S, Zhou X, Wappler T, Rust J, Misof B, Niehuis O. 2017. Evolutionary history of the Hymenoptera, *Current Biology* 27: 1–6.

Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard M a, Huelsenbeck JP. 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Ross HH. 1937. A Generic Classification of the Nearctic Sawflies (Hymenoptera: Symphyta). *Illinois biological monographs* Urbana 15(2): 1–173.

Ross HH. 1945. Sawfly genitalia: terminology and study techniques. *Entomological News* 61(10): 261–268.

Song SN, Tang P, Wei SJ, Chen XX. 2016. Comparative and phylogenetic analysis of the mitochondrial genomes in basal hymenopterans. *Scientific Reports* 6(1): 20972–20972.

Song SN, Wang ZH, Li Y, Wei SJ, Chen XX. 2015. The mitochondrial genome of *Tenthredo tienmushana* (Takeuchi) and a related phylogenetic analysis of the sawflies (Insecta: Hymenoptera). *Mitochondrial DNA*: 1–2.

Schmidt S, Walter GH, Grigg J, Moore CJ. 2006. Sexual communication and host plant associations of Australian pergid sawflies (Hymenoptera: Symphyta: Pergidae). *Recent Sawfly Research: Synthesis and Prospevts*, pages 173–194.

Schmidt S, Walter GH. 2014. Young clades in an old family: major evolutionary transitions and diversification of the eucalypt feeding pergid sawflies in Australia (Insecta, Hymenoptera, Pergidae). *Molecular Phylogenetics and Evolution*74: 111–121.

Sharkey MJ, Carpenter JM, Vilhelmsen L, Heraty J, Dowling APG, Schulmeister S, Murray D, Andrew R, Ronquist F, Krogmann L, Wheeler WC. 2012. Phylogenetic relationships among superfamilies of Hymenoptera Michael. *Cladistics* 28: 80–112

Siebert S, Backofen R. 2005. MARNA: multiple alignment and consensus structure prediction of RNAs based on sequence structure comparisons. *Bioinformatics* 21: 3352.

Schulmeister S.2003. Simultaneous analysis of basal Hymenoptera (Insecta): introducing robust-choice sensitivity analysis. *Biological Journal of the Linnean Society. Linnean Society of London*79: 245–275.

Taeger A, Blank SM, Liston AD. 2010. World catalog of Symphyta (Hymenoptera). *Zootaxa* 2580:1–1064.

Takeuchi K. 1952. A Generic Classification of the Japanese Tenthredinidae (Hymenoptera: Symphyta). *Kyoto*1–90.

Tang P, Zhu JC, Zheng BY, Wei SJ, Sharkey MJ, Chen XX, Vogler AP. 2019. Mitochondrial phylogenomics of the Hymenoptera. *Molecular Phylogenetics and Evolution* 131: 8–18

Viitasaari M. 2002. Sawflies (Hymenoptera, Symphyta) I. A review of the suborder, the Western Palaearctic taxa of Xyeloidea and Pamphilioidea. Helsinki *Tremex* Press Ltd.

~~Vilhelmesen~~ L. 1997. The phylogeny of lower Hymenoptera (Insecta), with a summary of the early evolutionary history of the order. *Journal of Zoological Systematics and Evolutionary Research*35:49–70.

Vilhelmsen L. 2001. Phylogeny and classification of the extant basal lineages of the Hymenoptera (Insecta), *Zoological Journal of the Linnean Society*131:393–442.

Vilhelmsen L. ~~2018.~~ Giant sawflies and their kin: morphological phylogeny of Cimbicidae (Hymenoptera). *Systematic Entomology*. DOI: 10.1111/syen.12314

Vaidya G, Lohman DJ, Meier R. 2011. Sequence Matrix: Concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27:171–180.

Vilhelmsen L.2015. Morphological phylogenetics of the Tenthredinidae (Insecta: Hymenoptera). *Invertebrate Systematics* 29:164–190.

Wei MC. Nie HY.1997. Studies on the biogeography of Tenthredinoidea (Hymenoptera) I. Analysis of the geographical distribution of Tenthredinoidea at family level. *Entomotaxonomia*, ~~pages~~ 127–132.

Wei SJ, Niu FF, and Du BZ. 2014. Rearrangement of trnQ-trnM in the mitochondrial genome of *Allantus luctifer* (Smith) (Hymenoptera: Tenthredinidae) Rearrangement of trnQ-trnM in the mitochondrial genome of *Allantus*. *Mitochondrial DNA A DNA Mapp Seq Anal* 27(2):856–858.

Wei SJ, Wu QL, Liu W. 2015. Sequencing and characterization of the *Monocellicampa pruni* (Hymenoptera: Tenthredinidae) mitochondrial genome. *Mitochondrial DNA* 26:157–158.

Wei SJ, Shi M, Chen XX, Sharkey MJ, Achterberg CV, Ye GY, He JH. 2010. New views on strand asymmetry in insect mitochondrial genomes. *PLOSONE*, 5(9):1–10.

Wolstenholme DR. 1992. Animal mitochondrial DNA: structure and evolution. *International Review of Cytology* 141:173.

Wu QL, Li Q, Gu Y, Shi BC, van Achterberg C, Wei SJ, Chen XX. 2014. The complete mitochondrial genome of *Taeniogonalos taihorina* (Bischoff) (Hymenoptera: Trigonalyidae) reveals a novel gene rearrangement pattern in the Hymenoptera. *Gene* 543:76–84.

Yan YC, Niu GY, Lan BC, Wei MC. 2018. Review of *Leptocimbex formosanus* group (Hymenoptera: Cimbicidae) with two new Chinese species, *Entomological Research* 48: 372–383.

Yan YC, Wei MC. 2010. The first record of *Orientabia* malaise form China with description of a new species (Hymenoptera, cimbicidae). *Acta Zootaxonomica Sinica*, 35 (3): 627–630

Yan YC, Wei MC. 2013. A new species of the genus *Leptocimbex* Semenov-Tian-Shanskij (Hymenoptera, Cimbicidae) from Hunan, China. *Acta Zootaxonomica Sinica* 38: 130–133.

Yan YC, Wei MC. 2016. A new species and a new record of *Leptocimbex* Semenov (Hymenoptera: Cimbicidae) from China. *Entomotaxonomia*38: 227–234.

Yan YC, Xiao W, Wei MC. 2014. Two new species of *Leptocimbex* Semenov (Hymenoptera: Cimbicidae) from China. *Entomotaxonomia*36: 311–320.

Zhao Y, Zhang H, Zhang Y. 2017. Complete mitochondrial genome of *Neochauliodes parasparsus* (Megaloptera: Corydalidae) with phylogenetic consideration. *Biochemical Systematics & Ecology* 70:192–199.

Notes

Table 1. Summary information of symphytan mitochondrial genomes used in phylogenetic analyses

Table 2. Mitochondrial genome characteristics of *L. sinica*.

Table 3. Nucleotide composition of *L. sinica* mitochondrial genome.

Table 4. Codon usage of PCGs in mitochondrial genome of *L. sinica.* No., frequency of each codon; RSCU, relative synonymous condon usage.

Figure 1. *Labriocimbex sinica* Yan & Wei sp. nov.
A. Adult female, dorsal view. B. Adult male, dorsal view.

Figure 2. *Labriocimbex sinica* Yan & Wei, gen. et sp. nov.
A. Head of female, front view; B. Head of female, dorsal view; C. Head of female, lateral view; D. Hind tibia and tarsus; E. Mesopleuron; F. Left mandibles; G. Right mandibles; H. Palp; I. Femur of hind leg; J. Claw; K. Ovipositor sheath of female, dorsal view; L. Genital plate of female, ventral view; M. Spurs of hind tibia; N. Metanotum and abdominal terga1–2; O. Lancet; P. Ovipositor sheath of female, lateral view; Q. Lance; R. Middle serrulae; S. Penis valve; T. Gonoforcep; U. Antenna of female.

Figure 3. Mitochondrial genome organization of three cimbicid species referenced with the ancestral insect mitochondrial genomes. Genes transcribed from the J and N strands are shown with green and orange colours, respectively. Overlapping and intergenic regions are marked in yellow and blue circles. tRNA genes are denoted by a one-letter symbol according to the IPUC-IUB single-letter amino acid codes A+ T-rich region is marked in blue and tRNA genes are labelled by the single-letter amino acid code.

Figure 4. Predicted secondary structures for the 22 typical tRNA genes of *L. sinica* and *T. anthracinum* mitogenomes.
Base-pairing is indicated as follows: Watson–Crick pairs by lines, wobble GU pairs by dots and other noncanonical pairs by circles. Variable regions are presented in boxes with red (*L. sinica*) and blue (*T. anthracinum*) colours.

Figure 5. Predicted secondary structures for the 22 tRNA genes of C. *lateralis.*
Dashes indicate Watson-Crick base pairing and dots indicate G-U base pairing.

Figure 6. The predicted secondary structures of *rrnL* of *L. sinica* and *T. anthracinum.*
Tertiary interactions and base triples are connected by continuous lines. The numbering of helix follows Gillespie et al. (2006). Roman numbers refer to domain names. Dashes indicate Watson-Crick base pairing and dots indicate G-U base pairing. The helical variation among cimbicid species are presented in boxes with red (*L. sinica*) and blue (*T. Anthracinum*) colours.

Figure 7. *Corynis lateralis rrnL*.
Predicted *rrnL* secondary structure in *C. lateralis.* The numbering of helix follows Gillespie et al. (2006).Roman numbers refer to domain names.

Figure 8. The predicted secondary structures of *rrnS* of *L. sinica* and *T. ~~A~~nthracinum.*
Tertiary interactions and base triples are connected by continuous lines. The numbering of helix follows Gillespie et al. (2006). Roman numbers refer to domain names. Dashes indicate Watson-Crick base pairing and dots indicate G-U base pairing. The helical variation among cimbicid species are presented in boxes with red (*L. sinica*) and blue (*T. Anthracinum*) colours.

Figure 9. *Corynis lateralis rrnS*.
Predicted *rrnS* secondary structure in *C. lateralis.* The numbering of helix follows Gillespie et al. (2006).Roman numbers refer to domain names.

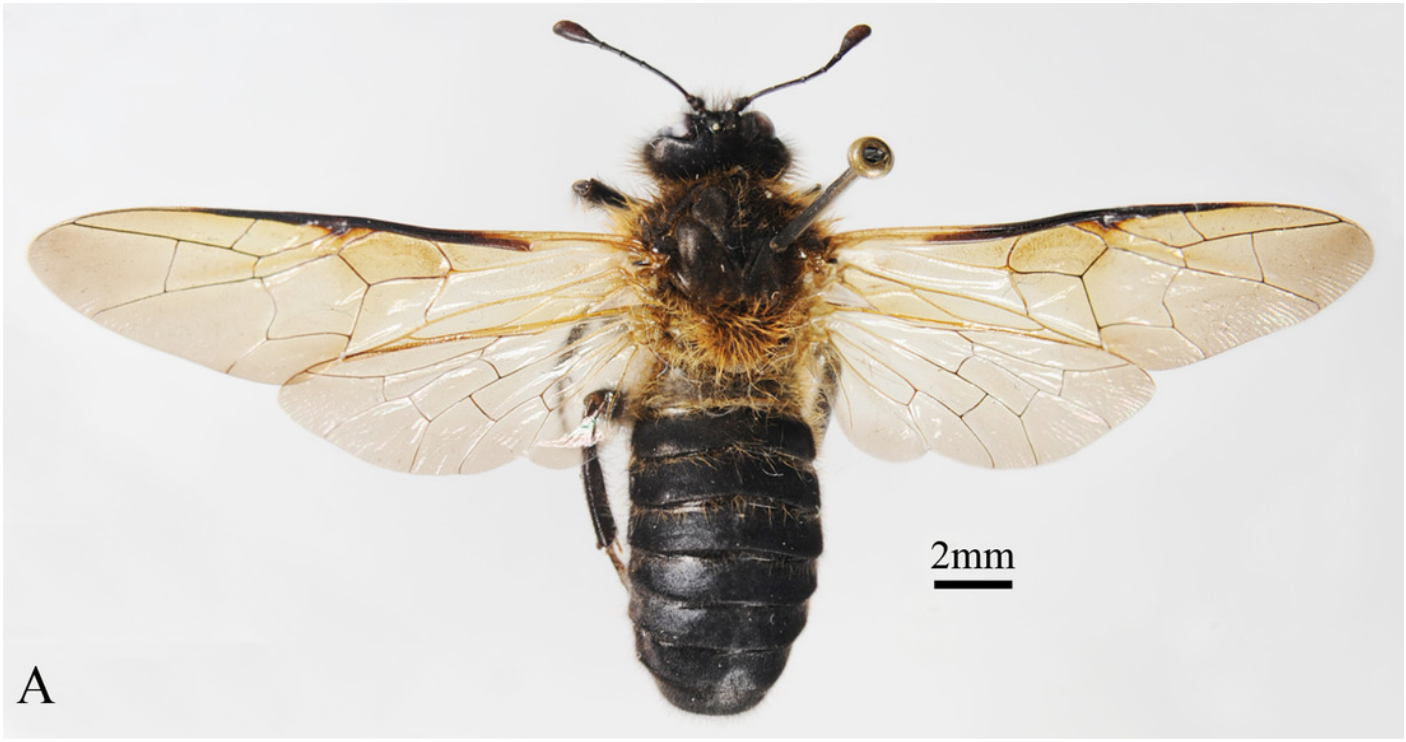
Figure 10. The phylogenetic tree of Symphyta ~~by iqtree~~.
Phylogenetic tree of Symphyta, based on ML analysis of the 13 PCGs (*ND2*, *COXI1*, *COX2*, *ATP8*, *ATP6*, *COX3*, *ND3*, *ND5*, *ND4*, *ND4L*, *ND6*, *CYT8*, and *ND1*) and 2 rRNAs (*rrnL* and *rrnS*) data set. Values branches represent maximum likelihood bootstrap clade frequency (CF) support. The scale bar indicates the number of substitutions per site.

Figure 11. The phylogenetic tree of Symphyta ~~by bayes~~.
Phylogenetic tree of Symphyta, based on BI analysis of the 13 PCGs (*ND2*, *COXI1*, *COX2*, *ATP8*, *ATP6*, *COX3*, *ND3*, *ND5*, *ND4*, *ND4L*, *ND6*, *CYT8*, and *ND1*) and 2 rRNAs (*rrnL* and *rrnS*) data set. Values branches represent BI posterior probability (PP) support. The scale bar indicates the numberof substitutions per site.

Figure 1

Labriocimbex sinica Yan & Wei sp. nov.

A. Adult female, dorsal view. B. Adult male, dorsal view.


2mm
A


2mm
B

Figure 2

Labriocimbex sinica Yan & Wei, gen. et sp. nov.

A. Head of female, front view; B. Head of female, dorsal view; C. Head of female, lateral view; D. Hind tibia and tarsus; E. Mesopleuron; F. Left mandibles; G. Right mandibles; H. Palp; I. Femur of hind leg; J. Claw; K. Ovipositor sheath of female, dorsal view; L. Genital plate of female, ventral view; M. Spurs of hind tibia; N. Metanotum and abdominal terga1–2; O. Lancet; P. Ovipositor sheath of female, lateral view; Q. Lance; R. Middle serrulae; S. Penis valve; T. Gonoforcep; U. Antenna of female.


A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U

Figure 3(on next page)

Mitochondrial genome organization of threecimbicid species referenced with the ancestral insect mitochondrial genomes.

Genes transcribed from the J and N stran ds are shown with green and orange colours, respectively. Overlapping and intergenic regions are marked in yellow and blue circles. tRNA genes are denoted by a one-letter symbol according to the IPUC-IUB single-letter amino acid codes A+ T-rich region is marked in blue and tRNA genes are labelled by the single-letter amino acid code.

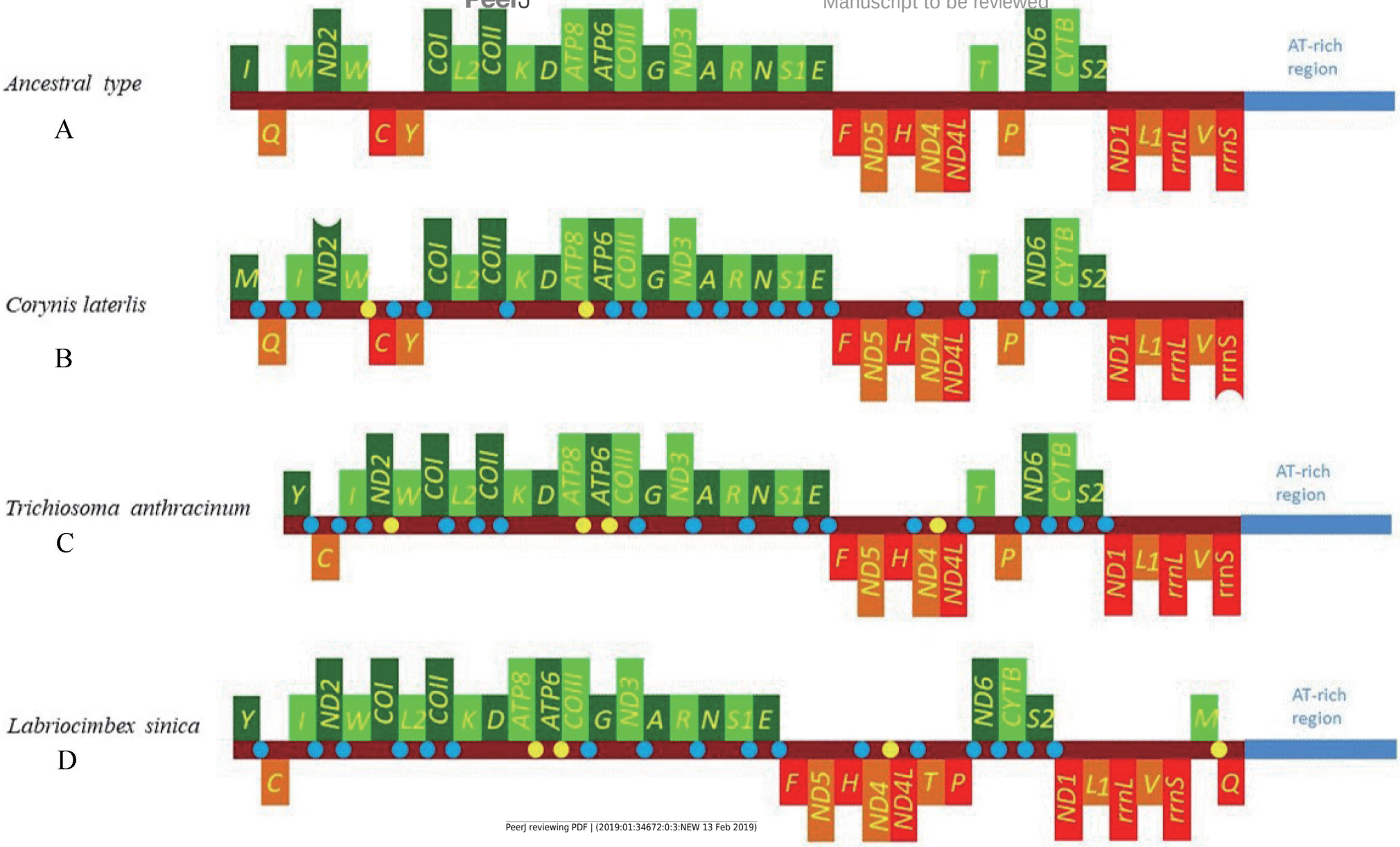

Ancestral type
A
I M ND2 W COI L2 COII K D ATP8 ATP6 COIII G ND3 A R N S1 E T ND6 CYTB S2
Q C Y F ND5 H ND4 ND4L P ND1 L1 rrnL V rrnS
AT-rich region
Corynis laterlis
B
M I ND2 W COI L2 COII K D ATP8 ATP6 COIII G ND3 A R N S1 E T ND6 CYTB S2
Q C Y F ND5 H ND4 ND4L P ND1 L1 rrnL V rrnS
Trichiosoma anthracinum
C
Y I ND2 W COI L2 COII K D ATP8 ATP6 COIII G ND3 A R N S1 E T ND6 CYTB S2
C F ND5 H ND4 ND4L P ND1 L1 rrnL V rrnS
AT-rich region
Labriocimbex sinica
D
Y I ND2 W COI L2 COII K D ATP8 ATP6 COIII G ND3 A R N S1 E ND6 CYTB S2 M
C F ND5 H ND4 ND4L T P ND1 L1 rrnL V rrnS Q
AT-rich region

Figure 4

Predicted secondarystructures for the 22 typical tRNA genes of *L.sinica* and *T.anthracinum* mitogenomes.

Base-pairing is indicated as follows: Watson–Crick pairs by lines, wobble GU pairs by dots and other noncanonical pairs by circles. Variable regions are presented in boxes with red (*L. sinica*) and blue (*T. anthracinum*) colours.

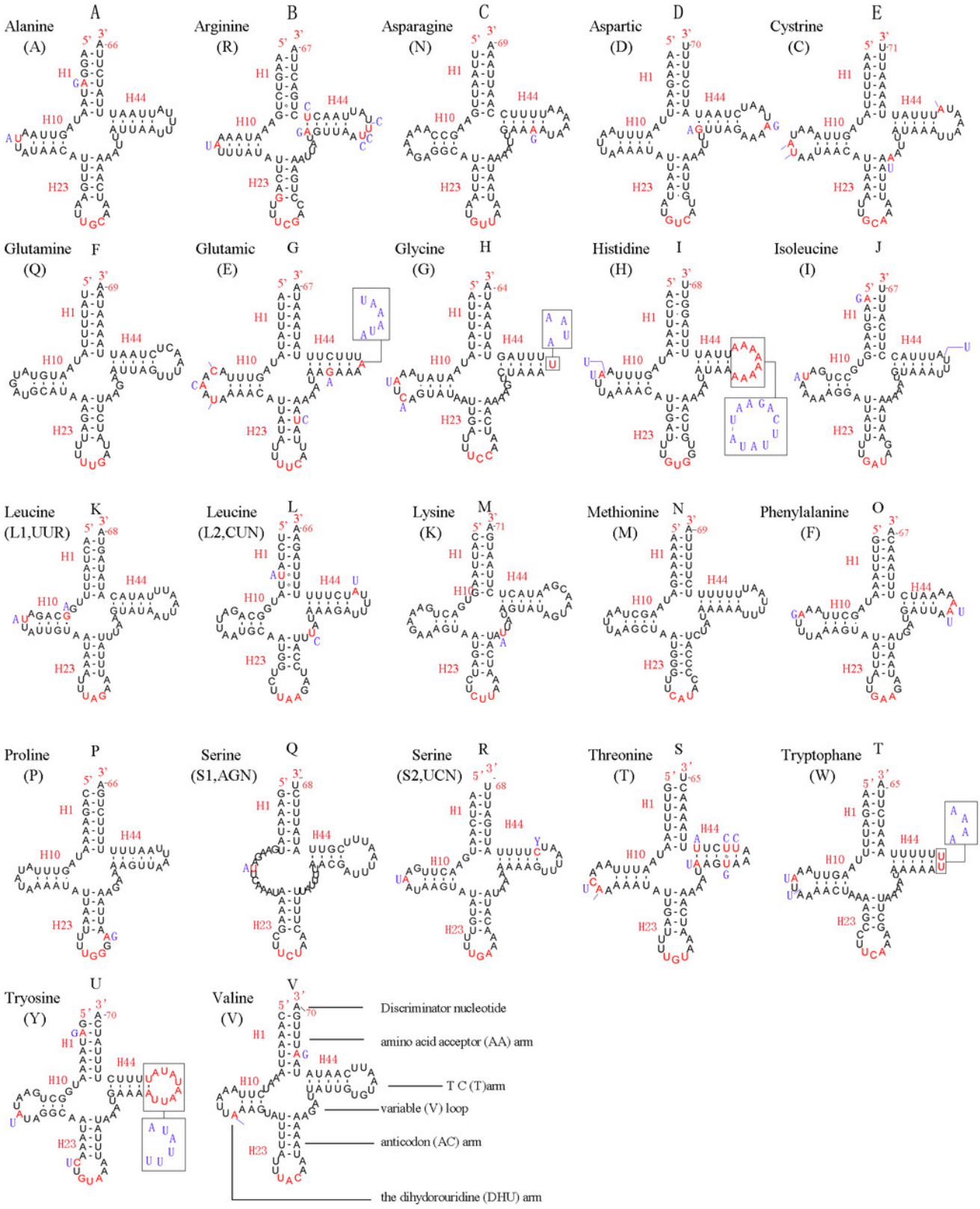

A
Alanine
(A)
B
Arginine
(R)
C
Asparagine
(N)
D
Aspartic
(D)
E
Cystrine
(C)
F
Glutamine
(Q)
G
Glutamic
(E)
H
Glycine
(G)
I
Histidine
(H)
J
Isoleucine
(I)
K
Leucine
(L1,UUR)
L
Leucine
(L2,CUN)
M
Lysine
(K)
N
Methionine
(M)
O
Phenylalanine
(F)
P
Proline
(P)
Q
Serine
(S1,AGN)
R
Serine
(S2,UCN)
S
Threonine
(T)
T
Tryptophane
(W)
U
Tryosine
(Y)
V
Valine
(V)
Discriminator nucleotide
amino acid acceptor (AA) arm
T C (T)arm
variable (V) loop
anticodon (AC) arm
the dihydorouridine (DHU) arm

Figure 5

Predicted secondarystructures for the 22 tRNA genes of C.*lateralis.*

Dashes indicate Watson-Crick base pairing and dots indicate G-U base pairing.

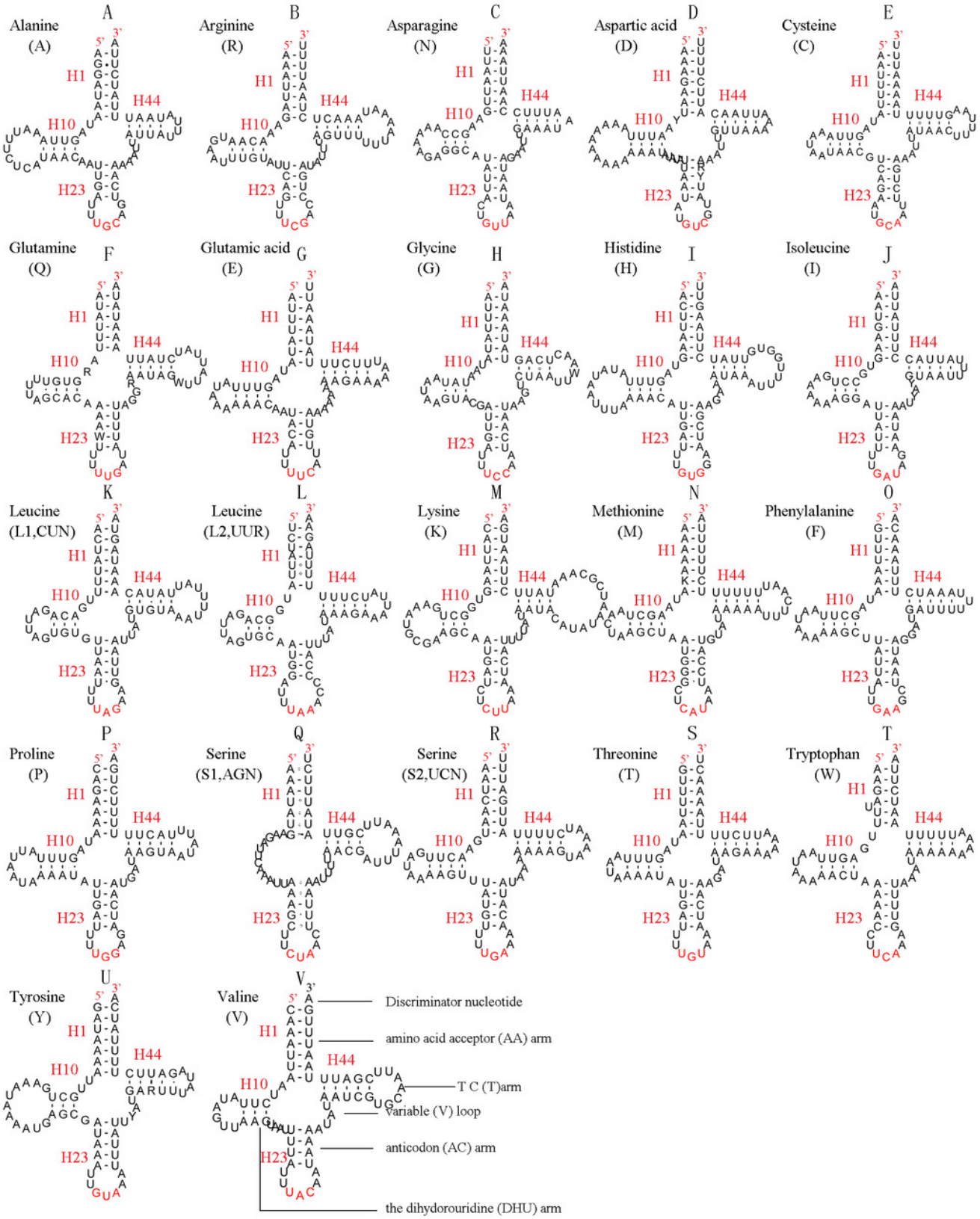

A
Alanine
(A)
B
Arginine
(R)
C
Asparagine
(N)
D
Aspartic acid
(D)
E
Cysteine
(C)
F
Glutamine
(Q)
G
Glutamic acid
(E)
H
Glycine
(G)
I
Histidine
(H)
J
Isoleucine
(I)
K
Leucine
(L1,CUN)
L
Leucine
(L2,UUR)
M
Lysine
(K)
N
Methionine
(M)
O
Phenylalanine
(F)
P
Proline
(P)
Q
Serine
(S1,AGN)
R
Serine
(S2,UCN)
S
Threonine
(T)
T
Tryptophan
(W)
U
Tyrosine
(Y)
V
Valine
(V)
H1
H10
H23
H44
5'
3'
Discriminator nucleotide
amino acid acceptor (AA) arm
T C (T)arm
variable (V) loop
anticodon (AC) arm
the dihydorouridine (DHU) arm

Figure 6

The predicted secondary structures of *rrnL* of *L. sinica* and *T. anthracinum.*

Tertiary interactions and base triples are connected by continuous lines. The numbering of helix follows Gillespie et al. (2006). Roman numbers refer to domain names. Dashes indicate Watson-Crick base pairing and dots indicate G-U base pairing. The helical variation among cimbicid species are presented in boxes with red (*L. sinica*) and blue (*T. Anthracinum*) colours.

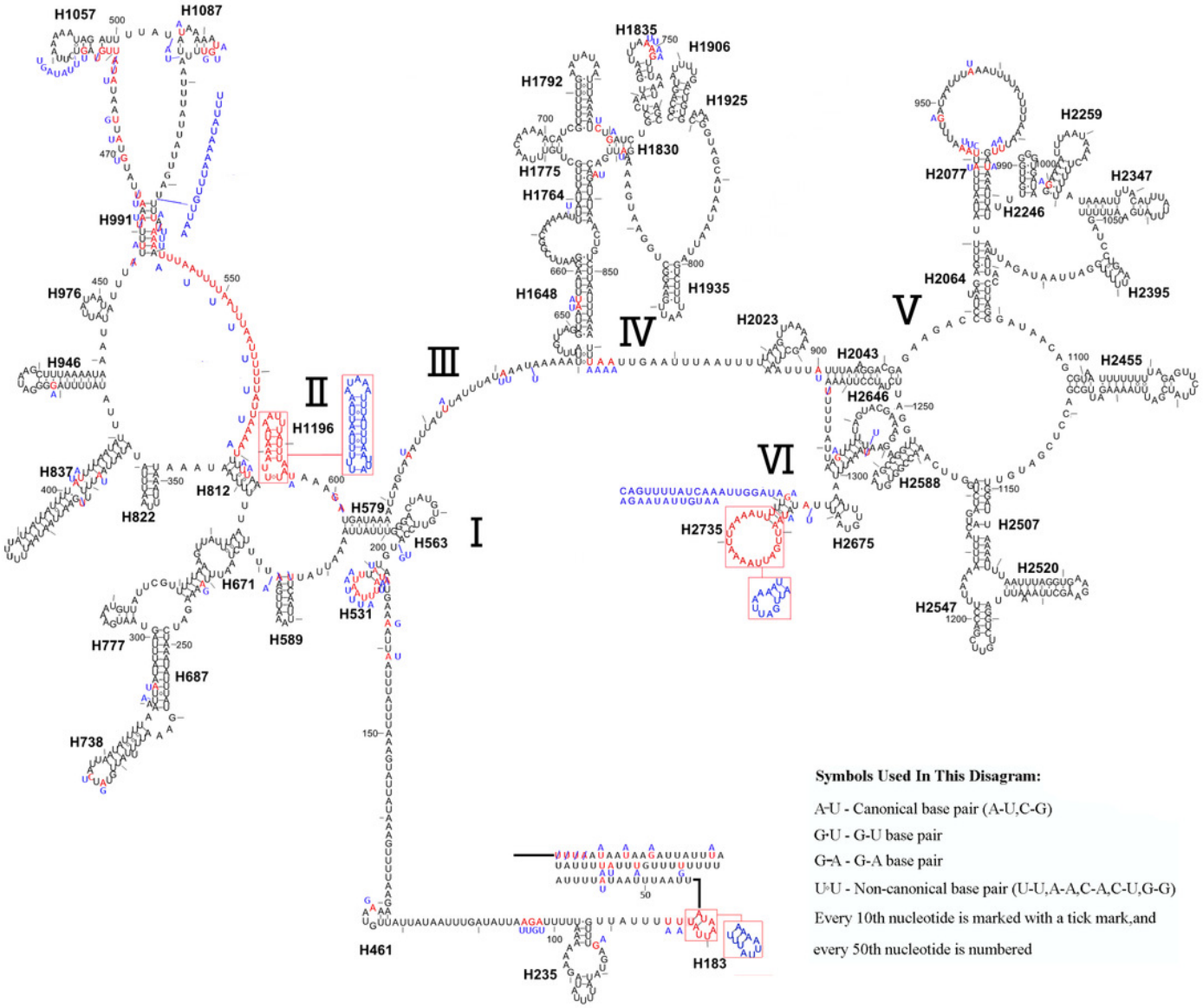


Figure 7

Corynislateralis rrnL .

Predicted *rrnL* secondary structure in *C. lateralis.* The numbering of helix follows Gillespie et al. (2006).Roman numbers refer to domain names.

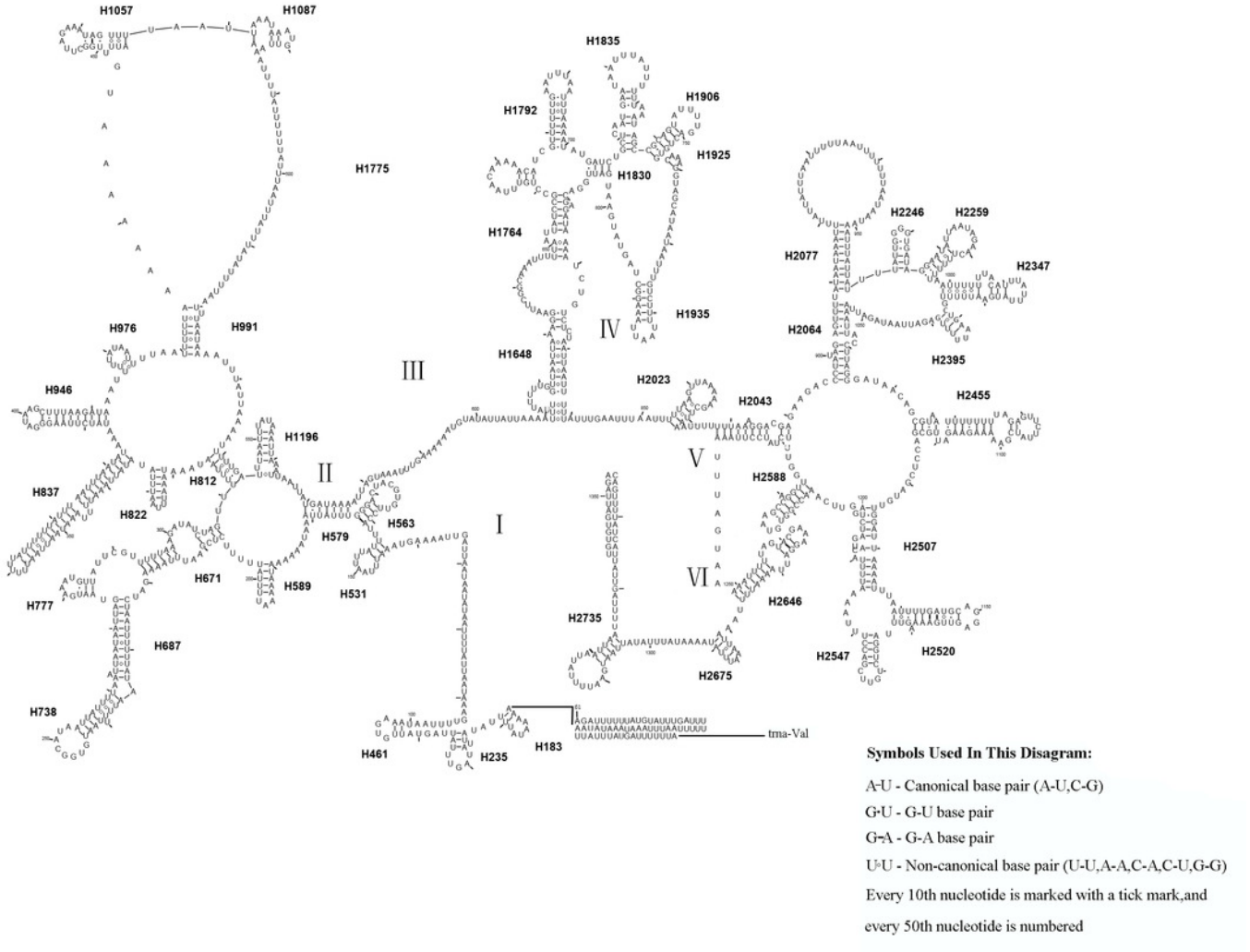


Figure 8

The predicted secondary structures of *rrnS* of *L. sinica* and *T. Anthracinum.*

Tertiary interactions and base triples are connected by continuous lines. The numbering of helix follows Gillespie et al. (2006). Roman numbers refer to domain names. Dashes indicate Watson-Crick base pairing and dots indicate G-U base pairing. The helical variation among cimbicid species are presented in boxes with red (*L. sinica*) and blue (*T. Anthracinum*) colours.

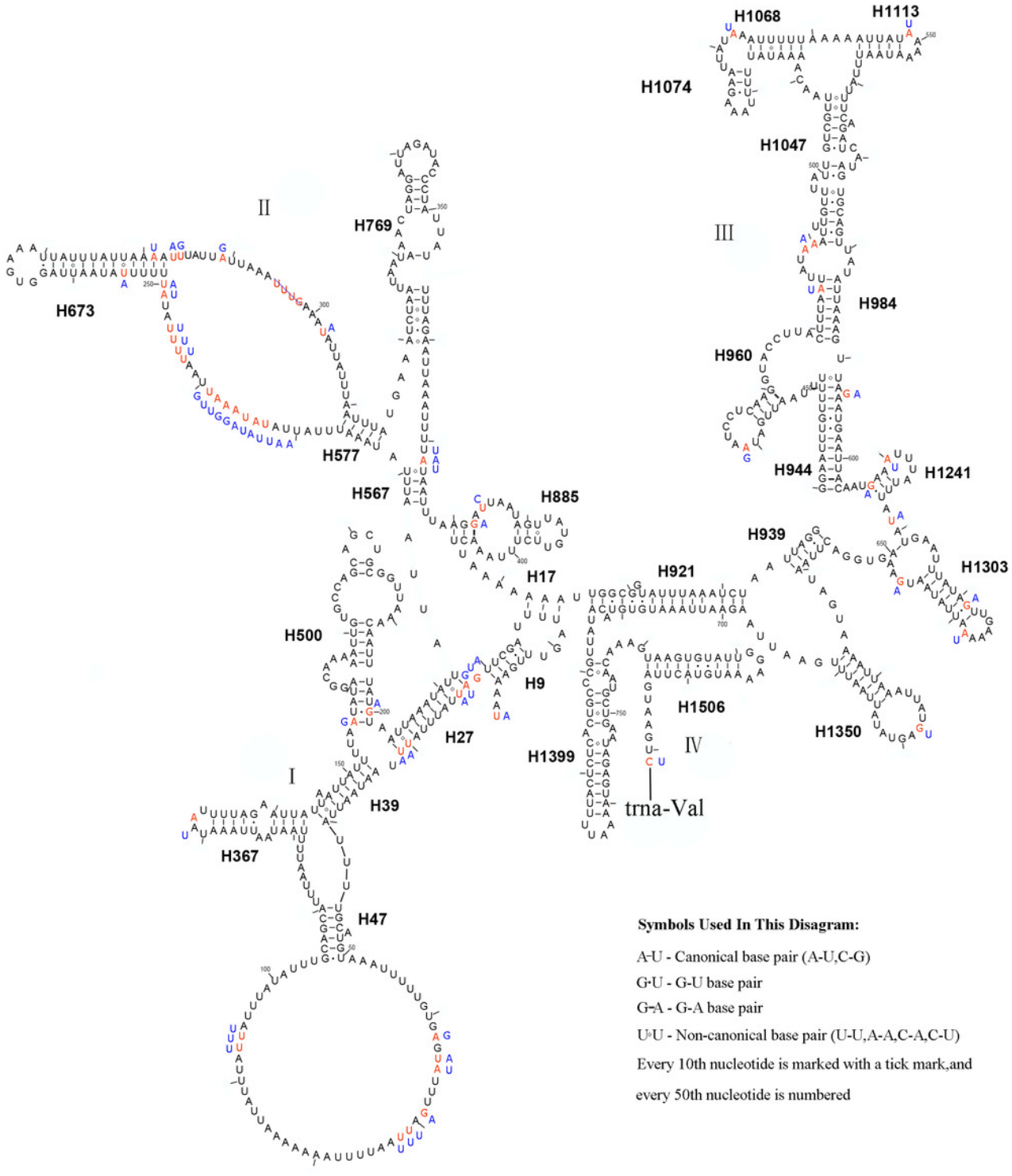
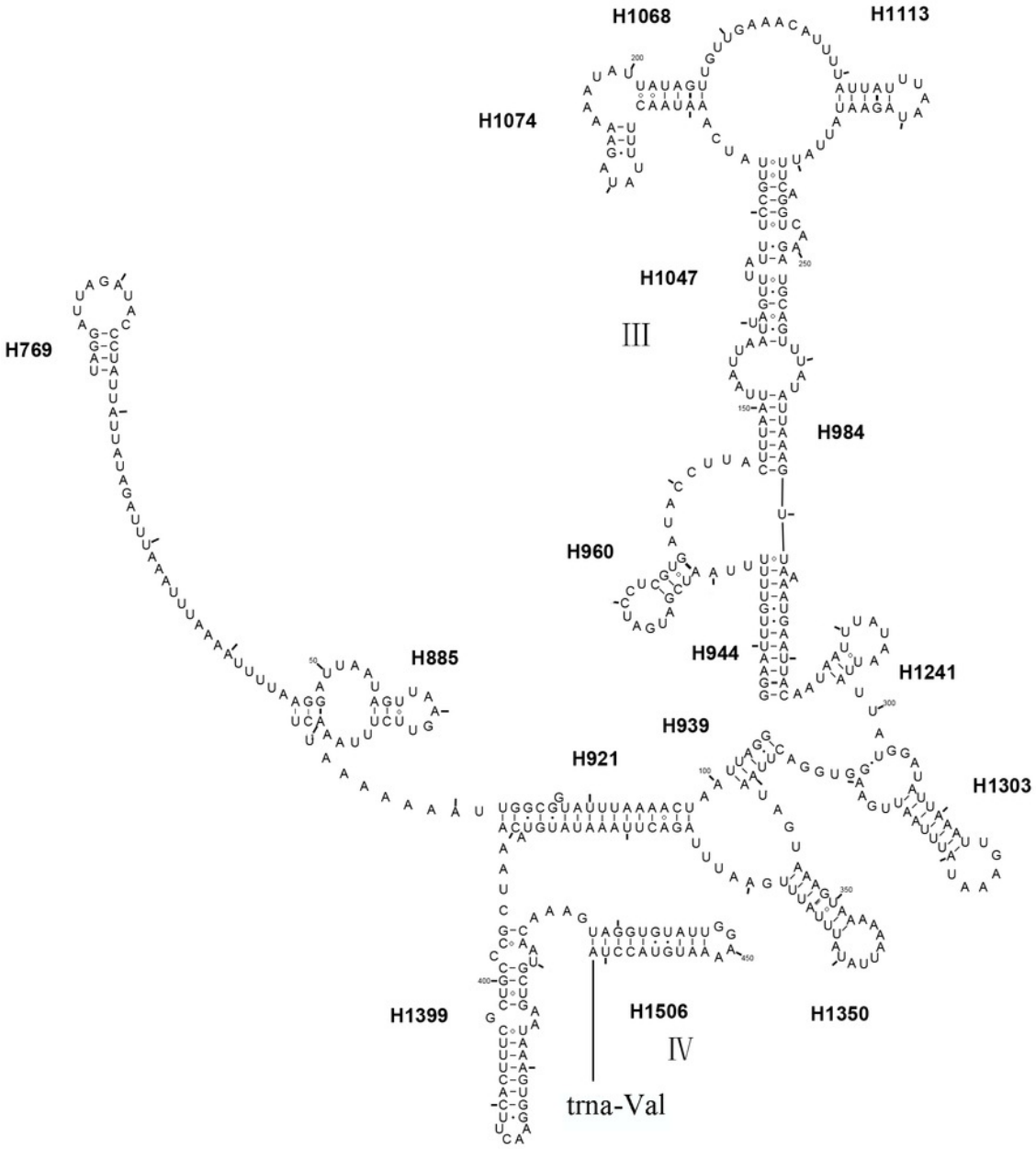

H1068
H1113
H1074
H1047
II
H769
III
H673
H984
H960
H577
H944
H1241
H567
H885
H939
H921
H1303
H17
H500
H9
H1506
H27
H1350
H1399
IV
I
trna-Val
H39
H367
H47
Symbols Used In This Disagram:
A-U - Canonical base pair (A-U,C-G)
G-U - G-U base pair
G-A - G-A base pair
U-U - Non-canonical base pair (U-U,A-A,C-A,C-U)
Every 10th nucleotide is marked with a tick mark,and
every 50th nucleotide is numbered

Figure 9

Corynislateralis rrnS .

Predicted *rrnS* secondary structure in *C. lateralis.* The numbering of helix follows Gillespie et al. (2006).Roman numbers refer to domain names.



Symbols Used In This Disagram:

A-U - Canonical base pair (A-U,C-G)

G·U - G-U base pair

G-A - G-A base pair

U◦U - Non-canonical base pair (U-U,A-A,C-A,C-U)

Every 10th nucleotide is marked with a tick mark,and every 50th nucleotide is numbered

Figure 10

The phylogenetic tree of Symphyta by iqtree.

Phylogenetic tree of Symphyta, based on ML analysis of the 13 PCGs (*ND2* , *COXI1* , *COX2* , *ATP8* , *ATP6* , *COX3* , *ND3* , *ND5* , *ND4* , *ND4L* , *ND6* , *CYT8* , and *ND1*) and 2 rRNAs (*rrnL* and *rrnS*) data set. Values branches represent maximum likelihood bootstrap clade frequency (CF) support. The scale bar indicates the number of substitutions per site.

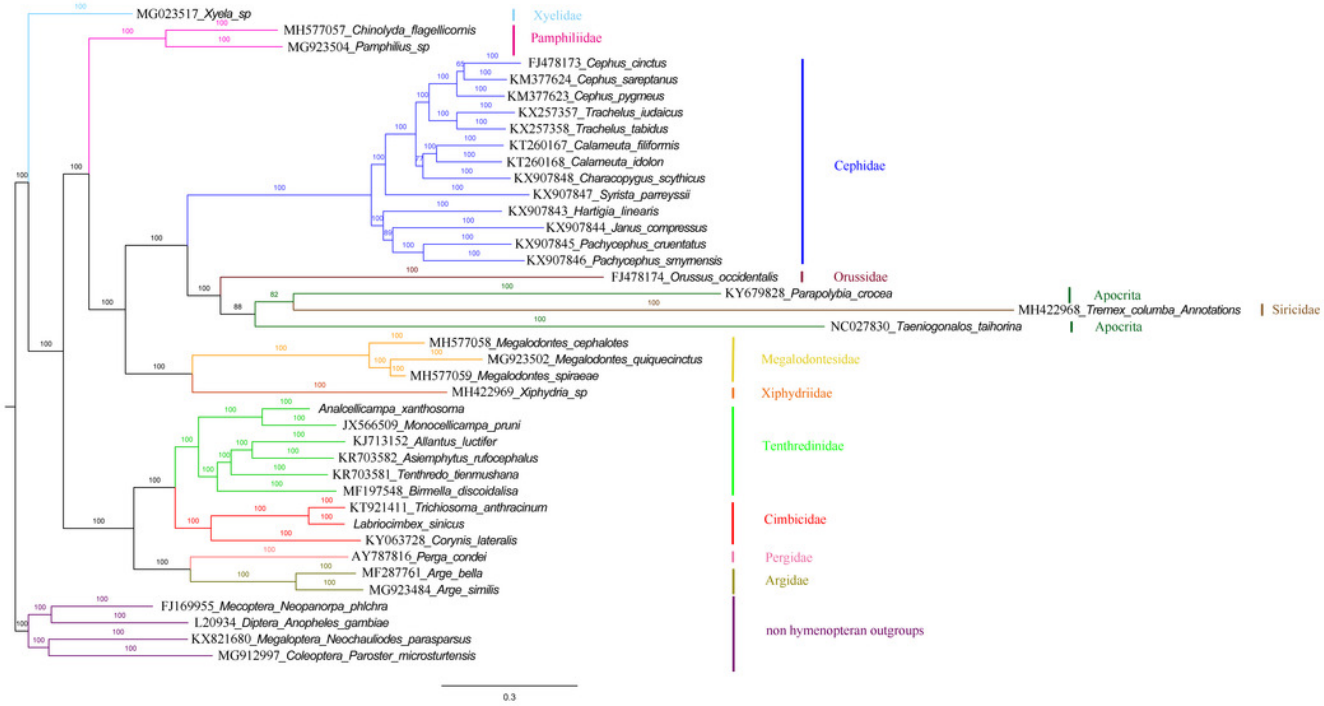


Figure 11

The phylogenetic tree ofSymphyta by bayes.

Phylogenetic tree of Symphyta, based on BI analysis of the 13 PCGs (*ND2* , *COXI1* , *COX2* , *ATP8* , *ATP6* , *COX3* , *ND3* , *ND5* , *ND4* , *ND4L* , *ND6* , *CYT8* , and *ND1*) and 2 rRNAs (*rrnL* and *rrnS*) data set. Values branches represent BI posterior probability (PP) support. The scale bar indicates the numberof substitutions per site.

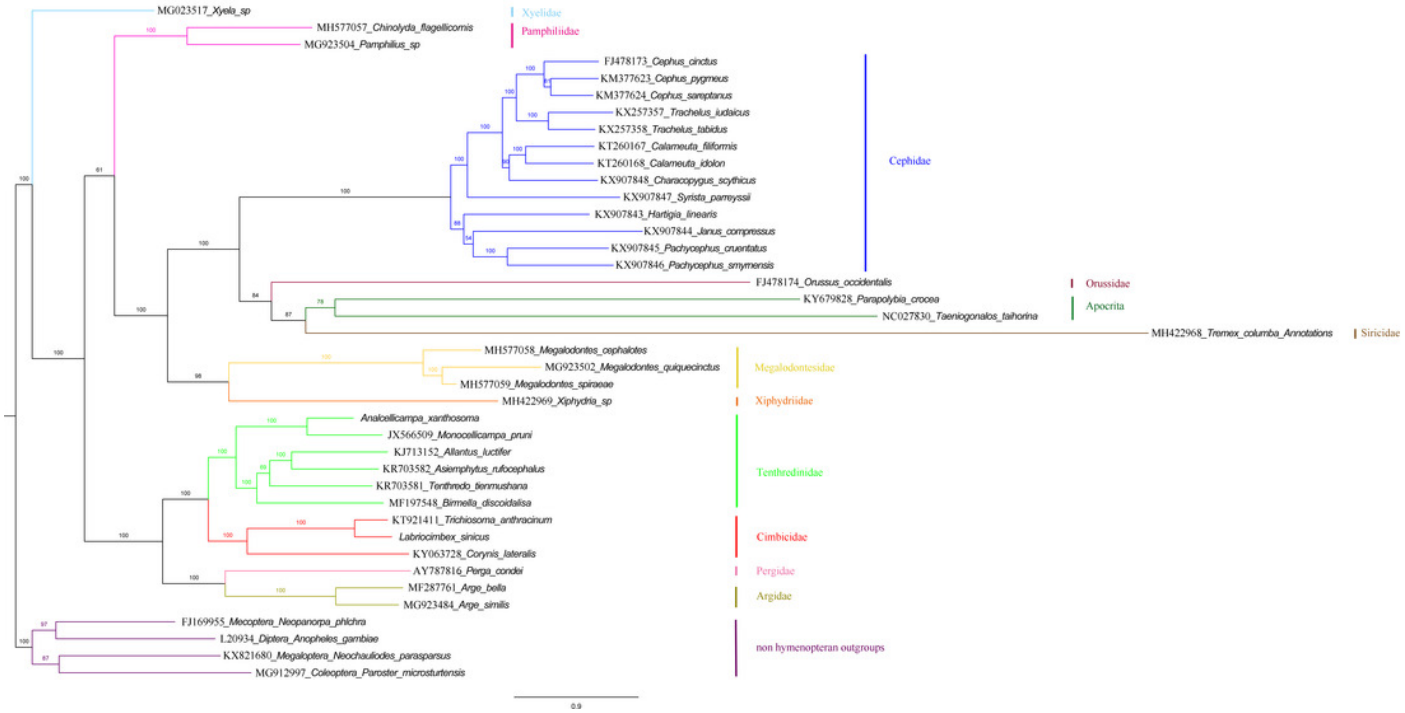


Table 1(on next page)

Summary information of symphytan mitochondrialgenomes used inphylogenetic analyses

1 Summary information of symphytan mitochondrial genomes used in phylogenetic
2 analyses

	Species	Family	Accesion number	References
Ingroup	*Labriocimbex sinica*	Cimbicidae	MH136623	this study
	Corynis lateralis	Cimbicidae	KY063728	Doğan and Korkmaz, 2017
	Trichiosoma anthracinum	Cimbicidae	NC029733	Song *et al*., 2016
	Megalodontes cephalotes	Megalodontesidae	MH577058	Niu *et al*., 2018
	Megalodontes spiraeae	Megalodontesidae	MH577059	Niu *et al*., 2018
	Megalodontes quinquecinctus	Megalodontesidae	MG923502	Tang *et al*., 2019
	Analcellicampa xanthosoma	Tenthredinidae	MH992752	Unpublished
	Allantus luctifer	Tenthredinidae	NC024664	Wei *et al*., 2014
	Asiemphytus rufocephalus	Tenthredinidae	KR703582	Song *et al*., 2016
	Monocellicampa pruni	Tenthredinidae	JX566509	Wei *et al*., 2015
	Tenthredo tienmushana	Tenthredinidae	KR703581	Song *et al*., 2015
	Birmella discoidalisa	Tenthredinidae	MF197548	Unpublished
	Xyela sp.	Xyelidae	MG923517	Tang *et al*., 2019
	Xiphydria sp.	Xiphydriidae	MH422969	Ma *et al*., 2018
	Tremex columba	Siricidae	MH422968	Ma *et al*., 2018
	Pamphilius sp.	Pamphiliidae	MG923504	Tang *et al*., 2019
	Chinolyda flagellicornis	Pamphiliidae	MH577057	Niu *et al*., 2018
	Orussus occidentalis	Orussidae	NC012689	Dowton *et al*., 2009
	Arge similes	Argidae	MG923484	Tang *et al*., 2019
	Arge bella	Argidae	MF287761	Du *et al*., 2018
	Calameuta filiformis	Cephidae	NC028445	Korkmaz *et al*., 2016
	Calameuta idolon	Cephidae	NC028446	Korkmaz *et al*., 2016
	Cephus cinctus	Cephidae	NC012688	Dowton *et al*., 2009
	Cephus pygmeus	Cephidae	KM377623	Korkmaz *et al*., 2015
	Cephus sareptanus	Cephidae	KM377624	Korkmaz *et al*., 2015
	Characopygus scythicus	Cephidae	KX907848	Korkmaz *et al*., 2018
	Hartigia linearis	Cephidae	KX907843	Korkmaz *et al*., 2018
	Janus compressus	Cephidae	KX907844	Korkmaz *et al*., 2018
	Pachycephus cruentatus	Cephidae	KX907845	Korkmaz *et al*., 2018
	Pachycephus smyrnensis	Cephidae	KX907846	Korkmaz *et al*., 2018
	Syrista parreyssi	Cephidae	KX907847	Korkmaz *et al*., 2018
	Trachelus iudaicus	Cephidae	KX257357	Korkmaz *et al*., 2017
	Trachelus tabidus	Cephidae	KX257358	Korkmaz *et al*., 2017
	Perga condei	Pergidae	AY787816	Castro and Dowton, 2005
	Taeniogonalos taihorina	Trigonalidae	NC027830	Wu *et al*., 2014
	Parapolybia crocea	Vespidae	KY679828	Peng *et al*., 2017
Outgroup	*Paroster microsturtensis*	Dytiscidae	MG912997	Hyde *et al*., 2018
	Neopanorpa phlchra	Panorpidae	FJ169955	Unpublished
	Neochauliodes parasparsus	Corydalidae	KX821680	Zhao *et al*., 2017

Anopheles gambiae	Culicidae	L20934	Beard *et al*., 1993

3

Table 2(on next page)

Mitochondrial genome characteristics of *L. sinica* .

1 Mitochondrial genome characteristics of *Labriocimbex sinica*

2

Gene	Strand	Start	Stop	Length(bp)	Start codon	Stop codon	Anticodon	IGN
trnI	J	1	67	67			GAU	
ND2	J	70	1113	1044	ATG	TAA		2
trnW	J	1117	1181	65			UCA	3
COI	J	1182	2720	1539	ATT	TAA		0
trnL2	J	2760	2825	66			UAA	39
COII	J	2827	3510	684	ATG	TAA		1
trnK	J	3532	3602	71			CUU	21
trnD	J	3603	3672	70			GUC	0
ATP8	J	3673	3834	162	ATC	TAA		0
ATP6	J	3828	4517	690	ATG	TAA		-7
COIII	J	4504	5289	786	ATG	TAA		-14
trnG	J	5310	5373	64			UCC	20
ND3	J	5374	5724	351	ATT	TAA		0
trnA	J	5732	5797	66		T	UGC	7
trnR	J	5798	5,864	67			UCG	0
trnN	J	5,866	5,934	69			GUU	1

trnS1	J	5,935	6,002	68			UGA	0
trnE	J	6,010	6,076	67			UUC	7
trnF	N	6,092	6158	67			AAG	15
ND5	N	6,159	7872	1714	ATT	T		0
trnH	N	7873	7940	68			GUG	0
ND4	N	7991	9343	1353	ATT	TAA		50
ND4L	N	9337	9618	282	ATT	TAA		-7
trnT	N	9621	9865	65			UGU	2
trnP	N	9686	9751	66			GGU	0
ND6	J	9753	10256	504	ATA	TAA		1
CYTB	J	10258	11391	1134	ATA	TAA		1
trnS2	J	11435	11502	68			UCU	43
ND1	N	11512	12462	951	ATT	TAA		9
trnL1	N	12463	12530	68			GAU	0
rrnL	N	12531	13871	1341				0
trnV	N	13872	13941	70			CAU	0
rrnS	N	13941	14731	791				-1
trnM	J	14777	14845	69			CAU	45
trnQ	N	14843	14911	69			GUU	-3
CR	none	14912	15261	350				0

trnY	J	15262	15331	70		GUA	0
trnC	N	15333	15403	71		ACG	1

3

4

Table 3(on next page)

Nucleotide compositionof *L. sinica* mitochondrialgenome.

1 Nucleotide composition of *Labriocimbex sinica* mitochondrial genome

2

Feature	Length(bp)	A%	C%	G%	T%	A+T%	AT-skew	GC-skew
Whole genome	15405	43.5	11.1	7.7	37.7	81.2	0.0714	-0.1809
Protein coding genes	12456	34.4	9.7	10.4	45.5	79.9	-0.1389	0.0348
First codon position	4152	36.9	9.5	15.1	38.5	75.4	-0.0212	0.2276
Second codon position	4152	20.9	16.2	13	49.9	70.8	-0.4096	-0.1096
Third codon position	4152	45.5	3.5	3	48	93.5	-0.0267	-0.0769
Protein coding genes-J	6840	37.8	12	9.3	40.9	78.7	-0.0394	-0.1268
First codon position	2280	40.4	11.9	14.6	33.1	73.5	0.0993	0.1019
Second codon position	2280	23	18.7	12	46.3	69.3	-0.3362	-0.2182
Third codon position	2280	50	5.3	1.4	43.3	93.3	0.0718	-0.5821
Protein coding genes-N	5616	30.3	7	11.6	51.1	81.4	-0.2555	0.2473
First codon position	1872	32.5	6.7	15.6	45.2	77.7	-0.1634	0.3991
Second codon position	1872	18.4	13.1	14.3	54.3	72.7	-0.4938	0.0438
Third codon position	1872	40.1	1.2	5	53.7	93.8	-0.1450	0.6129
ATP6	690	38.3	11.2	8	42.6	80.9	-0.0532	-0.1667
ATP8	162	45.1	9.3	2.5	43.2	88.3	0.0215	-0.5763
ND1	951	51.4	12.3	6.9	29.3	80.7	0.2739	-0.2813

ND2	1044	44.1	9.9	5.7	40.3	84.4	0.0450	-0.2692
ND3	351	35	10.5	9.7	44.7	79.7	-0.1217	-0.0396
ND4	1353	51.2	11.5	7.4	29.9	81.1	0.2626	-0.2169
ND4-BLASTP	1344	51.3	11.6	7.4	29.7	81	0.2667	-0.2211
ND4L	282	49.6	12.1	3.5	34.8	84.4	0.1754	-0.5513
ND5	1714	50.8	11.1	6.8	31.3	82.1	0.2375	-0.2402
ND6	504	42.1	8.7	5	44.2	86.3	-0.0243	-0.2701
COI	1539	35.1	13.5	12.8	38.7	73.8	-0.0488	-0.0266
COII	684	40.8	12.7	8	38.5	79.3	0.0290	-0.2271
COIII	786	33.5	13	12	41.6	75.1	-0.1079	-0.0400
CYTB	1134	35.4	13.1	10.4	41.2	76.6	-0.0757	-0.1149
12s	791	44	10.7	5.3	40.1	84.1	0.04637337	0.3375
16s	1341	46.8	11	4.9	37.4	84.2	0.111639	-0.383648

3

4

Table 4(on next page)

Codon usage of PCGs in mitochondrial genome of *L. sinica.* No., frequency of each codon; RSCU, relative synonymous condonusage.

1 Codon usage of PCGs in mitochondrial genome of *Labriocimbex sinica*

2

Amino acid	Codon	NO.	RSCU		Amino acid	Codon	NO.	RSCU	
Phe	TTT	409	1.9		Tyr	TAT	159	1.78	
	TTC	21	0.1	2		TAC	20	0.22	2
Leu	TTA	560	5.04		End	TAA	0	0	
	TTG	35	0.31	5.35		TAG	0	0	
Leu	CTT	37	0.33		His	CAT	68	1.79	
	CTC	0	0			CAC	8	0.21	2
	CTA	34	0.31	6	Gln	CAA	61	1.85	
	CTG	1	0.01	0.65		CAG	5	0.15	2
Ile	ATT	464	1.87		Asn	AAT	237	1.84	
	ATC	31	0.13	2		AAC	20	0.16	2
Met	ATA	314	1.91		Lys	AAA	135	1.88	
	ATG	15	0.09	2		AAG	9	0.13	2.01
Val	GTT	83	2.21		Asp	GAT	62	1.82	
	GTC	1	0.03			GAC	6	0.18	2
	GTA	65	1.73		Glu	GAA	72	1.85	
	GTG	1	0.03	4		GAG	6	0.15	2

Ser	TCT	134	2.67		Cys	TGT	37	1.95	
	TCC	4	0.08			TGC	1	0.05	2
	TCA	116	2.31		Trp	TGA	92	1.8	
	TCG	2	0.04	5.1		TGG	10	0.2	2
Pro	CCT	64	1.97		Arg	CGT	20	1.54	
	CCC	15	0.46			CGC	0	0	
	CCA	48	1.48			CGA	31	2.38	
	CCG	3	0.09	4		CGG	1	0.08	4
Thr	ACT	70	1.74		Ser	AGT	23	0.46	
	ACC	8	0.2			AGC	1	0.02	
	ACA	82	2.04			AGA	119	2.37	
	ACG	1	0.02	4		AGG	2	0.04	2.89
Ala	GCT	65	2.08		Gly	GGT	62	1.22	
	GCC	7	0.22			GGC	1	0.02	
	GCA	49	1.57			GGA	112	2.2	
	GCG	4	0.13	4		GGG	29	0.57	4.01

3

4