







## The mitogenome of the jumping bristletail *Trigoniophthalmus alternatus* (Insecta, Microcoryphia) and the phylogeny of insect early-divergent lineages

Chiara Leo , Francesco Nardi , Francesco Frati, Pietro Paolo Fanciulli , Claudio Cucini , Matteo Vitale ,  
Claudia Brunetti and Antonio Carapelli 

Department of Life Sciences, University of Siena, Siena, Italy

### ABSTRACT

The complete mitochondrial genome of the machilid *Trigoniophthalmus alternatus* (Silvestri 1904) is herein described and applied to phylogenetic analyses, inclusive of the most early-divergent lineages of hexapods. Both gene content and order generally conform with the organization of the arthropods' mitochondrial genome. One gene translocation involving *trnA* is the autapomorphic character observed in this species. Another peculiar molecular feature is the long size of the A + T-rich region, due to the occurrence of repeat units. The phylogenetic analyses support the typical placement, along the hexapods' tree, of Ectognatha, Monocondylia and Dicondylia, with Diplura as the adelphotaxon of all true insects.

### ARTICLE HISTORY

Received 12 June 2019  
Accepted 3 August 2019

### KEYWORDS

Basal hexapods; Ectognatha;  
Microcoryphia; Diplura;  
Zygentoma;  
Mitogenomics; Phylogeny

Total DNA of one specimen of *Trigoniophthalmus alternatus* (Voucher specimen ID: TAL\_05, preserved at Life Sciences Department of University of Siena; Collection site: Siena, Italy, 43°24'7.81"N 11°21'52.62"E) was isolated, amplified, and sequenced as described in Carapelli et al. (2007).

The complete mtDNA of *T. alternatus* was assembled using the software Sequencher 4.4.2 (Gene Codes Corporation, Ann Arbor, MI, USA). The consensus sequence was submitted to the tRNAs secondary structure prediction online tool ARWEN (Laslett and Canbäck 2008). Accordingly, the 22 tRNA-encoding sequences were then mapped on the genome contig. Boundaries of protein-encoding genes were also detected searching for their start and stop codons. The rDNA genes were identified drawing their putative secondary structure at the 5'- and 3'- ends of both *rnmS* and *rnmL*. The annotated genome sequence was deposited in GenBank under the accession number: NC010532.

The mitogenome of *T. alternatus* contains the common set of 37 genes, plus a non-coding DNA fragment (A + T-rich region) where the sequences that control both replication and transcription of the mtDNA are found. This complete genome sequence is the longest (16,197 bp in length) mtDNA among all Microcoryphia species so far investigated. In this taxon, the nucleotide content of the J-stand is biased toward A and T nucleotides (A = 38.8%, T = 32.6%, and G = 10.1%). All PCGs start with canonical Methionine codons (either ATA or ATG) except *cox1* (that starts with TTG, that encodes for Leucine); complete TAA stop codons are found in all genes apart from *cox1* and *nad4L* (both ending with truncated TA-).

Within the long A + T-rich region (1150 bp in size), there are three almost identical direct repeats of 217 bp (plus an

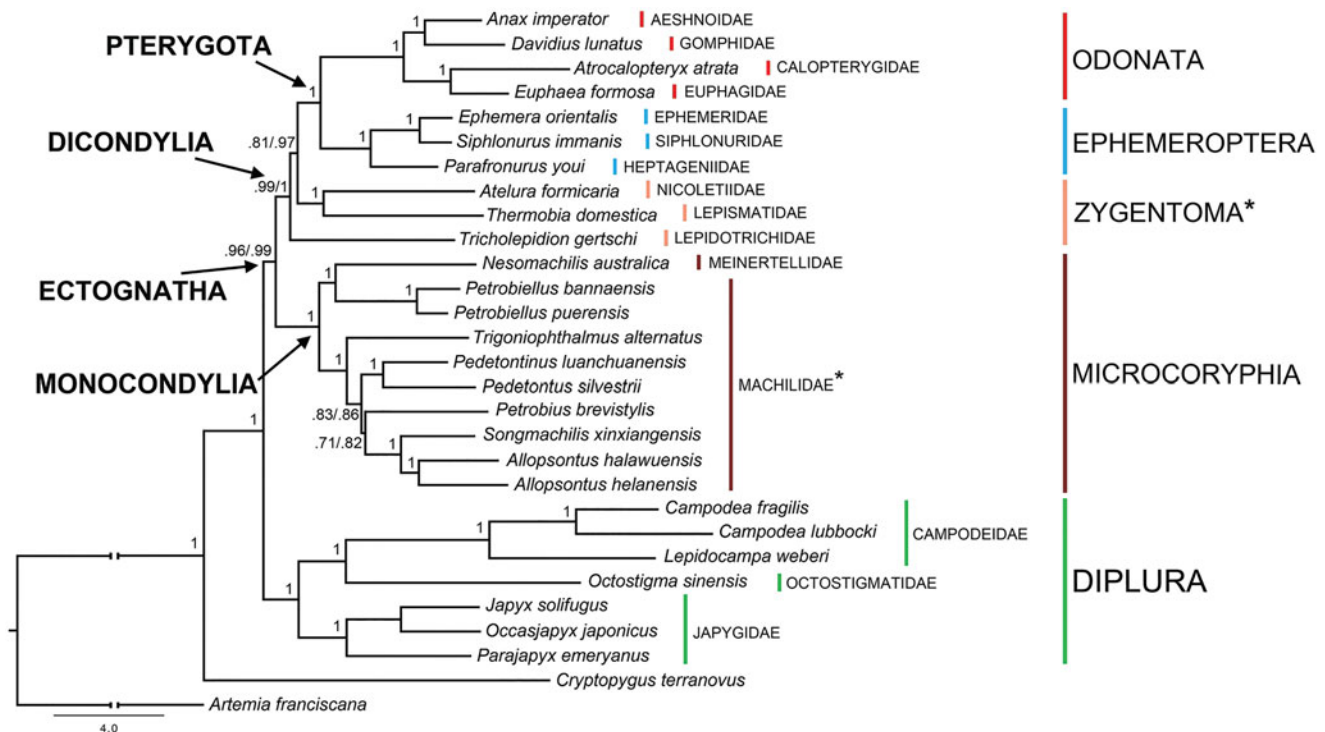
additional partial repeat unit), as previously observed in other mtDNAs (Nardi et al. 2012).

The gene order of Microcoryphia is similar to that ancestral for Pancrustacea (Boore et al. 1998) with the exception of the tRNA cluster A-R-N-S-E-F, that is rearranged to: A-R-N-E-F in *Nesomachilis australica*, A-R-N-S-E-Y-F in *Petrobius brevistylis* and R-N-S-E-A-F in *T. alternatus*. No shared gene rearrangement among Microcoryphia species is observed.

The *T. alternatus* mtDNA genes were aligned with those of other 28 species (Figure 1), concatenated and used for phylogenetic analyses. The final alignment, obtained with RevTrans 1.4 (Wernersson and Pedersen 2003), was deprived of positions of dubious homology with the online tool GBlock (Castresana 2000). The resulting nucleotide matrix was partitioned into 39 charsets (one for each of the three codon positions of the 13 PCGs) and tested with the software PartitionFinder 2.1.1 (GTR + I and GTR + I +  $\Gamma$  models were selected; Lanfear et al. 2016). Two independent runs (all nucleotides and first and second codon positions only) were processed with the software MrBayes 3.2 (Ronquist et al. 2012), with four chains for  $10^6$  generations and sampling one tree/1000 iterations (with 25% of burn-in). The resulting trees both support the monophyly of Ectognatha, Monocondylia, Dicondylia, Pterygota, Diplura, Ephemeroptera, Microcoryphia and Odonata (Figure 1). Zygentoma is the only high-ranking taxon that results paraphyletic (Figure 1).

### Disclosure statement

The authors declare no conflict of interest and are responsible for the content of the writing.



**Figure 1.** Bayesian phylogeny based on the concatenated set of 13 mitochondrial protein-encoding genes of the following basal hexapod lineages used in this study: *Campodea fragilis* DQ529236, *Campodea lubbocki* DQ529237 and *Lepidocampa weberi* JN990601 (Campodeidae; Diplura); *Octostigma sinensis* JN990598 (Octostigmatidae; Diplura); *Japyx solifugus* NC007214, *Occasjapyx japonicus* JN990600 and *Parajapyx emeryanus* JN990599 (Japygidae; Diplura); *Allopsontus helanensis* KJ754501, *Allopsontus halawuensis* KJ754500, *Nesomachilis australica* NC006895, *Pedetontinus luanchuanensis* KJ754502, *Pedetontus silvestrii* EU621793, *Petrobiellus bannaensis* KJ754503, *Petrobiellus puerensis* KJ754504, *Petrobius brevistylis* NC007688, *Songmachilis xinxiangensis* NC021384 and *Trigoniphthalmus alternatus* EU016193 (Microcoryphia); *Atelura formicaria* EU084035, *Thermobia domestica* AY639935 and *Tricholepidion gertschi* AY191994 (Zygentoma); *Ephemera orientalis* NC012645, *Parafonurus youi* NC011359 and *Siphonurus immanis* NC013822 (Ephemeroptera); *Atracalopteryx atrata* NC027181 and *Euphaea formosa* NC014493 (Zygotera; Odonata); *Anax imperator* NC031821 and *Davidius lunatus* NC012644 (Anisoptera; Odonata). The collembolan *Cryptopygus terranovus* NC037610 and the crustacean *Artemia franciscana* NC001620 have been used as outgroups. Double branch-support on nodes (i.e. when different among alternative data sets) refers to analyses based on complete set of nucleotide codon positions/first and second only, respectively. \*refers to paraphyletic groups.

## Funding

This research was supported by the University of Siena and Miur (2264-2011-FF-MIURPRIN\_001) to A.C. and F.F.

## ORCID

Chiara Leo <http://orcid.org/0000-0003-0510-2535>  
 Francesco Nardi <http://orcid.org/0000-0003-0271-9855>  
 Pietro Paolo Fanciulli <http://orcid.org/0000-0003-2999-6203>  
 Claudio Cucini <http://orcid.org/0000-0003-1918-0702>  
 Matteo Vitale <http://orcid.org/0000-0002-9687-8832>  
 Antonio Carapelli <http://orcid.org/0000-0002-3165-9620>

## References

- Boore JL, Lavrov DV, Brown WM. 1998. Gene translocation links insects and crustaceans. *Nature*. 392:667–668.
- Carapelli A, Liò P, Nardi F, van der Wath E, Frati F. 2007. Phylogenetic analysis of mitochondrial protein coding genes confirms the reciprocal paraphyly of Hexapoda and Crustacea. *BMC Evol Biol*. 7: 58.
- Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol*. 17: 540–552.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol Evol*. 34:772–773.
- Laslett D, Canbäck B. 2008. ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. *Bioinformatics*. 24:172–175.
- Nardi F, Carapelli A, Frati F. 2012. Repeated regions in mitochondrial genomes: distribution, origin and evolutionary significance. *Mitochondrion*. 12:483–491.
- Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Hohna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol*. 61:539–542.
- Wernersson R, Pedersen AG. 2003. RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res*. 31: 3537–3539.