


The complete mitochondrial genome of the springtail *Allacma fusca*, the internal phylogenetic relationships and gene order of Symphyleona

Francesco Nardi, Claudio Cucini, Chiara Leo, Francesco Frati, Pietro Paolo Fanciulli and Antonio Carapelli 

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

ABSTRACT

Symphyleona (*sensu stricto*) are a group of Collembola (=springtails) that, despite displaying some variation in gene order, have been poorly investigated under the phylomitogenomic perspective. How families and subfamilies of this taxon are evolutionary related is still partially unknown. For this reason we sequenced, and herein described, the complete mitochondrial genome sequence of *Allacma fusca* (Sminthuridae). This sequence, alongside others from the literature, is here used to study the phylogenetic relationships among Symphyleona.

ARTICLE HISTORY

Received 30 June 2020
Accepted 18 July 2020

KEYWORDS

Symphyleona;
Sminthuridae; springtails
mitogenomics; Collembola

Among Collembola, the order Symphyleona (*sensu stricto*) accounts for over 1200 species (<https://www.collembola.org/>). Traditionally, this taxon was grouped with Neelipleona because of the shared spherical body shape in Symphyleona *sensu lato*. Despite this similar morphological feature, recent molecular evidence established that Symphyleona *s.s.* are more closely related to springtails characterized by an elongated body (Entomobryomorpha and Poduromorpha) than to Neelipleona, supporting this latter as the basal group of springtails (Leo et al. 2019a, 2019b; Sun et al. 2020). However, knowledge on the systematic position of families and subfamilies within the order is still limited due, among other issues, to the restricted number of mitogenomes available. *Allacma fusca*, a member of family Sminthuridae, is a holarctic species and is strictly confined to humid soils (Fanciulli et al. 2009). In this study, we focus on the evolutionary relationships within Symphyleona *s.s.*, at family and subfamily level, describing and comparing the mitogenome of *A. fusca* with those of Symphyleona available on GenBank.

A pool of five individuals of *A. fusca* (voucher specimen ID: AFU_05, preserved at Life Sciences Department of the University of Siena), collected in the Feniglia dunes (Grosseto, Italy; 42°24'42"N 11°12'43"E) in November 2019, was used for the analysis. Total DNA was extracted using the QIAmp[®] UCP DNA kit and sequenced on a HiSeq 2500 platform (Illumina, San Diego, CA) at Macrogen Europe. Resulting reads were assembled using two *de novo* software: MEGAHIT version 1.2.9 (Li et al. 2015) on trimmed sequences and NOVOPlasty version 3.8.3 on untrimmed sequences (Dierckxsens et al. 2017). Single base ambiguities (resulting from the use of a pool of individuals) were resolved on a majority rules basis by remapping reads on interested

regions using bbmap version 38.84 (<https://www.sourceforge.net/projects/bbmap/>) and visualizing alignments in IGV version 2.8.2 (Robinson et al. 2011). The final mitochondrial DNA (mtDNA) sequence was annotated as described in Leo et al. (2019a). *Allacma fusca* protein-coding genes (PCGs) were aligned with homologous sequences from available Symphyleona complete mitogenomes plus two outgroups (Figure 1) as in Leo et al. (2019c). The resulting data matrix was partitioned in 39 subsets (one for each of the three codon positions of the 13 PCGs) and analyzed with PartitionFinder 2.1.1 (Lanfear et al. 2016). The selected models (HKY + Γ , HKY + I + Γ , GTR + I, and GTR + I + Γ) and partitioning scheme were used in MrBayes version 3.2 (Ronquist et al. 2012), through the CIPRES Science Gateway (Miller et al. 2010), for a Bayesian phylogenetic analysis using four chains for 10⁶ generations with a burn-in of 0.25 and a sampling frequency of one tree every 1000 iterations.

The complete mtDNA of *A. fusca* is a circular molecule of 15,111 bp in length. Overall, a bias in the nucleotide composition is observed toward As (39.4%) and Ts (32.6%), with respect to Cs (16.4%) and Gs (11.6%). It encodes for the set of 37 genes (13 PCGs, 22 tRNAs, and 2 rRNAs) commonly observed in metazoan mitochondrial genomes. Gene arrangement differs from both the ancestral pancrustacean gene order and from alternative gene orders previously described in springtails (Boore et al. 1998; Carapelli et al. 2007; Leo et al. 2019a; Sun et al. 2020). *Allacma fusca* shares with the other two representatives of family Sminthuridae, *Lipotrix lubbocki* and *Sminthurus viridis*, two translocations (*trnD* and exchange between *trnT* and *trnP*) that may be tentatively regarded as synapomorphies for the family. Moreover, at variance with *L. lubbocki* and *S. viridis*, *A. fusca* displays two additional translocations (*trnY* and *trnL1*) that appear

CONTACT Claudio Cucini  claudio.cucini@student.unisi.it  Department of Life Sciences, University of Siena, Siena, Italy

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

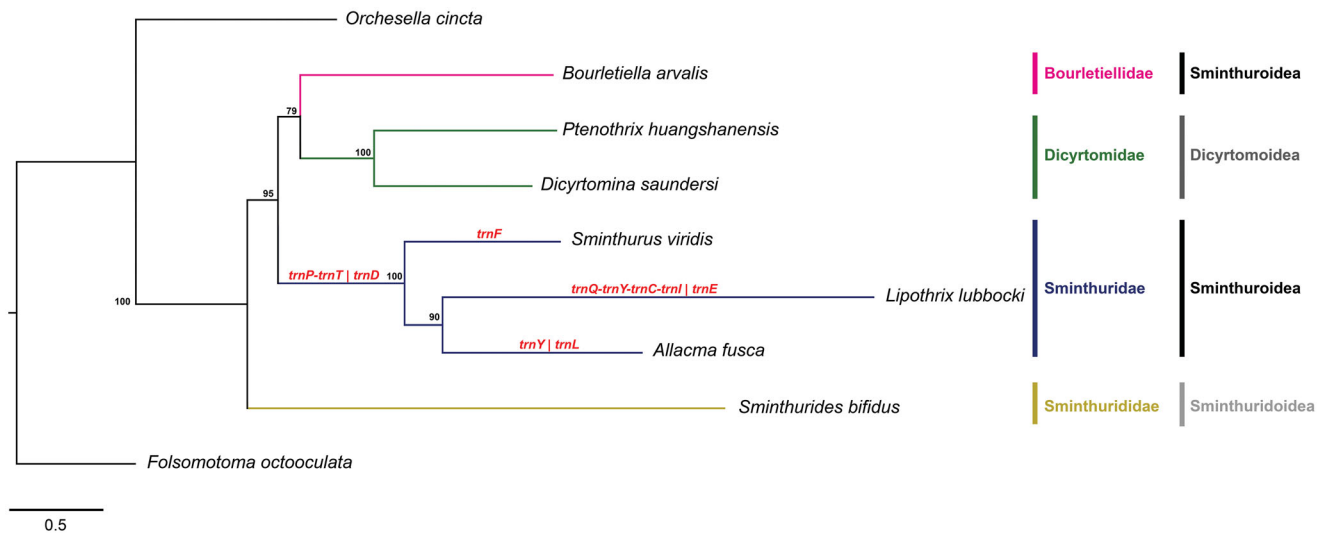


Figure 1. Bayesian phylogenetic tree inferred based on the 13 mitochondrial protein-coding genes sequences of *Folsomotoma octooculata* (NC024155, outgroup), *Orchesella cincta* (NC032283, outgroup), *Sminthurides bifidus* (MK423964), *Bourletiella arvalis* (NC039558), *Ptenothrix huangshanensis* (MK423965), *Dicyrtomina saundersi* (NC044134), *Sminthurus viridis* (NC010536), *Lipothrix lubbocki* (MK431899) and *Allacma fusca* (MT547779). Gene order translocations in Sminthuridae are shown in red. Posterior probability is shown at nodes.

autapomorphic. In particular, the *trnY* translocation occurs from *nad2-trnW-trnC-trnY-cox1* (ancestral pancrustacean GO) to *nad2-trnY-trnW-trnC-cox1*; the *trnL1* translocation occurs from *nad1-trnL1-rrnL-trnV-rrnS* (ancestral pancrustacean GO) to *nad1-rrnL-trnV-rrnS-trnL1* (Figure 1). Our analysis confirms that the ancestral pancrustacean GO model is plesiomorphic for the order Symphypleona s.s., but changed in the Sminthuridae family. The phylogenetic analysis, performed using aligned concatenated PCG sequences, produced the tree depicted in Figure 1. All nodes display high support with the exception of the one joining Bourletiellidae with Dicyrtomidae. The tree indicates that: (a) the Sminthuridae (represented by *Sminthurides bifidus*) are basal to all the other Symphypleona; (b) Dicyrtomidae and Sminthuridae (here represented by more than one species) are monophyletic groups; (c) the Sminthuroidea superfamily may be paraphyletic, because of the placement of *Bourletiella arvalis* (Bourletiellidae) as the sister group to the Dicyrtomidae clade. Nevertheless, this latter result receives only modest support and should be reevaluated based on an enlarged taxonomic coverage of Bourletiellidae.

Acknowledgments

We wish to thank the Department of Biotechnology, Chemistry, and Pharmacy for providing computational resources and Dr. Andrea Bernini for continuous support.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the University of Siena PAR (Piano di Ateneo per la Ricerca) program.

ORCID

Antonio Carapelli  <http://orcid.org/0000-0002-3165-9620>

Data availability statement

The data that support the findings of this study are openly available in GenBank of NCBI with reference number MT547779 at <https://www.ncbi.nlm.nih.gov/nuccore/MT547779>.

References

- Boore JL, Lavrov DV, Brown WM. 1998. Gene translocation links insects and crustaceans. *Nature*. 392(6677):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(2):S8.
- Dierckxens N, Mardulyn P, Smits G. 2017. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res*. 45(4):e18–e18.
- Fanciulli PP, Carapelli A, Belloni M, Dallai R, Frati F. 2009. Allozyme variation in the springtails *Allacma fusca* and *A. gallica* (Collembola, Sminthuridae). *Pedobiologia*. 52(5):309–324.
- Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B. 2016. Partition finder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Mol Biol E*. 34:772–773.
- Leo C, Carapelli A, Cicconardi F, Frati F, Nardi F. 2019a. Mitochondrial genome diversity in Collembola: phylogeny, dating and gene order. *Diversity*. 11(9):169.
- Leo C, Nardi F, Frati F, Fanciulli PP, Cucini C, Vitale M, Brunetti C, Carapelli A. 2019b. The mitochondrial genome of the springtail *Bourletiella arvalis* (Symphypleona, Collembola). *Mitochondrial DNA Part B*. 4(2):2978–2979.
- Leo C, Nardi F, Frati F, Fanciulli PP, Cucini C, Vitale M, Brunetti C, Carapelli A. 2019c. The mitogenome of the jumping bristletail *Trigoniophthalmus alternatus* (Insecta, Microcoryphia) and the phylogeny of insect early-divergent lineages. *Mitochondrial DNA Part B*. 4(2):2855–2856.
- Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 31(10):1674–1676. doi:10.1093/bioinformatics/btv033.

- Miller MA, Pfeiffer W, Schwartz T. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. Proceedings of the gateway computing environments workshop (GCE), 14 Nov. 2010. New Orleans (LA): p. 1–8.
- Robinson J T, Thorvaldsdóttir H, Winckler W, Guttman M, Lander E S, Getz G, Mesirov J P. 2011. Integrative genomics viewer. *Nat Biotechnol.* 29(1):24–26. doi:[10.1038/nbt.1754](https://doi.org/10.1038/nbt.1754).
- 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(3):539–542.
- Sun X, Yu D, Xie Z, Dong J, Ding Y, Yao H, Greenslade P. 2020. Phylomitogenomic analyses on collembolan higher taxa with enhanced taxon sampling and discussion on method selection. *PLoS One.* 15(4):e0230827.