



Research Article

The first complete mitochondrial genome in the family Attevidae (*Atteva aurea*) of the order Lepidoptera

Jun Seong Jeong^{‡,§}, Jeong Sun Park[§], Jae-Cheon Sohn[‡], Min Jee Kim[¶], Hyung Keun Oh[#], Iksoo Kim[§]

[‡] Division of Genetic Resources, Honam National Institute of Biological Resources, Mokpo, Republic of Korea

[§] Department of Applied Biology, College of Agriculture & Life Sciences, Chonnam National University, Gwangju, Republic of Korea

[‡] Department of Science Education, Gongju National University of Education, Chungnam, Republic of Korea

[¶] Experiment and Analysis Division, Honam Regional Office, Animal and Plant Quarantine Agency, Gunsan, Republic of Korea

[#] Crop Protection Team, Chamfield Co., Ltd, Hanam, Republic of Korea

Corresponding author: Iksoo Kim (ikkim81@chonnam.ac.kr)

Academic editor: Rodolphe Rougerie

Received: 05 Jul 2022 | Accepted: 29 Aug 2022 | Published: 12 Sep 2022

Citation: Jeong JS, Park JS, Sohn J-C, Kim MJ, Oh HK, Kim I (2022) The first complete mitochondrial genome in the family Attevidae (*Atteva aurea*) of the order Lepidoptera. Biodiversity Data Journal 10: e89982.

<https://doi.org/10.3897/BDJ.10.e89982>

Abstract

The superfamily Yponomeutoidea, one of the early-derived groups in the order Lepidoptera, consists of 11 families. However, mitochondrial genome (mitogenome) sequences, popularly used for phylogeny and evolutionary tracing, are available for only seven species across six genera and five families. Thus, a larger variety of mitogenome sequences in Yponomeutoidea are required to improve our understanding of lepidopteran phylogeny and genomic evolution. In this study, we present the complete mitogenome of *Atteva aurea* (Fitch, 1856), the first species in the family Attevidae (superfamily Yponomeutoidea, order Lepidoptera) to be sequenced. The complete mitogenome comprises 16,329 bp and contains a typical set of genes and one non-coding region. Within Yponomeutoidea, the mitogenome of *A. aurea* has a unique *trnI-trnM-trnQ* arrangement at the A + T-rich region and *ND2* junction and *trnA-ND3* arrangement at the *trnG* and *trnR* junction. Twelve of the 13 protein-coding genes (PCGs) of *A. aurea* have a typical ATN starting codon, whereas *COI* has the atypical CGA codon, which is frequently found in the starting region of lepidopteran *COI*. Phylogenetic analyses, based on the

concatenated sequences of 13 PCGs and two rRNA genes, using the Maximum Likelihood method, revealed a sister relationship between Attevidae and Praydidae with moderately low nodal support (bootstrap support = 64%).

Keywords

mitochondrial genome, *Atteva aurea*, phylogeny, Attevidae

Introduction

The superfamily Yponomeutoidea is one of the earliest groups to develop external feeding mechanisms in the order Lepidoptera and comprises ~ 1,800 species across 11 families (Sohn et al. 2013). However, only seven species in six genera across five families have available mitochondrial genome (mitogenome) sequences. Thus, the characterisation of the mitogenomes of more families will significantly contribute to the study of genomic evolution and subsequent phylogenetic analysis within this superfamily, as well as other early-derived lepidopteran clades.

The ailanthus webworm (*Atteva aurea* Fitch, 1856) is a small, colourful moth predominantly found north of Costa Rica, across the USA and in southern Quebec and Ontario, Canada (Wilson et al. 2010). Populations distributed south of Costa Rica in Uruguay and Argentina are known as *A. pustulella* (Fabricius, 1787), the former classification of *A. aurea* in North America (Wilson et al. 2010).

In this study, we present the complete mitogenome of *A. aurea*, the first species in the family Attevidae (superfamily Yponomeutoidea, order Lepidoptera) to be sequenced. The sequence was analysed in terms of its mitogenome characteristics and phylogenetic position within the superfamily Yponomeutoidea. Additionally, the DNA barcoding region of *A. aurea* was compared to that of previously-registered *A. aurea* and *A. pustulella*, which have been used for extensive phylogenetic analysis (Wilson et al. 2010), to further confirm sequence divergence between the two species.

Materials and methods

Sample collection, DNA extraction, PCR and sequencing

In 2011, a brood of *A. aurea* was collected from the Paint Branch Trail at the University of Maryland (College Park, MA, USA; 38°59'39"N, 76°56'5"W). In this study, DNA was extracted from the whole body of one adult male using the Wizard Genomic DNA Purification Kit (Promega, Madison, WI, USA). Using Lepidoptera-specific primers (Suppl. material 1, Kim et al. 2012), three overlapping long fragments (LFs; *COI* to *ND4*, *ND5* to *IrRNA* and *IrRNA* to *COI*) were amplified. These LFs were then used as templates for the amplification of 26 short fragments (SFs) using the same Lepidoptera-specific primers (Suppl. material 1, Kim et al. 2012). All products were sequenced in both forward and

reverse transcriptional directions by Sanger's methods. The whole body of the specimen was consumed in the process. Thus, other individuals of the brood were moved as voucher specimens to the Gongju National University of Education (Gongju, South Korea) and labelled with accession nos. GNUE-I-0001–GNUE-I-0003.

Boundary delimitation and annotation

Individual SF sequences were manually assembled into complete mitogenomes using SeqMan (DNASTAR, Madison, WI, USA). The identification and boundary delimitation of each gene and secondary structure folding of tRNAs were performed using the MITOS Web Server (<http://mitos.bioinf.uni-leipzig.de/index.py>) and using the default search mode, Mito/Chloroplast as the searching source and the genetic code of invertebrate mitogenomes for tRNA isotype prediction (Lowe and Chan 2016). Where necessary, mitogenome sequences of species in the superfamily Yponomeutoidea registered in GenBank were downloaded and aligned for improved annotation by following the protocols presented by Cameron 2014.

Phylogenetic analysis

Phylogenetic analysis was conducted using 25 available mitogenomes in 23 species (including *A. aurea*) in the superfamilies Gracillarioidea, Yponomeutoidea and Tineoidea. We selected Gracillarioidea and Tineoidea, along with Yponomeutoidea, because of the previously established sister-group relationship between Yponomeutoidea and Gracillarioidea and of the branching of Tineoidea as a lineage basal to these two superfamilies (Timmermans et al. 2014, Breinholt et al. 2018, Bao et al. 2019, Kawahara et al. 2019). Two species within the superfamily Nepticuloidea (*Stigmella roborella* and *Astrotischeria* sp.) were used as outgroups. Thirteen protein-coding genes (PCGs) and two rRNA genes (including those of two outgroup species) were aligned using RevTrans ver. 2.0 (Wernersson and Pedersen 2003) and concatenated using SequenceMatrix ver. 1.8 (Vaidya et al. 2011). The Maximum Likelihood method was applied using CIPRES Portal ver. 3.1 (Miller et al. 2010) for phylogenetic analyses, based on the GTR + Gamma + I model, which was selected using jModelTest (Posada 2008).

Data resources

Genome sequence data used in this study are openly available from the GenBank database of the National Center for Biotechnology Information (<https://www.ncbi.nlm.nih.gov>) under the accession no. ON480203. All datasets used in this study were published at Zenodo on 03 July 2022 (Zenodo. <https://doi.org/10.5281/zenodo.6791899>).

Results and discussion

Genome summary

The complete 16,392-bp mitogenome of *A. aurea* is composed of typical gene sets (two rRNAs, 22 tRNAs and 13 PCGs) and a major non-coding A + T-rich region (Table 1; GenBank accession no. ON480203). Twelve of the 13 PCGs have the typical ATN start codon, whereas *COI* has the atypical CGA codon, which is found in the majority of other available mitogenome sequences of Tineoidea, Gracillarioidea and Yponomeutoidea species (data not shown), as well as Lepidoptera species (Kim et al. 2010, Park et al. 2016, Kim et al. 2018, Jeong et al. 2021). Ten PCGs end with typical stop codon TAA, whereas *COII* and *ND4* have incomplete, single-thymine stop codons. The A/T content is 79.9% in PCGs, 82.4% in tRNAs, 82.5% in the whole genome, 86.5% in *lrRNA*, 87.6% in *srRNA* and 98.0% in the A + T-rich region (data not shown).

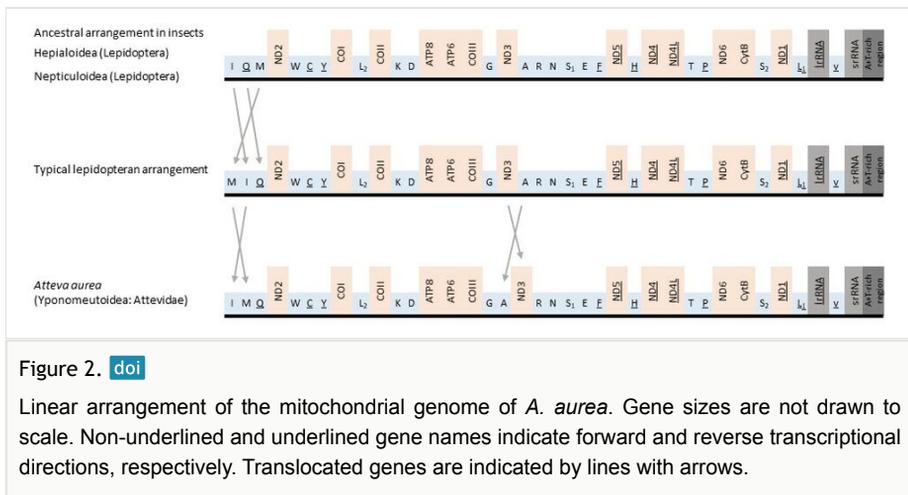
Table 1.

Summary of *Atteva aurea* mitochondrial genome.

Gene	Nucleotide number	Size	Anticodon	Start codon	Stop codon	O/S
<i>trnI</i>	1-64	64	GAT 29-31			
<i>trnM</i>	65-132	68	CAT 96-98			-8
<i>trnQ</i>	141-209	69	TTG 177-179			-41
<i>ND2</i>	251-1276	1026		ATT	TAA	+2
<i>trnW</i>	1275-1340	64	TCA 1305-1307			+8
<i>trnC</i>	1333-1401	69	GCA 1369-1371			-7
<i>trnY</i>	1409-1474	66	GTA 1441-1443			-5
<i>COI</i>	1480-3015	1536		CGA	TAA	-54
<i>trnL₂</i>	3070-3136	67	TAA 3100-3102			
<i>COII</i>	3137-3818	682		ATG	T-tRNA	
<i>trnK</i>	3819-3889	71	CTT 3849-3851			+1
<i>trnD</i>	3889-3954	66	GTC 3919-3921			
<i>ATP8</i>	3955-4116	162		ATC	TAA	+7
<i>ATP6</i>	4110-4781	672		ATG	TAA	+1
<i>COIII</i>	4781-5569	789		ATG	TAA	-2
<i>trnG</i>	5572-5640	69	TCC 5602-5604			-332
<i>trnA</i>	5973-6039	67	TGC 6002-6004			-103
<i>ND3</i>	6143-6517	375		ATA	TAA	-28
<i>trnR</i>	6546-6607	62	TCG 6573-6575			-4
<i>trnN</i>	6612-6677	66	GTT 6642-6644			+1
<i>trnS₁</i>	6677-6737	61	GCT 6698-6700			-23

Gene rearrangement

Compared with that of other Lepidoptera species, the *A. aurea* mitogenome has a very rare *trnI-trnM-trnQ* arrangement (underlining indicates gene inversion) at the A + T-rich region and *ND2* junction (Fig. 2). *Monopis longella* (Walker, 1863) (family Tineidae, superfamily Tineoidea) is the only species previously known to exhibit the *trnI-trnM-trnQ* arrangement in the Ditrysia clade, including the superfamilies Gracillarioidea, Yponomeutoidea, and Tineoidea (Jeong et al. 2021). Conversely, the majority of ditryisian Lepidoptera species have the gene order *trnM-trnI-trnQ* at the same junction (Fig. 2, Kim et al. 2014). This differs from the ancestral *trnI-trnQ-trnM* order found in most insects (Fig. 2, Boore 1999), including ancient, non-ditryisian lepidopteran groups, such as Hepialoidea and Nepticuloidea (Cao et al. 2012, Timmermans et al. 2014). Moreover, the *A. aurea* mitogenome has the *trnA-ND3* arrangement at the *trnG* and *trnR* junctions instead of the *ND3-trnA* arrangement found in almost all Lepidoptera species, including all those in Gracillarioidea, Yponomeutoidea, and Tineoidea (Fig. 2, Kim et al. 2014, Jeong et al. 2021). Thus far, only seven species in six genera across five families in Yponomeutoidea have had their mitogenomes sequenced. Thus, further analysis of this superfamily is required to make any conclusive remarks about the evolution of this rearrangement. Nevertheless, current analysis indicates that the arrangement of the family Attevidae is an autapomorphic characteristic of the superfamily Yponomeutoidea (data not shown).



Comparison of DNA barcoding sequence

The comparison between the DNA barcoding sequences of current *A. aurea* and those of *A. aurea* previously registered on GenBank, including those registered by Wilson et al. (2010), showed a 0.00–1.67% divergence. Compared with *A. pustulella* DNA barcoding sequences, there was a divergence of at least 3.95% (data not shown). This reflects the findings of a previous study that *A. aurea*, distributed between Costa Rica and southern Quebec and Ontario, is indeed *A. aurea* and that phylogenetic results demonstrate a clear

separation of *A. aurea* from other Attevidae species, including *A. pustulella* (Wilson et al. 2010).

Phylogenetic analysis

Phylogenetic analysis revealed overall lower nodal supports for familial relationships within Yponomeutoidea. A sister relationship between the families Attevidae and Praydidae, each of which is represented by a single species, was supported, but the nodal support for this relationship was not high (bootstrap support (BS) = 64%; Fig. 3). Within the Ditrysia clade, Gracillarioidea and Yponomeutoidea exhibit a sister relationship with the highest nodal support, placing Tineoidea sister to the two superfamilies with the highest support (Fig. 3). Previously, Sohn et al. (2013), using 8.0–18.9 kb of 8–27 genes from 11 families in Yponomeutoidea, also revealed a sister relationship between Attevidae and Praydidae and this relationship was supported with relatively high nodal support. In terms of relationships between superfamilies, our findings are consistent with previous studies based on mitogenomic, molecular, morphological, genomic and transcriptome data that proposed a sister-group relationship between Yponomeutoidea and Gracillarioidea, with Tineoidea diverging earlier within ditrysiian Lepidoptera (Heikkilä et al. 2014, Timmermans et al. 2014, Breinholt et al. 2018, Bao et al. 2019, Kawahara et al. 2019). Additional phylogenetic relationships within the early-derived groups of Lepidoptera could be determined, based on further taxonomic research with a wider scope.

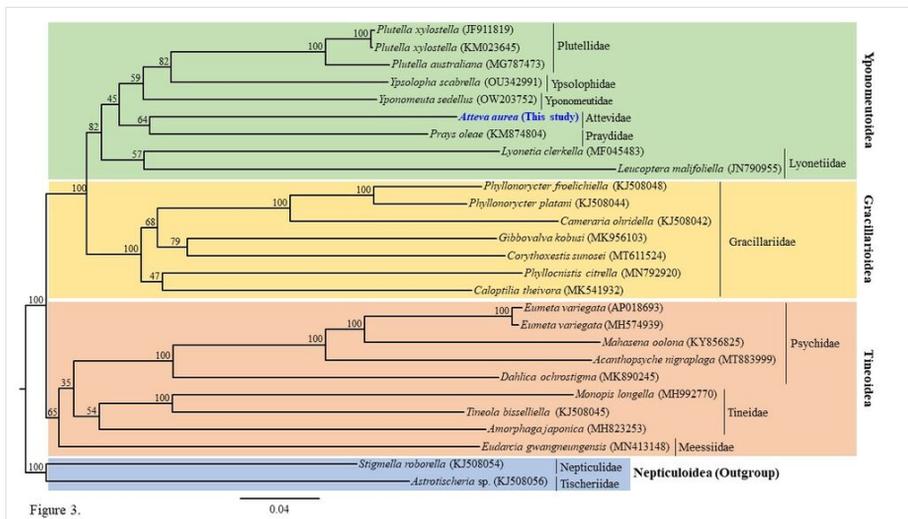


Figure 3.

Figure 3. [doi](#)

Phylogenetic tree of the three ditrysiian superfamilies included in this study: Tineoidea, Gracillarioidea and Yponomeutoidea. The number at each node indicates the bootstrap value. The scale bar indicates the number of substitutions per site. The GenBank accession number of each species is shown in brackets after its scientific name.

Conclusions

This mitogenome of *A. aurea* has a unique *trnI-trnM-trnQ* arrangement at the A + T-rich region and *ND2* junction and *trnA-ND3* arrangement at the *trnG* and *trnR* junction, which is unprecedented in Yponomeutoidea. Thus, additional mitogenome sequences are required from other genera, subfamilies and families to understand the taxonomic extent of this arrangement in Yponomeutoidea. Phylogenetic analysis revealed a sister relationship between Attevidae and Praydidae, consistent with the results of a previous large-scale molecular phylogenetic study, but nodal support was not high in this study. The result of our DNA barcoding sequence comparison supports the finding of a previous study that *A. aurea*, occurring north of Costa Rica in the USA and southern Quebec and Ontario, is genetically distinct from *A. pustulella* distributed from Costa Rica south to Uruguay and Argentina. Including that of *A. aurea*, only nine mitogenome sequences, representing seven genera across six families, are currently available for the superfamily Yponomeutoidea. Thus, more mitogenome sequences from the early-derived groups of Lepidoptera, including Yponomeutoidea, are essential for a greater understanding of mitogenome evolution and phylogenetic relationships in this order.

Acknowledgements

This study was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry through the Agriculture, Food and Rural Affairs Convergence Technologies Program and Educating Creative Global Leader Program, funded by the Ministry of Agriculture, Food and Rural Affairs (grant no. 321001-03).

Funding program

This study was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry through the Agriculture, Food and Rural Affairs Convergence Technologies Program and Educating Creative Global Leader Program, funded by the Ministry of Agriculture, Food and Rural Affairs (grant no. 321001-03).

Hosting institution

Chonnam National University, Gwangju, South Korea

Ethics and security

We declare that there are no violations of the guidelines of the authors' respective institutions and local, national and international regulations.

Author contributions

Conceptualisation, J.-C.S., J.S.J. and I.K.; fieldwork, J.-C.S.; data analysis and interpretation, J.S.J., J.S.P., H.K.O. and M.J.K.; writing—original draft, J.S.J., J.-C.S., M.J.K. and I.K.; writing—review and editing, J.-C.S., J.S.P., H.K.O. and I.K.; supervision, I.K.; project administration, J.S.P. and I.K.; funding acquisition, I.K.

Conflicts of interest

We declare no potential conflicts of interest.

References

- Bao L, Zhang Y, Gu X, Gao Y, Yu Y (2019) The complete mitochondrial genome of *Eterusia aedea* (Lepidoptera, Zygaenidae) and comparison with other zygaenid moths. *Genomics* 111 (5): 1043-1052. <https://doi.org/10.1016/j.ygeno.2018.06.007>
- Boore JL (1999) Animal mitochondrial genomes. *Nucleic Acids Research* 27 (8): 1767-1780. <https://doi.org/10.1093/nar/27.8.1767>
- Breinholt J, Earl C, Lemmon A, Lemmon EM, Xiao L, Kawahara A (2018) Resolving relationships among the megadiverse butterflies and moths with a novel pipeline for anchored phylogenomics. *Systematic Biology* 67 (1): 78-93. <https://doi.org/10.1093/sysbio/syx048>
- Cameron S (2014) How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Systematic Entomology* 39 (3): 400-411. <https://doi.org/10.1111/syen.12071>
- Cao Y, Ma C, Chen J, Yang D (2012) The complete mitochondrial genomes of two ghost moths, *Thitarodes renzhiensis* and *Thitarodes yunnanensis*: the ancestral gene arrangement in Lepidoptera. *BMC Genomics* 13 (1). <https://doi.org/10.1186/1471-2164-13-276>
- Heikkilä M, Mutanen M, Kekkonen M, Kaila L (2014) Morphology reinforces proposed molecular phylogenetic affinities: a revised classification for Gelechioidea (Lepidoptera). *Cladistics* 30 (6): 563-589. <https://doi.org/10.1111/cla.12064>
- Jeong SY, Park JS, Kim MJ, Kim SS, Kim I (2021) The complete mitochondrial genome of *Monopis longella* Walker, 1863 (Lepidoptera: Tineidae). *Mitochondrial DNA Part B* 6 (8): 2159-2161. <https://doi.org/10.1080/23802359.2021.1944389>
- Kawahara AY, Plotkin D, Espeland M, Meusemann K, Toussaint EFA, Donath A, Gimnich F, Frandsen PB, Zwick A, Dos Reis M, Barber JR, Peters RS, Liu S, Zhou X, Mayer C, Podsiadlowski L, Storer C, Yack JE, Misof B, Breinholt JW, et al. (2019) Phylogenomics reveals the evolutionary timing and pattern of butterflies and moths. *Proceedings of the National Academy of Sciences of the United States of America* 116 (45): 22657-22663. <https://doi.org/10.1073/pnas.1907847116>
- Kim JS, Park JS, Kim MJ, Kang PD, Kim SG, Jin BR, Han YS, Kim I (2012) Complete nucleotide sequence and organization of the mitochondrial genome of eri-silkworm, *Samia cynthia ricini* (Lepidoptera: Saturniidae). *Journal of Asia-Pacific Entomology* 15 (1): 162-173. <https://doi.org/10.1016/j.aspen.2011.10.002>

- Kim JS, Kim MJ, Jeong JS, Kim I (2018) Complete mitochondrial genome of *Saturnia jonasii* (Lepidoptera: Saturniidae): Genomic comparisons and phylogenetic inference among Bombycoidea. *Genomics* 110 (5): 274-282. <https://doi.org/10.1016/j.ygeno.2017.11.004>
- Kim MJ, Wan X, Kim KG, Hwang JS, Kim I (2010) Complete nucleotide sequence and organization of the mitogenome of endangered *Eumenis autonoe* (Lepidoptera: Nymphalidae). *African Journal of Biotechnology* 9 (5): 735-754. <https://doi.org/10.5897/ajb09.1486>
- Kim MJ, Wang AR, Park JS, Kim I (2014) Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperidae) and phylogenetic reconstruction of Lepidoptera. *Gene* 549 (1): 97-112. <https://doi.org/10.1016/j.gene.2014.07.052>
- Lowe T, Chan P (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Research* 44 <https://doi.org/10.1093/nar/gkw413>
- Miller M, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE) <https://doi.org/10.1109/gce.2010.5676129>
- Park JS, Kim MJ, Jeong SY, Kim SS, Kim I (2016) Complete mitochondrial genomes of two gelechioids, *Mesophleps albilinella* and *Dichomeris ustalella* (Lepidoptera: Gelechiidae), with a description of gene rearrangement in Lepidoptera. *Current Genetics* 62 (4): 809-826. <https://doi.org/10.1007/s00294-016-0585-3>
- Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* 25 (7): 1253-1256. <https://doi.org/10.1093/molbev/msn083>
- Sohn J, Regier J, Mitter C, Davis D, Landry J, Zwick A, Cummings M (2013) A Molecular Phylogeny for Yponomeutoidea (Insecta, Lepidoptera, Ditrysia) and its implications for classification, biogeography and the evolution of host plant use. *PLoS ONE* 8 (1). <https://doi.org/10.1371/journal.pone.0055066>
- Timmermans MT, Lees D, Simonsen T (2014) Towards a mitogenomic phylogeny of Lepidoptera. *Molecular Phylogenetics and Evolution* 79: 169-178. <https://doi.org/10.1016/j.ympev.2014.05.031>
- Vaidya G, Lohman D, Meier R (2011) SequenceMatrix: concatenation software for the fast assembly of multi-gene datasets with character set and codon information. *Cladistics* 27 (2): 171-180. <https://doi.org/10.1111/j.1096-0031.2010.00329.x>
- Wernersson R, Pedersen A (2003) RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Research* 31 (13): 3537-3539. <https://doi.org/10.1093/nar/gkg609>
- Wilson J, Landry J, Janzen D, Hallwachs W, Nazari V, Hajibabaei M, Hebert P (2010) Identity of the ailanthus webworm moth (Lepidoptera, Yponomeutidae), a complex of two species: evidence from DNA barcoding, morphology and ecology. *ZooKeys* 46: 41-60. <https://doi.org/10.3897/zookeys.46.406>

Supplementary material

Suppl. material 1: List of primers used to amplify and sequence the *Atteva aurea* mitochondrial genome [doi](#)

Authors: Jeong et al.

Data type: Primer list

[Download file](#) (33.10 kb)