Insect Systematics and Diversity, (2019) 3(6): 12; 1–8
doi: 10.1093/isd/ixz025
Research

OXFORD

Molecular phylogenetics, phylogenomics, and phylogeography

Phylogeny of the Hawkmoth Tribe Ambulycini (Lepidoptera: Sphingidae): Mitogenomes from Museum Specimens Resolve Major Relationships

Martijn J. T. N. Timmermans,[1,2,5] Sainab M. Daghmoumi,[1] Deborah Glass,[2] Chris A. Hamilton,[3,4] Akito Y. Kawahara,[4] and Ian J. Kitching[2]

[1]Department of Natural Sciences, Middlesex University, The Burroughs, London NW4 4BT, UK, [2]Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK, [3]McGuire Center for Lepidoptera and Biodiversity, Florida Museum of Natural History, University of Florida, Gainesville, FL 32611, [4]Department of Entomology, Plant Pathology and Nematology, University of Idaho, Moscow, ID 83844, and [5]Corresponding author, e-mail: m.timmermans@mdx.ac.uk

Subject Editor: Marko Mutanen

Received 9 September, 2019; Editorial decision 7 November, 2019

Abstract

Ambulycini are a cosmopolitan tribe of the moth family Sphingidae, comprised of 10 genera, 3 of which are found in tropical Asia, 4 in the Neotropics, 1 in Africa, 1 in the Middle East, and 1 restricted to the islands of New Caledonia. Recent phylogenetic analyses of the tribe have yielded conflicting results, and some have suggested a close relationship of the monobasic New Caledonian genus *Compsulyx* Holloway, 1979 to the Neotropical ones, despite being found on opposite sides of the Pacific Ocean. Here, we investigate relationships within the tribe using full mitochondrial genomes, mainly derived from dry-pinned museum collections material. Mitogenomic data were obtained for 19 species representing nine of the 10 Ambulycini genera. Phylogenetic trees are in agreement with a tropical Asian origin for the tribe. Furthermore, results indicate that the Neotropical genus *Adhemarius* Oiticica Filho, 1939 is paraphyletic and support the notion that *Orecta* Rothschild & Jordan 1903 and *Trogolegnum* Rothschild & Jordan, 1903 may need to be synonymized. Finally, in our analysis the Neotropical genera do not collectively form a monophyletic group, due to a clade comprising the New Caledonian genus *Compsulyx* and the African genus *Batocnema* Rothschild & Jordan, 1903 being placed as sister to the Neotropical genus *Protambulyx* Rothschild & Jordan, 1903. This finding implies a complex biogeographic history and suggests the evolution of the tribe involved at least two long-distance dispersal events.

Key words: next-generation sequencing, Bombycoidea, biogeography, paraphyletic

The Bombycoidea are one of the best studied lineages of Lepidoptera, and include several model organisms and families that are relevant for research in genetics, physiology, development and macroecology (Roe et al. 2009, Ballesteros-Mejia et al. 2017, Kitching et al. 2018). The superfamily includes many species that are economically important, either as crop pests, pollinators, human food, and silk production (Peigler 1993, Kitching and Cadiou 2000, Moré et al. 2005). Currently, there are 10 recognized families of bombycoid moths (Kitching and Rougerie et al. 2018), of which the hawkmoths (Sphingidae) may be the most spectacular. Many adult hawkmoths have the unique ability to hover while imbibing nectar from flowers with their long proboscides, whence they frequently attract the attention of the public. Hawkmoths are mostly strong and fast fliers, and many undertake long-distance dispersal flights (Beerli et al. 2019). Caterpillars are equally spectacular; they are often very large, with a characteristic curved horn at the rear end that has earned them the name 'hornworms', some even have blinking eyespots (Hossie et al. 2013, Ponce et al. 2015).

There are currently approximately 1,700 described species of sphingids, grouped into four subfamilies: Langiinae, Macroglossinae, Smerinthinae, and Sphinginae, the latter three of which are further divided into a number of tribes and subtribes (Kitching et al. 2018, Kitching 2019). Despite being the focus of numerous phylogenetic studies (e.g., Kawahara et al. 2009, Kawahara and Barber 2015), the phylogeny of the Sphingidae still remains to be fully elucidated. The current study aims to contribute to its phylogeny, focusing on the tribe Ambulycini.

Rothschild and Jordan (1903) were the first to formally recognize a phylogenetic relationship among the genera currently placed in Ambulycini. However, they did not unite them into a single monophyletic group but instead divided into three groups: 1) the genus *Ambulyx* Westwood, 1847 (as *Oxyambulyx* Rothschild and Jordan 1903); 2) a group comprising *Amplypterus* Hübner, [1819] (as *Compsogene* Rothschild and Jordan 1903), *Akbesia* Rothschild and Jordan 1903 and *Batocnema* Rothschild and Jordan 1903); and 3) a New World group comprising *Adhemarius* Oiticica Filho, 1939 (as *Amplypterus* Hübner, [1819]), *Orecta* Rothschild and Jordan 1903, *Protambulyx* Rothschild and Jordan 1903, and *Trogolegnum* Rothschild and Jordan 1903. From the first two of these groups, Rothschild and Jordan (1903) then considered the remaining smerinthine genera to have evolved.

The current concept of Ambulycini is based upon Kitching and Cadiou (2000) and includes 10 genera, of which 3 are restricted to South East Asia (*Ambulyx, Amplypterus*, and *Barbourion* Clark, 1934), 4 are Neotropical (*Adhemarius, Orecta, Protambulyx*, and *Trogolegnum*), 1 is Middle Eastern (*Akbesia*), 1 is tropical African/Madagascan (*Batocnema*), and 1 is restricted to New Caledonia (*Compsulyx* Holloway 1979). Kitching and Cadiou (2000) diagnosed the Ambulycini based primarily on the shared presence of an anterior, ventral notch on the pupal cremaster. However, they admitted that the presence of this structure had been confirmed in only four of the genera (*Akbesia*, *Ambulyx, Amplypterus, Protambulyx*), and associated the remaining genera based on general morphological similarity. Kitching and Cadiou (2000) also indicated they considered a subgroup excluding *Ambulyx, Amplypterus*, and *Barbourion* was monophyletic but did not provide any supporting evidence (in fact, the synapomorphy was the shared presence of a spinose gnathos in the male genitalia). Recent molecular phylogenetic studies (Kawahara et al. 2009, Kawahara and Barber 2015, Hamilton et al. 2019) have confirmed the monophyly of the tribe, although most did not include all described genera. Kawahara and Barber (2015) included six genera (*Adhemarius, Ambulyx, Amplypterus, Batocnema, Compsulyx*, and *Protambulyx*) and sequenced five nuclear genes (*pyrimidine biosynthesis*; *dopa-decarboxylase*; *elongation factor-1α*; *Period*; and *wingless*) and one mitochondrial gene (*cytochrome c oxidase subunit I*, COI), analyzing the data with both maximum likelihood and Bayesian inference methods. Both analyses recovered the same pattern of relationships among the six genera, with the Asian *Ambulyx* and *Amplypterus* forming the sister group of the remaining four, which were related as: *Adhemarius* (*Protambulyx* (*Batocnema* + *Compsulyx*)) (Fig. 1). Contemporaneously with Kawahara and Barber (2015), Cardoso (2015) undertook a combined molecular and morphological analysis of the Ambulycini, based on the nuclear genes CAD and *wingless*, the mitochondrial gene COI and 96 characters derived from the adult external morphology. A combined analysis using Bayesian inference (BI) recovered a monophyletic Ambulycini with the following phylogenetic relationships among the 10 genera: *Barbourion* (*Ambulyx* (*Amplypterus* ((*Compsulyx* (*Batocnema* (*Akbesia* + *Protambulyx*))) (*Adhemarius* (*Orecta*, *Adhemarius*, (*Adhemarius*, *Trogolegnum*))) (Fig. 1). In contrast, maximum parsimony (MP) analyses under both equal and implied weighting yielded a slightly different topology: *Barbourion* (*Ambulyx* (*Amplypterus* (*Protambulyx* ((*Compsulyx* + *Akbesia*) (*Batocnema* (*Orecta* (*Adhemarius* (*Adhemarius* (*Adhemarius*, *Trogolegnum*)))) (Fig. 1). These results differed from those of Kawahara and Barber (2015) in not grouping *Ambulyx* and *Amplypterus* together, nor *Batocnema* and *Compsulyx*. Cardoso (2015) also found that *Orecta* and *Trogolegnum* were both nested within *Adhemarius*, rendering this genus paraphyletic.

The phylogenetic results of both Kawahara and Barber (2015) and Cardoso (2015) raise interesting questions regarding the biogeography of the African/Madagascan genus *Batocnema* and New Caledonian genus *Compsulyx*. Although the former study grouped them together and the latter had them splitting off sequentially, both studies agreed in placing these Old World genera in a clade with the New World *Protambulyx* (in Cardoso's analysis, this clade also included *Akbesia*, a genus missing from the study of Kawahara & Barber) and placing these genera together as the sister-group of the New World genus *Adhemarius*.

In the present study, we use full mitochondrial genomes, derived from dried museum specimens as old as 28 yr, to elucidate further the phylogenetic relationships of the genera of Ambulycini.

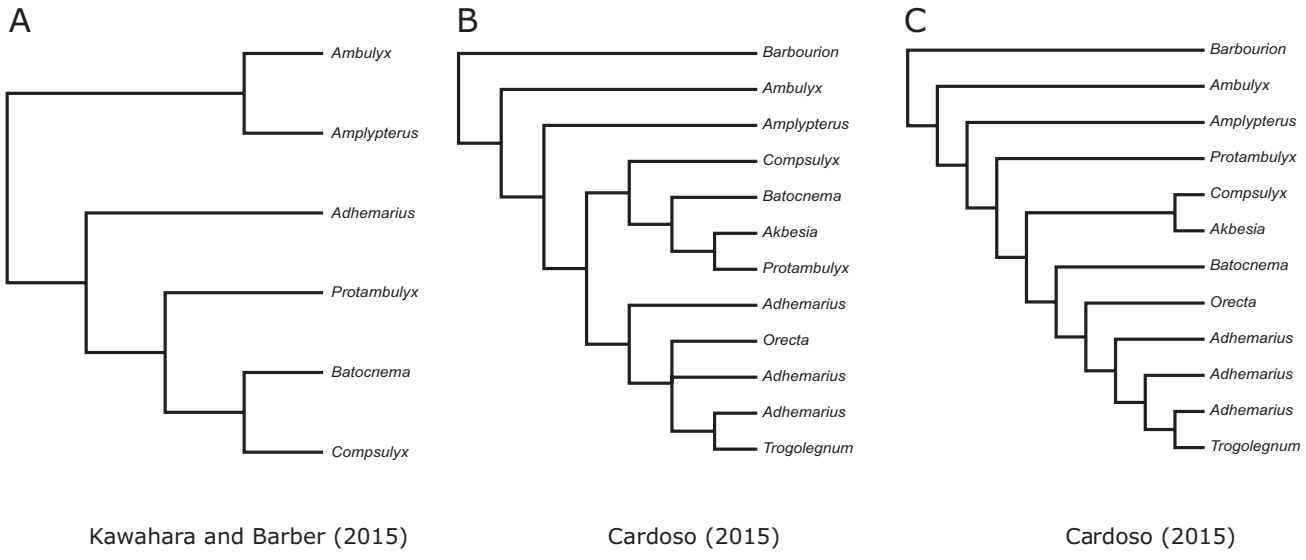

Fig. 1. Phylogenetic hypotheses for the hawkmoth tribe Ambulycini. (A) Based on Kawahara and Barber (2015), which used six genes and Maximum Likelihood and Bayesian Inference methods. (B and C) Based on Cardoso (2015), which used 3 genes and 96 morphological characters. For B, Bayesian Inference was used; for C, Maximum Parsimony.

Material and Methods

Samples, DNA Extraction, and Pooling

Mitochondrial genomes were sequenced from pooled genomic DNA samples (e.g., see Gillett et al. 2014, Timmermans et al. 2015). A single specimen from each of 22 species was selected for sequencing from the dry-pinned Sphingidae collection of the Natural History Museum, London (NHMUK) (Table 1), aiming for comprehensive generic-level taxonomic coverage and giving preference to specimens with more recent collection dates. Three species of the tribe Leucophlebiini, which is the putative sister-group of the Ambulycini (Kawahara and Barber 2015), were included as outgroup taxa: *Clanis bilineata* (Walker 1866), *Viriclanis kingstoni* Aarvik, 1999, and *Leucophlebia lineata* Westwood, 1847. Dorsal and ventral sides of these specimens and their data labels were digitally imaged using a Canon EOS 600D digital camera. Images have been uploaded into the NHM's Data Portal (https://data.nhm.ac.uk/) where they are available for open-access download (Table 1). A single hindleg from each specimen was taken for DNA extraction. Two different extraction methods were used: the Qiagen Blood and Tissue kit (Hilden, Germany) was used to extract DNA from *Adhemarius dariensis* and the specimens collected during or after 2004, with the exception of *Adhemarius sexoculata* and *Akbesia davidi*. For specimens of the latter two species, and for all other specimens collected prior to 2004, the method described by Thomsen et al. (2009) was used (Table 1). This approach was deemed more suitable for the older samples as it was specifically developed to extract degraded DNA from dried museum specimens. Extractions were performed on intact legs, ensuring the leg was completely submersed in lysis buffer. DNA purity and concentration were measured with a NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA) and a Qubit fluorometer (Thermo Fisher Scientific) using the Broad-Range (BR) assay kit. Genomic DNA (gDNA) was subsequently pooled (Table 1). To ensure sufficient genetic divergence within pooled samples (to assist correct assembly of individual mitochondrial genomes), the percentage of identical bases (% identity) were first calculated among the ambulycine species sampled using Cytochrome C Oxidase Subunit COI ('barcode region') sequences obtained from GenBank. Calculations of % identity were performed in Geneious (version 8) (https://www.geneious.com). Based on these similarity values (Supp Table 1 [online only]), we decided to construct two gDNA pools (Table 1). Equal amounts of gDNA from each sample were used for pooling.

Sequencing and Quality Control

Indexed TruSeq Nano Libraries (Illumina, San Diego, CA) were prepared at the NHMUK Sequencing Facility for both gDNA pools. The DNA was expected to be highly fragmented and therefore no further shearing of gDNA was performed. Libraries were sequenced on an Illumina MiSeq (PE; 2x250 bp). Sequencing data were preprocessed using Illumina's MiSeq Control Software (MCS), version 3.1 (Illumina).

Further processing largely followed Timmermans et al. (2015), which involved trimming low-quality bases at the start and end of reads (phred quality threshold 20) using TRIMMOMATIC (version 0.32) (Bolger et al. 2014), stitching paired-end reads using PEAR (default settings) (Zhang et al. 2014), and removing all stitched sequences with a minimum quality score of 20 from the dataset using prinseq-lite (Schmieder and Edwards 2011). Files were subsequently converted to fasta format using the Unix stream editor, sed. Finally, data were assembled using the de Bruijn graph assembler idba_ud (--mink=80, --maxk=150) (Peng et al. 2012).

The *Compsulyx* mitogenome assembly remained incomplete. To verify correct assembly of the partial mitochondrial genome, a 382 bp fragment of Cytochrome b (CYTB) was PCR amplified from sample NHMUK 010928590 (Table 1), using primers Sytb_F (5′-TGAGGNCAAATATCHTTYTGAGG-3′) and Sytb_R (5′-GCA AATARRAARTATCATTCDGG-3′) (Timmermans et al. 2010), and sequenced on an ABI 3730XL (Applied Biosystems, California). The resultant CYTB Sanger sequence and a COI sequence (GenBank accession number: KP720036; 657 bp) were mapped onto the mitogenome in Geneious and checked for discrepancies.

Data for *Batocnema africanus* (LEP31448d) were extracted from the raw Illumina sequencing reads of a previously sequenced anchored hybrid enrichment specimen (see Hamilton et al. (2019) for methodological details). Sequences were mapped onto the newly generated *Batocnema coquerelii* mitochondrial genome in Geneious (maximum gap size: 50, maximum mismatches per read: 30%).

Phylogenetic Analyses

Mitochondrial genomes were filtered from the assembly data using stand-alone BLAST (Altschul et al. 1997). The blastn searches used an *Orecta lycidas* COI sequence (GenBank accession number: GU703851) as the query sequence. Geneious was used to manually assemble partial genomes into full ones and to check whether the mitochondrial genomes were circular. Genomes were annotated by aligning them to the publicly available *Ampelophaga rubiginosa* mitogenome (GenBank accession number: NC_035431) and transferring across the annotations, which were visually inspected to ensure the correct start and stop codons were selected. Sequences for each of the 13 protein-coding genes were extracted and aligned using the codon-based aligner, MACSE (Ranwez et al. 2011) using default settings (gap penalty: 7, gap extension penalty: 1, stop codon penalty: 100, frameshift penalty: 30) and the mitochondrial genetic code. The 13 alignments were then concatenated into a single data string for each species using a custom PERL script (A.P. Vogler, personal communication). To investigate saturation in the dataset, plots of the p-distance against model corrected distance (TN93; Tamura and Nei 1993) were generated for each species pair using the ape package (Paradis et al. 2004) in R (R Core Team 2013). Maximum Likelihood phylogenetic analyses were performed using IQ-TREE (Nguyen et al. 2015) and Bayesian Inference in MrBayes (Ronquist and Huelsenbeck 2003) on a partitioned supermatrix dataset (six partitions; by strand and codon position). IQ-TREE was run with the following command: iqtree -spp partitions.txt -s <FASTA FILE> -m MFP+MERGE -nt AUTO -bb 1000 -alrt 1000. This command structure tells IQ-TREE to find the best model for each partition and subsequently merge partitions until an optimal partition scheme is found. The program then uses this for phylogenetic inference and performs an Ultrafast Bootstrap (Minh et al. 2013) and SH-like approximate likelihood ratio test (SH-aLRT) (Guindon et al. 2010) with 1,000 replicates each. MrBayes analyses were run for 1 million generations (two MCMC with four chains each; GTR+I+G model; unlinked model parameters across partitions). The first 25% of trees were discarded as burn-in and Posterior Probabilities calculated. Finally, the R library phytools (Revell 2012) was used to plot the tips of the Bayesian topology onto a world map.

Results

Mitogenome Similarity and Completeness

COI similarity between ambulycine species ranged between 86% (*Protambulyx astygonus* vs *A. dariensis*) and 93% (*Compsulyx*

Table 1. Sample details

Species	Species author and date	NHMUK specimen number	GenBank accession number	Year	Extraction method	DNA concentration (ng/µl)	Mitogenome length	Country	NHM Data Portal URL
Pool 1									
Protambulyx ockendeni	Rothschild and Jordan, 1903	NHMUK010928139	MK804162	2000	QIAquick	11.5	15395	Peru	https://data.nhm.ac.uk/object/eb6f2ab7-bcce-4a8a-92df-b4229380bb82/1559606400000
Protambulyx astygonus	Boisduval, [1875])	NHMUK010928137	MK804160	1993	QIAquick	7.13	15345	Paraguay	https://data.nhm.ac.uk/object/d7341bc5-4163-4aec-a439-e68c6d3fcb80/1559606400000
Compsulyx cochereaui	(Viette 1971)	NHMUK010928133		1984	Qiaquick	0.707	n.a.	New Caledonia	https://data.nhm.ac.uk/object/0b00aa07-1f96-4969-97a8-c377681480b5/1559606400000
Orecta lycidas	(Boisduval, [1875])	NHMUK010928136	MK804159	1991	QIAquick	15.1	15387	Brazil	https://data.nhm.ac.uk/object/003fb1ce-0ce7-49a0-b9e4-7d4d42f2678d/1559606400000
Batocnema africanus	(Distant, 1899)	NHMUK010928132		1999	QIAquick	8.93	n.a.	Tanzania	https://data.nhm.ac.uk/object/6aab9b7f-ec08-4a04-a982-29671651611a/1559606400000
Amplypterus panopus	(Cramer, 1779)	NHMUK010928129	MK804153	2013	Blood and Tissue	7.03	15370	Malaysia	https://data.nhm.ac.uk/object/f342c1a7-c387-4e96-859e-057a529bcf8d/1559606400000
Ambulyx substrigilis	Westwood, 1847	NHMUK010928146	MK804151	2014	Blood and Tissue	20.3	15333	Vietnam	https://data.nhm.ac.uk/object/22e8b379-c335-40bb-a946-bcc86e885bb6/1559606400000
Akbesia davidi	(Oberthür, 1884)	NHMUK010928126		2005	QIAquick	4.23	n.a.	Syria	https://data.nhm.ac.uk/object/f605a305-9ab2-493c-9c8b-6e11b0c57976/1559606400000
Adhemarius sexoculata	(Grote, 1865)	NHMUK010928125	MK804149	2005	QIAquick	20.8	12796*	Ecuador	https://data.nhm.ac.uk/object/57dfc885-37c0-4072-a340-f4cf34c2be3f/1559606400000
Adhemarius dariensis	(Rothschild & Jordan, 1916)	NHMUK010928124	MK784108	2003	Blood and Tissue	4.32	15676	Panama	https://data.nhm.ac.uk/object/4e39460a-2b3f-4383-baf3-a51a2b0873d5/1559606400000
Clanis bilineata	(Walker, 1866)	NHMUK010928120	MK804156	2014	Blood and Tissue	8.5	15426	Vietnam	https://data.nhm.ac.uk/object/7d7a8d4c-3f3d-48d4-bd59-1323bb7f9b64/1559606400000
Pool 2									
Protambulyx striglis	(Linnaeus, 1771)	NHMUK010928140	MK804163	2007	Blood and Tissue	8.35	15534	Peru	https://data.nhm.ac.uk/object/88fc4919-f24e-45a0-9919-2223afe535a0/1559606400000
Protambulyx eurycles	(Herrich-Schäffer, [1854])	NHMUK010928138	MK804161	2011	Blood and Tissue	5.41	15542	Peru	https://data.nhm.ac.uk/object/b3608737-63a4-4f29-b39c-67e19d455a10/1559606400000
Orecta acuminata	Clark, 1923	NHMUK010928135		1993	QIAquick	3.01	n.a.	Paraguay	https://data.nhm.ac.uk/object/1cbe91c4-292b-4be3-a009-31d230a922cf/1559606400000
Batocnema coquerelii	(Boisduval, [1875])	NHMUK010928134	MK804155	2007	Blood and Tissue	5.85	15361	Madagascar	https://data.nhm.ac.uk/object/86043377-620d-41fb-82a0-dac7a42c4b55/1559606400000
Barbourion lemaii	(Le Moult, 1933)	NHMUK010928130	MK804154	2016	Blood and Tissue	7.53	15366	Thailand	https://data.nhm.ac.uk/object/e362ac06-9f1b-4344-b86c-c53e03fd5a00/1559606400000
Ambulyx dohertyi	Rothschild, 1894	NHMUK010928128	MK804150	1996	QIAquick	14.3	15304	Papua New Guinea	https://data.nhm.ac.uk/object/31841297-31f8-46ca-a2e1-aefb6e51b00d/1559606400000
Amplypterus mansoni	(Clark, 1924)	NHMUK010928131	MK804152	2003	QIAquick	27.8	15394	Laos	https://data.nhm.ac.uk/object/94f9e9bc-b48f-4014-821a-85cfeeac980d/1559606400000
Trogolegnum pseudambulyx	(Boisduval, [1875])	NHMUK010928141	MK804164	2004	QIAquick	27.5	15387	Mexico	https://data.nhm.ac.uk/object/ca63639f-1cc8-43f1-ba17-aa40ef77bad2/1559606400000
Adhemarius dentoni	(Clark, 1916)	NHMUK010928123	MK804148	2000	QIAquick	17.5	15423	Unknown	https://data.nhm.ac.uk/object/a53ac9b6-3fbb-45e9-b61e-ae609aa47ca6/1559606400000

Table 1. Continued

Species	Species author and date	NHMUK specimen number	GenBank accession number	Year	Extraction method	DNA concentration (ng/µl)	Mitogenome length	Country	NHM Data Portal URL
Viriclanis kingstoni	Aarvik, 1999	NHMUK010928122		2004	QIAquick	3.47	n.a.	Tanzania	https://data.nhm.ac.uk/object/d5e66ace-f4e9-4be9-82f6-21069787399 8/1559606400000
Leucophlebia lineata	Westwood, 1847	NHMUK010928121	MK804158	1996	QIAquick	0.887	15454	Thailand	https://data.nhm.ac.uk/object/dbf461d0-2942-4378-b560-7a5aa543a633/1559606400000
Repeated *Compsulyx cochereaui*	(Viette 1971)	NHMUK010928590	MK804157	2001	Blood and Tissue		12882*	New Caledonia	https://data.nhm.ac.uk/object/0f693599-632a-4f66-83ba-8597f834947c/1566345600000

For library construction, DNA extracts were pooled as indicated (Pool 1, Pool 2). For each specimen, only the country of origin is given; further data can be found on the URL links to the specimen image pages on the NHM Data Portal. Species names not in bold face are those for which no mitochondrial genome sequence was recovered.

Mitogenome length, length of mitochondrial genome; n.a., not applicable.

* indicates that the mitochondrial genome is partial.

cochereaui vs *Batocnema africanus*) for pool 1, and 86% (*Adhemarius dentoni* vs *Protambulyx eurycles*) and 92% (*A. dentoni* vs *Barbourion lemaii* and *A. dentoni* vs *Orecta acuminata*) for pool 2 (Supp Table 1 [online only]).

A complete, circular mitochondrial genome was obtained for 7 of the 11 species in pool 1 and 9 of the 10 species in pool 2. For one additional species (*A. dentoni*), a contig of 13,731 bp was assembled. For five samples, no mitochondrial genome sequence was recovered: *A. davidi*, *B. africanus*, and *C. cochereaui* in pool 1, and *O. acuminata* and the outgroup species, *V. kingstoni* in pool 2. These were not consistently the oldest or the smallest samples, and it remains unclear why the sequencing and assembly failed to generate useable data for these specimens.

Of particular interest was the pool 1 specimen, *C. cochereaui*. The species was repeated using a different specimen in a different sequencing run and a contig of 13,347 bp obtained. This contig, like that of the above-mentioned *A. dentoni* sequence, contained all the protein-coding genes, but lacked information on the rRNAs and the d-loop region. To confirm correct assembly of the *C. cochereaui* genome, it was compared to independently derived COI and CYTB sequences. Sequences were aligned to the partial mitogenome (sequence position COI: 1,690–2,346, sequence position CYTB: 11,093–11,475) and shown to be 100% identical (i.e., no mismatch was observed).

The lengths of each of the 16 full mtDNA genomes ranged from 15,304 bp (*Ambulyx dohertyi*) to 15,676 bp (*A. dariensis*) (Table 1), slightly longer than the published genome of the sphingid *A. rubiginosa* (15,282 bp). As expected, gene order was highly conserved and matched the order typically observed in ditrysian Lepidoptera, with one exception in *A. dariensis*—a translocation of tRNA-Gln was observed from the 'tRNA-Met, tRNA-Ile, tRNA-Gln cluster' to a position in the d-loop region (Supp Fig. 1 [online only]).

Phylogenetic Analysis

Protein-coding genes were extracted, aligned, and concatenated into a single concatenated supermatrix of 11,235 bp for each species. The data matrix was supplemented with sequence data for *B. africanus* LEP31448d that had been extracted from the raw Illumina sequencing reads of a previously sequenced anchored hybrid enrichment specimen (94% complete). Saturation was investigated visually by plotting pairwise p-distances against pairwise TN93-corrected distances. The obtained plot revealed a strong linear relationship, suggesting saturation is negligible (Supp Fig. 2 [online only]). It was therefore decided not to recode or remove codon positions. Phylogenetic analyses performed on the partitioned dataset, using both Maximum Likelihood and Bayesian Inference, yielded identical topologies (Fig. 2).

The Ambulycini are recovered as monophyletic but with very poor support (SH-aLRT support = 53.4/Bootstrap support = 68.0; Posterior Probability = 0.74). The tropical Asian species, *B. lemaii*, is the first lineage to split from the rest of the samples, but its placement must be considered uncertain given that support values for the tribe as a whole on both trees are very low. In contrast, all other relationships were recovered with high support (BS ≥90.0; PP≥ 0.99), except for the pairing of *P. astygonus* and *P. strigilis* (SH-aLRT support = 78.9/Bootstrap support = 79.0; Posterior Probability = 0.99). The remaining two tropical Asian Ambulycini do not form a monophyletic group, as the genus *Amplypterus* splits off separately and after the genus *Ambulyx*. Nor were the four Neotropical genera collectively recovered as a monophyletic group. Rather, *Compsulyx* and *Batocnema* are together placed as sister to the genus *Protambulyx*. These three genera are sister to a group comprising the remaining

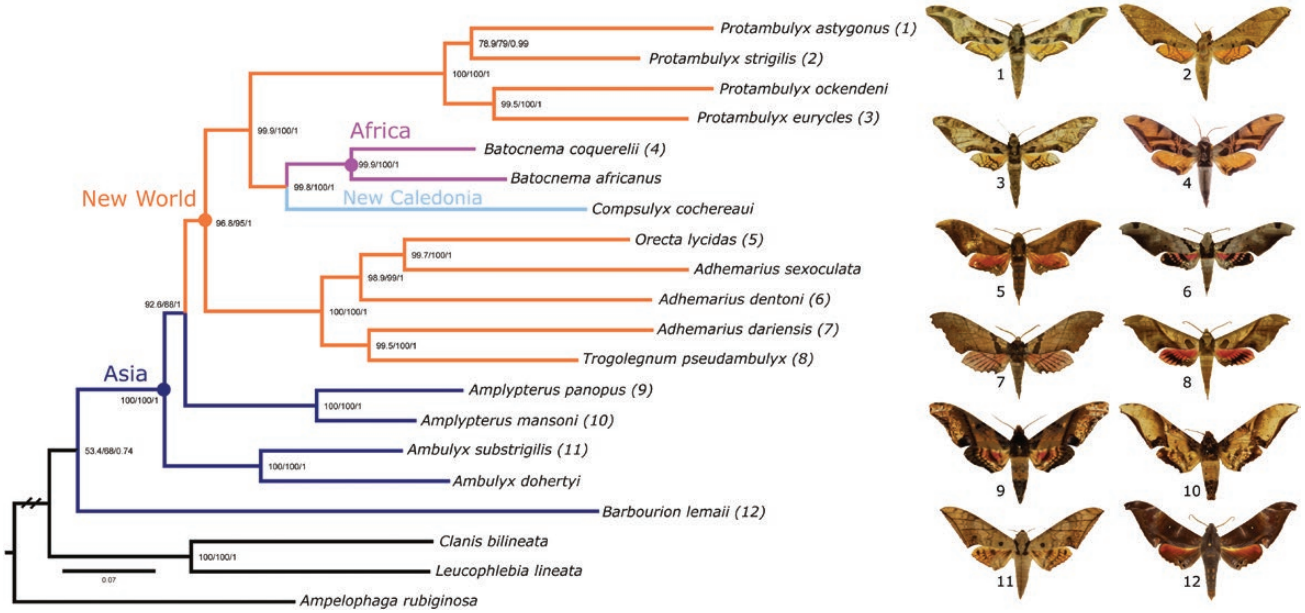

Fig. 2. Maximum Likelihood (ML) topology showing Ambulycini relationships inferred from mitochondrial genome data. Values at nodes indicate SH-aLRT/Ultrafast Bootstrap/Posterior probabilities. Posterior Probabilities were obtained using Bayesian Inference. ML and Bayesian inferences recovered the same phylogenetic relationships. Scale bar indicates number substitutions per site. Various species are represented by a photograph (indicated with a number behind the species name and next to the respective image). All images are available on the NHM Data Portal (see Table 1), except for 4) *B. coquerelii* which was taken by Laurel Kaminsky.

three Neotropical genera (*Trogolegnum*, *Orecta*, and *Adhemarius*). Within this latter group, *Adhemarius* is not recovered as monophyletic. Instead, the genera *Trogolegnum* and *Orecta* are placed within different branches of a paraphyletic *Adhemarius*, with *Trogolegnum pseudambulyx* as sister to *A. dariensis*, and *O. lycidas* being sister to *A. sexoculata*.

Discussion

Phylogeny of Ambulycini

Previous phylogenetic analyses of the Ambulycini have yielded conflicting patterns of relationships among the genera. In the present study, Ambulycini are recovered as monophyletic but with only weak support (Fig. 2), though in both analyses, *Barbourion* is the sister group to the rest of the tribe. In contrast, a clade comprising the remaining genera receives very strong support. This suggests that *Barbourion* might perhaps be misplaced in Ambulycini, although it could also be an artifact of our limited outgroup sampling. Next to split off is *Ambulyx*, then *Amplypterus*. These are followed by *Adhemarius*, in which *Orecta* and *Trogolegnum* are placed. Thus, *Adhemarius* is paraphyletic relative to the other two genera. The final clade comprises the remaining three ambulycine genera of the present analysis, with *Batocnema* placed as sister to *Compsulyx* and these two as sister to *Protambulyx*. Thus, our results are almost identical to those of the combined molecular and morphological BI analysis of Cardoso (2015; Fig. 1.7), except that *Batocnema* and *Compsulyx* are sisters, rather than arising separately, and the placement of *Orecta* is resolved.

Classification of Adhemarius, Orecta, and Trogolegnum

With regard to the phylogenetic relationships of *Adhemarius*, *Orecta*, and *Trogolegnum*, the present study (Fig. 2) found a topology in which the *Adhemarius donysa*-group + *Trogolegnum* split off first, followed by the *Adhemarius gannascus*-group, leaving a terminal sister-group pairing of the *A. sexoculata*-group and *Orecta*. This result contrasts with Cardoso (2015), who found different patterns of relationship among these groups, depending upon the analytical method and data set used: in their IW MP analysis (Cardoso 2015: Fig. 1.6), *Orecta* is first to branch off, followed by the *sexoculata*-group, then the *gannascus*-group, then finally *Trogolegnum* as sister to the *donysa*-group. In contrast, in the results of their BI analysis (Cardoso 2015: Fig. 1.7), the *sexoculata*-group branched off first, followed by a trichotomy comprising *Orecta*, the *gannascus*-group and the *donysa*-group + *Trogolegnum*. However, all analyses agree that *T. pseudambulyx* is simply a member of the *A. donysa* species-group, albeit one with a reduced proboscis and labial palps, and, like *Orecta*, spinulose abdominal tergites and nonspinose abdominal sternites (Rothschild and Jordan 1903). However, the phylogenetic relationships of *Orecta*, also considered by Rothschild and Jordan (1903) to be a derivative of *Adhemarius*, remain obscure, with each of the three analyses suggesting a different placement. We therefore consider it premature to make any formal changes to the classification of the three genera. In addition, although most of the relationships were recovered with high support, we should point out that mitochondrial genomes are maternally inherited, can introgress between hybridizing species and that the genes in the mitochondrial genome are tightly linked (Avise and Ellis 1986). It is therefore possible that the trees obtained here merely represent a deviating gene history and not the actual evolutionary history of the species involved (Ballard 2000). Phylogenomic studies currently in progress, which focus on the nuclear genome using anchored hybrid enrichment (Kawahara et al. in preparation) and ultra-conserved elements (Rougerie et al. in preparation), will show whether there is any discrepancy between the mitochondrial and nuclear genomes and are expected to unambiguously resolve the placement and relationships of ambulycine taxa, finally allowing taxonomic decisions to be made.

Biogeography of Ambulycini

Although this study is not intended to be a formal biogeographical study of the Ambulycini, it is possible to draw some preliminary conclusions based on the results presented here.

The first three genera to split from the rest of the ambulycine tree, *Barbourion*, *Ambulyx* and *Amplypterus*, are essentially tropical South-east Asian in distribution (although some *Ambulyx* species occur in more northern temperate regions) and it is likely that this region is where the tribe originated.

To date, no analysis has recovered a monophyletic group comprising only the four New World genera, *Adhemarius*, *Orecta*, *Protambulyx*, and *Trogolegnum*. Instead, in all cases, *Protambulyx* is placed in a clade together with the Old World genera, *Akbesia*, *Batocnema*, and *Compsulyx*, the sister-group of which is a clade comprising *Adhemarius*, *Orecta*, and *Trogolegnum*. If the Ambulycini originated in the Old World, then it is still unclear whether there were two independent dispersal events to the New World (the *Protambulyx* and *Adhemarius*/*Orecta*/*Trogolegnum* lineages), a single such dispersal event followed by a second back to the Old World by the *Akbesia*/*Batocnema*/*Compsulyx* group, or an even more complex scenario. The ambiguity currently surrounding the phylogenetic relationships of these genera precludes a more objective biogeographical analysis.

Taxonomy and Biogeography of Compsulyx

The monobasic genus *Compsulyx* is endemic to the main island of New Caledonia in the western Pacific, where it is particularly associated with ultramafic rainforest (Holloway 1979). Its only species, *C. cocheraeaui*, was originally described in the genus *Compsogene* (now *Amplypterus*), but Viette (1971) noted a resemblance to *Ambulyx*, although not in perfect agreement with either genus. A more thorough study by Holloway (1979) led to a conclusion that this species belonged in a separate genus, *Compsulyx*, but was nevertheless of Indo-Malayan origin: "the New Caledonian species bears closer resemblance to *Oxyambulyx* [*Ambulyx*] but certainly represents an offshoot of ambulycine [sic] stock prior to the main radiation within the other two genera [*Ambulyx* and *Amplypterus*]" (Holloway 1979: 351). Without giving explicit supporting characters, Kitching and Cadiou (2000) placed *Compsulyx* in a clade that also included *Akbesia*, *Batocnema*, and the four New World genera. In fact, the synapomorphy in question was a spinose gnathos, a character that was confirmed by Cardoso (2015) (but also recorded by him in the distantly related outgroup, *Parum colligata* [Walker 1856]), to which he added two further synapomorphies relating to the relative lengths of the diverticula on the vesica in the male genitalia and the degree of twisting of the antrum in the female genitalia, although both were rather homoplastic.

All analyses so far have recovered this clade of seven Old and New World genera (when included), but the placement of *Compsulyx* within it remains uncertain. Kawahara and Barber (2015) and the present study place *Compsulyx* as sister to *Batocnema*, whereas Cardoso (2015) placed it as either sister to *Akbesia* (MP analysis) or to a clade comprising *Akbesia, Batocnema*, and *Protambulyx* (BI analysis). It is unfortunate that we were not able to recover any mitochondrial genome sequence for *Akbesia*, as the absence of this genus makes a direct comparison with the results of Cardoso (2015) impossible, but all clearly reject Holloway's (1979) suggestion of a close relationship with *Ambulyx* and *Amplypterus*.

Regardless of its precise placement, the phylogenetic relationships of *Compsulyx* make for a highly enigmatic biogeography. To visualize this, our Bayesian topology plotted onto a map of the Earth (Fig. 3) highlights the discrepancy between the geographical distribution of *Compsulyx* and its phylogenetic placement, in which its closest relative would be *B. coquerelii* 12,000 km to the west, in Madagascar. Such long-distance sister-group relationships are rare in Lepidoptera, but not unknown. For example, Hundsdoerfer et al. (2017) reported a sister-group pairing in the hawkmoth genus *Hyles* Hübner, [1819], between *Hyles biguttata* (Walker 1856) from Madagascar and La Réunion and *Hyles livornicoides* (Lucas 1892) from Australia. However, further clarification of the biogeography of *Compsulyx* will require additional resolution of the phylogenetic relationships within the tribe.

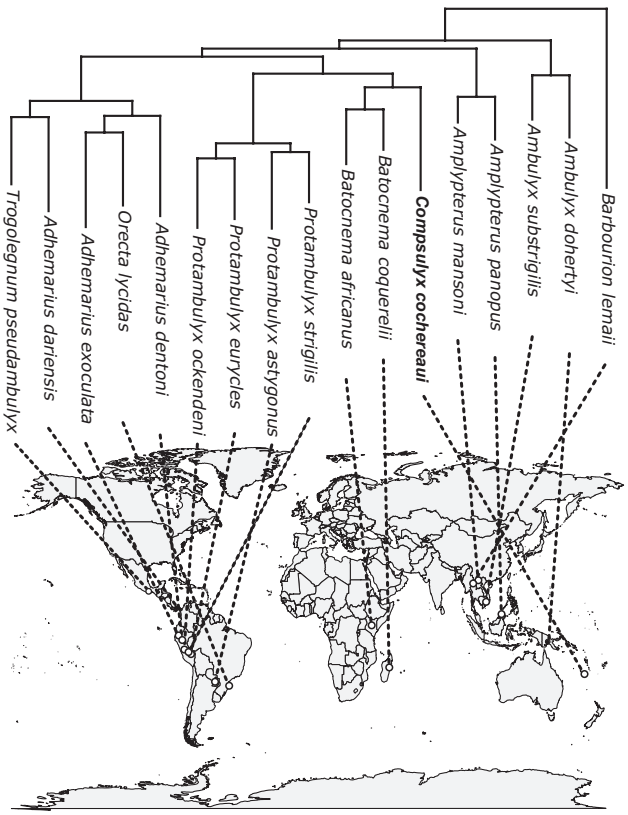

Fig. 3. The Bayesian topology plotted onto a map of the Earth.

Supplementary Data

Supplementary data are available at *Insect Systematics and Diversity* online.

Supplementary Table 1: Percentage COI similarity between Ambulycini species. For each pool, species names and GenBank accession numbers are given.

Supplementary Fig. 1: tRNA-Gln translocation in *Adhemarius dariensis*.

Supplementary Fig. 2: Saturation plot. P-distance plotted against TN93 (Tamura and Nei 1993) corrected distance.

Acknowledgments

We thank Ryan St Laurent for his discussions on Bombycoidea relationships. We thank Laurel Kaminsky and Brett Aiello for the image of *Batocnema coquerelii*. This study was funded by National Science Foundation (NSF) grant numbers DEB-1557007, IOS-1121739, IOS-1920895, DBI-1349345, DBI-1601369 to A.Y.K., and by Natural Environment Research Council grant number NE/P003915/1 to I.J.K. The Florida Museum of Natural History and the NSF Postdoctoral Fellowship award in Biology (PRFB) #1612862 supported C.A.H.

References Cited

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, W. Miller, and D. J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25: 3389–3402.

Avise, J. C. and D. Ellis. 1986. Mitochondrial DNA and the evolutionary genetics of higher animals [and discussion]. Philos. Trans. R. Soc. B Biol. Sci. 312: 325–342. doi:10.1098/rstb.1986.0011

Ballard, J. W. O. 2000. When one is not enough: introgression of mitochondrial DNA in *Drosophila*. Mol. Biol. Evol. 17: 1126–1130. doi:10.1093/oxfordjournals.molbev.a026394

Ballesteros-Mejia, L., I. J. Kitching, W. Jetz, and J. Beck. 2017. Putting insects on the map: near-global variation in sphingid moth richness along spatial and environmental gradients. Ecography 40: 698–708. doi:10.1111/ecog.02438

Beerli, N., F. Bärtschi, L. Ballesteros-Mejia, I. J. Kitching, and J. Beck. 2019. How has the environment shaped geographical patterns of insect body sizes? A test of hypotheses using sphingid moths. J. Biogeogr. jbi.13583. doi:10.1111/jbi.13583

Bolger, A. M., M. Lohse, and B. Usadel. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma. Oxf. Engl. 30: 2114–2120. doi:10.1093/bioinformatics/btu170

Cardoso, L. W. 2015. Contribuição de marcadores morfológicos e moleculares na elucidação da Sistemática de Ambulycini (Lepidoptera, Sphingidae, Smerinthinae) (Mestrado em Zoologia). Universidade de São Paulo, São Paulo. doi:10.11606/D.41.2015.tde-23092015-081044

Gillett, C. P. D. T., A. Crampton-Platt, M. J. T. N. Timmermans, B. H. Jordal, B. C. Emerson, and A. P. Vogler. 2014. Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionoidea). Mol. Biol. Evol. 31: 2223–2237. doi:10.1093/molbev/msu154

Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. 59: 307–321. doi:10.1093/sysbio/syq010

Hamilton, C. A., R. A. St Laurent, K. Dexter, I. J. Kitching, J. W. Breinholt, A. Zwick, M. J. T. N. Timmermans, J. R. Barber, and A. Y. Kawahara. 2019. Phylogenomics resolves major relationships and reveals significant diversification rate shifts in the evolution of silk moths and relatives. BMC Evol. Biol. 19: 182. doi:10.1186/s12862-019-1505-1

Holloway, J. D., 1979. A survey of the Lepidoptera, biogeography and ecology of New Caledonia. pp. 588. *In* E. Schmitschek and K. A. Spencer (eds.), Series entomologica 15. Dr. W. Junk B.V. Publishers, The Hague, Boston, London.

Hossie, T. J., T. N. Sherratt, D. H. Janzen, and W. Hallwachs. 2013. An eyespot that "blinks": an open and shut case of eye mimicry in *Eumorpha* caterpillars (Lepidoptera: Sphingidae). J. Nat. Hist. 47: 2915–2926. doi:10.1080/00222933.2013.791935

Hundsdoerfer, A. K., M. Päckert, C. Kehlmaier, P. Strutzenberger, and I. J. Kitching. 2017. Museum archives revisited: central Asiatic hawkmoths reveal exceptionally high late Pliocene species diversification (Lepidoptera, Sphingidae). Zool. Scr. 46: 552–570. doi:10.1111/zsc.12235

Kawahara, A. Y., and J. R. Barber. 2015. Tempo and mode of antibat ultrasound production and sonar jamming in the diverse hawkmoth radiation. Proc. Natl. Acad. Sci. USA 112: 6407–6412. doi:10.1073/pnas.1416679112

Kawahara, A. Y., A. A. Mignault, J. C. Regier, I. J. Kitching, and C. Mitter. 2009. Phylogeny and biogeography of hawkmoths (Lepidoptera: Sphingidae): evidence from five nuclear genes. PLoS ONE 4: e5719. doi:10.1371/journal.pone.0005719

Kitching, I. J. 2019. Sphingidae taxonomic inventory. http://sphingidae.myspecies.info/. Accessed 25 October 2019.

Kitching, I. J., and J.-M. Cadiou. 2000. Hawkmoths of the World: an annotated and illustrated revisionary checklist (Lepidoptera: Sphingidae). Ann. Entomol. Soc. Am. 93: 1195–1196. doi:10.1093/aesa/93.5.1195g

Kitching, I., R. Rougerie, A. Zwick, C. Hamilton, R. St Laurent, S. Naumann, L. Ballesteros Mejia, and A. Kawahara. 2018. A global checklist of the Bombycoidea (Insecta: Lepidoptera). Biodivers. Data J. 6: e22236. doi:10.3897/BDJ.6.e22236

Minh, B. Q., M. A. T. Nguyen, and A. von Haeseler. 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30: 1188–1195. doi:10.1093/molbev/mst024

Moré, M., I. J. Kitching, and A. Cocucci. 2005. Sphingidae: Esfíngidos de Argentina. Hawkmoths of Argentina. xvi + 166 pp. L.O.L.A. (Literature of Latin America), Buenos Aires, Argentina.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32: 268–274. doi:10.1093/molbev/msu300

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20: 289–290.

Peigler, R. S. 1993. Wild silks of the World. Am. Entomol. 39: 151–162. doi:10.1093/ae/39.3.151

Peng, Y., H. C. M. Leung, S. M. Yiu, and F. Y. L. Chin. 2012. IDBA-UD: a de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. Bioinforma. Oxf. Engl. 28: 1420–1428. doi:10.1093/bioinformatics/bts174

Ponce, F. V., J. W. Breinholt, T. Hossie, J. R. Barber, D. H. Janzen, W. Hallwachs, and A. Y. Kawahara. 2015. A molecular phylogeny of *Eumorpha* (Lepidoptera: Sphingidae) and the evolution of anti-predator larval eyespots. Syst. Entomol. 40: 401–408. doi:10.1111/syen.12111

R Core Team. 2013. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ranwez, V., S. Harispe, F. Delsuc, and E. J. P. Douzery. 2011. MACSE: multiple alignment of coding sequences accounting for frameshifts and stop codons. PLoS ONE 6: e22594. doi:10.1371/journal.pone.0022594

Revell, L. J. 2012. Phytools: an R package for phylogenetic comparative biology (and other things): phytools: R package. Methods Ecol. Evol. 3: 217–223. doi:10.1111/j.2041-210X.2011.00169.x

Roe, A., S. J. Weller, J. Baixeras, J. Brown, M. Cummings, D. Davis, A. Kawahara, C. Parr, J. C. Regier, and D. Rubinoff. 2009. Evolutionary framework for Lepidoptera model systems, pp. 1–24. *In* M. Goldsmith and F. Marec (eds.), Genetics and molecular biology of Lepidoptera. CRC Press, Boca Raton, FL.

Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.

Rothschild, L.W., and K. Jordan. 1903. A revision of the lepidopterous family Sphingidae. Novit. Zool. 9 (Suppl.): 1–972.

Schmieder, R., and R. Edwards. 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics 27: 863–864. doi:10.1093/bioinformatics/btr026

Tamura, K., and M. Nei. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol. Biol. Evol. 10: 512–526. doi:10.1093/oxfordjournals.molbev.a040023

Thomsen, P. F., S. Elias, M. T. P. Gilbert, J. Haile, K. Munch, S. Kuzmina, D. G. Froese, A. Sher, R. N. Holdaway, and E. Willerslev. 2009. Non-destructive sampling of ancient insect DNA. PLoS ONE 4: e5048.

Timmermans, M. J. T. N., S. Dodsworth, C. L. Culverwell, L. Bocak, D. Ahrens, D. T. J. Littlewood, J. Pons, and A. P. Vogler. 2010. Why barcode? High-throughput multiplex sequencing of mitochondrial genomes for molecular systematics. Nucleic Acids Res. 38:e197.

Timmermans, M. J. T. N., C. Viberg, G. Martin, K. Hopkins, and A. P. Vogler. 2015. Rapid assembly of taxonomically validated mitochondrial genomes from historical insect collections. Biol. J. Linn. Soc. 117: 83–95. doi:10.1111/bij.12552.

Viette, P. E. L. 1971. Un nouveau sphingide de Nouvelle-Caledonie. Entomops 23: 237–240.

Zhang, J., K. Kobert, T. Flouri, and A. Stamatakis. 2014. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics 30: 614–620. doi:10.1093/bioinformatics/btt593