NRC Research Press

ARTICLE

The mitogenome of a Malagasy butterfly *Malaza fastuosus* (Mabille, 1884) recovered from the holotype collected over 140 years ago adds support for a new subfamily of Hesperiidae (Lepidoptera)

Jing Zhang, David C. Lees, Jinhui Shen, Qian Cong, Blanca Huertas, Geoff Martin, and Nick V. Grishin

Abstract: *Malaza fastuosus* is a lavishly patterned skipper butterfly from a genus that has three described species, all endemic to the mainland of Madagascar. To our knowledge, *M. fastuosus* has not been collected for nearly 50 years. To evaluate the power of our techniques to recover DNA, we used a single foreleg of an at least 140-year-old holotype specimen from the collection of the Natural History Museum London with no destruction of external morphology to extract DNA and assemble a complete mitogenome from next generation sequencing reads. The resulting 15 540 bp mitogenome contains 13 protein-coding genes, 22 transfer RNA genes, two ribosomal RNA genes, and an A+T rich region, similarly to other Lepidoptera mitogenomes. Here we provide the first mitogenome also for Trapezitinae (*Rachelia extrusus*). Phylogenetic analysis of available skipper mitogenomes places *Malaza* outside of Trapezitinae and Barcinae + Hesperiinae, with a possible sister relationship to Heteropterinae. Of these, at least Heteropterinae, Trapezitinae, and almost all Hesperiinae have monocot-feeding caterpillars. *Malaza* appears to be an evolutionarily highly distinct ancient lineage, morphologically with several unusual hesperiid features. The monotypic subfamily **Malazinae** Lees & Grishin subfam. nov. (type genus *Malaza*) is proposed to reflect this morphological and molecular evidence.

Key words: next-generation sequencing, phylogeny, Madagascar, ancient DNA, Paul Mabille.

Résumé : Le *Malaza fastuosus* est une espèce de papillon aux motifs très colorés qui appartient à un genre qui compte trois espèces décrites, toutes indigènes de Madagascar. Au meilleur de la connaissance des auteurs, aucun spécimen du *M. fastuosus* n'a été collecté depuis près de 50 ans. Pour évaluer la puissance des techniques d'extraction d'ADN, les auteurs ont employé une seule patte avant sur un spécimen holotype vieux d'au moins 140 ans, conservé au sein de la collection du Natural History Museum London et qui ne présentait aucun signe morphologique de dommage extérieur, pour extraire de l'ADN et assembler un mitogénome complet à l'aide de séquençage de seconde génération. Le mitogénome obtenu compte 15 540 pb et contient 13 gènes codant pour des protéines, 22 gènes d'ARN de transfert, deux gènes codant pour des ARN ribosomiques et une région riche en A+T, tout comme les autres mitogénomes de lépidoptères. Les auteurs rapportent également le premier mitogénome au sein des Trapezitinae (*Rachelia extrusus*). Une analyse phylogénétique parmi les autres mitogénomes d'hespéridés place le genre *Malaza* à l'extérieur des Trapezitinae et Barcinae + Hesperiinae, mais avec une possible relation proche avec les Heteropterinae. Parmi celles-ci, les Heteropterinae, les Trapezitinae et presque tous les Hesperiinae présentent des chenilles qui se nourrissent de monocotylédones. Le genre *Malaza* semble former un groupe distinct et ancien sur le plan évolutif, mais qui présente plusieurs caractéristiques morphologiques inhabituelles présentes chez les Hesperiidae. Une sous-famille monotypique, **Malazinae** Lees & Grishin (genre type *Malaza*) est proposé pour refléter les évidences morphologiques et moléculaires. [Traduit par la Rédaction]

Mots-clés : séquençage de seconde génération, phylogénie, Madagascar, ADN ancient, Paul Mabille.

Received 8 November 2019. Accepted 8 January 2020.

J. Zhang* **and J. Shen.** Departments of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-8816, USA.
D.C. Lees,* **B. Huertas, and G. Martin.** Department of Life Sciences, Natural History Museum, London, SW7 5BD, United Kingdom.
Q. Cong. Institute for Protein Design and Department of Biochemistry, University of Washington, 1959 NE Pacific Street, HSB J-405, Seattle, WA 98195, USA.
N.V. Grishin. Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-9050, USA; Departments of Biophysics and Biochemistry, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, TX 75390-8816, USA.

Corresponding author: Nick V. Grishin (email: grishin@chop.swmed.edu).
*These authors contributed equally to this work.

Introduction

Its name translated from Malagasy as "famous", *Malaza* Mabille, 1904 is probably the most unusual skipper genus from Madagascar. Three Malagasy endemic species of *Malaza* have been described (Lees et al. 2003). One of them, aptly named *Malaza fastuosus* (Mabille, 1884), here dubbed Lavish Malaza, has not been collected to our knowledge for nearly 50 years. The study specimen (Fig. 1; whose sex Paul Mabille was unsure of: Mabille [1887]: 338), is a female, as ascertained by lack of androconial scale brushes on its ventral forewing and dorsal hindwing. Its taxonomic history is somewhat complicated. It was first described by Mabille in 1878 as his female syntype of *Cyclopides empyreus* (Mabille 1878). Its type locality is no more precise than Madagascar and Mabille subsequently (re-)described this specimen (Mabille 1884) as a distinct taxon, *Trapezites fastuosus*. The latter specimen (NHMUK010621353) is considered to be the holotype of *T. fastuosus*, and apart from its head and abdomen, is consistent in pattern with the illustration in pl. 13, figs. 9, 9*a* of Mabille ([1885]) (those parts were probably fabricated in this figure). This holotype is found in the collection of the Natural History Museum in London (NHMUK). Even at the time of its description as *T. fastuosus*, the specimen was "à corps incomplet" (Mabille 1884) or indeed "privé d'abdomen et de tête" (Mabille [1887]). As Mabille says in the last publication that it was lacking its head and abdomen, it may be inferred very likely that the abdomen, currently attached with dark red resin, is not the original one (it lacks an expected series of dark ventro-lateral stripes, and is slightly larger than normal for *Malaza carmides*). The current wingspan measures just short of 49 mm, but the tip of the left forewing is missing today. Mabille measured its wingspan at 52 mm in 1878 and 53 mm in 1884, both measurements being consistent with the specimen now numbered NHMUK010621353 being both the supposed female syntype of *C. empyreus* and the holotype of *T. fastuosus*. Once Mabille acquired a true female of *Malaza empyreus*, of similar size to its male, he apparently realized his mistake (Mabille [1887]: 337–338). While it is not known when and exactly where or by whom the specimen was collected, the collection event was prior to 15 December 1878, the dating of the page where the description of the female syntype of *C. empyreus* was published (Mabille 1878: 285), alongside other species typical of eastern lowland rainforest described there, notably *M. carmides* (as *Cyclopides catocalinus*), *Xanthodisca ariel*, and *Perrotia gillias*.

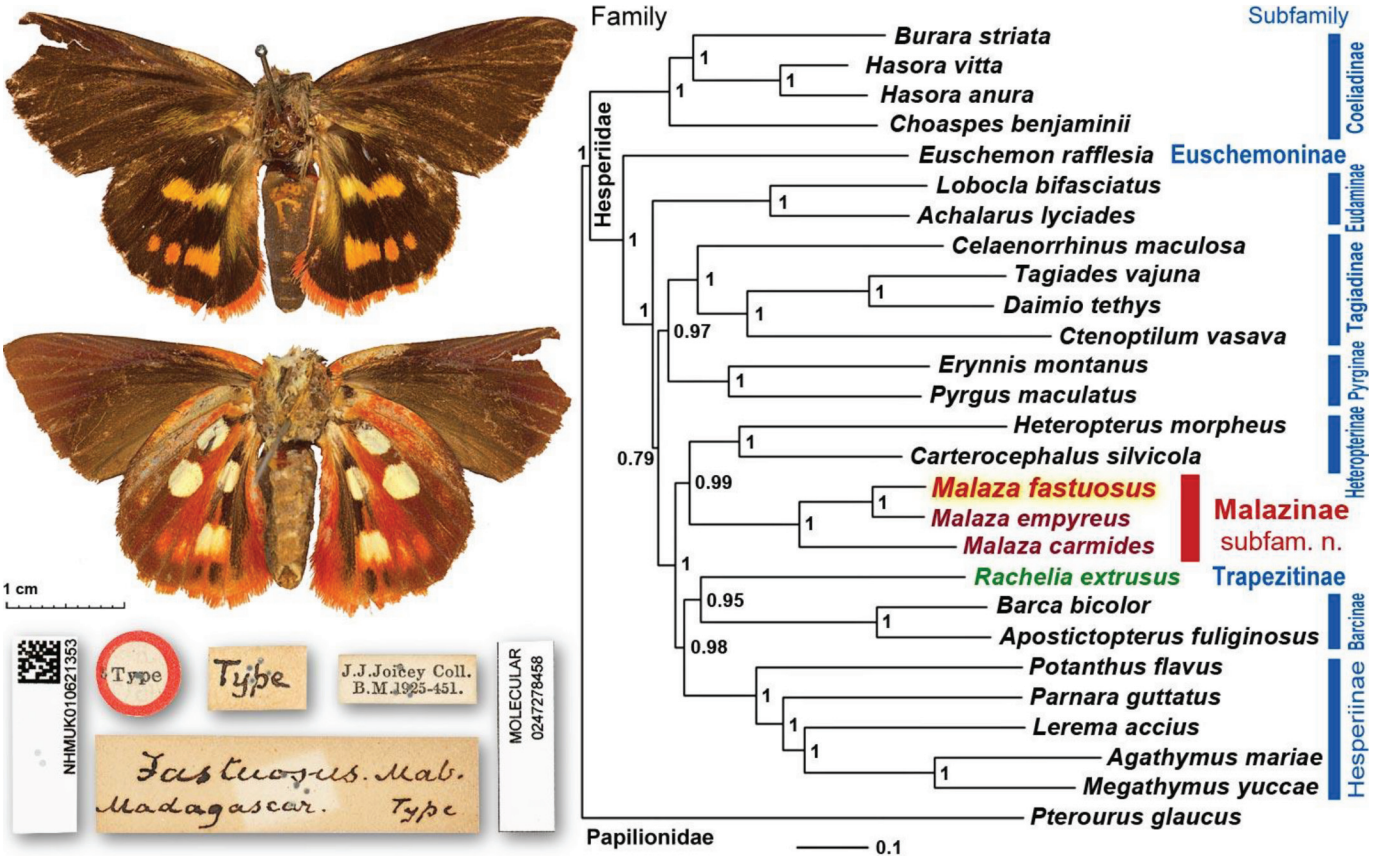
Malaza fastuosus is the largest and most sumptuously colored of the three species of *Malaza* and among the most stunning skippers in the Afrotropical region. The genus exhibits a rather stout, thorax and abdomen and both sexes are brown above with two orange bands on the hindwing and orange hindwing fringes. Ventrally, however, *M. fastuosus* resembles only the much smaller *M. empyreus*, the type species of *Manarina* Mabille, 1904 (Lucas 1905: 699). Richly colored in various shades of vermilion and brown, each hindwing of *M. fastuosus* is endowed with four large round and glimmering white spangles tinted with yellow, whereas *M. empyreus* has three orange spots embedded in a similar vermilion-red. *Malaza carmides* is less spectacular but sexually dimorphic, the female sporting a single silver spangle on the ventral hindwing underside.

Nothing is recorded about the habits of *M. fastuosus*, but it has been collected widely, from north, west, east, and southwest Madagascar (forêt d'Analalava, Ankarafantsika, Perinet, Ile Ste Marie, Fito, Forêt de Rogez, Imerimandroso Lac Alaotra, and Tongobory; museum data and Viette (1956)), sites varying from 0 to 900 m elevation in dry to wet forests, from September to December. Surprisingly, it was last collected in Ankarafantsika in 1971, whereas *M. empyreus* is occasionally collected or photographed in dense eastern rainforest interiors (Preston-Mafham 1991). *Malaza carmides* is found more often in littoral forest or scrub. It is probable that additional species remain to be described. Searches for ovipositions for any species have so far proved fruitless, but excepting *M. carmides*, *Malaza* encounters during fieldwork are rare and not easily predictable.

The systematics of *Malaza* is poorly understood. Evans (1937) placed *Malaza* in the first subgroup of his *Ploetzia* group, apparently largely based on wing venation "hindwing lower outer angle of cell not turned up, i.e., the median vein and vein 4 collinear" (an observation apparently not based on a wing preparation of *Malaza*, see Diagnosis). This group he considered contained also *Miraja* and *Perrotia*, two other genera endemic to Madagascar (see also Lees et al. 2003, where *Miraja* was sunk). An additional 12 Afrotropical genera were included in Evans' *Ploetzia* group. Viette (1956: 46–50) added a limited morphological study on the genus, illustrating the male and female genitalia of two species. Although these skippers are robust and striking, no authors have proposed any fundamentally outstanding morphological characters of *Malaza* in a comparative context, although Mabille (1904) mentioned the inflated hindleg first tarsal segment peculiar to *M. carmides*. Although the genus has not previously been included in molecular analyses, two species have been DNA barcoded by D.C.L. (*M. carmides*, *M. empyreus*) but neighbor-joining and ad hoc phylogenetic analyses of barcodes fails to place the genus reliably.

To learn more about these Malagasy skippers, we sequenced, assembled, and annotated the complete mitogenome of the *Malaza fastuosus* holotype. The foreleg was used for DNA work, because it remains to be verified if the abdomen currently affixed to NHMUK010621353 belongs either to this species or even to *Malaza*. Methods for genomic DNA extraction, library construction, next-generation sequencing, and computational procedures have been reported by us previously (Cong and Grishin 2016; Shen et al. 2015, 2016). Importantly, after being detached, the leg was incubated in the DNA extraction buffer as a whole, without being broken. This procedure allows for the preservation of body parts of unique specimens and

Fig. 1. Study specimen and phylogenetic tree. *Malaza fastuosus* holotype, female (specimen NHMUK010621353, molecular code 0247278458, sample NVG-18082C03) with its labels is shown on the left: dorsal view (above), ventral view (middle), and labels (below). Scale bar refers to all images. On the right, is the maximum likelihood tree of the coding regions in the mitogenomes of 26 Hesperiidae species. The tree is rooted with *Pterourus glaucus* (Papilionidae). MrBayes posterior probabilities are shown by each node. Species names for mitogenomes sequenced in this study are shown in color. GenBank accessions for mitogenome sequences are as follows: *Achalarus lyciades* NC_030602.1; *Agathymus mariae* KY630504.1; *Apostictopterus fuliginosus* NC_039946.1; *Barca bicolor* NC_039947.1; *Burara striata* NC_034676.1; *Carterocephalus silvicola* NC_024646.1; *Celaenorrhinus maculosa* NC_022853.1; *Choaspes benjaminii* NC_024647.1; *Ctenoptilum vasava* NC_016704.1; *Daimio tethys* NC_024648.1; *Erynnis montanus* NC_021427.1; *Euschemon rafflesia* NC_034231.1; *Hasora anura* NC_027263.1; *Hasora vitta* NC_027170.1; *Heteropterus morpheus* NC_028506.1; *Lerema accius* NC_029826.1; *Lobocla bifasciatus* NC_024649.1; *Malaza carmides* MN919191; *Malaza empyreus* MN919190; *Malaza fastuosus* MK301537; *Megathymus yuccae* KY630500.1; *Parnara guttatus* NC_029136.1; *Potanthus flavus* NC_024650.1; *Pterourus glaucus* NC_027252; *Pyrgus maculatus* NC_030192.1; *Rachelia extrusus* MN919192; *Tagiades vajuna* KX865091.1.



does not impair subsequent DNA work. From this leg, we extracted 0.78 ng of genomic DNA (gDNA). Due to the low amount of gDNA, we did not check DNA size using gel, but proceeded with pair-end DNA library construction for next generation sequencing. The final library was 7.1 ng/μL, 40 μL volume, with the average size of about 190 bp. Since sequencing adapters are about 125 bp, the library size of 190 bp suggests an average genomic DNA fragment of about 65 bp. Sequencing of the library on an Illumina HiSeq X10 instrument for 150 bp at both ends resulted in 55.6M raw reads and, after trimming adapters, 53.5M reads with lengths above 30 bp were left, totaling about 3 Gbp. Overall, more than 92% of reads had lengths from 20 to 70 bp, with the median around 47 bp, suggesting highly degraded genomic DNA. For samples field-preserved in a DNA-friendly way, we typically obtain from 100 ng to several milligrams of gDNA, and the read length after trimming adapters reaches a maximum, usually about 150 bp. The mitogenome of *Lerema accius* (Smith, 1797) (Cong and Grishin 2016) was used as a reference to search for ("bait") similar sequence reads using BWA (Li and Durbin 2009). Nearly 0.95% (520 325 out of 54 557 740) of *M. fastuosus* total genomic reads were extracted by BWA for mitogenome assembly (Hahn et al. 2013). Despite the DNA being heavily fragmented, a number of reads were longer than 100 bp, allowing us to confidently assemble the mitogenome. The complete mitogenome of *M. fastuosus* was assembled de novo using Platanus (Kajitani et al. 2014) followed by a manual gap-closing procedure.

The mitochondrial genome of *M. fastuosus* is 15 540 bp in length (GenBank: MK301537) and is AT rich, with a base composition of 40.1% A, 40.4% T, 7.4% G, and 12% C. It encodes a typical gene set for Lepidoptera mitogenomes,

including 13 protein-coding genes (ND1–ND6, COX1–COX3, ND4L, ATP8, ATP6, and CYTB), 22 tRNA genes (two for serine and leucine and one for each of the rest of the amino acids), two ribosomal RNAs (rrnL and rrnS), and an A+T rich D-loop control region. The typical start codon ATN is used in other genes except COX1, which has a GTG codon. The stop codon in COX2, ND4, and ND5 genes is incomplete (just T), and a complete TAA codon should be formed during mRNA maturation (Boore 1999; Ojala et al. 1981). The two rRNAs are 1479 bp (rrnL) and 773 bp (rrnS) and the tRNAs are from 66 to 72 bp in length.

Previously, by Sanger sequencing we had obtained DNA barcodes (COX1-5P) for the two other species of *Malaza*: *M. empyreus* and *M. carmides* (on BOLD: cluster numbers BOLD:AAK7349, BOLD:ACW7233, respectively, based on single specimens that D.C.L. collected in 2004 from Ambatovy, east Madagascar). Comparison with the barcode of *M. fastuosus* reveals that *M. empyreus* is more similar than is *M. carmides*: 92.7% sequence identity (6.3% pairwise divergence), in accord with morphological similarities (e.g., wing pattern, limited sexual dimorphism), as well as apparent phylogenetic closeness. *Malaza fastuosus* and *M. empyreus* were placed in the same genus *Manarina* by Mabille (1904: 95–96). The barcode of *M. carmides* is more distant at around 90.5% sequence identity (9.4% pairwise divergence between *M. carmides* and *M. fastuosus*; 9.6% between *M. carmides* and *M. empyreus*). The difference of less than 10% pairwise divergence for the three species of *Malaza* and strong similarities in wing venation (see below) suggests that they are close phylogenetic relatives and could be congeners. This small difference also indirectly validates our sequencing result, suggesting that the sequence we obtained is indeed of a third species of *Malaza*, and not some contaminant. This is an essential point, since above we mentioned the specimen sequenced might be chimaeric. However, when the foreleg was detached by N.V.G., there was no evidence that it had been glued there.

To better understand the phylogenetic affinities of *Malaza*, we additionally assembled mitogenomes of *M. carmides* and *M. empyreus* from the same vouchers used for DNA barcoding using the same procedures as for *M. fastuosus*. Furthermore, because no mitogenomes from the subfamily Trapezitinae were available, we assembled a mitogenome of *Rachelia extrusus* (C. & R. Felder, 1867) using the same approach. A Bayesian inference tree was constructed by MrBayes (Huelsenbeck and Ronquist 2001) from representative known Hesperiidae mitogenomes (Cao et al. 2016; Cong and Grishin 2016; Han et al. 2018; Hao et al. 2012; Kim et al. 2014; Liu et al. 2017; Shao et al. 2015; Shen et al. 2015, 2016; Wang et al. 2014; Wang et al. 2015; Wang et al. 2016; Zhang et al. 2017*a*, 2017*b*) with our four new mitogenomes added, and rooted with the *Pterourus glaucus* (Papilionidae) mitogenome (Shen et al. 2015) (Fig. 1). We used all protein coding regions, tRNAs (except tRNA-Cys), and two ribosomal RNAs. Protein-coding sequences were further partitioned by 1st codon (pos1), 2nd codon (pos2), and 3rd codon (pos3). We used the partitioning scheme that was developed for Hesperiidae mitogenomes in a dedicated study (Han et al. 2018). Accordingly, sequences were partitioned into 16 subsets. Subset 1: tRNA-Phe, tRNA-His, ND6-pos1, ND2-pos1, tRNA-Leu, ATP6-pos1, and ND3-pos1; subset 2: tRNA-Lys, tRNA-Leu, ATP6-pos2, COX2-pos2, COX3-pos2, and CYTB-pos2; subset 3: ATP6-pos3 and ND2-pos3; subset 4: ATP8-pos1 and ATP8-pos2; subset 5: ATP8-pos3, ND6-pos3, COX2-pos3, CYTB-pos3, COX1-pos3, COX3-pos3, and ND3-pos3; subset 6: COX1-pos1; subset 7: COX1-pos2; subset 8: tRNA-Gln and COX2-pos1; subset 9: COX3-pos1, CYTB-pos1, tRNA-Trp, and tRNA-Met; subset 10: ND4-pos2, tRNA-Ile, ND4L-pos1, ND1-pos1, and ND5-pos2; subset 11: ND4L-pos2, ND5-pos3, and ND1-pos2; subset 12: ND1-pos3; subset 13: ND3-pos2, ND6-pos2, and ND2-pos2; subset 14: ND4-pos3 and ND4-pos1; subset 15: ND5-pos1 and ND4L-pos3; subset 16: tRNA-Pro, tRNA-Tyr, tRNA-Val, 12S ribosomal RNA, 16S ribosomal RNA, tRNA-Asp, tRNA-Glu, tRNA-Arg, tRNA-Asn, tRNA-Gly, two tRNA-Ser, tRNA-Ala, and tRNA-Thr. The best substitution rate models for each partition were selected by MrBayes.

The resulting tree topology is consistent with previous phylogenetic studies that did not include *Malaza* (Sahoo et al. 2016; Warren et al. 2008, 2009; Zhang et al. 2017*a*, 2019). Coeliadinae are the sister to all other Hesperiidae; *Euschemon* is a sister to the rest of Hesperiidae except Coeliadinae; the relationship between Eudaminae and Pyrginae is unresolved (0.79 posterior probability) but with Pyrginae sister to Tagiadinae, and Heteropterinae are in the same clade with Hesperiinae (Zhang et al. 2019) (the last two subfamilies are monocot feeders as caterpillars, with secondary colonization of dicots by some Hesperiinae). *Malaza* is sister to Heteropterinae with posterior probability of 0.99. This relationship comes as a surprise and has not hithero been suggested by morphology. Therefore, additional studies, in particular those based on the analysis of nuclear genomes, are needed to firmly establish phylogenetic affinities of *Malaza*. However, even with the data at hand, it is clear that *Malaza* originated early during the radiation of monocot feeders and represents an ancient lineage, today surviving only in Madagascar. It is noteworthy that despite some phenotypic similarities with *Malaza* (e.g., bulky, robust bodies) that may have led to Mabille's (1884) placement in *Trapezites*, Trapezitinae (represented by *Rachelia extrusus*) is not closely related to it and is instead in the same clade with Barcinae and Hesperiinae (with 0.98 posterior probability support; Barcinae + Trapezitinae with marginal support at 0.95).

To probe its morphology further, D.C.L. examined specimens and preparations of *Malaza*, including genitalia dissections at NHMUK (H. 828, H. 901, H. 902, BMNH(E) 526064 - dry genitalia on card) and wing venation for *M. carmides* and *M. empyreus*, for characters of potential

relevance in the systematics of Hesperiidae. Reinforced by the mitogenomic tree, and its morphological distinctness from the other three subfamilies, we conclude that the *Malaza* lineage should have the same taxonomic rank as Heteropterinae, Trapezitinae, Barcinae, and Hesperiinae (see Zhang et al. 2019). To reflect the ancient origin of *Malaza* and its distinctness from other Hesperiidae, a new subfamily for it is proposed.

Malazinae Lees and Grishin, subfam. nov.

ZOOBANK NO.: http://zoobank.org/E4D5B92E-8E4A-4BEC-812B-5B1AA226FC0E.

TYPE GENUS: *Malaza* Mabille, 1904, by original designation.

DIAGNOSIS: Morphologically, distinguished from other Hesperiidae or other subfamilies, by the following characters or their combination (additional contrasts given in description): **Head.** 1. Compound eyes covered in short interommatidial setae, a very unusual feature in hesperiids, recorded in some coeliadines, notably *Choaspes* and a few *Hasora*, not known elsewhere (Warren et al. 2009) and also known in *Macrosoma* (Yack et al. 2007). 2. A short but prominent eyelash present, a synapomorphy for noncoeliadine hesperiids. 3. Posterior and anterior cephalic chaetosemata present, again a synapomorphy for noncoeliadine hesperiids, anterior pair concentrated into lateral bunches of radially oriented fine setae just anteriad of eyelash, posterior pair linearly radiating and almost meeting at midline. **Thorax.** 4. Forewing M_2 weak, curving slightly towards M_3 near and at cell; M_2 originating distinctly closer to M_3 than M_1 as in most Hesperiinae, but not so extreme as in most hesperiine cases (at cell, M_1-M_2:M_2-M_3 about 0.55:0.45: respectively); M_2 origin is not intermediate (Kitching et al. 1999) nor tending towards M_1 as in some Trapezitinae (Braby 2000: 88, although the opposite case is indicated on p. 22). 5. Hindwings with M_2 lost; discocellulars (cross-veins) also lost except for stumps visible after fork between M_1-Rs and M_3-CuA1, respectively. Extrapolated chord between stumps and their alignment with tornus also more or less perpendicular to costa and upper cross-vein [remnant] certainly not directed towards the apex, as has been supposed to be a synapomorphy of Trapezitinae, nor for the lower part of the M_2-M_3 discocellular, directed towards the body (Braby 2000; Simonsen et al. 2012; Warren et al. 2009). 6. Hindwing cell (or trace of it up to cross-vein stumps), about half wing length, not long (as typical of Heteropterinae) or fairly long (as in Trapezitinae). The hairy compound eye and loss of a tubular hindwing M_2 and associated discocellulars are the clearest presently known synapomorphies for Malazinae although probably they are also parallelisms exhibited elsewhere in Hesperiidae.

By DNA, a combination of the following base pairs in COI barcode is diagnostic: 38G, 319A, 451T, 475T, 479C, 481T, 512G, and 604A. To derive these DNA characters, we used a barcode alignment of 160 representative Hesperiidae specimens from all major phylogenetic groups included in Zhang et al. (2019). These base pairs are conserved in all three species of *Malaza* and uniquely place *Malaza* within the monocot-feeding Hesperiidae to diagnose Malazinae as distinct from the three other subfamilies, i.e., base pairs that are (*i*) the same as in Heteropterinae but different from most Hesperiinae and Trapezitinae (604); or (*ii*) the same as in most Hesperiinae but different from most Heteropterinae and Trapezitinae (319, 481); or (*iii*) the same as in Trapezitinae, but different from most Hesperiinae and Heteropterinae (475); or (*iv*) the same as most Trapezitinae and Heteropterinae but mostly different from Hesperiinae (38, 512); or (*v*) the same as in few or most Hesperiinae and Trapezitinae, but different from Heteropterinae (451, 475, 479). None of these characters is a unique synapomorphy of *Malaza*, but it may be noted that 38 is a first position change coding for valine (whereas leucine in Hesperiinae) and 512 is a first position change coding for alanine (but serine in Hesperiinae).

DESCRIPTION: **Generalities**. Medium to large hesperiids, forewing length 12.5–24 mm, wingspan (apex–apex) 30–50 mm, wings broad and rounded, slightly narrower and more acute in male, very sexually dimorphic or little so, abdomen similar length to hindwings, body very robust, head relatively broad. Wings held over back, downwards, at rest (basking posture not observed). Proboscis quite long, flower visiting, diurnal, life history unknown, but expressed egg of *M. carmides* is quite large (around 1 mm diameter) and domed, hemispherical, orange, with around 17 radial ribs. **Head**. Compound eyes covered in short interommatidial setae. Eyelash and posterior and anterior cephalic chaetosemata present (see diagnosis). Labial palps erect, partly straight, second segment irregularly bulbous but relatively flattened (quadrate would not be a precise description (Evans 1937; Viette 1956)) rather than porrect and hairy as in Heteropterinae, collapsed inwards in their proximal part towards head, third segment in line with second segment (i.e., erect; observed also in natural position), relatively short and stout (about twice as long as maximum width), slightly fatter than conical, pointed towards tip. Antennae about 1/2 costal length, as, e.g., in trapezitines (not shorter than mid-costa, as in many Heteropterinae), strongly downbent well after commencement of club (strongly hooked), long apiculus tapered to pointed tip. **Thorax**. Foreleg epiphysis present, well developed, not reduced as suggested for most heteropterines (Warren et al. 2009). Mesothoracic tibia with one pair of spines and metathoracic tibia with one pair of spines (in *M. carmides*, *M. fastuosus*; (Mabille 1904)) or two, as in some Trapezitinae (Kitching et al. 1999: 83) with mid tibial pair shorter (*M. empyreus*). Hindleg first tarsal segment strongly inflated (not the tibia as misstated by Evans 1937: 132) in male of *M. carmides* (Mabille 1904: 95), excavate lengthwise towards the body like a hollowed out boat, not itself clothed internally with obvious androconial scales but harboring the dark tip of a brush originating along the

length of the femur, bundled with a larger recumbent hair pencil originating at the proximal end of the tibia; otherwise, hindtibiae unmodified except fringed with orange hairs. The presence in at least one Malazinae of a recumbent metatibial hair tuft (common in Coeliadinae) stands in contrast to its apparent absence all three prior monocot feeding subfamilies (Warren et al. 2009). **Wings.** Forewing with costal yellow-orange streak in basal half on both surfaces, except in *M. carmides*. In *M. empyreus* and *M. fastuosus*, hindwing dorsum with orange spot at wing base and end of cell the latter a continuation of orange crescent at base of cubitus, two-three orange spots more distally in space CuA_2, with two extra orange flecks in space M_3 in *M. carmides* making a linear series of four. Hindwing ventrally with pale spangles in similar positions, two along space CuA_2 and 1–2 either end of cell, as silver spangles in *M. fastuosus* or as similarly shaped (but larger) orange marks in *M. empyreus*, which lacks the silvery basal cell spangle present in *M. fastuosus*. These marks are set in vermilion-red areas to dark orange red areas, the latter much more extensive in the hindwing in *M. empyreus*, which has more defined black spots, three along the cell region and two along space CuA_2, and submarginal black spots in the spaces of the ventral hindwing not conspicuous in the other species. Hindwing fold, widely channeled ventrally, in male tergad of 1A+2A bearing dorsally an extensive tuft of yellowish scales in both sexes. *Malaza fastuosus* has particularly pronounced orange cilia along the hindwing margin, other species with narrower fringe. Patterning similar in sexes of *M. empyreus* and *M. fastuosus*, whereas striking sexual dimorphism in *M. carmides*, the female's hindwing ventral surface with a single large silver spangle only in the cell, matched by an orange crescent dorsally, and an arc of four orange arrowheads in space M_3-CuA_2. Unlike in most Trapezitinae and Hesperiinae where they tend to be well developed, hyaline patches in forewing lost, or they are reduced in the case of the male of *M. carmides*, except in the female where a single hyaline patch in the forewing cell and another two in space M_3 and space CuA_1, are clearly defined on both surfaces as quadrate markings (sometimes also weakly expressed in the female of *M. empyreus*). In the ventral forewing of *M. carmides*, there is an extensive medial pale area diffusely in spaces CuA_2-1A+2A. Apart from this and the hyaline traces, the ventral surface of the male of *M. carmides* is more or less unmarked, except for a pale trace of the large hindwing cell spot on the dorsum. The ventral surface of *M. carmides* has a violet sheen missing in the other species. **Wing venation.** Forewings with CuA_2 originating relatively close to wing base (primary fork with stem of M and CuA_1 at about 1/8 of total length), as, e.g., in some coeliadines (e.g., *Bibasis oedipodia* (Swainson, 1820) (Bascombe et al. 1999: 76)). Forewing M_2 as for diagnosis. Hindwings with M_2 lost and cell end cross-veins (discocellular veins) reduced to stumps (see diagnosis). No trace of median fork before cell end (as in trapezitines and a few other taxa, see Parsons 1999: 152) and, where stumps may be substantial, top cross-vein stump oriented towards costa rather than towards body or apex. Extrapolated chord between stumps and their alignment also more or less perpendicular to costa in line with tornus (although inaccurate to state: end of cell straight (Viette 1956)); and upper cross-vein [remnant] certainly not directed towards the apex (or termination of Rs), as has been supposed to be a synapomorphy of Trapezitinae (Simonsen et al. 2012; Warren et al. 2009). Hindwing cell (or trace of it up to cross-vein stumps), about half wing length, not long (as typical of Heteropterinae). Rs fork with stem originating well before fork of CuA_2 with its stem (Viette 1956). **Sex scales.** See hindleg of *M. carmides*, above. Ventral forewing brush of long fine pale bristles towards base of space CuA_1 (*M. empyreus*, *M. fastuosus*) (Evans 1937: 132; Viette 1956). Also on dorsal hindwing of male, extensive (dark in *M. empyreus* or light yellowish in *M. fastuosus*, absent in *M. carmides*) brush of scales at base of (originating mainly above) $Sc+R_1$ and another brush, less dense, at base of cell (black in males of *M. fastuosus* and *M. empyreus*, yellowish in both sexes in *M. carmides*), interspersed with lenticular patch of minute grey relatively shiny scales (latter in all three species). Patches of fine thin pale scales in ventral forewing at base of space 1A+2A and at base of space CuA_1 (*M. empyreus*, *M. carmides*). **Abdomen**. As long as hindwings, in *M. empyreus* and *M. fastuosus*, underside orange flanked with black streaks, in *M. carmides*, plain orange. **Male genitalia**. Uncus much shorter than tegumen, appearing slightly hooked at tip from lateral view, spatulate-truncate from dorsal view and marginally bifid (in *M. carmides*) to distinctly bifid (in *M. fastuosus* and *M. empyreus*), similar to some trapezitines (Kitching et al. 1999: 82; Viette 1956: fig. 51). Gnathos large to very large (*M. fastuosus* preparation H. 828, #18516), sclerite somewhat rectangular from lateral view but expanded laterally (in all three species; see also fig. 51 in Viette (1956), for *M. carmides*, gnathos apparently incomplete). Valves at the apex in the costal region and the distal region of the sacculus (harpe) darkly sclerotized, with several short teeth. In *M. carmides*, the costal and distal lobes are rounded and the distal lobe of the sacculus bears about 6–8 teeth, whereas in *M. empyreus* and *M. fastuosus* the main valve is squared off apically with a series of about 20–30 much finer teeth with further smaller ones to their side and the more rounded sacculus lobe bears a smaller number of rather finer teeth (Viette 1956: figs. 51, 52). Saccus very short. Aedeagus straight, without unusual modifications. **Female genitalia.** Hairy papillae anales variably shaped, bearing a pair of long apophyses posteriores that extend dorsally well beyond the shield-like postvaginal median lobe which has a medially indented slot distad, as far as a pair of narrow ventro-lateral sclerites (shaped like curly braces) bearing flaps mediad. In

between these sclerites ventrally is a median antevaginal plate in *M. carmides*, absent in *M. empyreus* and *M. fastuosus*. Further proximad are paired sclerotized lobes (absent in *M. carmides*) densely covered in long setae that continue to the intersegmental membrane of A6/A7 imparting a "silky" appearance (in *M. fastuosus* and *M. empyreus*) (Viette 1956: fig. 53). Corpus bursae linked by a more or less narrow ductus bursae. Ostium bursae of female in all species with extensive paired spiculose signa consisting of two narrow lateral weakly sclerotized areas bearing numerous minute triangular spines on internal walls leading up to a sclerotized "neck" distad before the entrance of the ductus bursae. Warren et al. (2009) did not score this character, so its distribution among Hesperiidae subfamilies is not fully clear, but it occurs in some Heteropterinae (Steinhauser 1991). Appendix bursae clearly present and non-spiculose (at least in *M. fastuosus* preparation H.901, #18568; apparently also present in *M. empyreus*, but available slide preparation H.902, #18569 was not adequate to confidently score this character) nor was it clear from constrictions indicated in diagrams of *M. fastuosus* and *M. carmides*, figs. 53, 54 in Viette (1956), but this is a structure well developed in Heteropterinae and Trapezitinae in contrast to Hesperiinae, with a few exceptions (e.g., *Ceratrichia clara* and *Meza meza*: Warren et al. 2009).

GENERA INCLUDED: *Malaza* and its junior subjective synonym *Manarina* Mabille, 1904.

The Malazinae is the third monotypic subfamily of Hesperiidae after the Indian Chamundinae (Zhang et al. 2019) and the Euschemoninae. The latter consists of a single species, *Euschemon rafflesia* (MacLeay, [1826]). It is an Australian endemic and the only butterfly with a frenulum and retinaculum (apart from all but three species of Hedylidae), in males only in both families (Braby 2000; Scoble 1986), structures commonly present in moths to couple the wings. *Euschemon rafflesia* is also a large showy skipper with black, yellow- or white-marked wings and a red abdomen tip. It is interesting that these ancient monotypic lineages survive in locations characterised by high endemism (India, Australia, and Madagascar). This reinforces the idea that Hesperiidae as a family might have originated and diversified around the Indian Ocean (while the sister family Hedylidae is neotropical only). The earliest diverging hesperiid subfamily, Coeliadinae, is also palaeotropical and suggested to be of possible oriental origin (de Jong 2007), but Madagascar actually contains the earliest diverging member in one of the two clades in this subfamily, *Tekliades ramanatek* (Boisduval, 1833) (Li et al. 2019: S1 Appendix, p. 91). Therefore, Hesperiidae as a family might have originated and diversified around the Indian Ocean. Malazinae is the third supratribal endemic lineage of Lepidoptera so far identified for Madagascar (after Whalleyanidae and Callidulidae: Griveaudiinae).

Data deposition

The mitogenomes reported in this study were deposited in GenBank with the accessions MK301537, MN919190–MN919192. This publication has been registered with ZooBank as http://zoobank.org/8E2333AE-FEBC-4025-AA02-DE4BFC28E3F7 and the taxonomic act (new subfamily) was registered as http://zoobank.org/E4D5B92E-8E4A-4BEC-812B-5B1AA226FC0E.

Disclosure statement

The authors declare no conflicts of interest.

Funding information

This work was supported in part by the grants to N.V.G. from the National Institutes of Health (GM094575 and GM127390) and the Welch Foundation (I-1505). D.C.L. acknowledges Leverhulme Trust (F/00696/I), for fieldwork funding, ERC-EMARES (#250325), and Malagasy authorities including MICET for assistance in logistics and permits (No. 0 – MINEV.EF/SG/DGEF/DPB/SCBLF).

References

Bascombe, M.J., Johnston, G., and Bascombe, F.S. 1999. Butterflies of Hong Kong. Academic Press, London, United Kingdom.

Boore, J.L. 1999. Animal mitochondrial genomes. Nucleic Acids Res. **27**(8): 1767–1780. doi:10.1093/nar/27.8.1767. PMID:10101183.

Braby, M.F. 2000. The butterflies of Australia: their identification, biology and distribution. CSIRO Publishing, Melbourne.

Cao, L., Wang, J., James John, Y., Yau Shiu, M., Yuan, X., Liu, J., and Cao, T. 2016. The complete mitochondrial genome of *Hasora vitta* (Butler, 1870) (Lepidoptera: Hesperiidae). Mitochondrial DNA A DNA Mapp. Seq. Anal. **27**(4): 3020–3021. doi:10.3109/19401736.2015.1063048. PMID:26152350.

Cong, Q., and Grishin, N.V. 2016. The complete mitochondrial genome of *Lerema accius* and its phylogenetic implications. PeerJ, **4**: e1546. doi:10.7717/peerj.1546. PMID:26788426.

de Jong, R. 2007. Estimating time and space in the evolution of the Lepidoptera. Tijdschrift voor Entomologie, **150**: 319–346. doi:10.1163/22119434-900000233.

Evans, W.H. 1937. A catalogue of the African Hesperiidae indicating the classification and nomenclature adopted in the British Museum. The Trustees of the BMNH, London, United Kingdom.

Hahn, C., Bachmann, L., and Chevreux, B. 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads–a baiting and iterative mapping approach. Nucleic Acids Res. **41**(13): e129. doi:10.1093/nar/gkt371. PMID:23661685.

Han, Y., Huang, Z., Tang, J., Chiba, H., and Fan, X. 2018. The complete mitochondrial genomes of two skipper genera (Lepidoptera: Hesperiidae) and their associated phylogenetic analysis. Sci. Rep. **8**(1): 15762. doi:10.1038/s41598-018-34107-1. PMID:30361496.

Hao, J., Sun, Q., Zhao, H., Sun, X., Gai, Y., and Yang, Q. 2012. The complete mitochondrial genome of *Ctenoptilum vasava* (Lepidoptera: Hesperiidae: Pyrginae) and its phylogenetic implication. Comp. Funct. Genomics, **2012**: 328049. doi:10.1155/2012/328049. PMID:22577351.

Huelsenbeck, J.P., and Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics, **17**(8): 754–755. doi:10.1093/bioinformatics/17.8.754. PMID:11524383.

Kajitani, R., Toshimoto, K., Noguchi, H., Toyoda, A., Ogura, Y., Okuno, M., et al. 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short

reads. Genome Res. **24**(8): 1384–1395. doi:10.1101/gr.170720.113. PMID:24755901.

Kim, M.J., Wang, A.R., Park, J.S., and Kim, I. 2014. Complete mitochondrial genomes of five skippers (Lepidoptera: Hesperiidae) and phylogenetic reconstruction of Lepidoptera. Gene, 549(1): 97–112. doi:10.1016/j.gene.2014.07.052. PMID: 25058696.

Kitching, R.L., Scheermeyer, E., Jones, R.E., and Pierce, N.E. 1999. Biology of Australian butterflies. CSIRO Publishing, Melbourne.

Lees, D.C., Kremen, C., and Raharitsimba, H. 2003. Classification, diversity and endemism of the butterflies (Papilionoidea and Hesperioidea): a revised species checklist. *In* The natural history of Madagascar. *Edited by* S.M. Goodman and J.P. Benstead. University of Chicago Press, Chicago, Ill. pp. 762–793.

Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, **25**(14): 1754–1760. doi:10.1093/bioinformatics/btp324. PMID:19451168.

Li, W., Cong, Q., Shen, J., Zhang, J., Hallwachs, W., Janzen, D.H., and Grishin, N.V. 2019. Genomes of skipper butterflies reveal extensive convergence of wing patterns. Proc. Natl. Acad. Sci. USA. **116**(13): 6232–6237. doi:10.1073/pnas.1821304116. PMID: 30877254.

Liu, F.-F., Li, Y.-P., Jakovlic, I., and Yuan, X.-Q. 2017. Tandem duplication of two tRNA genes in the mitochondrial genome of *Tagiades vajuna* (Lepidoptera: Hesperiidae). Eur. J. Entomol. **114**(1): 407–415. doi:10.14411/eje.2017.052.

Lucas, R. 1905. Lepidoptera für 1904. Archiv für Naturgeschichte Jahrg.71: Bd.2: Heft.2: pt.1: 515–786.

Mabille, P. 1878. Diagnoses de Lépidoptères de Madagascar. Petites Nouvelles Entomologiques, **2**(10): 285.

Mabille, P. 1884. Descriptions de Lepidopteres Exotiques. Annales de la Société Entomologique de Belgique, **28**: CLXXXIV–CXCI.

Mabille, P. [1885]. Histoire naturelle des Lépidoptères. *In* Histoire physique naturelle et politique du Madagascar. *Edited by* A. Grandidier. Imprimerie Nationale, Paris. 55 Pls.

Mabille, P. [1887]. Histoire naturelle des Lépidoptères. *In* Histoire physique naturelle et politique du Madagascar. *Edited by* A. Grandidier. Imprimerie Nationale, Paris. 364 pp.

Mabille, P. 1904. Lepidoptera Rhopalocera. Fam. Hesperidae. *In* Genera Insectorum. *Edited by* P. Wytsman.

Ojala, D., Montoya, J., and Attardi, G. 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature, **290**(5806): 470–474. doi:10.1038/290470a0. PMID:7219536.

Parsons, M. 1999. Butterflies of New Guinea. Academic Press, London, United Kingdom.

Preston-Mafham, K. 1991. Madagascar, a natural history. Facts on file Inc., Oxford, United Kingdom and New York.

Sahoo, R.K., Warren, A.D., Wahlberg, N., Brower, A.V., Lukhtanov, V.A., and Kodandaramaiah, U. 2016. Ten genes and two topologies: an exploration of higher relationships in skipper butterflies (Hesperiidae). PeerJ, **4**: e2653. doi:10.7717/peerj.2653. PMID:27957386.

Scoble, M.J. 1986. The structure and affinities of the Hedyloidea: a new concept of the butterflies. Bull Br. Mus. **53**(5): 251–286.

Shao, L., Sun, Q., and Hao, J. 2015. The complete mitochondrial genome of *Parara guttata* (Lepidoptera: Hesperiidae). Mitochondrial DNA, **26**(5): 724–725. doi:10.3109/19401736.2013.845759. PMID:24460163.

Shen, J., Cong, Q., and Grishin, N.V. 2015. The complete mitochondrial genome of *Papilio glaucus* and its phylogenetic implications. Meta Gene, **5**: 68–83. doi:10.1016/j.mgene.2015.05.002. PMID:26106582.

Shen, J., Cong, Q., and Grishin, N.V. 2016. The complete mitogenome of *Achalarus lyciades* (Lepidoptera: Hesperiidae). Mitochondrial DNA B Resour. **1**(1): 581–583. doi:10.1080/23802359.2016.1197070. PMID:28367503.

Simonsen, T.J., de Jong, R., Heikkilä, M., and Kaila, L. 2012. Butterfly morphology in a molecular age. Does it still matter in butterfly systematics? Arthropod. Struct. Dev. **41**: 307–322. doi:10.1016/j.asd.2012.04.006. PMID:22583793.

Steinhauser, S.R. 1991. Taxonomic notes and descriptions of new taxa in the Neotropical Hesperiidae. Part II, Heteropterinae and Hesperiinae, *Vinius* group. Bull. Allyn Mus. **132**: 1–47.

Viette, P. 1956. Insectes: Lépidoptères Hesperiidae. *In* Faune de Madagascar. IRSM Tananarive, Tsimbazaza, Madagascar. pp. 1–85.

Wang, A.R., Jeong, H.C., Han, Y.S., and Kim, I. 2014. The complete mitochondrial genome of the mountainous duskywing, *Erynnis montanus* (Lepidoptera: Hesperiidae): a new gene arrangement in Lepidoptera. Mitochondrial DNA, **25**(2): 93–94. doi:10.3109/19401736.2013.784752. PMID:23586336.

Wang, J., James John, Y., Xuan, S., Cao, T., and Yuan, X. 2016. The complete mitochondrial genome of the butterfly *Hasora anura* (Lepidoptera: Hesperiidae). Mitochondrial DNA Part A, **27**(6): 4401–4402. doi:10.3109/19401736.2015.1089543. PMID: 26466192.

Wang, K., Hao, J., and Zhao, H. 2015. Characterization of complete mitochondrial genome of the skipper butterfly, *Celaenorrhinus maculosus* (Lepidoptera: Hesperiidae). Mitochondrial DNA, **26**(5): 690–691. doi:10.3109/19401736.2013.840610. PMID:24102602.

Warren, A.D., Ogawa, J.R., and Brower, A.V.Z. 2008. Phylogenetic relationships of subfamilies and circumscription of tribes in the family Hesperiidae (Lepidoptera: Hesperioidea). Cladistics, **24**(5): 642–676. doi:10.1111/j.1096-0031.2008.00218.x.

Warren, A.D., Ogawa, J.R., and Brower, A.V.Z. 2009. Revised classification of the family Hesperiidae (Lepidoptera: Hesperioidea) based on combined molecular and morphological data. Syst. Entomol. **34**(3): 467–523. doi:10.1111/j.1365-3113.2008.00463.x.

Yack, J.E., Johnson, S.E., Brown, S.G., and Warrant, E.J. 2007. The eyes of *Macrosoma* sp. (Lepidoptera: Hedyloidea): A nocturnal butterfly with superposition optics. Arthropod. Struct. Dev. **36**: 11–22. doi:10.1016/j.asd.2006.07.001. PMID:18089084.

Zhang, J., Cong, Q., Fan, X.L., Wang, R., Wang, M., and Grishin, N.V. 2017*a*. Mitogenomes of Giant-Skipper Butterflies reveal an ancient split between deep and shallow root feeders. F1000Res. **6**: 222. doi:10.12688/f1000research.10970.1. PMID:28408977.

Zhang, J., Cong, Q., Shen, J., Fan, X.-L., Wang, M., and Grishin, N.V. 2017*b*. The complete mitogenome of *Euschemon rafflesia* (Lepidoptera: Hesperiidae). Mitochondrial DNA Part B, **2**(1): 136–138. doi:10.1080/23802359.2017.1292478. PMID:30931394.

Zhang, J., Cong, Q., Shen, J., Brockmann, E., and Grishin, N.V. 2019. Three new subfamilies of skipper butterflies (Lepidoptera, Hesperiidae). Zookeys, **861**: 91–105. doi:10.3897/zookeys.861.34686. PMID:31333327.