



HAL
open science

Australian Sphingidae – DNA Barcodes Challenge Current Species Boundaries and Distributions

Rodolphe Rougerie, Ian J. Kitching, Jean Haxaire, Scott E. Miller, Axel Hausmann, Paul D. N. Hebert

► **To cite this version:**

Rodolphe Rougerie, Ian J. Kitching, Jean Haxaire, Scott E. Miller, Axel Hausmann, et al.. Australian Sphingidae – DNA Barcodes Challenge Current Species Boundaries and Distributions. PLoS ONE, 2014, 9 (7), 11 p. 10.1371/journal.pone.0101108 . hal-02634031

HAL Id: hal-02634031

<https://hal.inrae.fr/hal-02634031>

Submitted on 27 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Australian Sphingidae – DNA Barcodes Challenge Current Species Boundaries and Distributions

Rodolphe Rougerie^{1*}, Ian J. Kitching², Jean Haxaire³, Scott E. Miller⁴, Axel Hausmann⁵, Paul D. N. Hebert¹

1 University of Guelph, Biodiversity Institute of Ontario, Guelph, Ontario, Canada, **2** Natural History Museum, Department of Life Sciences, London, United Kingdom, **3** Honorary Attaché, Muséum National d'Histoire Naturelle de Paris, Le Roc, Laplume, France, **4** National Museum of Natural History, Smithsonian Institution, Washington, DC, United States of America, **5** Bavarian State Collection of Zoology, Section Lepidoptera, Munich, Germany

Abstract

Main Objective: We examine the extent of taxonomic and biogeographical uncertainty in a well-studied group of Australian Lepidoptera, the hawkmoths (Sphingidae).

Methods: We analysed the diversity of Australian sphingids through the comparative analysis of their DNA barcodes, supplemented by morphological re-examinations and sequence information from a nuclear marker in selected cases. The results from the analysis of Australian sphingids were placed in a broader context by including conspecifics and closely related taxa from outside Australia to test taxonomic boundaries.

Results: Our results led to the discovery of six new species in Australia, one case of erroneously synonymized species, and three cases of synonymy. As a result, we establish the occurrence of 75 species of hawkmoths on the continent. The analysis of records from outside Australia also challenges the validity of current taxonomic boundaries in as many as 18 species, including *Agrius convolvuli* (Linnaeus, 1758), a common species that has gained adoption as a model system. Our work has revealed a higher level of endemism than previously recognized. Most (90%) Australian sphingids are endemic to the continent (45%) or to Australia, the Pacific Islands and the Papuan and Wallacean regions (45%). Only seven species (10%) have ranges that extend beyond this major biogeographical boundary toward SE Asia and other regions of the Old World.

Main Conclusions: This study has established that overlooked cryptic diversity and inaccurate species delineation produced significant misconceptions concerning diversity and distribution patterns in a group of insects that is considered well known taxonomically. Because DNA barcoding represents a straightforward way to test taxonomic boundaries, its implementation can improve the accuracy of primary diversity data in biogeography and conservation studies.

Citation: Rougerie R, Kitching IJ, Haxaire J, Miller SE, Hausmann A, et al. (2014) Australian Sphingidae – DNA Barcodes Challenge Current Species Boundaries and Distributions. PLoS ONE 9(7): e1011108. doi:10.1371/journal.pone.0101108

Editor: Ben J. Mans, Onderstepoort Veterinary Institute, South Africa

Received: March 19, 2014; **Accepted:** May 31, 2014; **Published:** July 2, 2014

Copyright: © 2014 Rougerie et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. Specimen and sequences data are publicly available from the Barcode of Life Database (BOLD, www.boldsystems.org) and from GenBank (all accession numbers are given in supplementary tables).

Funding: This work was supported by funding from the government of Canada through Genome Canada (www.genomecanada.ca) and the Ontario Genomics Institute (www.ontariogenomics.ca) in support of the International Barcode of Life project (www.iBOL.org), and by the Natural Sciences and Engineering Research Council of Canada (NSERC, www.nserc-crsng.gc.ca). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: rrougeri@gmail.com

‡ Current address: INRA, UR633 Zoologie Forestière, Orléans, France

Introduction

“The crux of the problem is that if we don’t know what is out there or how widely species are distributed, how can we convince people about the reality and form of the biodiversity crisis?” Riddle et al. [1].

As in many other groups of Australian organisms, the insect fauna of this continent is renowned for its diversity, uniqueness, and many iconic endemics [2]. However, in contrast to plants and vertebrates, the diversity of Australian insects remains poorly documented. There are about 62,000 described species of insects on the continent, representing from 15% to 50% of the total fauna depending on estimates [3–5]. The level of species endemism is

difficult to evaluate and undoubtedly varies among insect orders [2], but is generally considered to be very high. The overall value may reach 90% [3], and many families include only endemics [2]. However, much effort is still needed to fill gaps in taxonomy and to extend understanding of taxon distributions, deficits that have come to be termed respectively the Linnean and Wallacean shortfalls [6].

The integration of molecular methods into taxonomic practice has improved our capacity to understand and describe diversity [7,8] and has been proposed as a way to help increase the pace of species discovery and description [9,10]. Although the application of integrative approaches to the Australian fauna is still in its infancy, studies have already revealed spectacular levels of overlooked and cryptic species in some groups of vertebrates

[11,12]. Similar cases have also been reported in invertebrates, typically revealing species that raise the already high level of endemism in Australia [13–15]. Particularly comprehensive efforts have been directed toward Australian Lepidoptera. Much of this information is available in an indexed public reference library of DNA barcode sequences [16,17] (database accessible in BOLD at www.boldsystems.org) and via the current taxonomic system for the complete Australian fauna [18] (see also <http://www.lepbarcoding.org/australia>).

The order Lepidoptera includes about 130 families whose diversity varies dramatically – from a single to more than 24,000 species [19]. Because the current count of approximately 160,000 described species is thought to reflect less than one-third of the total [20], diversity estimates for some families, especially smaller-bodied taxa, are very uncertain. However, other families, such as the Sphingidae, have attracted so much taxonomic interest that present estimates of diversity are thought to be comprehensive except for ‘frontier’ regions. The last global conspectus on the Sphingidae recognized 1278 species [21], but this count has now risen by about 200 species [19] with most new taxa deriving from hyperdiverse tropical faunas. In Australia, Europe and North America, the pace of species discovery has slowed, suggesting that species counts are complete or nearly so.

The Sphingidae of Australia

Australia provides a good example of a continental fauna of sphingids with a mature taxonomy. Collection programs have been wide ranging, morphological studies have been careful, and species discovery has slowed: 53 species were described by the end of the 19th century, 62 by 1927, but only two more taxa were added during the next 70 years [22]. Although many families of Australian Lepidoptera are dominated by endemic species [2], the sphingid fauna is thought to include a mix of endemics and species with distributions extending into the Pacific islands, Southeast Asia and even into Europe and Africa.

The descriptions of some wide-ranging species in the Australian fauna date back to the launch of Linnaean nomenclature (e.g. *Agrus convolvuli* (Linnaeus, 1758) and *Hippotion celerio* (Linnaeus, 1758)). Other species were described only slightly later from Asia (e.g. *Theretra nessus* (Drury, 1773) and *Nephela hespera* (Fabricius, 1775), both described from India). The first sphingids to be described from Australia were the endemics *Coequosa australasiae* (Donovan, 1805), *Coequosa triangularis* (Donovan, 1805) and *Cizara ardeniae* (Lewin, 1805). Moulds [22] provided an historical overview of work on Australian sphingids, listing 64 species in 22 genera for the continent. Although only two species (*Imber tropicus* (Moulds, 1983) and *Psilogramma argos* Moulds & Lane, 1999) were described from 1930–2000, eleven species and one subspecies have been described since then, including seven species of *Psilogramma*, a complex genus whose species count exploded globally from four to more than 60 over the past decade [21,23–25], although perhaps only about half of these are actually valid species (I.J. Kitching, unpublished data). Three new genera, *Imber* Moulds *et al.*, 2010, *Pseudoangonyx* Eitschberger, 2011, and *Cerberonoton* Zolotuhin & Ryabov, 2012 were recently erected for rather distinctive species, with a fourth new genus being established for a newly encountered species (*Zacria vojtechii* Haxaire & Melichar, 2003) demonstrating that striking discoveries can still occur in this fauna. Recent nomenclatural changes within the genus *Macroglossum* also affected the names of two Australian species [26]. In his checklist of Australian Sphingidae, Moulds [22] omitted subspecies names, concealing the distinctiveness of the Australian fauna from those of Neighbouring Southeast Asia and Pacifica. In actuality, many Australian sphingids have been

assigned to an endemic subspecies because of their morphological divergence from the nominotypical subspecies. At least some of these cases may reflect cases of long separation that merit recognition as distinct species, but their status has not yet been investigated in detail.

The delineation of species/subspecies and the existence of cryptic or overlooked undescribed species can be a serious impediment for biogeography and conservation studies [6]. This ‘Linnean shortfall’ represents a formidable challenge for research in these disciplines [27]. Building on recent access to molecular tools for species identification and discrimination [28,29] and the success of integrative approaches in taxonomy [30–32], the present study employs sequence diversity in the DNA barcode region of the mitochondrial gene COI to investigate the diversity of Australian sphingids and their distributions. We begin by examining patterns of sequence diversity within the Australian fauna to ascertain the efficacy of DNA barcodes to discriminate known species and reveal overlooked diversity. We then place these results in a broader geographic context by examining patterns of sequence divergence between Australian populations and their conspecific lineages outside the continent.

Material and Methods

Specimen Sampling

More than 1200 Australian sphingids were sampled, most from three collections – the Biodiversity Institute of Ontario (361), the Zoologische Staatssammlung München (300), and the Australian National Insect Collection (112). The remaining specimens derive from 17 other collections (see details in Table S1). All institutions and collections granted us permission to access and study the material used in this work. Sampling aimed to maximize the geographic coverage for each species or subspecies to examine the extent of barcode variation across its distribution. In the most common species, a large number of specimens were sampled (more than 30 for 11 species, 113 for *Agrus convolvuli*), but fewer than five specimens were examined for 20 species/subspecies that are rare. Each record was given a unique specimen identifier (SampleID) and sequence identifier (ProcessID) to provide a direct link between voucher specimens and DNA barcode records. Collection data, a photograph and ancillary information, such as the collection holding each specimen, were compiled in BOLD for each individual record.

DNA Sequencing

The methods used for generation of DNA sequences follow standard protocols designed for amplification and sequencing of the standard DNA barcode. In four species, the same DNA extracts were also used for the sequencing of the D2 expansion segment of the 28S rDNA gene. Details of these methods are given as Supporting Information (Appendix S1).

Data Analysis

A nearest Neighbour analysis was conducted with BOLD using either a dataset restricted to sphingid records from Australia or the full database for world sphingids, including all COI barcode sequences longer than 500 bp. The latter database includes about 23 K sequences, 1.8 K species and subspecies (as of September 12th, 2013). BOLD was also used for sequence alignment and calculation of genetic divergences and Neighbour Joining (NJ) trees using a K2P distance model [33] after alignment. These trees were imported in iTOL [34] to exploit its capacity for the visualization of large trees. In two cases where DNA barcode analysis suggested potential cases of overlooked diversity, we also

Table 1. Sampling coverage and genetic distances for the Sphingidae of Australia.

#	Species	Type Locality	N	Australia	EM	Pi	WM	SEA	OW	ΔA	sd	ΔG	dG	sd	
1	<i>Acosmeryx anceus</i> (Stoll, 1781)	Indo.: Maluku	59	47 [QLD,NSW]	12	-	-	-	-	6.41	2.86	0	6.61	2.86	0
2	<i>Acosmeryx miskini</i> (Murray, 1873)	Austr.: QLD	13	13 [QLD,NSW]	-	-	-	-	-	0.93	0.35	0	N/A	N/A	N/A
3	<i>Angonyx papuana</i> Rothschild & Jordan, 1903	Austr.: QLD	11	4 [QLD]	7	-	-	-	-	0.15	0.03	0.01	1.08	0.38	0
4	<i>Cephonodes hylas australis</i> Kitching & Cadiou, 2000	Austr.: QLD	4	4 [QLD,NT]	-	-	-	-	-	0.51	0.25	0.03	N/A	N/A	N/A
5	<i>Cephonodes janus</i> Miskin, 1891	Austr.: QLD	2	2 [QLD]	-	-	-	-	-	0.17	N/A	N/A	N/A	N/A	N/A
6	<i>Cephonodes kingii</i> (W.S. Macleay, 1826)	Austr.	6	6 [QLD]	-	-	-	-	-	0.33	0.18	0.01	N/A	N/A	N/A
7	<i>Cephonodes picus</i> (Cramer, 1777)	India	11	5 [QLD]	1	0	3	2	-	1.39	0.57	0.05	3.59	1.37	0.02
8	<i>Cizara ardeniae</i> (Lewin, 1805)	Austr.: NSW	4	4 [QLD]	-	-	-	-	-	0.85	0.4	0.04	N/A	N/A	N/A
9	<i>Daphnis dohertyi</i> Rothschild, 1897	Indo.: Papua Barat	4	1 [QLD]	3	-	-	-	-	N/A	N/A	N/A	0.38	0.29	0.02
10	<i>Daphnis moorei</i> (Macleay, 1866)	Austr.: QLD	16	10 [QLD,WA]	5	1	-	-	-	0.46	0.18	0	1.03	0.27	0
11	<i>Daphnis placida</i> (Walker, 1856)	Indo.: Sumatra	25	12 [QLD,NSW]	5	1	4	3	-	0.77	0.22	0	2.13	0.68	0
12	<i>Daphnis protrudens</i> Felder, C. & Felder R., 1874	Indo.: Maluku	6	3 [QLD]	3	-	-	-	-	0	0	0	0.61	0.26	0.01
13	<i>Eupanacra splendens</i> (Rothschild, 1894)	Austr.: QLD	27	15 [QLD]	10	2	-	-	-	2.19	0.97	0.01	6.87	2.17	0
14	<i>Gnathothlibus australiensis</i> Lachlan, 2004	Austr.: QLD	5	5 [QLD]	-	-	-	-	-	0.61	0.34	0.02	N/A	N/A	N/A
15	<i>Gnathothlibus eras</i> (Boisduval, 1832)	French Polynesia	39	30 [QLD,NT]	3	6	-	-	-	0.92	0.28	0	3.47	0.58	0
16	<i>Hippotion biennus</i> (Stoll, 1782)	Indo.: Maluku	38	8 [QLD,NSW]	30	-	-	-	-	1.39	0.9	0.02	3.96	1.93	0
17	<i>Hippotion celestia</i> (Linnaeus, 1758)	?	92	42 [QLD,CT,NSW,WA]	7	1	5	0	37	1.09	0.07	0	1.58	0.34	0
18	<i>Hippotion rosetta</i> (Swinhoe, 1892)	Indo.: Maluku	32	27 [QLD]	2	-	2	1	-	0.77	0.30	0	1	0.35	0
19	<i>Hippotion scrofa</i> (Boisduval, 1832)	Austr.	66	63 [QLD,NSW,CT,WA,NT]	-	3	-	-	-	0.68	0.08	0	1.49	0.16	0
20	<i>Hippotion velox</i> (Fabricius, 1793)	"India orientali"	24	7 [QLD]	10	2	1	4	-	0.31	0.09	0	1.55	0.22	0
21	<i>Hyles livornicoides</i> (Lucas, 1892)	Austr.: QLD	39	39 [QLD,NSW,WA]	-	-	-	-	-	0.66	0.21	0	N/A	N/A	N/A
22	<i>Macroglossum alcedo</i> Boisduval, 1832	Irian Jaya	2	2 [QLD]	0	-	-	-	-	0.48	N/A	N/A	N/A	N/A	N/A
23	<i>Macroglossum corythus</i> Walker, 1856	Sri Lanka	34	1 [QLD]	15	6	3	9	-	N/A	N/A	N/A	3.48	1.66	0
24	<i>Macroglossum dohertyi doddii</i> Clark, 1922	Austr.: QLD	7	5 [QLD]	2	-	-	-	-	0.33	0.13	0.01	0.33	0.13	0
25	<i>Macroglossum divergens queenslandi</i> Clark, 1927 ¹	Austr.: QLD	-	0	-	-	-	-	-	-	-	-	N/A	N/A	N/A
26	<i>Macroglossum hirundo errans</i> Walker, 1856	Austr.	13	11 [Q,NSW]	0	2	-	-	-	1.03	0.33	0	1.37	0.53	0
27	<i>Macroglossum joannisi</i> Rothschild & Jordan, 1903	Austr.: QLD	1	1 [QLD]	-	-	-	-	-	N/A	N/A	N/A	N/A	N/A	N/A
28	<i>Macroglossum micacea</i> Walker, 1856	Austr.: QLD	6	6 [QLD,NSW]	0	-	-	-	-	0.86	0.48	0.02	N/A	N/A	N/A
29	<i>Macroglossum nubilum</i> Rothschild & Jordan, 1903	PNG: Milne Bay	4	3 [QLD]	1	-	-	-	-	0.21	0.13	0.03	1.04	0.38	0.05
30	<i>Macroglossum prometheus lineata</i> Lucas, 1891	Austr.: QLD	8	4 [QLD]	4	-	-	-	-	0.48	0.34	0.02	1.34	0.7	0.01
31	<i>Macroglossum rectans</i> Rothschild & Jordan, 1903	Indo.: Maluku	9	5 [QLD]	4	-	-	-	-	1.04	0.5	0.03	1.04	0.53	0.01
32	<i>Macroglossum tenebrosa</i> Lucas, 1891	Austr.: QLD	10	4 [QLD]	6	-	-	-	-	0.46	0.23	0.04	2.98	1.17	0.02
33	<i>Macroglossum troglodytes</i> Boisduval, [1875] ²	India: Darjeeling	1	1 [QLD]	-	-	-	-	-	N/A	N/A	N/A	N/A	N/A	N/A
34	<i>Macroglossum vacillans</i> Walker, 1865	"Timor"	22	13 [QLD,NT,WA]	9	-	-	-	-	0.83	0.27	0	1.24	0.36	0
35	<i>Nephele hespera</i> (Fabricius, 1775)	"India orientali"	11	1 [QLD]	-	-	-	10	-	N/A	N/A	N/A	2.32	0.48	0.01

Table 1. Cont.

#	Species	Type Locality	N	Australia	EM	Pi	WM	SEA	OW	ΔA	sd	ΔG	dG	sd
36	<i>Nephele subvaria</i> (Walker, 1856)	Austr.	15	15 [QLD,WA,NT]	-	0	-	-	-	0.26	0.042	0	N/A	N/A
37	<i>Pseudoangonyx excellens</i> (Rothschild, 1911)	Indo.: Papua	4	2 [QLD]	2	-	-	-	-	0	N/A	N/A	0.53	0.4
38	<i>Theretra celata</i> (Butler, 1877)	Austr.: QLD	25	17 [QLD, NSW]	8	-	-	-	-	0.59	0.14	0	0.79	0.16
39	<i>Theretra indistincta</i> (Butler, 1877)	Austr.: QLD	12	12 [QLD, NT]	-	-	-	-	-	0.36	0.06	0	N/A	N/A
40	<i>Theretra inornata</i> (Walker, 1865)	"N. Austr."	12	12 [QLD, NT]	-	-	-	-	-	1.24	0.47	0.01	N/A	N/A
41	<i>Theretra latreillii</i> (W.S. Macleay, 1826)	Austr.	58	56 [QLD,NSW,WA]	2	-	-	-	-	0.47	0.025	0	0.62	0.04
42	<i>Theretra margarita</i> (Kirby, 1877)	Austr.: QLD	30	30 [QLD,WA,NT]	-	-	-	-	-	0.89	0.22	0	N/A	N/A
43	<i>Theretra nessus</i> (Drury, 1773)	India: Tamil Nadu	31	16 [QLD,NSW]	6	1	0	8	-	0.37	0.15	0	4.97	2.2
44	<i>Theretra oldenlandiae lewini</i> (Thon, 1828)	Austr.: NSW	82	81 [QLD,NSW,NT,WA]	1	-	-	-	-	0.86	0.24	0	0.86	0.24
45	<i>Theretra queenslandi</i> (Lucas, 1891)	Austr.: QLD	15	15 [QLD,NSW]	-	-	-	-	-	0.18	0.02	0	N/A	N/A
46	<i>Theretra radiosa</i> Rothschild & Jordan, 1916	PNG	4	0	4	-	-	-	-	-	-	-	4.42	2.73
47	<i>Theretra silhetensis intersepta</i> (Butler, [1876])	Austr.: QLD	24	20 [QLD,NT,WA]	0	2	2	-	-	0.31	0.03	0	0.31	0.05
48	<i>Theretra tryoni</i> (Miskin, 1891)	Austr.: QLD	10	5 [QLD,NSW]	5	-	-	-	-	0.31	0.093	0.01	0.84	0.44
49	<i>Theretra turneri</i> (Lucas, 1891)	Austr.: QLD	7	7 [QLD]	-	-	-	-	-	0.63	0.22	0.01	N/A	N/A
50	<i>Zacria vojtechii</i> Haxaire & Melichar 2003	Austr.: WA	4	4 [WA]	-	-	-	-	-	0.17	0.06	0.01	N/A	N/A
51	<i>Ambulyx dohertyi queenslandi</i> Clark, 1928	Austr.: QLD	11	4 [QLD]	7	-	-	-	-	0.15	0.02	0.01	0.82	0.31
52	<i>Ambulyx wildei</i> Miskin, 1891	Austr.: QLD	20	11 [QLD]	9	-	-	-	-	0	0	0	0.78	0.28
53	<i>Coequosa australasiae</i> (Donovan, 1805)	Austr.	8	8 [QLD,NSW,NT]	-	-	-	-	-	1.17	0.62	0.01	N/A	N/A
54	<i>Coequosa triangularis</i> (Donovan, 1805)	Austr.	5	5 [QLD,NSW]	-	-	-	-	-	0.5	0.32	0.02	N/A	N/A
55	<i>Imber tropicus</i> Moulds, 1983	Austr.: NT	7	7 [QLD,WA]	-	-	-	-	-	2.02	1.21	0.03	N/A	N/A
56	<i>Agrilus convolvuli</i> (Linnaeus, 1758)	?	174	113 [QLD,CT,NSW,WA,SA,NT]	4	3	0	14	40	0.52	0.06	0	5.37	1.79
57	<i>Agrilus godarti</i> (W.S. Macleay, 1826)	Austr.	16	16 [QLD,NSW]	-	-	-	-	-	0.98	0.23	0	N/A	N/A
58	<i>Cerberonoton rubescens severina</i> (Miskin, 1891)	Austr.: QLD	8	6 [QLD]	2	-	-	-	-	0.17	0.05	0	1.08	0.39
59	<i>Coenotes eremophilae</i> (Lucas, 1891) ³	Austr.: QLD	17	17 [QLD,WA,NT]	-	-	-	-	-	1.28	0.31	0	N/A	N/A
60	<i>Hoplocnema brachycera</i> (Lower, 1897)	Austr.: NSW	13	13 [NT,WA,NSW,SA,VIC]	-	-	-	-	-	0.5	0.15	0	N/A	N/A
61	<i>Hoplocnema lacunosa</i> Tuttle et al. 2012	Austr.: WA	5	5 [WA]	-	-	-	-	-	0.93	0.4	0.03	N/A	N/A
62	<i>Hoplocnema ochra</i> Tuttle et al. 2012	Austr.: WA	6	6 [WA]	-	-	-	-	-	0.93	0.54	0.01	N/A	N/A
63	<i>Leucomania bethia</i> (Kirby, 1877)	Austr.: QLD	6	6 [QLD,WA,NT]	-	-	-	-	-	0.31	0.16	0.01	N/A	N/A
64	<i>Psilogramma argos</i> Moulds & Lane, 1999	Austr.: QLD	3	3 [NT]	-	-	-	-	-	0.61	0.41	0.1	N/A	N/A
65	<i>Psilogramma casuarinae</i> (Walker, 1856) ⁴	Austr.: NSW	32	32 [QLD,NSW,VIC]	-	-	-	-	-	1.24	0.43	0	N/A	N/A
66	<i>Psilogramma exigua</i> Brechlin et al. 2010	Austr.: WA	16	16 [QLD,WA,NT]	-	-	-	-	-	0.46	0.13	0	N/A	N/A
67	<i>Psilogramma maxmouldsi</i> Eitschberger, 2001	Austr.: QLD	9	9 [QLD,NSW]	-	-	-	-	-	0.31	0.15	0	N/A	N/A
68	<i>Psilogramma menephron</i> (Cramer, 1780) ⁵	Indo.: Maluku	114	34 [QLD,NSW]	30	-	21	29	-	0.78	0.2	0	8.25	3.97
69	<i>Psilogramma papuensis</i> Brechlin, 2001 ⁶	PNG: Western	37	25 [QLD]	12	-	-	-	-	0.35	0.11	0	1.17	0.34
70	<i>Psilogramma penumbra</i> Lane et al. 2011	Austr.: NT	4	4 [NT]	-	-	-	-	-	0	0	0	N/A	N/A

Table 1. Cont.

#	Species	Type Locality	N	Australia	EM	Pi	WM	SEA	OW	δA	dA	sd	δG	dG	sd
71	<i>Synoecha marmorata</i> (Lucas, 1891) ³	Austr.: QLD	14	14 [QLD,NSW,NT]	-	-	-	-	-	9.74	2.72	0.04	N/A	N/A	N/A
72	<i>Tetrachroa edwardsi</i> (Olliff, 1890)	Austr.: QLD/NSW	4	4 [QLD]	-	-	-	-	-	0	0	0	N/A	N/A	N/A

Taxonomic checklist, type localities, sample size (N), geographical origin and DNA barcode variations for analysed samples; underlined names highlight taxa with no DNA barcode available from Australia. CT = Capital Territory, EM = Eastern Malaysia (Wallacea + Papuan region), NSW = New South Wales, NT = Northern Territory, OW = other regions of the world, Pi = Pacific islands, QLD = Queensland, SA = South Australia, SEA = continental Southeast Asia, VIC = Victoria, WA = Western Australia, WM = Western Australia, WM = West of the Wallacean region; δA, dA, δG and dG = maximum (δ) and mean (d) intraspecific K2P distance (%) within Australia (A) and globally (G); sd = standard deviation; numbers in the first column match numbered clusters in Figs 2 and 3. The δ distances over 2% are highlighted in bold characters in cases involving at least one Confirmed Candidate Species (CCS), and in italics in cases where there is no complementary evidence available (Deep Conspecific Lineages, DCL). Species are given a null sample size in areas where they are known to occur but from where no sample was analyzed whereas empty cells mark areas where the species are considered absent in our current state of knowledge. Taxa are sorted by subfamily as follows: 1–50 = Macroglossinae, 51–55 = Smerinthinae, 56–72 = Sphinginae.

¹This species is only known to us from specimens bred or collected in the 1910's. Previously referred as *M. heliophila queenslandi*. See Holloway (2011)

²Previously referred as *M. insipida* Butler, 1875. See Holloway (2011)

³A new species, *Coenotes arida* Moulds & Melichar, [2014], was described during the review of this study and does not appear in this table. Its original generic and specific identifications as *Synoecha marmorata* were erroneous, and the affinity of these specimens with the genus *Coenotes* has been confirmed by morphology.

⁴*P. hausmanni* Eitschberger, 2001 syn. nov. (holotype sequenced) is here confirmed a synonym of *P. casuarinae*. It is neither morphologically nor genetically distinct from this species.

⁵*P. gloriosa* Eitschberger, 2001 syn. nov. (holotype sequenced) is here considered a synonym of *P. menephron*. It is neither morphologically nor genetically distinct. *P. menephron* represents a vast species complex in Asia and the Indo-Australian region, and a revision is needed to sort out the divisions within this complex (e.g. DNA barcodes reveal nine distinct clusters).

⁶*P. koalae* Eitschberger, 2001 syn. nov. (holotype sequenced) is here considered a synonym of *P. papuensis* (holotype sequenced). It is neither morphologically nor genetically distinct from this species.

doi:10.1371/journal.pone.0101108.t001

generated “geo-phylogenies” using GenGIS 2.2.0 [35] to better visualize the geographical structure of genetic variation. MEGA 5.2.2 [36] was also used to produce character-based trees using the maximum parsimony (MP) criterion (all settings at default, with the complete deletion option for missing data) in cases of species-complexes where genetic distances are shallow and require a more rigorous method for reconstructing relationships between terminals. The topology of these DNA barcode “gene trees” was compared with the “species tree” inferred from current taxonomy based on morphology. Attention was directed toward an examination of the incidence of reciprocal monophyly rather than relationships between species. The division of a species into two or more clusters of specimens was considered a potential indicator of cryptic diversity (candidate species) when the species included two or more monophyletic clusters with more than 2% divergence [37,38] in their DNA barcode sequences. This threshold was used as an operational criterion to trigger in-depth study (comparative morphology, sequencing of a nuclear marker), but cases of shallower divergences are also reported and discussed when the genetic clusters match existing or previously proposed taxonomic divisions (at species or subspecies level). Candidate species were considered as Confirmed Candidate Species (CCS) when independent evidence (e.g. morphology, nuclear gene) also supported their distinction, or as Deep Conspecific Lineages (DCL) when no morphological differences or divergence at 28S rDNA were detected [39,40]. We note, however, that there is no objective way to confirm the status of any geographically isolated CCS as distinct species (or subspecies). As pointed out by Mutanen et al. [38], taxonomists working on Lepidoptera when dealing with closely related but allopatric populations have traditionally assigned them to different subspecies when slight morphological differences are apparent (usually in wing patterns), but to distinct species when morphological differences are greater (usually both in wing patterns and genitalia). The consideration of genetic differences as a basis for species delineation in Lepidoptera is recent. As a result there is no established practice regarding the taxonomic treatment of units currently diagnosed only through DNA data, whether their divisions are derived from the analyses of a single or multiple DNA markers. Our strategy has involved close collaboration with taxonomists studying Australian hawkmoths, highlighting all potential or confirmed candidate species for further morphological, ecological and biogeographical investigation.

Data Accessibility (BOLD, GenBank)

Data on specimen vouchers, sequence data including trace files and registered primer pairs are available on BOLD within the datasets SPH01AUS (dx.doi.org/10.5883/DATASET-SPH01AUS) for Australian records, and SPH02AUS (dx.doi.org/10.5883/DS-SPH02AUS) for conspecifics and most closely related taxa outside Australia. All sequences were also deposited in GenBank for COI DNA barcodes and for 28S rDNA sequences (see Tables S1–S3).

Results

Diversity of Australian Sphingids

In total, DNA barcodes were obtained from 1054 specimens representing 70 of the 72 valid sphingid species currently recognized from Australia (Table 1). Belonging to 25 genera, these specimens were sampled from across the continent with the majority from Queensland (56%) and New South Wales (24%), where diversity is the highest and collecting efforts have been greatest (Fig. 1). This total includes four species (*Psilogamma exigua*,

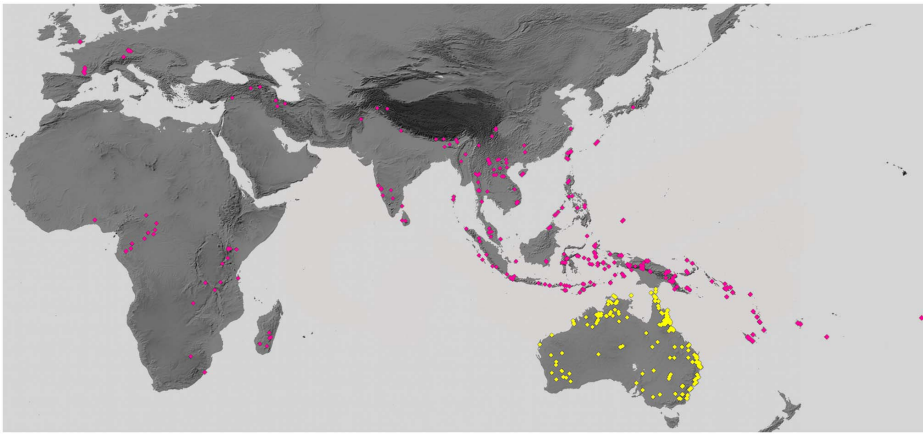


Figure 1. Geographical sampling. Distribution of the 1054 specimens of Australian Sphingidae with DNA barcodes analysed in this study (yellow) and of the 735 additional specimens of conspecifics (including hetero-subspecific taxa) from outside Australia and relevant closely related species with DNA barcodes (see details of record lists in Table S1–S3). doi:10.1371/journal.pone.0101108.g001

P. penumbra, *Hopliocnema ochra*, *H. lacunosa*) that were recently described on the basis of their DNA barcode divergence and correlated morphological differences [23,41,42]. The genus *Psilogramma* requires a global taxonomic and nomenclatural revision [41]. Nevertheless, but for the possible exceptions of *P. menephron* and *P. papuensis*, the nomenclature of the seven species found in Australia should remain stable if considering the three cases of synonymy revealed here by the analysis of the relevant type specimens (*P. hausmanni* syn. nov., *P. gloriosa* syn. nov., *P. koalae* syn. nov.; see Table 1, footnotes 4–6).

NJ analysis of the barcode sequences indicated that representatives of each of the 70 species formed a cohesive and reciprocally monophyletic cluster (Fig. 2; see Fig. S1 for details). The mean within-genus divergence (7.13%, sd. = 0.007%) is considerably higher than the mean within-species distance (0.3%, sd. = 0.007%), matching the pattern of variation already reported in other families of Lepidoptera [43]. The lowest minimum interspecific distance was 1.19% between *Macroglossum rectans* and *M. hirundo errans*, but the clusters for these species were still separate, and both taxa possessed a maximum intraspecific divergence of 1.0%. The next closest pair, *Hippotion celerio* and *H. velox*, had 2.7% divergence, while *Coequosa australasiae* and *C. triangularis* had the greatest nearest-Neighbour divergence (7.2%). Intraspecific variation among the Australian samples (Fig. 2 and Table 1) was generally low, even in some of the most heavily sampled species such as *Agrius convolvuli* (Dmax = 0.5%, N = 113) or *Theretra latreillii* (Dmax = 0.5%, N = 56). However, intraspecific divergence did exceed 2.0% in four species (Table 1): *Imber tropicus* (2%), *Eupanacra splendens* (2.2%), *Acosmeryx anceus* (6.4%), and *Synoecha marmorata* (9.7%) with each taxon represented by two clusters (Fig. 2, Fig. S1). The first two cases involve intraspecific distances of approximately 2%, and we treat them as cases of DCL (Fig. 2, highlighted in pale yellow), a conclusion supported by their lack of morphological divergence and the absence in *E. splendens* of congruent variation in 28S rRNA sequences (Fig. S2). In the other two cases, intraspecific divergence was substantially higher (Fig. 2, highlighted in red), and other lines of evidence support their status as CCS. Morphological investigation of the split in *S. marmorata* [44] revealed that one of its lineages represents an undescribed species (formally described as *Coenotes* Moulds & Melichar, [2014] during the review of this study). The original generic and specific identifications were erroneous, and the affinity of these specimens with the genus *Coenotes* is confirmed by morphology and is

supported by the DNA barcode data (representatives of the new taxon show just 3.3% minimum genetic distance from *C. eremophilae* (Fig. S1)). The fourth species, *A. anceus*, includes two lineages with 6.4% divergence, one broadly distributed in eastern Australia, the other restricted to northern Queensland. Interestingly, this split matches a tentative separation already proposed in the 1970s by taxonomists Alan Hayes and Jean-Marie Cadiou (I.J. Kitching, personal communication) and materialized through specimens sorted accordingly in collections at the Natural History Museum, London. Re-examination of series of specimens and rearing revealed several morphological characters (differences in size and wing colouration, as well as larval and pupal morphology (J.P. Tuttle & M.S. Moulds, personal communication)) and there is also one diagnostic nucleotide substitution between the taxa at position 550 in the analysed segment of 28S rRNA (Fig. S2). Furthermore, while the two DNA barcode clusters of *A. anceus* appear sympatric in northern Queensland (Fig. S3), they have micro-allopatric distributions; one cluster is restricted to sclerophyll forests along the east coast of Australia, and the other has only been collected in rain forests in Eastern Cape York Peninsula and Papua New Guinea. Lastly, a shallower split (1.4%) among Australian records of *Hippotion brennus* suggests the validity (in agreement with similar suggestions by Darge [45] and Riotte [46]) of the synonymized taxon, *johanna* (Kirby, 1877), because the three sequenced specimens with a typical *johanna*-like morphology form a distinct DNA barcode group (Figs S1 and S4) that is also separated geographically. Overall, considering this last and the two additional new species in the genera *Acosmeryx* and *Coenotes*, these results bring the total of Australian sphingid species to 75.

A Global Perspective on Australian Sphingids

Of the 75 species of sphingids that occur in Australia, 30 of them (40%) are endemic to the continent (Table S4). Another 45 species then occur in both Australia and another geographic region, of which four are represented by an endemic subspecies in Australia (the nominal subspecies in two cases). This brings to 34 (45%) the number of endemic sphingid taxa in Australia, at species or subspecies level. For all of the 45 species distributed in Australia and outside the continent, we examined the patterns of DNA barcode variation after including specimens from other regions. This involved the combined analysis of sequences from 1054 Australian specimens, 599 conspecifics from outside Australia, and

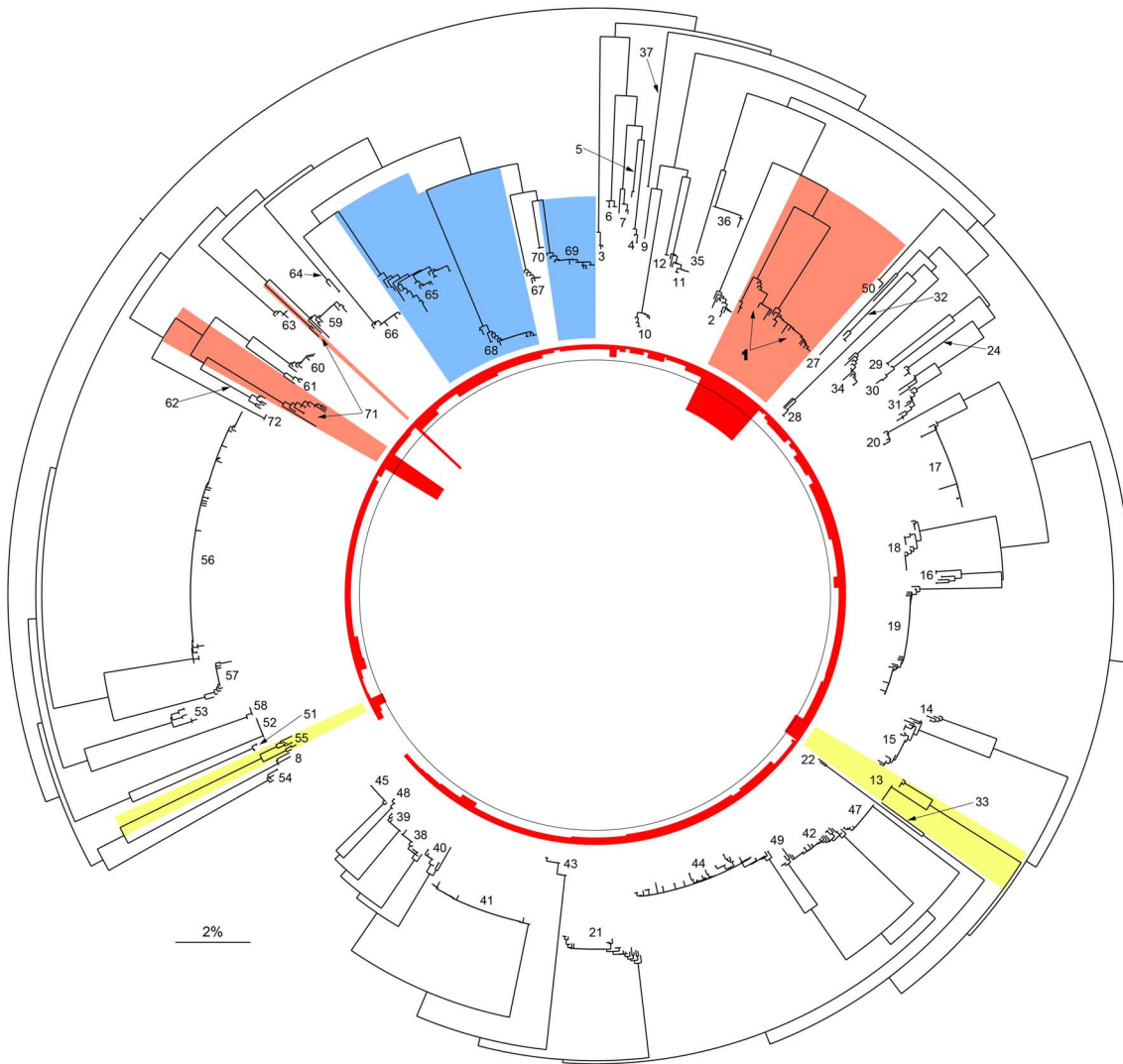


Figure 2. DNA barcode variation in Australian Sphingidae. Neighbour Joining tree based on K2P distances for 1054 DNA barcodes of 70 species of Australian sphingid moths. For each species the maximum intraspecific distance is given in the central part of the tree as histograms at the tip of branches (ranging from 0 to 9.7%, the circle line mark the 2% threshold); colour ranges highlight cases of synonymy (in blue), cryptic diversity (CCS, in red) as well as these species with 2% or more intraspecific distance (DCL, yellow) but not proven to represent two different species. Numbers refer to species numbers as listed in Table 1. An interactive and fully explorable version of the tree is available at <http://itol.embl.de/shared/rodroug>. (See also Fig. S1).

doi:10.1371/journal.pone.0101108.g002

136 records for closely related species from other regions (Table 1 and Tables S1–S3). The 599 conspecific records include 112 con-specific records and 140 hetero-subspecific records for 18 polytypic Australian species, and 347 records for 26 monotypic Australian species or species with no available record for distinct subspecies. The results of the analysis of this combined dataset are given in Fig. 3 (see Fig. S5 for details).

This analysis revealed an additional very deep split in *Eupanacra splendens* and ten more cases of deep ($>2\%$) intraspecific divergences involving *Agrilus convolvuli*, *Cephonodes picus*, *Daphnis placida*, *Gnathothlibus eras*, *Hippotion brennus*, *Macroglossum corythus*, *M. tenebrosa*, *Nephele hespera*, *Psilogramma menephron* and *Theretra nessus* (Table 1, Fig. 3). The two DNA barcode clusters in *Agrilus convolvuli* show deep divergence ($D_{max} = 5.37\%$), and clear geographical segregation (Fig. S6). Preliminary comparisons of wing patterns and male genitalia have been inconclusive and further morphological study is required, but the two lineages show three

diagnostic substitutions (positions 317, 486 and 505) in sequences for the 28S rDNA gene (Fig. S2), supporting their status as different species. Three other species with deep splits (*E. splendens* (the West New Britain record only), *T. nessus*, *H. brennus*) showed morphological differences between representatives of the divergent DNA barcode clusters (e.g. Fig. S4), suggesting that they too represent cases of overlooked diversity. Other cases also merit deeper investigation. For example, the observed genetic divergences in *C. picus* challenge the proposed synonymy of *C. cunninghami* (Walker, 1856) [21] described from Australia. In this case, as well as three others (*G. eras*, *D. placida*, *N. hespera*), the genetic splits are associated with geographical distances of several thousand kilometres across major biogeographical boundaries. In *M. tenebrosa*, DNA barcodes reveal the genetic distinctness of the representatives from Sulawesi (Fig. S5); morphological comparisons are yet to be conducted, and no variation was observed in the few available sequences of the D2 fragment of the 28S rDNA gene

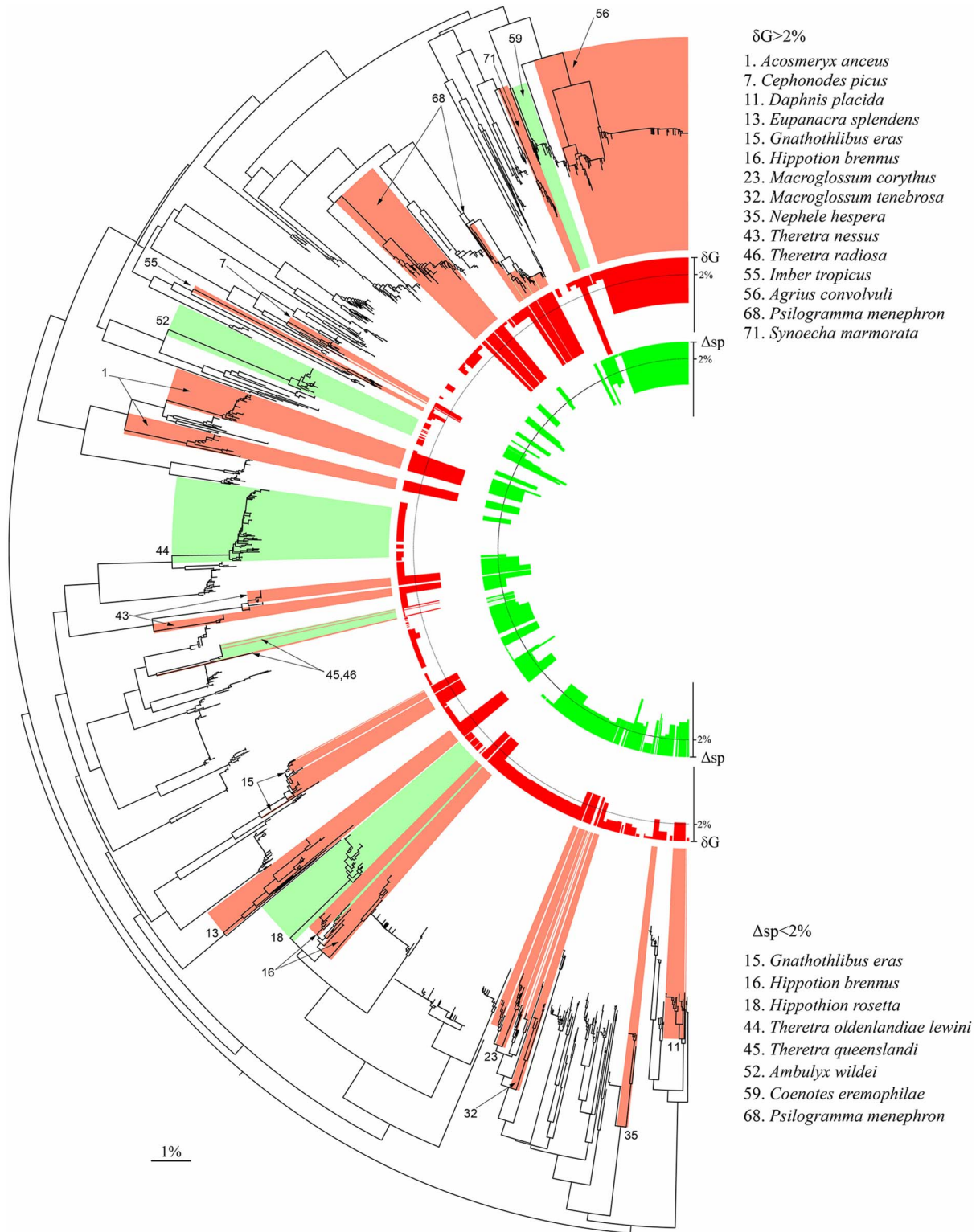


Figure 3. DNA barcode variation in the geographically extended dataset. Neighbour Joining tree based on K2P distances for the 1054 Australian sphingid records analysed in Fig. 2 augmented by 735 records for conspecifics, co-subspecifics and closely related species from outside Australia. In the centre of the tree, the inner (red) and outer (green) histograms represent the maximum intraspecific distance (δG) and the distance to the nearest heterospecific Neighbour on BOLD (Δsp). The species with $\delta G > 2\%$ and $\Delta sp < 2\%$ are highlighted in the tree with red and green colour ranges, respectively; those species are also listed with numbers following species numbers in Table 1. An interactive and fully explorable and searchable version of this tree can be accessed at <http://itol.embl.de/shared/rodroug>. (See also Fig. S5.). doi:10.1371/journal.pone.0101108.g003

(Fig. S2). Finally, both *M. corythus* and *P. menephron* represent species complexes requiring a comprehensive revision to establish the validity of the many available names; there are nine valid subspecies names [21] for the former, while we found that the sequenced holotypes of 10 species within genus *Psilogramma* fall within our set of DNA barcode clusters forming the *P. menephron* complex (Fig. S5).

In contrast to these cases of deep sequence divergence within species, the extended dataset revealed eight cases where Australian taxa showed less than 2% divergence (Table S5) from non-Australian specimens assigned to a different species (Fig. 3, Fig. S5, green boxes). Further analysis is required to evaluate the significance of these cases of similarity in DNA barcodes despite differing taxonomic assignments. Such cases can have several origins: identification errors, overlooked synonymy, incomplete lineage sorting or mitochondrial introgression through hybridization [47], or mitogenome replacement induced by *Wolbachia* infections [48]. Identification errors could not be ruled out to explain our results in the pair *Hippotion rosetta*/*H. boerhaviae*, two species long treated as one and often indistinguishable without genitalia dissection, and in the triplet *Gnathothlibus eras/saccoi/vanuatuensis*, which are difficult to distinguish from wing morphology. Cases of overlooked synonymy account for many of the taxa in the *Psilogramma menephron* complex, and probably also the triplet *Ambulyx wildei/ceramensis/rudloffii* (holotype sequenced) and possibly the pair *Coenotes eremophilae/jakli* (paratype sequenced). Finally, hybridization and associated introgression, or mitogenome replacement, may account for the barcode similarity between the morphologically very distinctive species pairs *Hippotion brennus/joiceyi* (Fig. S4) and *Theretra oldenlandiae/insignis* (Fig. S7). In certain cases such as the triplet *Theretra radiosa/queenslandi/muricolour*, barcode similarity may reflect the impact of two or more of these causal agents.

These analyses also indicated that Australian subspecies are reciprocally monophyletic for 16 of the 18 species examined (Table S5 and Fig. S5). In ten of these cases, the Australian subspecies diverges by more than 2% from its closest conspecific subspecies with the most distant pair being *Eupanacra splendens splendens/paradoxa* with a minimum pairwise distance of 5.96%. We found only one pair and one triplet of subspecies that share highly similar DNA barcodes: (1) *Theretra oldenlandiae oldenlandiae/lewini*, for which we found no shared haplotype, despite dense sampling (82 samples of *lewini* and 6 of *oldenlandiae*); and (2) *Theretra indistincta indistincta/manuselensis/papuensis*, a triplet possibly representing a case of synonymy.

Discussion

Identification and discrimination of Australian sphingid moths

The barcode reference library developed in this study provides coverage for 70 of the 72 currently valid Australian sphingid species and each of these species possesses a diagnostic array of DNA barcodes on the continent. Our results revealed seven species overlooked by past taxonomic work and five of these have recently been described, raising the species count for *Psilogramma* from five to seven, *Hopliocnema* from one to three, and *Coenotes* from one to two [23,41,42,44]. One species of *Acosmeryx* and *Hippotion johanna* need revalidation. These changes raise the species count for Australian sphingids to 75, a 17% increase from the last checklist for the family [22]. Our work also led to the unexpected discovery that the Convolvulus Hawkmoth, *Agrius convolvuli*, an emerging model species [49] is almost certainly two species. Its division into two species is relevant to all studies on this species (e.g. the

identification of its sex pheromones [50]) and may be of practical importance as it is a pest on sweet potatoes, a major crop in Papua New Guinea [51].

Extent of the Linnean shortfall

The Linnean shortfall [6] refers to the fact that most species on Earth are yet to be formally described, and implies that our current census of species (and other meaningful evolutionary units used by taxonomists, such as subspecies) is incomplete. In the case of Australian sphingids, our acquisition of DNA barcodes revealed seven overlooked species. While this result was perhaps unexpected for a fauna considered as taxonomically mature, its significance is twofold. On one hand, it must be noted that all of the newly revealed species are highly similar morphologically to previously known taxa, and thus the integration of DNA barcodes played a key role in circumscribing these species and revealing their distinctness. On the other hand, these discoveries also reflect insufficient sampling effort as most of the new species originate from areas in the outback, where limited collecting efforts have occurred. In addition, our results question the taxonomic status of 13 other species whose Australian lineages show marked sequence divergence from their counterparts outside Australia. Finally, eight other Australian species show close barcode congruence with lineages in other settings that are assigned to different species (Fig. 3). In total, 18 species (three show both splitting and lumping) – almost a quarter of the continental fauna – possess a taxonomic assignment that requires further study. The resolution of each case will require a thorough and critical revision of species diagnoses using an integrative approach, and will demand consideration of names currently listed as junior synonyms. Tackling the nomenclatural part of this work is a particular challenge because sphingids have attracted so much attention from taxonomists resulting, for example, in 14 synonyms for *A. convolvuli* [21]. Fortunately, recent progress in the recovery of whole or partial DNA barcodes from old type specimens will aid resolution of these nomenclatural conundrums [32,52].

Extent of the Wallacean shortfall

Knowledge of species distributions is fundamental to the design of sound conservation strategies and gaps in such information have been termed the “Wallacean shortfall” by biogeographers and conservation biologists [6]. Besides the scarcity of distributional data for most species, incorrect or incomplete delineation of taxonomic units will, in most cases, lead to incorrect information on species distributions, impeding efforts to understand the historical and contemporaneous processes underlying the occurrence patterns of species or subspecies [6,53]. This study has revealed a significant mismatch between past understanding of the taxonomic diversity of Australian sphingids and that revealed through molecular analysis, especially when results are placed in a broader biogeographical context. The recent description of two Australian species each of *Psilogramma* and *Hopliocnema* [23,41,42], of one new species of *Coenotes* [44], and the separation by Vaglia & Haxaire [54] of the Papuan *A. miskinooides* from its Australian counterpart, *A. miskini*, were all motivated by DNA barcode data coupled subsequently with morphological analysis. These updates raise the number of endemic Australian species and subspecies from 26 to 34 (Fig. 4, Table S4). Further analysis of *N. hespera* may add another endemic species, while the possible synonymy of subspecies of *T. indistincta* may change its status from one polytypic species with an endemic Australian subspecies to a monotypic species distributed over Australia and the Malesian region.

Our results additionally revealed that seven species with broad ranges (namely *Cephonodes picus*, *Macroglossum corythus*, *Theretra nessus*,

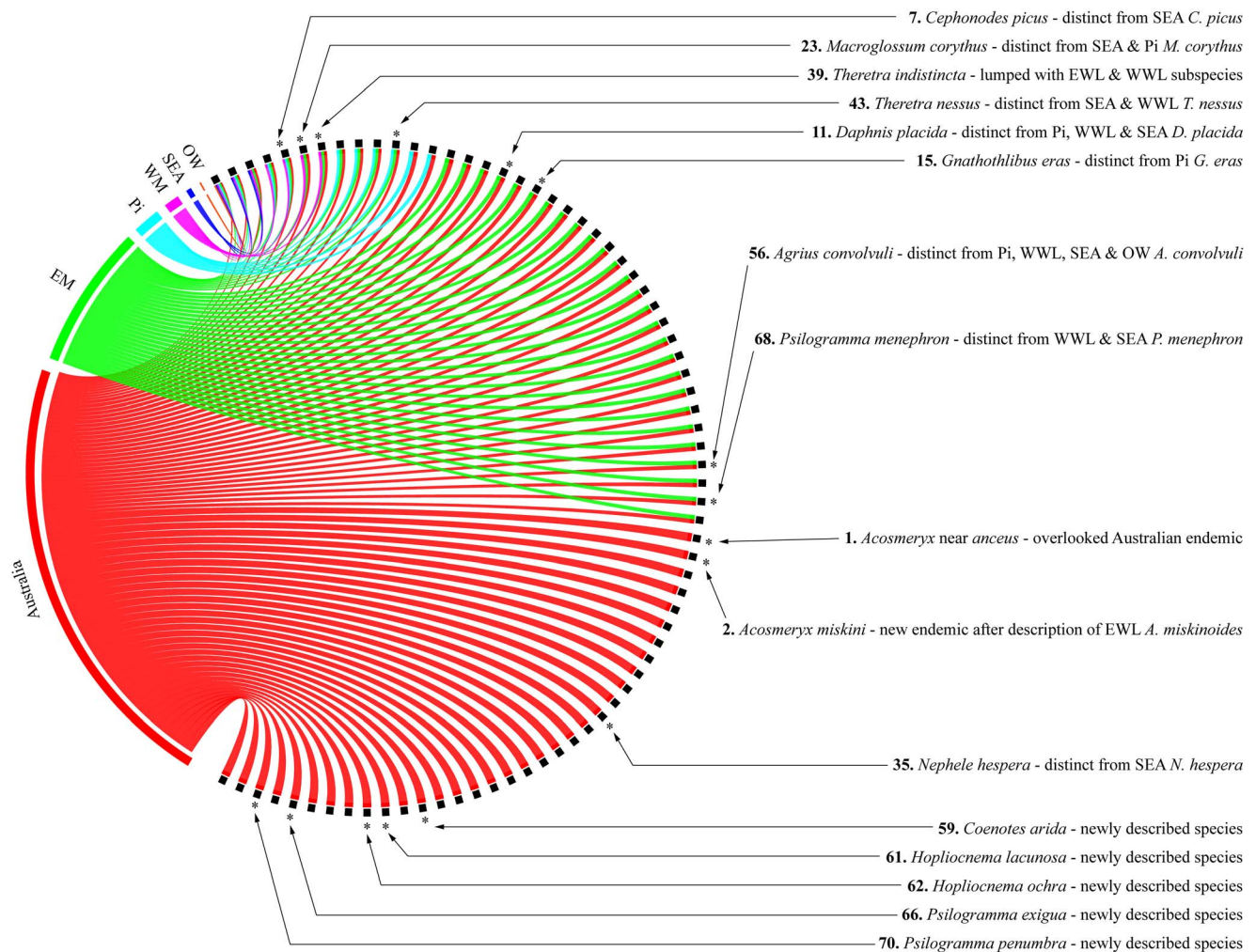


Figure 4. Distribution of Australian sphingid species. All 75 recognized species are linked by one or more coloured ribbons to the boundaries of their distributions (abbreviations as in Table 1). Species are ordered by distribution type, from the Australian endemics at the bottom left toward taxa with broader ranges at the top. Stars mark all species whose distributions were altered by our work, while text provides details about the cause of their distributional shift. The figure was assembled with the online application of Circos [57]. doi:10.1371/journal.pone.0101108.g004

Daphnis placida, *Nephela hespera*, *Agrius convolvuli*, and *Psilogramma menephron*; see Fig. 4, and Fig. S5) actually comprise two (or more) genetically differentiated lineages likely to represent distinct species. In fact, in all of these cases, the genetic split is associated to a geographical disjunction. Although our current sampling of specimens with a DNA barcode remains too sparse in the region to precisely delimit the ranges of these newly circumscribed units, it is worth noting that in all cases the genetic lineage including the Australian representatives never extend westward beyond the Wallacean region. This result fits previous observations of the region as representing a major faunal discontinuity for Malesian sphingids, although it remains unclear where precisely the disjunctions occur and whether they match the modified course of Wallace's Line proposed by Beck *et al.* [55] for sphingids. Overall, our results indicate that the Australian sphingid fauna comprises two large biogeographical subsets (Fig. 4), one endemic to the continent (30 species and 4 subspecies, 45% of the fauna) and one including taxa also occurring on Pacific Islands and/or in the Papuan and Wallacean region (34 species/subspecies, 45%). A third small subset includes just seven Australian species (10%) whose distributions extend westward beyond the Wallacean region

(Fig. 4). The situation of those taxa found in Australia and on Pacific Islands deserves further attention. It is, for instance, interesting that the two species found in Australia and in the Pacific, but not in Papua New Guinea, occur in New Caledonia (*N. subvaria*) or New Caledonia and Fiji (*Hippotion scrofa*) but not in the Solomon Islands. In contrast, the three Australian species occurring on the Solomon Islands (*D. moorei*, *T. nessus* and *E. splendens*) are all also present on New Guinea. Overall, among the 11 Australian species or subspecies that also occur on Pacific islands (Table 1; Fig. 4), eight were sampled from these islands and six revealed genetic divergences ranging from 1 to 2% (e.g. 1% for *D. moorei* in Solomon Islands, 1.5% for *H. scrofa* in New Caledonia and Fiji, 1.4% for *A. convolvuli* in New Caledonia), while two (*H. velox* and *T. silhetensis intersecta*) were genetically undifferentiated. These relatively low divergences likely reflect the recent history of colonization of these islands from Australia or Papua New Guinea during the Pleistocene, when periods of lowered sea level facilitated dispersal across sea gaps. Climatic events or human-mediated colonization might also have caused contemporaneous range expansion in Pacific islands; *H. scrofa* was for instance unrecorded from Fiji before 1975 despite sustained collecting

efforts before then (G. Robinson, personal communication), and the genetic identity of the specimen sampled here with one from New Caledonia (Fig. S5) supports this hypothesis of a recent colonization event.

The taxonomic implications of these genetic divergences between allopatric populations require further examination, but these island “populations” certainly represent important evolutionary units, yet another layer of the Wallacean shortfall concealed by their current taxonomic treatment. We did examine some subspecies already described from Pacific islands and they show comparable levels of genetic divergence (e.g. 1.9% between *D. p. placida* and *D. p. salomonis* from the Solomon Islands) or even less (e.g. only 0.6% between *T. n. nessus* and *T. n. albata* from New Caledonia and Vanuatu). Judging from genetic evidence alone, the island lineages of some Australian species may deserve at least subspecific status, and confirmation by morphology should be sought.

Conclusions

The Linnean and Wallacean shortfalls have been stressed as two critical gaps in our knowledge of global biodiversity. Overlooked cryptic diversity and inadequate delineation of species boundaries inevitably create erroneous distribution patterns impeding the development of conservation strategies that focus on the right objects and the right places. In groups that have received little taxonomic attention, such as tropical and small-bodied invertebrates, these shortfalls have been recognized as substantial [9,56]. However, this study reveals that a taxonomically mature insect fauna can also be strongly impacted by these shortfalls. Here, we have illustrated a solution involving the integration of DNA barcodes with an established taxonomic system based largely on morphology. This approach provides a robust means to identify and correct possible discrepancies in our current understanding of species and their distributions.

Supporting Information

Figure S1 NJ phylogram for Australian sphingid records.
(PDF)

Figure S2 28S–D2 rDNA sequence alignments.
(PDF)

Figure S3 Overlooked diversity within *Acosmeryx anceus*.
(PDF)

References

- Riddle BR, Ladle RJ, Lourie SA, Whittaker RJ (2011) Basic biogeography: estimating biodiversity and mapping nature. In: Ladle RJ, Whittaker RJ, editors. *Conservation Biogeography*. Chichester, UK: Wiley-Blackwell. pp. 47–92.
- Austin AD, Yeates DK, Cassis G, Fletcher MJ, La Salle J, et al. (2004) Insects ‘Down Under’ - Diversity, endemism and evolution of the Australian insect fauna: examples from select orders. *Aust J Entomol* 43: 216–234.
- Williams J, Read C, Norton A, Dovers S, Burgman M, et al. (2001) Biodiversity, Australia State of the Environment Report 2001 (Theme Report). Canberra, ACT: CSIRO. 227 p.
- Cranston PS (2010) Insect biodiversity and conservation in Australasia. *Annu Rev Entomol* 55: 55–75.
- Chapman AD (2009) Numbers of living species in Australia and the world. Canberra, ACT: Australian Biological Resources Study (ABRS). 80 p.
- Lomolino MV (2004) Frontiers of biogeography: new directions in the geography of nature. In: Lomolino MV, Heaney LR, editors. *Conservation Biogeography*. Sunderland, MA: Sinauer Ass. pp. 293–296.
- Riedel A, Sagata K, Suhardjono YR, Tanzler R, Balke M (2013) Integrative taxonomy on the fast track - towards more sustainability in biodiversity research. *Front Zool* 10: 15.
- Goldstein PZ, DeSalle R (2011) Integrating DNA barcode data and taxonomic practice: determination, discovery, and description. *BioEssays* 33: 135–147.
- Scheffers BR, Joppa LN, Pimm SL, Laurance WF (2012) What we know and don't know about Earth's missing biodiversity. *Trends Ecol Evol* 27: 501–510.
- Hebert PDN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. *Syst Biol* 54: 852–859.
- Oliver PM, Adams M, Lee MS, Hutchinson MN, Doughty P (2009) Cryptic diversity in vertebrates: molecular data double estimates of species diversity in a radiation of Australian lizards (*Diplodactylus*, Gekkota). *Proc Roy Soc B-Biol Sci* 276: 2001–2007.
- Marin J, Donnellan SC, Hedges SB, Puillandre N, Aplin KP, et al. (2013) Hidden species diversity of Australian burrowing snakes (*Ramphotyphlops*). *Biol J Linn Soc* 110: 427–441.
- Sharma P, Kotov AA (2013) Molecular approach to identify sibling species of the *Ceriodaphnia cornuta* (Cladocera: Daphniidae) from Australia with notes on the continental endemism of this group. *Zootaxa* 3702: 79–89.
- Landry JF, Hebert PDN (2013) *Plutella australiana* (Lepidoptera, Plutellidae), an overlooked diamondback moth revealed by DNA barcodes. *Zookeys* 327: 43–63.

Figure S4 Consensus tree (MP) for records of *Hippotion brennus* and *H. joiceyi*.
(PDF)

Figure S5 NJ phylogram for Australian and non-Australian records.
(PDF)

Figure S6 Geographical structure of genetic variation in the *Convolvulus* Hawkmoth.
(PDF)

Figure S7 Consensus tree (MP) for records within the *Theretra oldenlandiae/insignis* complex.
(PDF)

Table S1 List of Australian records.
(PDF)

Table S2 List of conspecific records from outside Australia.
(PDF)

Table S3 List of heterospecific records from outside Australia.
(PDF)

Table S4 List of endemic species and subspecies of Sphingidae in Australia.
(PDF)

Table S5 DNA barcode matches at subspecies and species levels.
(PDF)

Appendix S1 Sequencing methods.
(DOCX)

Acknowledgments

This work was enabled by the generosity of the many collaborators involved in the DNA barcoding campaign on Lepidoptera. We are especially grateful to Ron Brechlin, Graeme Cocks, Ted Edwards, Ulf Eitschberger, Yves Estradel, Egbert Friedrich, Marianne Horak, David Lane, John La Salle, Ian McMillan, Tomas Melichar, Andrew Mitchell, Max Moulds, Jim Tuttle, and Thierry Vaglia. We also thank colleagues at the Canadian Centre for DNA Barcoding.

Author Contributions

Conceived and designed the experiments: RR PDNH. Performed the experiments: RR JH SEM AH PDNH. Analyzed the data: RR IJK JH PDNH. Contributed reagents/materials/analysis tools: RR JH SEM AH PDNH. Contributed to the writing of the manuscript: RR IJK PDNH.

15. Hausmann A, Hebert PDN, Mitchell A, Rougerie R, Sommerer M, et al. (2009) Revision of the Australian *Oenochroma vinaria* Guenée, 1858 species-complex (Lepidoptera: Geometridae, Oenochrominae): DNA barcoding reveals cryptic diversity and assesses status of type specimen without dissection. *Zootaxa* 2239: 1–21.
16. Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System. Available: <http://www.barcodinglife.org>. *Mol Ecol Notes* 7: 355–364.
17. Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: the Barcode Index Number (BIN) System. *PLoS ONE* 8: e66213.
18. Hebert PD, Dewaard JR, Zakharov EV, Prosser SW, Sonens JE, et al. (2013) A DNA 'barcode blitz': rapid digitization and sequencing of a natural history collection. *PLoS ONE* 8: e68535.
19. Nieukerken EJV, Kaila L, Kitching IJ, Kristensen NP, Lees DC, et al. (2011) Order Lepidoptera. *Zootaxa* 3148: 212–221.
20. Kristensen NP, Scoble MJ, Karsholt O (2007) Lepidoptera phylogeny and systematics: the state of inventoring moth and butterfly diversity. *Zootaxa* 1668: 699–747.
21. Kitching IJ, Cadiou JM (2000) Hawkmoths of the World. An annotated and illustrated revisionary checklist. Ithaca: The Natural History Museum, London & Cornell University Press. 227 p.
22. Moulds MS (1996) Sphingidae. In: Nielsen ES, Edwards ED, Rangsi TV, editors. Checklist of the Lepidoptera of Australia. Canberra, Australia: CSIRO. 529.
23. Brechlin R, Kitching IJ (2010) Four new species of the genus *Psilogamma* Rothschild & Jordan, 1903 (Lepidoptera, Sphingidae). *Entomo-Satsphingia* 3 (3): 66–74.
24. Brechlin R (2001) Einige Anmerkungen zur Gattung *Psilogamma* Rothschild & Jordan, 1903 mit Beschreibung neuer Arten (Lepidoptera, Sphingidae). *Arthropoda* 9: 6–47.
25. Eitschberger U (2001) Revision der Gattung *Psilogamma* Rothschild & Jordan, 1903 mit der Beschreibung weiterer neuer Arten (Lepidoptera, Sphingidae). *Neue Entomologische Nachrichten* 52: 3–174.
26. Holloway JD (2011) The moths of Borneo: families Phaulidae, Himanopteridae and Zygaenidae; revised and annotated checklist. *Malayan Nature Journal* 63: 1–548.
27. May RM, Harvey PH (2009) Species Uncertainties. *Science* 323: 687.
28. Godfray HCJ (2006) To boldly sequence. *Trends Ecol Evol* 21: 603–604.
29. Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proc Roy Soc B-Biol Sci* 270: 313–321.
30. James SW, Porco D, Decaens T, Richard B, Rougerie R, et al. (2010) DNA barcoding reveals cryptic diversity in *Lumbricus terrestris* L., 1758 (Clitellata): resurrection of *L. herculeus* (Savigny, 1826). *PLoS ONE* 5: e15629.
31. Padial JM, De la Riva I (2007) Integrative taxonomists should use and produce DNA barcodes. *Zootaxa* 1586: 67–68.
32. Rougerie R, Naumann S, Nässig WA (2012) Morphology and molecules reveal unexpected cryptic diversity in the enigmatic genus *Sinobirna* Bryk, 1944 (Lepidoptera: Saturniidae). *PLoS ONE* 7: e43920.
33. Kimura M (1980) A simple method for estimating evolutionary rate of base substitution through comparative studies of nucleotide sequences. *J Mol Evol* 16: 111–120.
34. Letunic I, Bork P (2007) Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics* 23: 127–128.
35. Parks DH, Mankowski T, Zangooei S, Porter MS, Armanini DG, et al. (2013) GenGIS 2: Geospatial analysis of traditional and genetic biodiversity, with new gradient algorithms and an extensible plugin framework. *PLoS ONE* e69885.
36. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Mol Biol Evol* 24: 1596–1599.
37. Hebert PDN, Ratnasingham S, DeWaard JR (2003) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Biol Letters* 270: S96–S99.
38. Mutanen M, Hausmann A, Hebert PDN, Landry JF, DeWaard JR, et al. (2012) Allopatry as a Gordian knot for taxonomists: patterns of DNA barcode divergence in Arctic-Alpine Lepidoptera. *PLoS ONE* 7: e47214.
39. Gehring PS, Tolley KA, Eckhardt FS, Townsend TM, Ziegler T, et al. (2012) Hiding deep in the trees: discovery of divergent mitochondrial lineages in Malagasy chameleons of the *Calumma nasutum* group. *Ecol Evol* 2: 1468–1479.
40. Vieites DR, Wollenberg KC, Andreone F, Kohler J, Glaw F, et al. (2009) Vast underestimation of Madagascar's biodiversity evidenced by an integrative amphibian inventory. *Proc Natl Acad Sci USA* 106: 8267–8272.
41. Lane D, Moulds MS, Tuttle JP (2011) A new species of *Psilogamma* Rothschild & Jordan, 1903 from northern Australia (Lepidoptera: Sphingidae). *NEVA (NF)* 31: 227–230.
42. Tuttle JP, Moulds MS, Lane DA (2012) A new species of the hawkmoth genus *Hoplocnema* Rothschild and Jordan, with the description of two new species (Lepidoptera: Sphingidae). *The European Entomologist* 1: 65–77.
43. Hausmann A, Haszprunar G, Hebert PDN (2011) DNA Barcoding the geometrid fauna of Bavaria (Lepidoptera): successes, surprises, and questions. *PLoS ONE* 6: e17134.
44. Moulds MS, Melichar T (2014) A new species of hawk moth in the genus *Coenotes* Rothschild & Jordan from Australia (Sphingidae: Smerinthinae: Sphingulini). *The European Entomologist* 5: 11–17.
45. Darge P (1975) Note sur *Hippotion brunus* (Lep., Sphingidae). *Bull Soc Ent Fr* 80: 179–180.
46. Riotte JCE (1979) Zur Gattungszugehörigkeit von "*Theretia*" *brunnea* (Lep.: Sphingidae). *Entomologische Zeitschrift* 89: 225–230.
47. Rougerie R, Haxaire J, Kitching IJ, Hebert PDN (2012) DNA barcodes and morphology reveal a hybrid hawkmoth in Tahiti (Lepidoptera: Sphingidae). *Inv Syst* 26: 445.
48. Smith MA, Bertrand C, Crosby K, Eveleigh ES, Fernandez-Triana J, et al. (2012) *Wolbachia* and DNA Barcoding Insects: Patterns, Potential, and Problems. *PLoS ONE* 7: e36514.
49. Roe AD, Weller SJ, Baixeras J, Brown J, Cummings MP, et al. (2010) Evolutionary framework for Lepidoptera model systems. In: Goldsmith MR, Marek F, editors. *Molecular Biology and Genetics of the Lepidoptera*. Boca Raton, FL, USA: CRC Press. 1–24.
50. Wakamura S, Yasuda T, Wanatabe M, Kiguchi K, Shimoda M, et al. (1996) Sex pheromone of the sweetpotato hornworm, *Agrius convolvuli* (L.) (Lepidoptera: Sphingidae): identification of a major component and its activity in a wind tunnel. *Appl Entomol Zool* 31: 171–174.
51. Peters J (2001) Local human-sweet potato-pig systems characterization and research in Irian Jaya, Indonesia. With limited reference to Papua New Guinea. International Potato Center (CIP), Lima, Peru. 79 p.
52. Lees DC, Rougerie R, Zeller-Lukashort C, Kristensen NP (2010) DNA mini-barcode in taxonomic assignment: a morphologically unique new homoneurous moth clade from the Indian Himalayas described in *Micropterix* (Lepidoptera, Micropterigidae). *Zool Scr* 39: 642–661.
53. Romero D, Olivero J, Márquez AL, Báez JC, Real R (2014) Uncertainty in distribution forecasts caused by taxonomic ambiguity under climate change scenarios: a case study with two newt species in mainland Spain. *J Biogeogr* 41: 111–121.
54. Vaglia T, Haxaire J (2007) Un nouveau Sphingidae du genre *Acosmeryx* Boisduval, 1875: *Acosmeryx miskinoïdes* sp. n. (Lepidoptera, Sphingidae). *The European Entomologist* 1: 7–13.
55. Beck J, Kitching IJ, Linsenmair KE (2006) Wallace's line revisited: has vicariance or dispersal shaped the distribution of Malesian hawkmoths (Lepidoptera: Sphingidae)? *Biol J Linn Soc* 89: 455–468.
56. Lees DC, Kawahara AY, Rougerie R, Oshima I, Kawakita A, et al. (2014) DNA barcoding reveals a largely unknown fauna of Gracillariidae leaf mining moths in the Neotropics. *Mol Ecol Resour* 14: 286–296.
57. Krzywinski MI, Schein JE, Birol I, Connors J, Gascoyne R, et al. (2009) Circos: An information aesthetic for comparative genomics. *Genome Res* 19: 1639–1645.