Molecular Phylogenetics and Evolution 120 (2018) 129–143

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

A preliminary molecular phylogeny of shield-bearer moths (Lepidoptera: Adeloidea: Heliozelidae) highlights rich undescribed diversity

Liz Milla[a,1,*], Erik J. van Nieukerken[b,1], Ruben Vijverberg[b], Camiel Doorenweerd[b], Stephen A. Wilcox[a], Mike Halsey[d], David A. Young[c], Therésa M. Jones[a], Axel Kallies[a], Douglas J. Hilton[a]

[a] School of BioSciences, The University of Melbourne, Parkville, Victoria 3010, Australia
[b] Naturalis Biodiversity Center, PO Box 9517, 2300 RA Leiden, The Netherlands
[c] D'Estrees Entomology & Science Services, Kingscote 5223, Australia
[d] Faculty of Health and Life Sciences, Oxford Brookes University, England, UK

ARTICLE INFO

Keywords:
Lepidoptera
Multilocus phylogeny
Taxonomy
Family-level phylogeny
Australia

ABSTRACT

Heliozelidae are a widespread, evolutionarily early diverging family of small, day-flying monotrysian moths, for which a comprehensive phylogeny is lacking. We generated the first molecular phylogeny of the family using DNA sequences of two mitochondrial genes (COI and COII) and two nuclear genes (H3 and 28S) from 130 Heliozelidae specimens, including eight of the twelve known genera: *Antispila*, *Antispilina*, *Coptodisca*, *Heliozela*, *Holocacista*, *Hoplophanes*, *Pseliastis*, and *Tyriozela*. Our results provide strong support for five major Heliozelidae clades: (i) a large widespread clade containing the leaf-mining genera *Antispilina*, *Coptodisca* and *Holocacista* and some species of *Antispila*, (ii) a clade containing most of the described *Antispila*, (iii) a clade containing the leaf-mining genus *Heliozela* and the monotypic genus *Tyriozela*, (iv) an Australian clade containing *Pseliastis* and (v) an Australian clade containing *Hoplophanes*. Each clade includes several new species and potentially new genera. Collectively, our data uncover a rich and undescribed diversity that appears to be especially prevalent in Australia. Our work highlights the need for a major taxonomic revision of the family and for generating a robust molecular phylogeny using multi-gene approaches in order to resolve the relationships among clades.

1. Introduction

Heliozelidae are an evolutionarily early diverging family of small, day-flying monotrysian moths, found on all continents except Antarctica. Worldwide, there are twelve Heliozelidae genera comprising 125 described species (van Nieukerken et al., 2011, 2012; van Nieukerken and Geertsema, 2015). The four most speciose genera, *Antispila*, *Coptodisca*, *Heliozela*, and *Hoplophanes*, contain over 90% of the described species, with the highest described species diversity found in North America and Australia (van Nieukerken et al., 2012).

Most described Heliozelidae larvae are leaf miners of trees and vines, while a few species are known to mine petioles, midribs, twigs or initiate galls (Davis, 1998). Additionally, flower and seed mining appears to be prevalent in many Australian species (our unpublished observations). Leaf-mining heliozelids cut distinctive shield-shaped cases from the leaf surface, which they carry to the ground to pupate, leaving behind a characteristic pattern of scattered holes – hence the term "shield-bearers" used to describe the family. Some Heliozelidae species are well known pests of important commercial crops, notably vines, cranberry and walnut (Maier, 1988; van Nieukerken et al, 2012; van Nieukerken and Geertsema, 2015; Bernardo et al., 2015). However, the current lack of comprehensive taxonomic, molecular and ecological data has hindered adequate species identification, as shown by the discovery of an invasive heliozelid species from North America on Italian walnut trees (Bernardo et al., 2012). Overall, fewer than twenty species have been described in the last 50 years, mainly from Japan and the Americas (Opler, 1971; Lafontaine, 1974; Kuroko, 1982; Karsholt and Kristensen, 2003; Lee et al., 2006a, 2006b; van Nieukerken et al., 2012; Lee and Hirowatari, 2013; van Nieukerken and Geertsema, 2015).

Heliozelidae, together with Adelidae, Incurvariidae, Cecidosidae, and Prodoxidae, comprise the primitive superfamily Adeloidea (van Nieukerken et al., 2011; Regier et al., 2015). Heliozelidae are estimated to have diverged from their putative sister family Adelidae ("fairy" or

* Corresponding author.
E-mail address: emilla@student.unimelb.edu.au (L. Milla).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.ympev.2017.12.004
Received 4 June 2017; Received in revised form 24 November 2017; Accepted 4 December 2017
Available online 08 December 2017

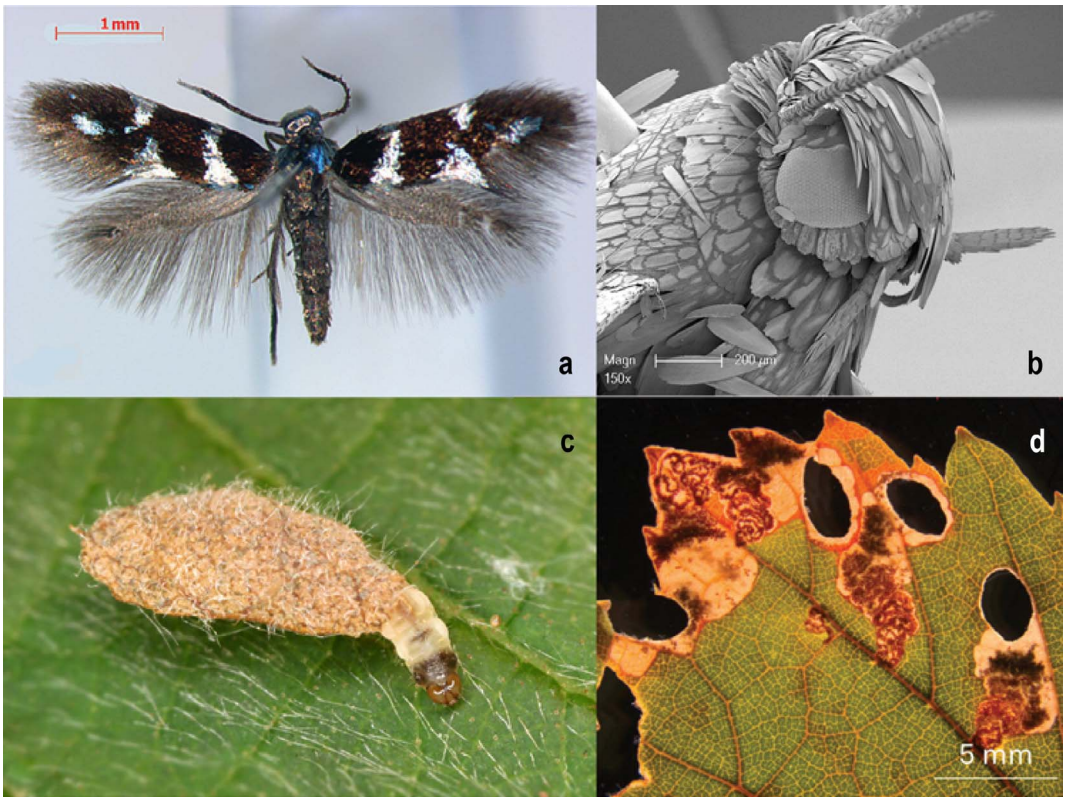

Fig. 1. (a) Adult habitus of *Antispila hydrangaeella*, female (Photo: E.J. van Nieukerken); (b) Head close up of undescribed heliozelid species showing flattened scales (SEM: Q. Wang); (c) Fully grown *Heliozela resplendella* larva with shield, mine in *Alnus incana* (Photo: R. Bryner); (d) Characteristic heliozelid leaf mines, produced by *Holocacista capensis* (Photo: E.J. van Nieukerken).

"longhorn" moths) approximately 80 million years ago (Wahlberg et al., 2013). A number of taxonomic and multi-gene molecular studies of Lepidoptera have proposed the Heliozelidae family as monophyletic (Nielsen, 1980; Friedlander et al., 2000; Wahlberg et al., 2013; Regier et al., 2015; Fig. 2, nt123 analyses). Synapomorphies for Heliozelidae include dorsally curved anterior tentorial arms, the lack of a hindwing M-Cua crossvein and minute mandibles (Nielsen and Davis, 1985). More generally, Heliozelidae can be distinguished from the other Adeloidea by their shiny, overlapping, lamellar head scales (Fig. 1), a characteristic found in all genera except *Plesiozela*, the putative sister group to all other Heliozelidae (Karsholt and Kristensen, 2003).

Molecular phylogenies of Heliozelidae have been constructed from COI sequences (van Nieukerken et al., 2012; Bernardo et al., 2015; van Nieukerken and Geertsema, 2015). However, these analyses were mostly limited to northern hemisphere species and comparable phylogenies for southern hemisphere species are lacking. Thus, a substantial gap remains in our knowledge of the family. In order to fill this gap, over the last decade we have systematically collected Heliozelidae specimens from areas known or predicted to harbour high heliozelid diversity (Common, 1990; Heppner, 1991), focusing on the Palaearctic and Nearctic regions and southern Australia. Here, we use sequence data obtained for two mitochondrial (COI and COII) and two nuclear genes (H3 and 28S) from representatives of eight of the twelve described Heliozelidae genera to generate a preliminary molecular phylogeny of the Heliozelidae. We have included representatives from the most diverse groups, the widespread genera *Antispila* and *Heliozela*, the Nearctic genus *Coptodisca*, and the Australian endemic genus *Hoplophanes*. We also included representatives of another Australian endemic genus, *Pseliastis*, the widespread genus *Holocacista*, and two monotypic genera, *Antispilina* and *Tyriozela*. Based on our phylogeny, we propose five major monophyletic clades within Heliozelidae and discuss the monophyly of the genera they contain.

2. Materials and methods

2.1. Ingroup selection

A total of 130 specimens belonging to eight Heliozelidae genera were selected for sequencing. These specimens represented a total of 79 species within the currently described genera: 20 described and eight putative species of *Antispila*, one described and one putative species of *Antispilina*, seven described and four putative species of *Coptodisca*, four described and 11 putative species of *Heliozela*, three described and nine putative species of *Holocacista*, one described and four putative species of *Hoplophanes*, two described and three putative species of *Pseliastis*, and one described species of *Tyriozela*. We included an additional 15 putative species, some of which may be placed in potentially new genera. Full names and authorities for all sampled taxa are given on Table 1. Identifiers for putative species were formed by a combination of the genus name and an "epithet" formed by hostplant genus (or unknown), with country or region of origin. New or unknown genera are indicated as "*heliozelidgenus*". We were unable to obtain suitable material from four other heliozelid genera: *Plesiozela* from the Patagonian region and putative sister-group to all other Heliozelidae (Karsholt and Kristensen, 2003), and the small genera *Ischnocanaba* from Solomon Islands, *Phanerozela* from Brazil and *Microplitica* from India and Indonesia. Similarly, we lacked specimens of *Lamprozela* from Guyana, which was originally allocated to Heliozelidae, but removed by Nielsen (1980) as possible Heliodinidae based on taxonomic re-examination. For DNA extraction material, we used larvae from collected leaf mines and adult specimens either from existing museum collections, collected from the field using sweep nets or reared from late instars. Detailed methods for collecting and rearing heliozelids have been published elsewhere (van Nieukerken et al., 2012, Bernardo et al., 2015; van Nieukerken and Geertsema, 2015). All specimen data with their COI sequences are provided in the BOLD dataset DS-HELIPHYL (https://doi.

Table 1
Species names with authority, voucher codes (specimen ID), country of origin, biogeographic region, host family, sex (where known), life stage and collecting method for adult specimens used in this study.

Species and author	Voucher	Country	Region	Host Family	Sex, stage, adult collection method (el = ex larva, oh = on host, dc = day collecting)
***Antispila* Hübner, 1825**					
Antispila ampelopsia Kuroko, 1961	RMNH.INS.30326	Japan	East Palaearctic	Vitaceae	Larva
Antispila ampelopsifoliella Chambers, 1874	RMNH.INS.18588	United States	Nearctic	Vitaceae	Larva
Antispila ampelopsifoliella Chambers, 1874	RMNH.INS.24377	United States	Nearctic	Vitaceae	Male, adult, el
Antispila Ampelopsis.brevTaiwan	RMNH.INS.29431	Taiwan	East Palaearctic	Vitaceae	Larva
Antispila Ampelopsis.brevTaiwan	RMNH.INS.24365	Taiwan	East Palaearctic	Vitaceae	Male, adult, el
Antispila Ampelopsis.cantTaiwan	RMNH.INS.29436	Taiwan	East Palaearctic	Vitaceae	Larva
Antispila Ampelopsis.cantTaiwan	RMNH.INS.24364	Taiwan	East Palaearctic	Vitaceae	Male, adult, el
Antispila argentifera Braun, 1927	RMNH.INS.18566	United States	Nearctic	Myricaceae	Larva
Antispila cleyerella Lee, 2006	RMNH.INS.24532	Japan	East Palaearctic	Pentaphylacaceae	Female, adult, el
Antispila corniella Kuroko, 1961	RMNH.INS.24535	Japan	Nearctic	Cornaceae	Male, adult
Antispila cornifoliella Clemens, 1860	RMNH.INS.24366	United States	Nearctic	Cornaceae	Male, adult, el
Antispila cornifoliella Clemens, 1860	RMNH.INS.18270	United States	Nearctic	Cornaceae	Larva
Antispila CornusJapan	RMNH.INS.29809	Japan	East Palaearctic	Cornaceae	Larva
Antispila CornusJapan	RMNH.INS.29810	Japan	East Palaearctic	Cornaceae	Larva
Antispila distyliella Lee, 2006	RMNH.INS.24537	Japan	East Palaearctic	Hamamelidaceae	Female, adult, el
Antispila hikosana Kuroko, 1961	RMNH.INS.24540	Japan	East Palaearctic	Cornaceae	Female, adult
Antispila hydrangaeella Chambers, 1874	RMNH.INS.18416	United States	Nearctic	Hydrangeaceae	Larva
Antispila hydrangaeella Chambers, 1874	RMNH.INS.18206	United States	Nearctic	Hydrangeaceae	Larva
Antispila Hydrangea.chinensisTaiwan	RMNH.INS.29423	Taiwan	East Palaearctic	Hydrangeaceae	Larva
Antispila HydrangeaTaiwan	RMNH.INS.29377	Taiwan	East Palaearctic	Hydrangeaceae	Larva
Antispila isabella Clemens, 1860	RMNH.INS.18504	United States	Nearctic	Vitaceae	Larva
Antispila isabella Clemens, 1860	RMNH.INS.18127	United States	Nearctic	Vitaceae	Larva
Antispila metallella (Denis and Schiffermüller, 1775)	RMNH.INS.24371	Netherlands	West Palaearctic	Cornaceae	Male, adult, el
Antispila nysaefoliella Clemens, 1860	RMNH.INS.29113	United States	Nearctic	Cornaceae	Larva
Antispila nysaefoliella Clemens, 1860	RMNH.INS.18303	United States	Nearctic	Cornaceae	Larva
Antispila oinophylla van Nieukerken & Wagner, 2012	RMNH.INS.18643	United States	Nearctic	Vitaceae	Larva
Antispila oinophylla van Nieukerken & Wagner, 2012	RMNH.INS.23920	Italy	West Palaearctic[a]	Vitaceae	Male, adult, el
Antispila petryi Martini, 1898	RMNH.INS.24372	Greece	West Palaearctic	Cornaceae	Male, adult, el
Antispila petryi Martini, 1898	RMNH.INS.29304	Netherlands	West Palaearctic	Cornaceae	Larva
Antispila purplella Kuroko, 1961	RMNH.INS.24529	Japan	East Palaearctic	Cornaceae	Male
Antispila tateshinensis Kuroko, 1987	RMNH.INS.29735	Japan	East Palaearctic	Vitaceae	Larva
Antispila tateshinensis Kuroko, 1987	RMNH.INS.29823	Japan	East Palaearctic	Vitaceae	Larva
Antispila treitschkiella (Fischer von Röslerstamm, 1843)	RMNH.INS.11856	Netherlands	East Palaearctic	Cornaceae	Larva
Antispila treitschkiella (Fischer von Röslerstamm, 1843)	RMNH.INS.29300	Netherlands	West Palaearctic	Cornaceae	Larva
Antispila uenoi Kuroko, 1987	RMNH.INS.29237	Korea	East Palaearctic	Vitaceae	Larva
Antispila viticordifoliella Clemens, 1860	RMNH.INS.24207	United States	Nearctic	Vitaceae	Female, adult, el
Antispila viticordifoliella Clemens, 1860	RMNH.INS.18508	United States	Nearctic	Vitaceae	Larva
Antispila Vitis1USA	RMNH.INS.18126	United States	Nearctic	Vitaceae	Larva
Antispila Vitis1USA	RMNH.INS.24205	United States	Nearctic	Vitaceae	Female, adult, el
Antispila Vitis2USA	RMNH.INS.18656	United States	Nearctic	Vitaceae	Larva
Antispila Vitis2USA	RMNH.INS.18131	United States	Nearctic	Vitaceae	Larva
Antispila VitisKorea	RMNH.INS.29232	Korea	East Palaearctic	Vitaceae	Larva
Antispila voraginella Braun, 1927	RMNH.INS.23917	United States	Nearctic	Vitaceae	Male, adult, el
***Antispilina* Hering, 1941**					
Antispilina ludwigi Hering, 1941	RMNH.INS.24677	France	West Palaearctic	Polygonaceae	Female, adult, el
Antispilina ludwigi Hering, 1941	RMNH.INS.11853	Poland	West Palaearctic	Polygonaceae	Larva
Antispilina PersicariaVietnam	RMNH.INS.24363	Vietnam	Oriental	Polygonaceae	Male, adult, el
Antispilina PersicariaVietnam	RMNH.INS.17990	Vietnam	Oriental	Polygonaceae	Larva
***Coptodisca* Walsingham, 1895**					
Coptodisca AmelanchierUSA	RMNH.INS.29176	United States	Nearctic	Rosaceae	Larva
Coptodisca arbutiella Busck, 1904	RMNH.INS.24466	Canada	Nearctic	Ericaceae	Male, adult, el
Coptodisca CaryaGeorgia	RMNH.INS.24369	United States	Nearctic	Juglandaceae	Male, adult, el
Coptodisca juglandiella (Chambers, 1874)	RMNH.INS.18240	United States	Nearctic	Juglandaceae	Larva
Coptodisca lucifluella (Clemens, 1860)	RMNH.INS.24368	Italy	West Palaearctic[a]	Juglandaceae	Male, adult, el
Coptodisca ostryaefoliella (Clemens, 1861)	RMNH.INS.18337	United States	Nearctic	Betulaceae	Larva
Coptodisca PopulusCalifornia	RMNH.INS.29635	United States	Nearctic	Salicaceae	Larva
Coptodisca quercicolella Braun, 1927	RMNH.INS.29191	United States	Nearctic	Fagaceae	Larva
Coptodisca saliciella (Clemens, 1861)	RMNH.INS.29630	United States	Nearctic	Salicaceae	Larva
Coptodisca saliciella (Clemens, 1861)	RMNH.INS.29933	United States	Nearctic	Salicaceae	Larva
Coptodisca splendoriferella (Clemens, 1859)	RMNH.INS.18637	United States	Nearctic	Rosaceae	Larva
Coptodisca splendoriferella (Clemens, 1859)	RMNH.INS.18341	United States	Nearctic	Rosaceae	Larva
Coptodisca VacciniumUSA	RMNH.INS.18522	United States	Nearctic	Ericaceae	Larva
Coptodisca VacciniumUSA	RMNH.INS.18713	United States	Nearctic	Ericaceae	Larva
***Heliozela* Herrich-Schäffer, 1853**					
Heliozela castaneella Kuroko, 1982	RMNH.INS.24523	Japan	East Palaearctic	Fagaceae	Male, adult

(*continued on next page*)

Table 1 (*continued*)

Species and author	Voucher	Country	Region	Host Family	Sex, stage, adult collection method (el = ex larva, oh = on host, dc = day collecting)
Heliozela eucarpa Meyrick, 1897	RMNH.INS.24527	Australia	Australian		Male, adult, at light
Heliozela KunzeaAusVIC	RMNH.INS.24525	Australia	Australian	Myrtaceae	Male, adult, oh
Heliozela MelaleucaAusSA	RMNH.INS.24526	Australia	Australian	Myrtaceae	Male, adult, oh
Heliozela MelastomaKalimantan	RMNH.INS.24161	Indonesia	Oriental	Melastomataceae	Female, adult, el
Heliozela resplendella Stainton, 1851	RMNH.INS.24359	Finland	West Palaearctic	Betulaceae	Female, adult
Heliozela sericiella Haworth, 1828	RMNH.INS.24358	Finland	West Palaearctic	Fagaceae	Male, adult
Heliozela Syzygium.acuminatissimum1Indo	RMNH.INS.24156	Indonesia	Oriental	Myrtaceae	Female, adult, el
Heliozela Syzygium.acuminatissimum1Indo	RMNH.INS.24157	Indonesia	Oriental	Myrtaceae	Male, adult, el
Heliozela Syzygium.acuminatissimum2Indo	RMNH.INS.11977	Indonesia	Oriental	Myrtaceae	Larva
Heliozela SyzygiumVietnam	RMNH.INS.24362	Vietnam	Oriental	Myrtaceae	Male, adult, el
Heliozela unknownAusWA.EvN4519	RMNH.INS.24519	Australia	Australian		Male, adult, dc
Heliozela unknownAusWA.EvN4521	RMNH.INS.24521	Australia	Australian		Male, adult, dc
Heliozela unknownAusWA.EvN4528	RMNH.INS.24528	Australia	Australian		Male, adult, dc
Heliozela unknownLesbos.EvN4606	RMNH.INS.24606	Greece	West Palaearctic		Male, adult
Heliozela unknownVietnamVuQuang	RMNH.INS.17970	Vietnam	Oriental	unidentified	Larva
***Holocacista* Walsingham & Durrant, 1909**					
Holocacista capensis van Nieukerken & Geertsema, 2015	RMNH.INS.24260	South Africa	Afrotropical	Vitaceae	Male, adult, el
Holocacista capensis van Nieukerken & Geertsema, 2015	RMNH.INS.24622	South Africa	Afrotropical	Vitaceae	Male, adult, el
Holocacista DyerophytumUAE	RMNH.INS.24628	UAE	West Palaearctic	Plumbaginaceae	Male, adult, el
Holocacista DyerophytumUAE	RMNH.INS.18451	UAE	West Palaearctic	Plumbaginaceae	Larva
Holocacista ImpatiensVietnam	RMNH.INS.17961	Vietnam	Oriental	Balsaminaceae	Larva
Holocacista LasianthusBorneo	RMNH.INS.24159	Indonesia	Oriental	Rubiaceae	Female, adult, el
Holocacista LasianthusSabah	RMNH.INS.30054	Malaysia	Oriental	Rubiaceae	Larva
Holocacista LasianthusSabah	RMNH.INS.30250	Malaysia	Oriental	Rubiaceae	Larva
Holocacista LeeaBorneo	RMNH.INS.24158	Indonesia	Oriental	Vitaceae	Female, adult, el
Holocacista PaederiaTaiwan	RMNH.INS.29541	Taiwan	East Palaearctic	Rubiaceae	Larva
Holocacista PsychotriaAusQLD	RMNH.INS.24367	Australia	Australian	Rubiaceae	Male, adult, el
Holocacista Rhoicissus.tomentosaSthAfrica	RMNH.INS.30313	South Africa	Afrotropical	Vitaceae	Larva
Holocacista Rhoicissus.tomentosaSthAfrica	RMNH.INS.29566	South Africa	Afrotropical	Vitaceae	Larva
Holocacista Rhoicissus.tridentataSthAfrica	RMNH.INS.29659	South Africa	Afrotropical	Vitaceae	Larva
Holocacista rivillei (Stainton, 1855)	RMNH.INS.24626	Italy	West Palaearctic	Vitaceae	Male, adult, el
Holocacista varii (Mey, 2012)	RMNH.INS.29600	South Africa	Afrotropical	Geraniaceae	Larva
Holocacista varii (Mey, 2012)	RMNH.INS.24625	South Africa	Afrotropical	Geraniaceae	Female, adult, el
***Hoplophanes* Meyrick, 1897**					
Hoplophanes Leucopogon1AusWA	HLZ.00757	Australia	Australian	Ericaceae	Male, adult, oh
Hoplophanes Leucopogon1AusWA	HLZ.00758	Australia	Australian	Ericaceae	Female, adult, oh
Hoplophanes Leucopogon2AusWA	HLZ.00751	Australia	Australian	Ericaceae	Male, adult, oh
Hoplophanes Leucopogon3AusWA	HLZ.00752	Australia	Australian	Ericaceae	Male, adult, oh
Hoplophanes Leucopogon5AusWA	HLZ.00750	Australia	Australian	Ericaceae	Female, adult, oh
Hoplophanes niphochalca Meyrick, 1897	HLZ.00659	Australia	Australian		Male, adult, oh
***Pseliastis* Meyrick, 1897**					
Pseliastis CorreaAusVIC	HLZ.00300	Australia	Australian	Rutaceae	Female, adult, oh
Pseliastis CorreaAusVIC	HLZ.00231	Australia	Australian	Rutaceae	Male, adult, oh
Pseliastis MicrocybeAusSA	HLZ.00506	Australia	Australian	Rutaceae	Female, adult, oh
Pseliastis MicrocybeAusSA	HLZ.00524	Australia	Australian	Rutaceae	Male, adult, oh
Pseliastis RutaceaeAusWA	HLZ.00767	Australia	Australian	Rutaceae	Male, adult, oh
Pseliastis RutaceaeAusWA	HLZ.00766	Australia	Australian	Rutaceae	Male, adult, oh
Pseliastis spectropa Meyrick, 1897	HLZ.00220	Australia	Australian	Rutaceae	Female, adult, oh
Pseliastis xanthodisca Meyrick, 1897	HLZ.02568	Australia	Australian	Rutaceae	Male, adult, oh
***Tyriozela* Meyrick, 1931**					
Tyriozela porphyrogona Meyrick, 1931	HLZ.02690	Japan	East Palaearctic		Adult, museum specimen
Undescribed genera					
heliozelidgenus ConostegiaCostaRica	RMNH.INS.24670	Costa Rica	Neotropical	Melastomataceae	Male, adult, el
heliozelidgenus HibbertiaAusWA	RMNH.INS.24522	Australia	Australian	Dilleniaceae	Male, adult, el
heliozelidgenus Tetracera1Kalimantan	RMNH.INS.24160	Indonesia	Oriental	Dilleniaceae	Female, adult, el
heliozelidgenus Tetracera2Kalimantan	RMNH.INS.11987	Indonesia	Oriental	Dilleniaceae	Larva
heliozelidgenus1 Boronia5AusWA	HLZ.02533	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus1 Boronia5AusWA	HLZ.02530	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus1 unknownAusWA	HLZ.02569	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus2 Boronia1AusWA	HLZ.02554	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus2 Boronia2AusWA	HLZ.02536	Australia	Australian	Rutaceae	Female, adult, oh
heliozelidgenus2 Boronia2AusWA	HLZ.02537	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus2 Boronia3AusWA	HLZ.02546	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus2 Boronia3AusWA	HLZ.02547	Australia	Australian	Rutaceae	Female, adult, oh
heliozelidgenus2 Boronia4AusWA	HLZ.02529	Australia	Australian	Rutaceae	Female, adult, oh
heliozelidgenus3 SpyridiumAusVIC	HLZ.02622	Australia	Australian	Rhamnaceae	Adult, oh
heliozelidgenus3 SpyridiumAusVIC	HLZ.02621	Australia	Australian	Rhamnaceae	Adult, oh
heliozelidgenus4 DodonaeaAusSA	HLZ.02627	Australia	Australian	Sapindaceae	Adult, oh
heliozelidgenus4 DodonaeaAusSA	HLZ.02628	Australia	Australian	Sapindaceae	Adult, oh

(*continued on next page*)

Table 1 (*continued*)

Species and author	Voucher	Country	Region	Host Family	Sex, stage, adult collection method (el = ex larva, oh = on host, dc = day collecting)
heliozelidgenus5 GeleznowiaAusWA	HLZ.00678	Australia	Australian	Rutaceae	Female, adult, oh
heliozelidgenus5 GeleznowiaAusWA	HLZ.00663	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus6 BoroniaAusWA	HLZ.02566	Australia	Australian	Rutaceae	Male, adult, oh
heliozelidgenus7 ZieriaAusVIC	HLZ.00770	Australia	Australian	Rutaceae	Male, adult, oh
Outgroups[b]					
Ectoedemia AcerTaiwan (Nepticulidae)	RMNH.INS.29364	Taiwan	Palearctic, Afrotropical, Nearctic	Various	Larva
Nematopogon adansoniella (Villers, 1789) (Adelidae)	RMNH.INS.538731	Netherlands	Palearctic, Afrotropical, Oriental	Various	Adult
Nemophora degeerella (Linnaeus, 1758) (Adelidae)	SRR921621	Austria	Palearctic, Afrotropical, Oriental	Various	Adult
Perthida EucalyptusAusACT (Incurvariidae)	RMNH.INS.11848	Australia	Palearctic, Australia	Various	Larva
Perthida EucalyptusAusACT (Incurvariidae)	RMNH.INS.11847	Australia	Palearctic, Australia	Various	Larva
Tegeticula yuccasella Riley, 1873 (Prodoxidae)	SRR3180626	United States	Nearctic, Palearctic	Various	Adult

[a] Introduced from North America.
[b] For outgroups, we have included the biogeographical regions hosting the most diversity for each family according to Heppner (1991).

org/10.5883/DS-HELIPHYL).

2.2. Outgroup selection

Both cladistic (Mutanen et al., 2010) and molecular studies (Regier et al., 2015) place the Heliozelidae within the superfamily Adeloidea. Based on these conclusions, we have sourced sequences from three of the four sister families in the Adeloidea, namely Adelidae, Prodoxidae, and Incurvariidae, as well as sequences from another family, Nepticulidae, as outgroup taxa. We sequenced two individuals from the genus *Perthida* (Incurvariidae), and a specimen of *Nematopogon adansoniella* (Adelidae), using the methods described in Section 2.3. For additional outgroup sequences, we searched the published transcriptomes of *Tegeticula yuccasella* (Prodoxidae, NCBI SRR3180626) and *Nemophora degeerella* (Adelidae, NCBI SRR921621). We also included the COI, 28S, and COII sequences of putative species *Ectoedemia AcerTaiwan* voucher RMNH.INS.29364 (Nepticulidae) (Doorenweerd et al., 2016), publically available in NCBI as *Ectoedemia olvina*, details available in Table 1.

2.3. DNA library preparations

For the Australian specimens, total genomic DNA was extracted non-destructively from the abdomens using a Macherey-Nagel NucleoSpin® Tissue XS kit following the manufacturer's protocol, using 40 µl of nuclease-free water as the elution buffer. Molecular profiles and concentrations of DNA were quantified on an Agilent 2200 Tapestation™. For other specimens, total genomic DNA was extracted non-destructively from adult abdomens or larvae frozen in ethanol > 95%, or occasionally from larvae that had been dried inside their leaf mines, using a Macherey-Nagel NucleoMag 96® Tissue magnetic bead kit on a Thermo Fisher KingFisher flex system. Primers used in previous studies of Lepidoptera (see Table 2) were used to PCR amplify fragments of four genes (COI, COII, 28S and H3), which were then sequenced using the Sanger method or using next-generation sequencing (NGS) protocols. For samples that yielded high quality DNA, we used PCR primers to amplify target genes and prepared NGS libraries from amplified products as per the Illumina TruSeq Nano protocol. For samples that yielded poor quality DNA, we used an RNA bait and capture method, based on a published protocol (Carpenter et al., 2013). The RNA baits were created from PCR products using "founder" moths from which high quality DNA could be extracted. Specimens from different genera across four families (Heliozelidae, Micropterigidae, Aenigmatineidae and Oecophoridae) were selected to derive the bait pool and maximise the chances of finding sequence homology. Fragments captured by baits were made into NGS libraries using the Illumina TruSeq Nano protocol. A full description of the RNA bait and capture protocol can be found in Appendix A.

2.4. Next generation sequencing and assembly

NGS libraries were sequenced using the Illumina MiSeq 300 cycle kit as paired end reads according to manufacturer's instructions. Paired FASTQ files were trimmed and quality filtered (minimum of Q20 and length 35). Filtered reads were corrected using Musket 1.1 (Liu et al., 2013), a multi-stage k-mer based corrector that corrects substitution errors based on consensus of reads. The corrected reads for each sample were assembled *de novo* using Velvet 1.2.08 (Zerbino and Birney, 2008) with a range of k-mers between 79 and 33, and Spades v3.0.0 (Bankevich et al., 2012) using built-in k-mer values of 77, 55, 33 and 21. All resulting contigs of at least 300 base pairs were searched against a BLAST database of heliozelid reference genes with a minimum e-value of 10^{-9}. Reads were mapped against the top three matching contigs for each sample and gene. The contig coverage was checked visually using IGV v2.3.32 (Thorvaldsdóttir et al., 2013) and regions with poor coverage and the primer binding sites were trimmed. The best matching

Table 2
Gene amplification primers used for PCR.

Gene	Primer name	Direction	Primer sequence	Reference
COI	LepF1	Forward	5′-ATTCAACCAATCATAAAGATATTGG-3′	Hebert et al. (2003)
COI	LepR1	Reverse	5′-TAAACTTCTGGATGTCCAAAAAATCA-3′	Hebert et al. (2003)
COII	F	Forward	5′-GGAGCATCTCCTTTAATAGAACA-3′	Sperling et al. (1995)
COII	Eva	Reverse	5′-GAGACCATTACTTGCTTTCGATCATCT-3′	Caterino and Sperling (1999)
28S	S3660 F	Forward	5′-GAGAGTTMAASAGTACGTGAAAC-3′	Dowton and Austin (1998)
28S	A335 R	Reverse	5′-TCGGARGGAACCAGCTACTA-3′	Whiting et al. (1997)
H3	HexAF	Forward	5′-ATGGCTCGTACCAAGCAGACGGC-3′	Ogden and Whiting (2003)
H3	HexAR	Reverse	5′-ATATCCTTGGGCATGATGGTGAC-3′	Ogden and Whiting (2003)

Table 3
NCBI Genbank IDs for sequences used in analyses.

Voucher code	Species name	COI	COII	28S	H3
RMNH.INS.30326	*Antispila ampelopsia*	MF118352	MF118441	MF118266	MF118536
RMNH.INS.18588	*Antispila ampelopsifoliella*	JQ412550.1	MF118375	MF118194	MF118463
RMNH.INS.24377	*Antispila ampelopsifoliella*	MF118305	MF118397	MF118221	MF118490
RMNH.INS.24365	*Antispila Ampelopsis.brevTaiwan*	MF118298		MF118214	MF118483
RMNH.INS.29431	*Antispila Ampelopsis.brevTaiwan*	MF118336	MF118425	MF118250	MF118520
RMNH.INS.24364	*Antispila Ampelopsis.cantTaiwan*	MF118297	MF118390	MF118213	MF118482
RMNH.INS.29436	*Antispila Ampelopsis.cantTaiwan*	MF118337	MF118426	MF118251	MF118521
RMNH.INS.18566	*Antispila argentifera*	MF118282	MF118374	MF118193	MF118462
RMNH.INS.24532	*Antispila cleyerella*	MF118316	MF118407	MF118231	MF118500
RMNH.INS.24535	*Antispila corniella*	MF118317	MF118408	MF118232	MF118501
RMNH.INS.18270	*Antispila cornifoliella*	MF118276	MF118366	MF118184	MF118453
RMNH.INS.24366	*Antispila cornifoliella*	MF118299	MF118391	MF118215	MF118484
RMNH.INS.29809	*Antispila CornusJapan*	MF118345	MF118434	MF118259	MF118529
RMNH.INS.29810	*Antispila CornusJapan*	MF118346	MF118435	MF118260	MF118530
RMNH.INS.24537	*Antispila distyliella*	MF118318		MF118233	MF118502
RMNH.INS.24540	*Antispila hikosana*	MF118319	MF118409	MF118234	MF118503
RMNH.INS.18206	*Antispila hydrangaeella*	JQ412516.1	MF118365	MF118182	MF118451
RMNH.INS.18416	*Antispila hydrangaeella*	JQ412515.1	MF118370	MF118188	MF118457
RMNH.INS.29423	*Antispila Hydrangea.chinensisTaiwan*	MF118335	MF118424		MF118519
RMNH.INS.29377	*Antispila HydrangeaTaiwan*	MF118334	MF118423	MF118249	MF118518
RMNH.INS.18127	*Antispila isabella*	JQ412510.1	MF118363	MF118180	
RMNH.INS.18504	*Antispila isabella*	MF118280	MF118372	MF118190	MF118459
RMNH.INS.24371	*Antispila metallella*	MF118303	MF118395	MF118219	MF118488
RMNH.INS.18303	*Antispila nysaefoliella*	JQ412523.1	MF118367	MF118185	MF118454
RMNH.INS.29113	*Antispila nysaefoliella*	MF118327	MF118416	MF118242	MF118511
RMNH.INS.18643	*Antispila oinophylla*	MF118284		MF118196	MF118465
RMNH.INS.23920	*Antispila oinophylla*	JQ412536.1		MF118200	MF118468
RMNH.INS.24372	*Antispila petryi*	MF118304	MF118396	MF118220	MF118489
RMNH.INS.29304	*Antispila petryi*	MF118333	MF118422	MF118248	MF118517
RMNH.INS.24529	*Antispila purplella*	MF118315		MF118230	MF118499
RMNH.INS.29735	*Antispila tateshinensis*	MF118344	MF118433	MF118258	MF118528
RMNH.INS.29823	*Antispila tateshinensis*	MF118347	MF118436	MF118261	MF118531
RMNH.INS.11856	*Antispila treitschkiella*	JQ412555.1	MF118356	MF118174	MF118443
RMNH.INS.29300	*Antispila treitschkiella*	MF118332	MF118421	MF118247	MF118516
RMNH.INS.29237	*Antispila uenoi*	MF118331	MF118420	MF118246	MF118515
RMNH.INS.18508	*Antispila viticordifoliella*	JQ412547.1		MF118191	MF118460
RMNH.INS.24207	*Antispila viticordifoliella*	JQ412546.1	MF118385	MF118207	MF118476
RMNH.INS.18126	*Antispila Vitis1USA*	JQ412557.1	MF118362	MF118179	MF118449
RMNH.INS.24205	*Antispila Vitis1USA*	JQ412556.1		MF118206	MF118475
RMNH.INS.18131	*Antispila Vitis2USA*	JQ412558.1	MF118364	MF118181	MF118450
RMNH.INS.18656	*Antispila Vitis2USA*	JQ412559.1	MF118377	MF118197	MF118466
RMNH.INS.29232	*Antispila VitisKorea*	MF118330	MF118419	MF118245	MF118514
RMNH.INS.23917	*Antispila voraginella*	JQ412562.1	MF118378	MF118199	MF118467
RMNH.INS.11853	*Antispilina ludwigi*	MF118269	MF118355	MF118173	MF118442
RMNH.INS.24677	*Antispilina ludwigi*	MF118326	MF118415	MF118241	MF118510
RMNH.INS.17990	*Antispilina PersicariaVietnam*	MF118274	MF118361	MF118178	MF118448
RMNH.INS.24363	*Antispilina PersicariaVietnam*	MF118296	MF118389	MF118212	MF118481
RMNH.INS.29176	*Coptodisca AmelanchierUSA*	MF118328	MF118417	MF118243	MF118512
RMNH.INS.24466	*Coptodisca arbutiella*	MF118306	MF118398	MF118222	MF118491
RMNH.INS.24369	*Coptodisca CaryaGeorgia*	MF118302	MF118394	MF118218	MF118487
RMNH.INS.18240	*Coptodisca juglandiella*	MF118275		MF118183	MF118452
RMNH.INS.24368	*Coptodisca lucifluella*	MF118301	MF118393	MF118217	MF118486
RMNH.INS.18337	*Coptodisca ostryaefoliella*	MF118277	MF118368	MF118186	MF118455
RMNH.INS.29635	*Coptodisca PopulusCalifornia*	MF118342	MF118431	MF118256	MF118526
RMNH.INS.29191	*Coptodisca quercicolella*	MF118329	MF118418	MF118244	MF118513
RMNH.INS.29630	*Coptodisca saliciella*	MF118341	MF118430	MF118255	MF118525
RMNH.INS.29933	*Coptodisca saliciella*	MF118348	MF118437	MF118262	MF118532
RMNH.INS.18341	*Coptodisca splendoriferella*	MF118278	MF118369	MF118187	MF118456
RMNH.INS.18637	*Coptodisca splendoriferella*	MF118283	MF118376	MF118195	MF118464
RMNH.INS.18522	*Coptodisca VacciniumUSA*	MF118281	MF118373	MF118192	MF118461
RMNH.INS.18713	*Coptodisca VacciniumUSA*	MF118285		MF118198	
RMNH.INS.24523	*Heliozela castaneella*	MF118310	MF118402		MF118495
RMNH.INS.24527	*Heliozela eucarpa*	MF118313	MF118405	MF118228	MF118498
RMNH.INS.24525	*Heliozela KunzeaAusVIC*	MF118311	MF118403	MF118226	MF118496
RMNH.INS.24526	*Heliozela MelaleucaAusSA*	MF118312	MF118404	MF118227	MF118497
RMNH.INS.24161	*Heliozela MelastomaKalimantan*	MF118291	MF118384	MF118205	MF118474
RMNH.INS.24359	*Heliozela resplendella*	MF118294	MF118388	MF118210	MF118479
RMNH.INS.24358	*Heliozela sericiella*	MF118293	MF118387	MF118209	MF118478
RMNH.INS.24156	*Heliozela Syzygium.acuminatissimum1Indo*	MF118286	MF118379	MF118201	MF118469
RMNH.INS.24157	*Heliozela Syzygium.acuminatissimum1Indo*	MF118287	MF118380		MF118470
RMNH.INS.11977	*Heliozela Syzygium.acuminatissimum2Indo*	MF118270	MF118357	MF118175	MF118444
RMNH.INS.24362	*Heliozela SyzygiumVietnam*	MF118295		MF118211	MF118480
RMNH.INS.24519	*Heliozela unknownAusWA.EvN4519*	MF118307	MF118399	MF118223	MF118492

(*continued on next page*)

Table 3 (*continued*)

Voucher code	Species name	COI	COII	28S	H3
RMNH.INS.24521	*Heliozela unknownAusWA.EvN4521*	MF118308	MF118400	MF118224	MF118493
RMNH.INS.24528	*Heliozela unknownAusWA.EvN4528*	MF118314	MF118406	MF118229	
RMNH.INS.24606	*Heliozela unknownLesbos.EvN4606*	MF118320	MF118410	MF118235	MF118504
RMNH.INS.17970	*Heliozela unknownVietnamVuQuang*	MF118273	MF118360	MF118177	MF118447
RMNH.INS.24260	*Holocacista capensis*	MF118292	MF118386	MF118208	MF118477
RMNH.INS.24622	*Holocacista capensis*	MF118321	MF118411	MF118236	MF118505
RMNH.INS.18451	*Holocacista DyerophytumUAE*	MF118279	MF118371	MF118189	MF118458
RMNH.INS.24628	*Holocacista DyerophytumUAE*	MF118324	MF118414	MF118239	MF118508
RMNH.INS.17961	*Holocacista ImpatiensVietnam*	MF118272	MF118359	MF118176	MF118446
RMNH.INS.24159	*Holocacista LasianthusBorneo*	MF118289	MF118382	MF118203	MF118472
RMNH.INS.30054	*Holocacista LasianthusSabah*	MF118349	MF118438	MF118263	MF118533
RMNH.INS.30250	*Holocacista LasianthusSabah*	MF118350	MF118439	MF118264	MF118534
RMNH.INS.24158	*Holocacista LeeaBorneo*	MF118288	MF118381	MF118202	MF118471
RMNH.INS.29541	*Holocacista PaederiaTaiwan*	MF118338	MF118427	MF118252	MF118522
RMNH.INS.24367	*Holocacista PsychotriaAusQLD*	MF118300	MF118392	MF118216	MF118485
RMNH.INS.29566	*Holocacista Rhoicissus.tomentosaSthAfrica*	MF118339	MF118428	MF118253	MF118523
RMNH.INS.30313	*Holocacista Rhoicissus.tomentosaSthAfrica*	MF118351	MF118440	MF118265	MF118535
RMNH.INS.29659	*Holocacista Rhoicissus.tridentataSthAfrica*	MF118343	MF118432	MF118257	MF118527
RMNH.INS.24626	*Holocacista rivillei*	MF118323	MF118413	MF118238	MF118507
RMNH.INS.24625	*Holocacista varii*	MF118322	MF118412	MF118237	MF118506
RMNH.INS.29600	*Holocacista varii*	MF118340	MF118429	MF118254	MF118524
HLZ.00757	*Hoplophanes Leucopogon1AusWA*	MF062295	MF062333	MF062395	MF062366
HLZ.00758	*Hoplophanes Leucopogon1AusWA*	MF062296	MF062334		
HLZ.00751	*Hoplophanes Leucopogon2AusWA*	MF062291	MF062330	MF062392	MF062363
HLZ.00752	*Hoplophanes Leucopogon3AusWA*	MF062292	MF062331	MF062393	MF062364
HLZ.00750	*Hoplophanes Leucopogon5AusWA*	MF062290	MF062329	MF062391	MF062362
HLZ.00659	*Hoplophanes niphochalca*	MF062287	MF062326		
HLZ.00231	*Pseliastis CorreaAusVIC*	MF062281	MF062321	MF062386	MF062357
HLZ.00300	*Pseliastis CorreaAusVIC*	MF062282	MF062322	MF062387	MF062358
HLZ.00506	*Pseliastis MicrocybeAusSA*	MF062284	MF062323	MF062388	MF062359
HLZ.00524	*Pseliastis MicrocybeAusSA*	MF062285	MF062324	MF062389	MF062360
HLZ.00766	*Pseliastis RutaceaeAusWA*	MF062297	MF062335		
HLZ.00767	*Pseliastis RutaceaeAusWA*	MF062298	MF062336		
HLZ.00220	*Pseliastis spectropa*	MF062280	MF062320		
HLZ.02568	*Pseliastis xanthodisca*	MF062314		MF062411	MF062382
HLZ.02690	*Tyriozela porphyrogona*	MF171066	MF171067	MF171068	MF171069
RMNH.INS.24670	*heliozelidgenus ConostegiaCostaRica*	MF118325		MF118240	MF118509
RMNH.INS.24522	*heliozelidgenus HibbertiaAusWA*	MF118309	MF118401	MF118225	MF118494
RMNH.INS.11987	*heliozelidgenus Tetracera2Kalimantan*	MF118271	MF118358		MF118445
RMNH.INS.24160	*heliozelidgenus Tetracera1Kalimantan*	MF118290	MF118383	MF118204	MF118473
HLZ.02530	*heliozelidgenus1 Boronia5AusWA*	MF062301	MF062339	MF062398	MF062369
HLZ.02533	*heliozelidgenus1 Boronia5AusWA*	MF062302	MF062340	MF062399	MF062370
HLZ.02569	*heliozelidgenus1 unknownAusWA*	MF062315	MF062352	MF062412	MF062383
HLZ.02554	*heliozelidgenus2 Boronia1AusWA*	MF062312	MF062350	MF062409	MF062380
HLZ.02536	*heliozelidgenus2 Boronia2AusWA*	MF062304	MF062342	MF062401	MF062372
HLZ.02537	*heliozelidgenus2 Boronia2AusWA*	MF062305	MF062343	MF062402	MF062373
HLZ.02546	*heliozelidgenus2 Boronia3AusWA*	MF062309	MF062347	MF062406	MF062377
HLZ.02547	*heliozelidgenus2 Boronia3AusWA*	MF062310	MF062348	MF062407	MF062378
HLZ.02529	*heliozelidgenus2 Boronia4AusWA*	MF062300	MF062338	MF062397	MF062368
HLZ.02621	*heliozelidgenus3 SpyridiumAusVIC*	MF062316	MF062353		
HLZ.02622	*heliozelidgenus3 SpyridiumAusVIC*	MF062317	MF062354	MF062413	MF062384
HLZ.02627	*heliozelidgenus4 DodonaeaAusSA*	MF062318	MF062355	MF062414	MF062385
HLZ.02628	*heliozelidgenus4 DodonaeaAusSA*	MF062319	MF062356		
HLZ.00663	*heliozelidgenus5 GeleznowiaAusWA*	MF062288	MF062327		
HLZ.00678	*heliozelidgenus5 GeleznowiaAusWA*	MF062289	MF062328	MF062390	MF062361
HLZ.02566	*heliozelidgenus6 BoroniaAusWA*	MF062313	MF062351	MF062410	MF062381
HLZ.00770	*heliozelidgenus7 ZieriaAusVIC*	MF062299	MF062337	MF062396	MF062367
Outgroups					
RMNH.INS.29364	*Ectoedemia AcerTaiwan*	KM077659.1	KM078269.1	KM078456.1	
RMNH.INS.538731	*Nematopogon adansoniella*	KX048872.1	[a]	[a]	[a]
SRR921621	*Nemophora degeerella*	[a]	[a]	[a]	[a]
RMNH.INS.11847	*Perthida EucalyptusAusACT*	MF118267	MF118353	MF118171	
RMNH.INS.11848	*Perthida EucalyptusAusACT*	MF118268	MF118354	MF118172	
SRR3180626	*Tegeticula yuccasella*	KX232884	[a]	[a]	[a]

[a] Sequences available for download from Dryad.

contig with highest coverage for each gene was chosen as the final gene sequence. Genbank IDs are listed in Table 3.

2.5. Analysis of published transcriptome data

For additional outgroup sequences, we searched the published transcriptomes of *Tegeticula yuccasella* (Prodoxidae, NCBI SRR3180626) and *Nemophora degeerella* (Adeloidea, NCBI SRR921621). The FASTQ

files were downloaded from NCBI and the transcriptomes assembled *de novo* using Trinity r20131110 (Grabherr et al., 2011) using the default parameter values. The contig abundance was calculated using RSEM within Trinity. We used Geneious R11.02 (Biomatters Ltd.) to map the assembled contigs in each transcriptome to a set of reference sequences from other voucher specimens downloaded from NCBI. As reference sequences for the *T. yuccasella* transcriptome, we used COI: *T. yuccasella* KX232884.1, COII: *T. synthethica* AY327144.1:1583-2104, 28S: *Adela reaumurella* AY.230752.1, H3: *Bombyx mori* DQ443228.1. As reference sequences for the *N. degeerella* transcriptome, we used COI: *N. degeerella* KX061994.1, COII: *Adela septentrionella* EU884115.1:1563-2084, 28S: *A. reaumurella* AY.230752.1, H3: *Bombyx mori* DQ443228.1. The best matching contig with highest abundance for each gene was chosen as the final gene sequence. The selected sequences are available for download at Dryad (https://doi.org/10.561/dryad.r51c7).

2.6. Phylogenetic analyses

Sequence alignments for individual genes were created using the MAFFT v7.3.09 (Katoh and Standley, 2013) plugin within Geneious R11.02 (Biomatters Ltd.) using the default Auto option. The resulting alignments were concatenated using the Concatenate Sequences or Alignments function in Geneious R11.0.2 (Biomatters Ltd.). The resulting alignments are available for download at Dryad (https://doi.org/10.561/dryad.r51c7). We analysed three data sets using maximum likelihood (ML): (a) all four genes, (b) mitochondrial genes (COI and COII) and (c) nuclear genes (28S and H3). We further analysed the four-gene data set using Bayesian methods. The data sets were partitioned by gene, and the protein coding genes further partitioned by codon position, with the first two positions estimated independently of the third. For all ML analyses we used the nucleotide substitution model GTR + I + G for the RAxML (Stamatakis, 2014) MPI version, which we ran on the Melbourne Bioinformatics (University of Melbourne) cluster. We ran 100 ML searches and 1000 bootstraps for the mitochondrial and nuclear gene alignments, and 1000 ML searches and 1000 bootstraps for the combined four-gene alignments. For the Bayesian analysis of the four-gene alignment we used ExaBayes (Kozlov et al., 2015) and ran a minimum of 1 million generations and two independent runs, with a 25% burn in proportion and the same partitioning as RAxML. The ExaBayes run stopped after 88,90,000 generations. To check for run convergence, we ran the ExaBayes sdsf tool to ensure that the average standard deviation of the split frequencies was < 1% (average deviation was 0.999937%).

2.7. Biogeography and host associations

Host association and biogeography information was mainly from our own observations, since we collected the majority of specimens. Most non-Australian specimens were collected as larvae, sometimes reared into the adult stage, so we have primary host information. Most Australian specimens were adults collected by sweeping the vegetation, and observed on particular plant species on which oviposition was often observed. In some cases we assumed host plants on circumstantial evidence. The method used to collect each specimen is indicated in Table 1. In addition, we used the literature for additional information on hostplant data and biogeography (Bernardo et al., 2015; Emmet, 1976; Kuroko, 1961; Kuroko, 1982, 1987; Lafontaine, 1973, 1974; Lee and Hirowatari, 2013; Lee et al., 2006a, 2006b; Maier, 1988; Mutanen et al., 2007; Nielsen, 1980; van Nieukerken and Geertsema, 2015; van Nieukerken et al., 2012, submitted for publication; Opler, 1971; Robinson et al., 2008).

3. Results

The alignments of 136 concatenated sequences of four genes resulted in a 2453 bp long matrix with 13.1% missing data. The mitochondrial gene sets (COI and COII) resulted in a 1269 bp matrix with 7.9% missing data. The nuclear gene sets (28S and H3) resulted in an 1184 bp matrix with 13.5% missing data, after excluding eight samples without any data. We examined all of the resulting ML phylogenies to assess the evolutionary history of each gene combination. In the nuclear gene phylogeny (Supplementary material, Fig. 1), four major clades were recovered with moderate to strong support (more details on clades below) but the placement of the outgroups could not be resolved. The groups within the cosmopolitan clade were recovered but poorly supported by the nuclear gene phylogeny, with the exception of *Antispila* group II. By contrast, the mitochondrial gene phylogeny (Supplementary material, Fig. 2), separated the Heliozelidae taxa from the outgroup taxa, and recovered the five major clades with moderate to strong support (80–100%), but could not recover three of the groups within the main cosmopolitan clade. By combining the sequences from all four genes we found significant improvement in the resolution of both the major clades and the groups within them, with Heliozelidae separated from the outgroups, and five major clades recovered with strong support. Thus, we chose to focus on the four-gene data set, which we analysed with both Bayesian and maximum likelihood (ML) methods, and our discussion below refers to the four-gene analyses unless specified. The ML topology labelled with significant support values for the major groups (ML bootstrap > = 80%, Bayesian posterior probability > = 0.8) is illustrated in Fig. 3, and the groups recovered are described in detail below, listing the support received in brackets (ML bootstrap %/Bayesian posterior probability). The full results for the ML and Bayesian analyses are in the Supplementary material, Figs. 3 and 4 respectively.

Both the ML and Bayesian analyses retrieved five main groups with strong support. The largest cosmopolitan clade (100/1) contains all species of three of the described leaf-mining genera (*Antispilina*, *Coptodisca*, *Holocacista*) as well as two of the groups currently combined in *Antispila* (*A. ampelopsifoliella* group and *A.* group II), plus a number of unplaced taxa (*heliozelidgenus*, species "*Tetracera1Kalimantan*", "*Tetracera2Kalimantan*", "*ConostegiaCostaRica*" and "*HibbertiaAusWA*"). The species of *Coptodisca* examined in this study formed a monophyletic clade with strong support (100/1). *Holocacista* was split into multiple groups with weak support in the mitochondrial gene set (Supplementary material, Fig. 2). In the nuclear gene set, this group was recovered as a monophyletic clade with poor support (Supplementary material, Fig. 1). When all four genes were combined, *Holocacista* was recovered with weak support from the ML analysis (57) but strong support in the Bayesian analysis (1). The single named species of *Antispilina* (*A. ludwigi*) included in this study formed a clade with strong support (100/1) with an unnamed *Antispilina* from Vietnam. The sister group to the remaining species in the main cosmopolitan clade, Undescribed group I, comprised solely of undescribed taxa, received weak support in the ML analysis (64), but strong support in the Bayesian analysis (0.98). This group was not recovered as monophyletic in either the nuclear or mitochondrial gene analysis (Supplementary material, Figs. 1 and 2).

Our analyses consistently split *Antispila* into several groups: a large monophyletic group (*Antispila* group I) and two smaller groups (*Antispila ampelopsifoliella* group and *Antispila* group II), suggesting that the genus, as currently described, is not monophyletic. The largest group, *Antispila* group I, is strongly supported (100/1) and comprises the majority of described *Antispila* species, including the type species, *Antispila stadtmuellerella* (junior synonym of *Antispila metallella*), as well as *A. ampelopsia*, *A. corniella*, *A. cornifoliella*, *A. cleyerella*, *A. distyliella*, *A. hikosana*, *A. isabella*, *A. nysaefolliella*, *A. petryi*, *A. purplella*, *A. tateshinensis*, *A. treitschkiella*, *A. uenoi* (identification provisional), and seven undescribed species. This group was consistently placed as sister to *Heliozela* + *Tyriozela* group in both the Bayesian and ML analyses, although neither analysis provided strong support. The smaller *Antispila* groups were placed within the large cosmopolitan clade and were most closely associated with *Coptodisca*. *Antispila* group II was not strongly

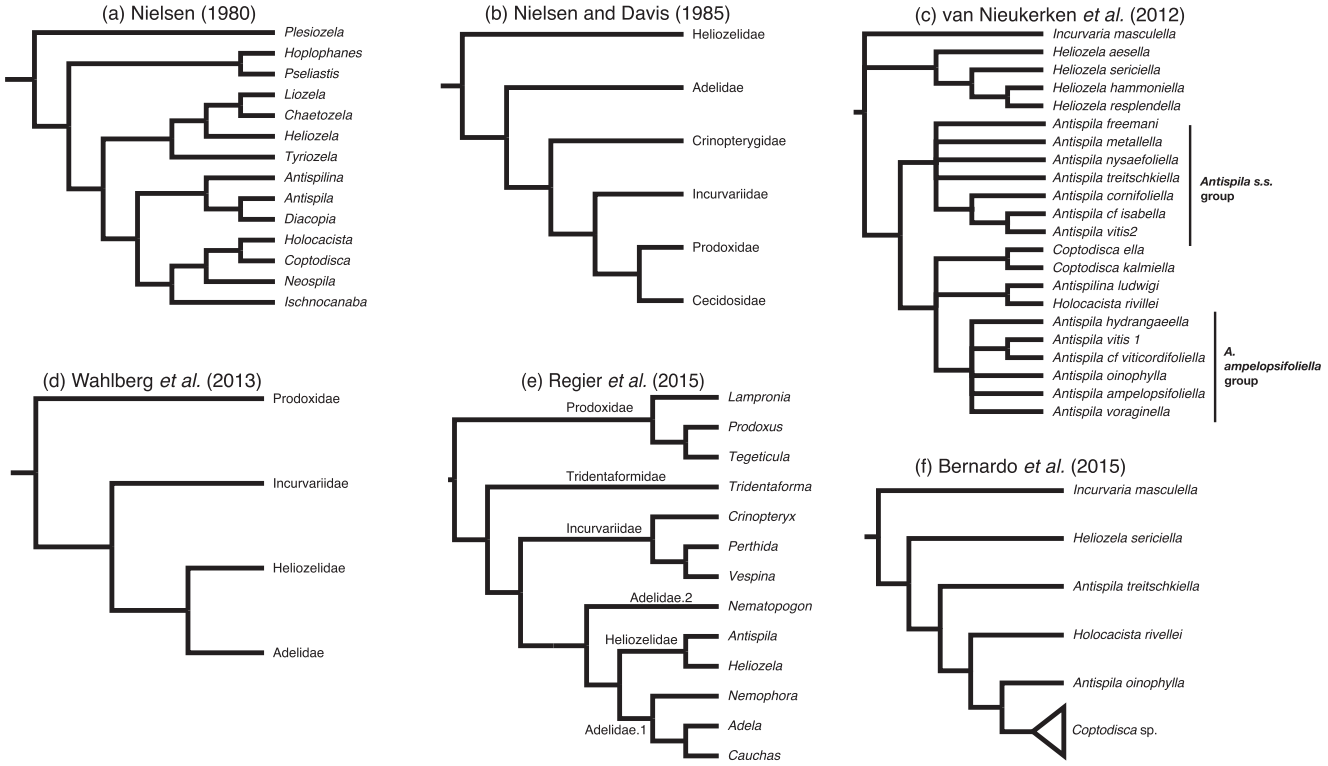

Fig. 2. Previous hypotheses regarding Heliozelidae. (a) Cladogram of Heliozelidae (after Nielsen 1980). *Liozela*, *Chaetozela* and *Neospila* are unpublished manuscript names for genera proposed by Nielsen. *Diacopia* is a synonym of *Antispila* that Nielsen regarded as separate genus. (b) Cladogram of Incurvarioidea (Adeloidea) including Heliozelidae (after Nielsen and Davis, 1985). Crinopterigidae has been subsumed into Incurvariidae by van Nieukerken et al. (2011). (c) Cladogram, 50% majority rule consensus tree from maximum parsimony analysis of COI sequences after van Nieukerken et al. (2012). (d) Cladogram based on phylogeny of Lepidoptera showing the position of Heliozelidae in relation to other families in Adeloidea after Wahlberg et al. (2013). (e) Cladogram based on phylogeny of non-dytrisian lineages after Regier et al. (2015) showing the split of *Nematopogon* from the rest of Adelidae seen in some analyses. (f) Cladogram based on maximum likelihood (ML) tree for COI data after Bernardo et al. (2015).

supported and included *A. viticordifoliella* along with one undescribed species. The single species *A. argentifera* was placed as sister to these two species in the ML phylogeny, but as sister to *Coptodisca* in the Bayesian phylogeny. The *Antispila ampelopsifoliella* group (100/1) was strongly supported and comprised *A. ampelopsifoliella*, *A. hydrangaeella*, *A. oinophylla* and *A. voraginella*.

A large group containing the two named *Pseliastis* species and several undescribed taxa was strongly supported (100/1). This group included the Australian endemic species *P. spectropa* and *P. xanthodisca*, as well as several other undescribed Australian species, some of which have been tentatively assigned to *Pseliastis*, while others may represent as yet unnamed genera.

The *Heliozela* + *Tyriozela* group is strongly supported (100/1) and comprises all species of the genus *Heliozela* that were included in this study, namely *H. castaneella*, *H. eucarpa*, *H. resplendella* and *H. sericiella*, as well as 11 undescribed species. *Tyriozela* was nested within this clade, rendering *Heliozela* paraphyletic in its current form.

The *Hoplophanes* group, also strongly supported (99/1), represents another endemic Australian clade, of which we included one described (*H. niphochalca*) and four undescribed species. This clade was consistently placed as the sister group to the remaining Heliozelidae examined in this study.

4. Discussion

4.1. Support for described genera

Overall, our results strongly support the monophyly of five currently described genera (*Coptodisca*, *Holocacista*, *Antispilina*, *Pseliastis* and *Hoplophanes*). One of the largest genera, *Antispila*, is broken into three separate groups, rendering it polyphyletic. *Heliozela* was recovered as a paraphyletic clade, but would become monophyletic with the inclusion of the monotypic genus *Tyriozela*. While the relationship between the major clades lacks resolution in our results, the general pattern is strikingly similar to the first Hennigian cladistic analysis of Heliozelidae performed by Nielsen (1980), with the exception of the placement of *Antispilina* (Fig. 2a). We also recognise two of the clades that were apparent in a previous limited analysis of COI barcode data; namely *Antispila sensu stricto* and *A. ampelopsifoliella* group, as well as a *Coptodisca* group comprising two species (van Nieukerken et al., 2012).

The monophyly of *Coptodisca* is well supported in our study, which is consistent with a previous phylogeny of COI sequences from several *Coptodisca* species. The study (Bernardo et al., 2015, Fig. 2f) was aimed at establishing the source of *C. lucifluella* in Italian walnuts, and recovered a similar *Coptodisca* clade using sequences from *C. arbutiella*, *C. juglandiella*, *C. lucifluella*, *C. negligens* (here identified as *Coptodisca* "*VacciniumUSA*"), *C. ostryaefoliella*, *C. quercicolella*, *C. saliciella* and *C. splendoriferella*. Morphologically, *Coptodisca* differs from other Heliozelidae by its distinct forewing colour pattern, more closely resembling unrelated leafmining genera such as *Leucoptera* (Lyonetiidae), that includes leaf miners of crops such as coffee and apples, and *Phyllocnistis* (Gracillariidae), of which the now global citrus leafminer *P. citrella* is commonly known as pest of citrus and other Rutaceae. Thus, even though Nielsen (1980) considered this characteristic forewing pattern as an apomorphy for *Coptodisca*, the pattern may have evolved independently in multiple unrelated families.

In all our analyses, the genus *Antispila* was consistently broken up into multiple groups. *Antispila sensu stricto* (*Antispila* group I) formed a strongly supported monophyletic clade. A study by van Nieukerken et al. (2012) recovered a similar clade of 'true' *Antispila* formed by *A. metallella*, *A. nysaefoliella*, *A. petryi* and *A. treitschkiella* (Fig. 2c). Morphologically, this group is defined by the extensive venation with the

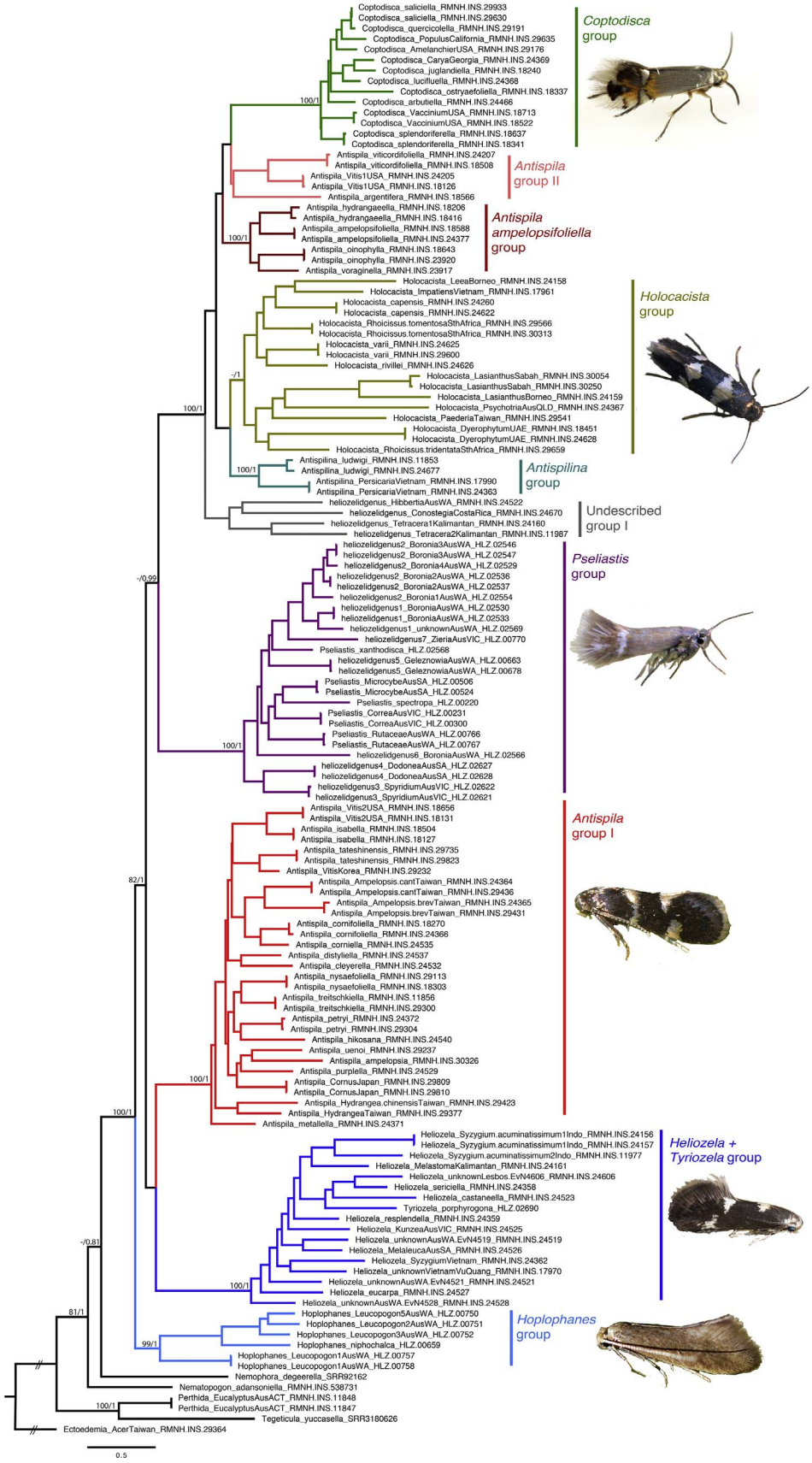

Fig. 3. ML tree (ln = −51259.825874) inferred from four genes. Branch lengths are proportional to ML estimated branch lengths. The numbers above the branches are MP bootstrap supports/Bayesian posterior probabilities calculated using gene-partitioned models. Only support values at or above 80% bootstrap and 0.8 posterior probabilities for the major clades are shown. Adult representatives of various Heliozelidae genera are shown next to corresponding group. Species names and photo credits: *Antispila* group I: *A. treitschkiella*, Switzerland (R. Bryner); *Coptodisca* group: *C. splendoriferella*, USA (C. Eiseman); *Holocacista* group: *H. capensis*, male, South Africa (E.J. van Nieukerken); *Heliozela* + *Tyriozela* group: *Heliozela* sp., Australia (D. Carman); *Pseliastis* group: *Pseliastis* sp., Australia (L. Milla); *Hoplophanes* group: *Hoplophanes* sp., Australia (D.A. Young).

discoidal cell present in the forewing, which separates it from other species placed in *Antispila*. Nielsen (1980) suggested the presence of an interapodemal process in the female as an apomorphy for *Antispila*, although a reduced similar structure occurs in *Antispilina* as well. Within *Antispila sensu stricto*, male androconial structures on the forewings and hindwings are also frequently present (van Nieukerken

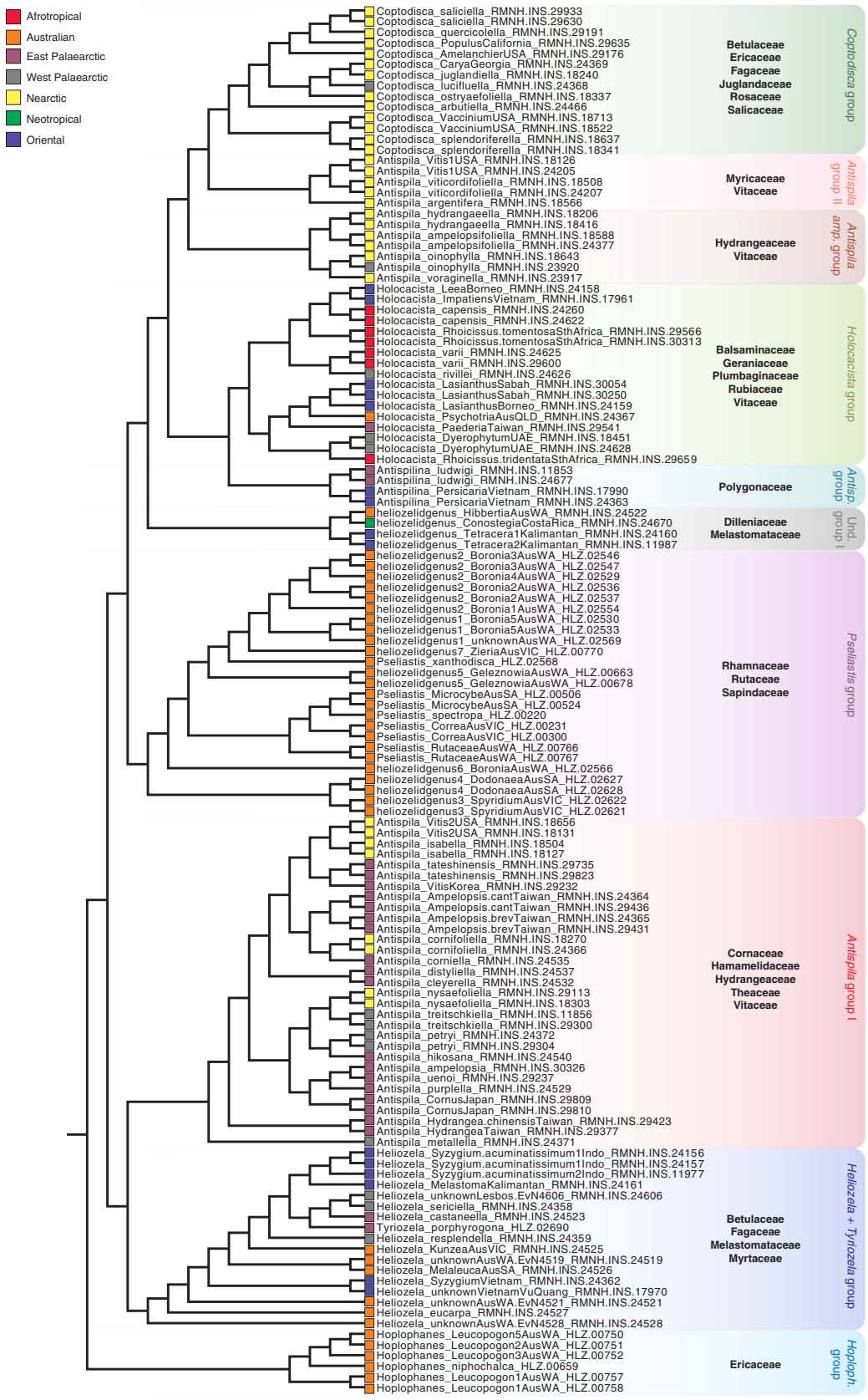

Fig. 4. ML cladogram inferred from four genes with biogeographical region of each Heliozelidae specimen indicated by the colour of rectangle at the branch tip. Clades recovered in highlight. Major host plant families are listed next to each clade.

et al., 2012).

The species assigned to the *Antispila ampelopsifoliella* group in an earlier study (van Nieukerken et al., 2012, Fig. 2c) fall into two strongly supported clades, one including *A. viticordifoliella* and *A.* "*Vitis1USA*", and the other including the remaining named species in the *A. ampelopsifoliella* group. The species in the *A. ampelopsifoliella* clade differ from the first by the presence of an apical spot on the forewing. Additional morphological and molecular analyses of these clades are

necessary to see whether they form one or two new genera. The placement of the specimen identified as *A. argentifera* is inconsistent between the Bayesian and ML analyses. In the ML phylogeny it groups with *A. viticordifoliella* and *A.* "*Vitis1USA*", while in the Bayesian phylogeny it groups with *Coptodisca*, but neither placement is strongly supported. Specimens reared from *Myrica* and *Comptonia* (Myricaceae) have been identified as *A. argentifera* on the basis of their external features, in agreement with Braun's (1927) description and photographs of the type (provided by C. Eiseman). However, the leaf-mines most closely resemble those of *Coptodisca*, of which one unnamed species also feeds on *Myrica* and *Morella*. A detailed morphological analysis of *A. argentifera* is needed to establish whether it will be placed in a new genus. We have previously suggested that *Antispila* has been used as a 'wastebasket' for heliozelids with similar wing markings (van Nieukerken et al., 2012; van Nieukerken and Geertsema, 2015), a conclusion supported by the results of our current study.

Holocacista is a widespread genus with seven described species (van Nieukerken and Geertsema, 2015). In our present results, the main *Holocacista* group had good support, similar to a previous COI study that recovered a clade consisting of *H. capensis*, *H. rivillei*, *H. varii* and a number of putative new species (van Nieukerken and Geertsema, 2015). The *Holocacista* clade was split into two groups, one including all of the described species and three putative species, and another containing six putative species. Detailed morphological analyses are required to determine whether these species belong to *Holocacista* or require the erection of another genus. Morphological characters for *Holocacista* include the typical, often curved, appendix of the phallic tube and the small epiphysis, which is absent in all related genera (van Nieukerken and Geertsema, 2015).

The *Antispilina* group consisting of two species received strong support. Morphologically, the genus is challenging to define, as it differs from its sister group *Holocacista* mainly by a lack of apomorphies. A clade combining *Holocacista* and *Antispilina* has strong support in the Bayesian analysis (1), but much less so in the ML tree (58). Several undescribed species are known, which may belong to *Antispilina*, but could not be included in this study.

A clade comprising species of the Australian endemic genus *Pseliastis*, including *P. spectropa*, *P. xanthodisca* and other putative species, was strongly supported in our analyses. The species within the putative genera labelled *heliozelidgenus1-7* are all superficially similar to *Pseliastis*, and have been placed within the *Pseliastis* group by every analysis. However, our observations from preliminary morphological examinations suggest that the *Pseliastis* group represents more than one genus. For example, the type species *P. trizona* (not included in this study), *P. spectropa* and other putative *Pseliastis* species share characteristics such as forewings with white fasciae; however, *P. xanthodisca* lacks these typical fasciae and groups most closely with undescribed species of possible new genera, labelled *heliozelidgenus1*, *2* and *7*, although this grouping has very low support. Further detailed examination is required to in order to determine apomorphies that would support the erection of one or more new genera.

The strongly supported *Heliozela* + *Tyriozela* group contains all described and putative *Heliozela* species, as well as the only known *Tyriozela* species, *T. porphyrogona*. *Heliozela* are mostly leaf miners, although the larvae of a few species mine the petiole or midrib of the leaf, or feed in a gall (Davis, 1998). Nielsen (1980) placed *Tyriozela* as sister to *Heliozela* based on several synapomorphies, such as the spear-shaped ovipositor of the females, and suggested a strongly developed scent organ in the male abdomen as a possible autapomorphy for the group. Thus, taxonomic revision of this group will be necessary to determine whether it contains multiple genera or whether *Heliozela* and *Tyriozela* should be collapsed into one genus.

Hoplophanes is another genus endemic to Australia. Based on our analyses of five species, it is a strongly supported monophyletic group that appears to be sister to all other Heliozelidae. The *Hoplophanes* group contains one described (*H. niphochalca*) and four putative species that share a number of characters. Morphologically, *Hoplophanes* are quite distinct; species of this genus have the largest wingspan of all Heliozelidae (up to 16 mm), and females are characterised by a long, pointed ovipositor. Nielsen (1980) proposed the distinct ovipositor shape and large size as possible apomorphies for this group.

Holocacista, *Antispilina*, *Coptodisca* and some North American species of *Antispila*, form a large and strongly supported clade. This clade resembles Nielsen's *Holocacista* group, although he did not include any of these *Antispila*, which he likely never studied, nor *Antispilina*. Morphologically all species in this clade share reduced venation, with five to six terminal branches in the forewing, a strong apomorphy shown in our previous studies (Bernardo et al., 2015; van Nieukerken and Geertsema, 2015; van Nieukerken et al., 2012). On the basis of this character, the genus *Ischnocanaba* also belongs here.

An additional result from our analyses was the inconsistent placement of the outgroup taxa *Nematopogon adansoniella* and *Nemophora degeerella*, both belonging to Adelidae, sister family to Heliozelidae, but to different subfamilies, respectively Nematopogoninae and Adelinae. These two taxa formed a clade in the mitochondrial gene phylogeny, but became separated in the nuclear and four-gene phylogenies. A study of non-dytrisian lineages by Regier et al. (2015) found a similar conflict with *Nematopogon* and the rest of the Adelidae, which occurred when comparing non-synonymous (Degen1 dataset) versus all nucleotide changes (nt123 dataset). In their results, non-synonymous changes placed *Nematopogon* outside the Adelidae. They reported that this conflict was not due to compositional heterogeneity in the three-nucleotide dataset. Our results suggest a conflict in phylogenetic signal between mitochondrial and nuclear genes, and that a more in-depth analysis is required to fully resolve the relationship of *Nematopogon* to Adelidae: Adelinae and Heliozelidae.

4.2. Biogeography and host associations

Antispila is a widespread genus within the family, occurring in several biogeographical regions (Fig. 4). *Antispila sensu stricto*, here labelled *Antispila* group I, includes the West Palaearctic type species *A. metallella* along with various European, North American, Asian and one African species (van Nieukerken and Geertsema, 2015). By contrast, the species within *Antispila* group II and *A. ampelopsifoliella* group occur exclusively in North America and are most closely related to *Coptodisca*, which, apart from one recent introduction to Europe (Bernardo et al., 2015), is endemic to the Nearctic and Neotropical regions. These Nearctic "*Antispila*", plus *Coptodisca*, form a clade with modest support in our Bayesian analysis.

Several of the currently described cosmopolitan or Palaearctic genera exhibit greater diversity in the southern hemisphere and in Asia than previously known. This includes *Heliozela*, a widespread genus, occurring on most continents. Many of the recently discovered *Heliozela* species occur in the Oriental and Australian regions, for example, the putative species *H.* "*SyzygiumVietnam*" from Vietnam, *H.* "*MelastomaKalimantan*" from Borneo, and *H.* "*KunzeaAusVIC*" from Australia. Similarly, the monotypic genus *Antispilina*, represented by *A. ludwigi*, currently only known to occur in central Europe, appears to be more widespread in eastern and South-East Asia. One undescribed species included in our analysis, *A.* "*PersicariaVietnam*", is found in Vietnam, and another unnamed one in Japan, indicating that the distribution of this genus is broader than currently appreciated. While the type species of *Holocacista*, *H. rivillei*, and a few others occur in Europe and Central Asia, much of the recently discovered diversity is in South Africa, South-East Asia and Australia (van Nieukerken and Geertsema, 2015).

Two of the currently known heliozelid genera, *Pseliastis* and *Hoplophanes*, occur only in Australia. *Pseliastis*, with just three described species, was thought to be endemic to the state of Tasmania (Common, 1990). However, based on our field collections over the last few years, *Pseliastis* species are widespread and diverse in the southern half of

Australia, extending from the alpine regions of Tasmania and Victoria to sub-tropical Queensland and the dry inland of Western Australia. *Hoplophanes* is also far more diverse than currently described, with many of the recently discovered species occurring in the floristically diverse south-western region of Western Australia.

Heliozelidae feed on a wide variety of plant families. *Antispila sensu stricto* feed mainly on Cornaceae and Vitaceae, but also Hydrangeaceae, Pentaphylacaceae and Hamamelidaceae. The species assigned to the *Antispila ampelopsifoliella* group feed on Vitaceae, with the exception of *A. hydrangaeella*, which produces mines on *Hydrangea arborea* (Hydrangeaceae). Braun (1927) suggested that *A. argentifera* was responsible for long linear mines on *Betula*. We think, however, that these mines belong to the incurvariid *Phylloporia bistrigella* and that *A. argentifera* is responsible for the mines in several Myricaceae. The major host families of *Heliozela* species are Myrtaceae, Fagaceae and Betulaceae, but also include Melastomataceae and Vitaceae (van Nieukerken et al., 2012, and our unpublished data). The host plant and feeding mode of *Tyriozela* remain unknown. However, in our study, *Tyriozela* consistently grouped with the Palaearctic Fagaceae feeders, which form a well-supported clade in our analysis, while other clades within *Heliozela* were not well resolved. *Holocacista* species are all leaf miners, feeding mainly on Vitaceae and Rubiaceae, but host records also include Anacardiaceae, Balsaminaceae, Geraniaceae and Plumbaginaceae. The Rubiaceae feeders form a relatively well-supported clade in *Holocacista*, which is sister to the species pair *Holocacista* "*Rhoicissus.-tridentataSthAfrica*" and *Holocacista* "*DyerophytumUAE*". The undescribed group I, which forms a weakly supported clade within the main cosmopolitan clade, includes four unnamed species from Borneo, Costa Rica and Western Australia and was previously considered to belong to *Holocacista* (van Nieukerken and Geertsema, 2015). However, our current results suggest that this placement needs to be reconsidered. This clade contains multiple species that occur in the southern part of Australia; most, if not all, of which are associated with Dilleniaceae (our unpublished data). The species from Borneo also feed on a Dilleniaceae, whereas the single Costa Rican species feed on Melastomataceae. These species appear not to belong to any known genera, further highlighting the need for additional studies.

Notably, more geographically restricted genera appear to feed on only one or two plant families. For example, all three species in *Antispilina* feed on herbaceous Polygonaceae. Similarly, our observations suggest that *Hoplophanes* species are restricted to southern Australia, and are predominantly associated with plants in the Ericaceae family. Several species were found to have larvae feeding in galls in the growing tips of Ericaceae. Likewise, all of the examined *Pseliastis* species (including *P. xanthodisca*) and most species grouped in the same clade appear to be associated with host plants in the family Rutaceae. The exceptions are a few species found on *Dodonaea* (Sapindaceae) and *Spyridium* (Rhamnaceae). However, *Coptodisca*, which is restricted to Nearctic and Neotropic regions, feeds on a wide range of Eudicot hostplant families, including Rosaceae, Betulaceae, Fagaceae, Salicaceae, Juglandaceae, Rhamnaceae, Combretaceae, Rhizophoraceae and Ericaceae. Interestingly, *C. lucifluella* has shifted hosts from *Carya* to *Juglans* (both Juglandaceae) since its introduction to Italy (Bernardo et al., 2015), providing a recent example of Heliozelidae expanding their host range.

Overall, it is remarkable that several host families have been colonized several times by Heliozelidae. Notably, Vitaceae serve as hostplants for at least one species of *Heliozela* (not sampled here, van Nieukerken et al, 2012), and several species in *Antispila* group I, *Holocacista*, *Antispila ampelopsifoliella* group and *Antispila* group II. At least two heliozelid genera have species feeding on Fagaceae, Rhamnaceae, Hydrangeaceae, Myricaceae, Ericaceae and Balsaminaceae. This pattern suggests that these plant families share properties that make them suitable to be colonized by Heliozelidae. Although our results do not support their conclusion, van Nieukerken et al. (2012) suggested that Vitaceae could be the ancestral hostplants of Heliozelidae. However, it is important to note that several Heliozelidae are potential pests for grapevines, and that local Vitaceae feeding species may colonize commercial grapevines (van Nieukerken and Geertsema, 2015). The patterns we have observed in Heliozelidae resemble those in the larger family of leafminers Nepticulidae (Doorenweerd et al., 2016), with the notable exception of Vitaceae being completely absent from the host record of Nepticulidae. Like in Nepticulidae, the phylogeny of Heliozelidae in no way mirrors that of angiosperms.

5. Conclusions

Although our study consistently recovered five major Heliozelidae clades, the relationships between these groups lacked statistical support and remain unresolved. This is partly due to the number and nature of the genes used. Two mitochondrial and two nuclear genes provide a good first estimate of the phylogenetic relationships within Heliozelidae but are insufficient to resolve relationships between clades that diverged as early as the Late Cretaceous, based on estimates by Wahlberg et al. (2013). Increasing the number of nuclear genes is likely to provide more phylogenetic information to resolve these older nodes. A previous study based on 19 genes was found to be insufficient to determine the relationships outside the family level (Regier et al., 2015), therefore, to resolve Heliozelidae phylogeny and to confidently place the family, a broader gene sampling method is recommended. Advanced phylogenomic methods such as transcriptome sequencing (Bazinet et al., 2017) or anchored hybrid enrichment (Breinholt et al., 2017) could well be suitable for resolving this issue.

Based on the results of our study, the majority of the undescribed diversity both at the genus and species level appears to occur in the southern hemisphere, which may suggest a southern origin of the family. Four of the genera (*Holocacista*, *Heliozela*, *Pseliastis* and *Hoplophanes*) have been recorded in Australia, with the two described genera (*Hoplophanes* and *Pseliastis*) showing high species diversity in the southern part of the continent. *Holocacista* is another genus with undescribed high diversity predominantly in South Africa and South-east Asia. Three of the genera not included in this study (*Plesiozela*, *Phanerozela* and *Ischnocabana*) are recorded only from the southern hemisphere, while a fourth genus (*Microplitica*) occurs in both India and Indonesia. The placement of *Plesiozela*, the putative sister group to all other Heliozelidae and endemic to South America, would provide important evidence towards establishing the origin of Heliozelidae.

Our study presents the first phylogenetic framework of the Heliozelidae at the global level. It provides strong evidence for major evolutionary clades, creating a preliminary framework and a starting point to fully resolving the relationships within the family. Our study also highlights unexpected heliozelid diversity, in particular in the southern hemisphere, and the need for a broad taxonomic revision of the family.

Acknowledgements

The authors acknowledge the support of many people and organisations during this study. For material: Mario Baldessari (San Michele al'Adige, Italy), Charley Eiseman (Northfield, MA, USA), Henk Geertsema (Stellenbosch, South Africa), Toshiya Hirowatari (Fukuoka, Japan), Dave G. Holden (Vancouver, Canada), Marianne Horak, Ted Edwards, Youning Su and Andreas Zwick (ANIC, Canberra, Australia), Michael Hull (Caernarfon, UK), Bongwoo Lee (Pocheon City, S. Korea), Anna Mazurkiewicz (Warszawa, Poland), Marko Mutanen (Oulu, Finland), Kenji Nishida (San José, Costa Rica), Arnold Schreurs (Kerkrade, Netherlands), Michael Stiller (Roodeplaat, South Africa), Paolo Triberti (Verona, Italy), David Wagner (Storrs, CT, USA). For lab support: Kevin Beentjes, Frank Stokvis (Leiden, Netherlands), Tom McConville (Melbourne, Australia). For general support, rearing and dissection: Kees van den Berg (Leiden, Netherlands), Terry and Jen Haddon (Nannup, Australia). For computational resources: Melbourne

Bioinformatics at the University of Melbourne (UOM0024). For SEM: Qike Wang (Melbourne, Australia). We would like to thank the staff of the following Australian state government departments for their assistance obtaining research permits and for their generous advice and help they have provided in the field: Department of Parks and Wildlife (Western Australia), Parks Victoria (Victoria, Australia), Department of Primary Industries, Parks, Water and Environment (Tasmania, Australia) and Office of Environment and Heritage (NSW, Australia). We acknowledge Charley Eiseman (Northfield, MA, USA), Rudolf Bryner (Biel, Switzerland), and Darren Carman (Australia) for allowing us to use their high quality photographs, which remain copyright to their respective owners. This work was supported by the following grants: Hermon Slade Foundation (HSF15/6), The Holsworth Wildlife Research Endowment - Equity Trustees Charitable Foundation (HOLSW2015-1-F067 and HOLSW2016-R1-F009), and Australian Lepidoptera Research Endowment. L. Milla has been supported by an Australian Postgraduate Award (APA), provided by the Australian government.

Appendix A. RNA bait and capture protocol

A.1. RNA bait creation

The PCR products obtained for the COI, COII, H3, and 28S from 12 "founder" specimens were sheared on a Covaris S220 instrument at 200 cycles, peak power 140, duty 10 for 900 sec, generating fragments in the 100–200 base pair range. T7 oligo adapter libraries were constructed using these fragments as a template and followed the same procedure described in Carpenter et al. (2013). Except where indexed adapters are described, these were replaced with the following T7 adapters: 5′-GAT CTTAGGCTAGAGTACTAATACGACTCACTATAGGGT-3′ and 5′-CTAGAA TCCGATCTCATGATTATGCTGAGTGATATCCC-3′.

RNA baits from these PCR libraries were generated via an *in vitro* transcription (IVT) reaction. The IVT was set up as follows: 500 ng of input T7 adapter library was used in 50 µl reaction consisting of 5× NASBA buffer (185 mM Tris-HCl pH 8.5, 93 mM $MgCl_2$, 185 mM KCl, 46% DMSO), 5 mM DTT, BSA (100ug/ml), 2.5 mM NTP mix (10 mM CTP/GTP/ATP, 6.5 mM UTP, 3.5 mM biotin-16-UTP), 0.6 units T7 RNA polymerase, 0.0006 units pyrophosphatase, and 1 unit Superase-In RNase inhibitor. The reaction was incubated at 37 °C for 16 h, then treated with 0.04 units of TURBO DNase. The IVT generated biotinylated RNA baits were purified using the Macherey-Nagel NucleoSpin RNA XS purification kits, as described by the manufacturer. All IVT libraries were assessed using the RNA screentape on the Agilent Tapestation. The COI and COII bait libraries were pooled in equimolar ratios, as were the H3 and 28S libraries to be used in the subsequent DNA capture procedure.

A.2. DNA capture

100 ng from each RNA bait pool was combined with 100 ng of a specimen library pool for a final volume of 16 µl, and the solution was incubated for a minimum of 20 h at 60 °C. During this process, DNA fragments annealed to RNA baits with homologous sequences. The resulting RNA/DNA hybrid molecules were bound with Dynabeads© Streptavidin C1 beads on a magnetic stand. Another 50 µl of Dynabeads were washed twice with bead wash buffer (1M NaCl, 10 mM Tris-HCl pH7.5, 1 mM EDTA, 0.01% Tween20) and resuspended in a final wash volume of 100 µl. The washed beads were added to each of the library pools at a 2:1 volume ratio. The mixes were left for 30 min at room temperature to allow the Streptavidin/biotin interaction to form. The library pools were then placed on a magnetic stand to separate the RNA/DNA hybrids. The supernatant was removed and 200 µl low stringency wash solution (1× SSC, 0.1% SDS, 0.01% Tween20) added and the pools vortexed. This solution was left for 5 min and the beads separated on a magnetic stand. Preheated low stringency wash solution (60 °C) was used for 3 further washes of 5 min each. After the final separation on the magnetic stand the beads were incubated with 50 µl of 0.2 M NaOH for 10 min to denature the RNA/DNA hybrid molecules. The reaction was neutralised with an equal volume of 1 M Tris-HCl pH8.0. The reaction was then placed on a magnetic stand for a final time. 100 µl of single stranded DNA solution were removed for clean up using 1.8× AMPure beads as previously described. The clean single stranded DNA was then amplified using the following conditions: 95 °C for 3 min, followed by 16 cycles of 98 °C for 30 s, 60 °C for 15 s, 72 °C for 30 s with a final extension step of 72 °C for 5 min. The size distribution of the captured molecules was ascertained using the D1000 screentape on the Agilent Tapestation. The captured products were made into sequencing libraries following the Illumina TruSeq DNA Sample Preparation - Low Sample protocol.

Appendix B. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2017.12.004.

References

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.

Bazinet, A.L., Mitter, K.T., Davis, D.R., van Nieukerken, E.J., Cummings, M.P., Mitter, C., 2017. Phylotranscriptomics resolves ancient divergences in the Lepidoptera. Syst. Entomol. 42, 305–316. http://dx.doi.org/10.1111/syen.12217.

Bernardo, U., Sasso, R., Gebiola, M., Viggiani, G., 2012. First record of a walnut shield bearer Coptodisca (Lepidoptera: Heliozelidae) in Europe. J. Appl. Entomol. 136, 638–640. http://dx.doi.org/10.1111/j.1439-0418.2011.01693.x.

Bernardo, U., van Nieukerken, E., Sasso, R., Gebiola, M., Gualtieri, L., Viggiani, G., 2015. Characterization, distribution, biology and impact on Italian walnut orchards of the invasive North-American leafminer *Coptodisca lucifluella* (Lepidoptera: Heliozelidae). Bull. Entomol. Res. 105, 210–224. http://dx.doi.org/10.1017/S0007485314000947.

Braun, A.F., 1927. Descriptions of new Microlepidoptera. Trans. Am. Entomol. Soc. 53, 191–199.

Breinholt, J.W., Earl, C., Lemmon, A.R., Lemmon, E.M., Xiao, L., Kawahara, A.Y., 2017. Resolving relationships among the megadiverse butterflies and moths with a novel pipeline for Anchored Phylogenomics. Syst. Biol. syx048. http://dx.doi.org/10.1093/sysbio/syx048.

Carpenter, M.L., Buenrostro, J.D., Valdiosera, C., Schroeder, H., Allentoft, M.E., Sikora, M., Rasmussen, M., Gravel, S., Guillen, S., Nekhrizov, G., 2013. Pulling out the 1%: whole-genome capture for the targeted enrichment of ancient DNA sequencing libraries. Am. J. Hum. Genet. 93, 852–864.

Caterino, M.S., Sperling, F.A., 1999. *Papilio* phylogeny based on mitochondrial Cytochrome Oxidase I and II genes. Mol. Phylogenet. Evol. 11, 122–137. http://dx.doi.org/10.1006/mpev.1998.0549.

Common, I.F.B., 1990. Moths of Australia. Melbourne University Press, Carlton.

Davis, D.R., 1998. The Monotrysian Heteroneura. In: Kristensen, N.P. (Ed.), Lepidoptera: Moths and Butterflies. 1. Evolution, Systematics, and Biogeography. Handbuch der Zoologie/Handbook of Zoology 4(35). Walter De Gruyter Inc, Berlin, pp. 65–90.

Dowton, M., Austin, A.D., 1998. Phylogenetic relationships among the microgastroid wasps (Hymenoptera: Braconidae): combined analysis of 16S and 28S rDNA genes and morphological data. Mol. Phylogenet. Evol. 10, 354–366. http://dx.doi.org/10.1006/mpev.1998.0533.

Doorenweerd, C., van Nieukerken, E.J., Hoare, R.J.B., 2016. Phylogeny, classification and divergence times of pygmy leafmining moths (Lepidoptera: Nepticulidae): the earliest lepidopteran radiation on Angiosperms? Syst. Entomol. 42, 267–287. http://dx.doi.org/10.1111/syen.12212.

Emmet, A.M., 1976. Heliozelidae. In: Heath, J. (Ed.), The moths and butterflies of Great Britain and Ireland, 1. Micropterigidae – Heliozelidae. Blackwell Scientific Publications Ltd, Oxford & London, pp. 300–306.

Friedlander, T.P., Regier, J.C., Mitter, C., Wagner, D.L., Fang, Q., 2000. Evolution of heteroneuran Lepidoptera (Insecta) and the utility of dopa decarboxylase for Cretaceous-age phylogenetics. Zool. J. Linn. Soc. 130, 213–234. http://dx.doi.org/10.1006/zjls.1999.0217.

Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-seq data without a reference genome. Nat. Biotechnol. 29 (7), 644–652. http://dx.doi.org/10.1038/nbt.1883. PubMed PMID: 21572440.

Hebert, P.D., Cywinska, A., Ball, S.L., 2003. Biological identifications through DNA barcodes. Proc. R. Soc. Lond. B Biol. Sci. 270, 313–321. http://dx.doi.org/10.1098/rspb.2002.2218.

Heppner, J.B., 1991. Faunal regions and the diversity of Lepidoptera. Trop. Lepidopt. 2.

Karsholt, O., Kristensen, N.P., 2003. *Plesiozela*, gen. nov. from temperate South America:

apparent sister-group of the previously known Heliozelidae (Lepidoptera: Incurvarioidea: Heliozelidae). Invertebr. Syst. 17, 39–46. http://dx.doi.org/10.1071/IS02047.
Katoh, K., Standley, D.M., 2013. MAFFT Multiple Sequence Alignment Software Version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Kozlov, A.M., Aberer, A.J., Stamatakis, A., 2015. ExaML version 3: a tool for phylogenomic analyses on supercomputers. Bioinformatics btv184.
Kuroko, H., 1961. The genus *Antispila* from Japan, with descriptions of seven new species (Lepidoptera, Heliozelidae). Esakia: occasional papers of the Hikosan Biological Laboratory in Entomology 3, 11–24.
Kuroko, H., 1982. Heliozelidae. In: Inoue, H., Sugi, S., Kuroko, H., Moriuti, S., Kawabe, A., Owada, M. (Eds.), Moths of Japan. Kodansha, Tokyo, Japan, pp. 50–51 448.
Kuroko, H., 1987. Three new species of the genus *Antispila* (Lepidoptera: Heliozelidae) from Japan (Collection of Papers presented to celebrate Prof. Hiroshi Inoue's Seventieth Birthday). Tinea 12, 109–117.
Lafontaine, J.D., 1973. Eastern North American species of *Antispila* (Lepidoptera: Heliozelidae) feeding on *Nyssa* and *Cornus*. Canad. Entomol. 105, 991–994.
Lafontaine, J.D., 1974. A new species of *Coptodisca* (Heliozelidae) from Mississippi on farkleberry (Vaccinium arboreum). J. Lepidopterists' Soc. 28, 126–130.
Lee, B.-W., Hirowatari, T., 2013. Heliozelidae. In: Hirowatari T, N.Y., Sakamaki Y, Kishida Y (Ed.), The standard of moths in Japan. Gakken Kyoiku Shuppan, Tokyo, pp. 99–102.
Lee, B.-W., Hirowatari, T., Kuroko, H., 2006a. A new species of *Antispila* (Lepidoptera: Heliozelidae) from the Ryukyus, Japan, with descriptions of immature stages. Ann. Entomol. Soc. Am. 99, 1019–1027.
Lee, B.-W., Hirowatari, T., Kuroko, H., Arita, Y., 2006b. A new species of the genus *Antispila* Huebner (Lepidoptera, Heliozelidae), feeding on *Cleyera japonica* (Theaceae) from Japan. Trans. Lepidopterol. Soc. Japan 57, 7–12.
Liu, Y., Schröder, J., Schmidt, B., 2013. Musket: a multistage k-mer spectrum-based error corrector for Illumina sequence data. Bioinformatics 29, 308–315.
Maier, C.T., 1988. Life cycle of *Coptodisca negligens* (Lepidoptera: Heliozelidae) on cranberry. J. Econ. Entomol. 81, 497–500.
Mutanen, M., Itamies, J., Kaila, L., 2007. *Heliozela resplendella* (Stainton, 1851) and *H. hammoniella* Sorhagen, 1885: two valid species distinguishable in the genitalia of both sexes and life histories (Heliozelidae). Nota Lepidopterol. 30, 79–92.
Mutanen, M., Wahlberg, N., Kaila, L., 2010. Comprehensive gene and taxon coverage elucidates radiation patterns in moths and butterflies. Proc. Roy. Soc. Lond. B: Biol. Sci. 277, 2839–2848. http://dx.doi.org/10.1098/rspb.2010.0392.
Nielsen, E.S., 1980. The Incurvarioid genera of the world (Lepidoptera: Incurvarioidea). University of Copenhagen unpublished thesis, typescript.
Nielsen, E.S., Davis, D.R., 1985. The first southern hemisphere prodoxid and the phylogeny of the Incurvarioidea (Lepidoptera). Syst. Entomol. 10, 307–322.
Ogden, T.H., Whiting, M.F., 2003. The problem with “the Paleoptera problem:” sense and sensitivity. Cladistics 19, 432–442. http://dx.doi.org/10.1016/S0748-3007(03)00083-5.
Opler, P.A., 1971. Seven new lepidoterous leaf-miners associated with *Quercus agrifolia* (Heliozelidae, Gracilariidae). J. Lepidopter. Soc. 25, 194–211.
Regier, J.C., Mitter, C., Kristensen, N.P., Davis, D.R., van Nieukerken, E.J., Rota, J., Simonsen, T.J., Mitter, K.T., Kawahara, A.Y., Yen, S.H., 2015. A molecular phylogeny for the oldest (nonditrysian) lineages of extant Lepidoptera, with implications for classification, comparative morphology and life-history evolution. Syst. Entomol. 40, 671–704. http://dx.doi.org/10.1111/syen.12129.
Robinson, G.S., Ackery, P.R., Kitching, I.J., Beccaloni, G.W., Hernández, L.M., 2008. HOSTS - a database of the world's Lepidopteran hostplants. The Natural History Museum, London http://www.nhm.ac.uk/research-curation/projects/hostplants/.
Sperling, F.A., Landry, J.-F., Hickey, D.A., 1995. DNA-based identification of introduced ermine moth species in North America (Lepidoptera: Yponomeutidae). Ann. Entomol. Soc. Am. 88, 155–162.
Stamatakis, A., 2014. RAxML Version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 1312–1313.
Thorvaldsdóttir, H., Robinson, J.T., Mesirov, J.P., 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Briefings Bioinf. 14, 178–192.
van Nieukerken, E.J., Geertsema, H., 2015. A new leafminer on grapevine and *Rhoicissus* (Vitaceae) in South Africa within an expanded generic concept of *Holocacista* (Insecta, Lepidoptera, Heliozelidae). ZooKeys 41–97.
van Nieukerken, E., Kaila, L., Kitching, I., Kristensen, N.P., Lees, D., Minet, J., Mitter, J., Mutanen, M., Regier, J., Simonsen, T., Wahlberg, N., Yen, S., Zahiri, R., Adamski, D., Baixeras, J., Bartsch, D., Bengtsson, B., Brown, J., Bucheli, R., Davis, D., Prins, J.D., Prins, W.D., Epstein, M., Gentili-Poole, P., Gielis, C., Hättenschwiler, P., Hausmann, A., Holloway, J., Kallies, A., Karsholt, O., Kawahara, A., Koster, S., Kozlov, M., Lafontaine, J., Lamas, G., Landry, J., Lee, S., Nuss, M., Park, K.-T., Penz, C., Rota, J., Schmidt, B., Schintlmeister, A., Sohn, J., Solis, M., Tarmann, G., Warren, A., Weller, S., Yakovlev, Y., Zolotuhin, V., Zwick, A., 2011. Order Lepidoptera Linnaeus, 1758. In: Zhang, Z.-Q. (Ed.), Animal biodiversity: An outline of higher-level classification and survey of taxonomic richness. Zootaxa, 3148, pp. 212–221.
van Nieukerken, E.J., Wagner, D., Baldessari, M., Mazzon, L., Angeli, G., Girolami, V., Duso, C., Doorenweerd, C., 2012. *Antispila oinophylla* new species (Lepidoptera, Heliozelidae), a new North American grapevine leafminer invading Italian vineyards: taxonomy, DNA barcodes and life cycle. ZooKeys 170, 29–77.
van Nieukerken, E.J., Lees, D.C., Doorenweerd, C., Koster, J.C., Bryner, R., Schreurs, A., Timmermans, M.J.T.N., Sattler, K., 2017. Two European Cornus feeding leafmining moths, *Antispila petryi* sp. rev. and *A. treitschkiella* (Lepidoptera, Heliozelidae): an unjustified synonymy and overlooked range expansion. Nota Lepidopterol submitted.
Wahlberg, N., Wheat, C.W., Peña, C., 2013. Timing and Patterns in the Taxonomic Diversification of Lepidoptera (Butterflies and Moths). PLoS ONE 8, e80875.
Whiting, M.F., Carpenter, J.C., Wheeler, Q.D., Wheeler, W.C., 1997. The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18S and 28S ribosomal DNA sequences and morphology. Syst. Biol. 46, 1–68.
Zerbino, D.R., Birney, E., 2008. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 18, 821–829. http://dx.doi.org/10.1101/gr.074492.107.