Hiding in plain sight: DNA barcoding suggests cryptic species in all ‘well-known’ Australian flower beetles (Scarabaeidae: Cetoniinae)

Andrew Mitchell Corresp., 1 , **Christian H Moeseneder** 1, 2 , **Paul M Hutchinson** 3

1 Australian Museum Research Institute, Australian Museum, Sydney, New South Wales, Australia
2 Oceans and Atmosphere Flagship, Commonwealth Scientific and Industrial Research Organisation, St Lucia, Queensland, Australia
3 Quarantine Western Australia, Department of Agriculture and Food Western Australia, Perth, Western Australia, Australia

Corresponding Author: Andrew Mitchell
Email address: andrew.mitchell@austmus.gov.au

DNA barcode data is presented for Australian cetoniine flower beetles to aid with species discovery and guide revisionary taxonomy. Sequences of the COI gene’s DNA barcode region were acquired from 284 cetoniine specimens, covering 68 described species and 33 genera. This equates to 48% of the known species and 83% of the genera which occur in Australia. Results suggest up to 27 putative undescribed species in our sample, only 11 of which were suspected to be undescribed before this study, leaving 16 unexpected (“cryptic”) species. The Australian cetoniine fauna may hence be increased by up to 19%. An unanticipated result of the work is that each of the five most visible and commonly collected Australian cetoniine species, *Eupoecila australasiae* (Donovan, 1805), *Neorrhina punctatum* (Donovan, 1805), *Glycyphana (Glycyphaniola) stolata* (Fabricius, 1871), *Chondropyga dorsalis* (Donovan, 1805) and *Bisallardiana gymnopleura* (Fischer, 1823), have unexpectedly high diversity in DNA barcode sequences and were consequently split into multiple taxa, possibly indicating the presence of cryptic species.

Hiding in plain sight: DNA barcoding suggests cryptic species in all ‘well-known’ Australian flower beetles (Scarabaeidae: Cetoniinae)

Andrew Mitchell[1], Christian H. Moeseneder[1,2], Paul M. Hutchinson[3]

[1] Australian Museum Research Institute, Australian Museum, Sydney, New South Wales, Australia

[2] Oceans and Atmosphere Flagship, Commonwealth Scientific and Industrial Research Organisation, St Lucia, Queensland, Australia

[3] Quarantine Western Australia, Department of Agriculture and Food Western Australia, Perth, Western Australia, Australia

Corresponding Author:

Andrew Mitchell[1]

Australian Museum, 1 William Street, Sydney, NSW, 2010, Australia

Email address: andrew.mitchell@austmus.gov.au

Abstract

DNA barcode data is presented for Australian cetoniine flower beetles to aid with species discovery and guide revisionary taxonomy. Sequences of the COI gene’s DNA barcode region were acquired from 284 cetoniine specimens, covering 68 described species and 33 genera. This equates to 48% of the known species and 83% of the genera which occur in Australia. Results suggest up to 27 putative undescribed species in our sample, only 11 of which were suspected to be undescribed before this study, leaving 16 unexpected (“cryptic”) species. The Australian cetoniine fauna may hence be increased by up to 19%. An unanticipated result of the work is that each of the five most visible and commonly collected Australian cetoniine species, *Eupoecila australasiae* (Donovan, 1805), *Neorrhina punctatum* (Donovan, 1805), *Glycyphana*

(Glycyphaniola) stolata (Fabricius, 1781), *Chondropyga dorsalis* (Donovan, 1805) and *Bisallardiana gymnopleura* (Fischer, 1823), have unexpectedly high diversity in DNA barcode sequences and were consequently split into multiple ~~taxa~~, possibly indicating the presence of cryptic species.

Introduction

The cosmopolitan scarab beetle subfamily Cetoniinae, or flower beetles, comprises 4,273 species worldwide in 485 genera (Krajčík, 2012). The common species are well-represented in institutional and private collections in Australia, and one early collector, F.P. Dodd, arranged hundreds of colourful specimens in large display frames for exhibition (Monteith, 2010). Despite their visibility, the taxonomy of Australian cetoniines has been somewhat neglected until recent times, with only 10 works in scientific literature and 16 species described in the 65 years from 1944–2009. Previous taxonomic work on Australian fauna is detailed in Moeseneder *et al.* (2019).

Approximately 75% of the country's cetoniine species are anthophagous. The adults are pollinators of many tree and shrub species (Williams & Adam, 1998) and feed on nectar, pollen (Moore 1987), fruit and honey. A few occasionally consume flower petals (Moore 1987) or sap (unpublished data, P.M.H.). In the remainder of the species males are often in flight, females are sedentary, and adults are rarely or never found on flowers (Hutchinson & Moeseneder, 2013a; Hutchinson & Moeseneder, 2013b; Moeseneder & Hutchinson, 2016; Moeseneder *et al.*, 2019). Before observations of adults of two species as pests in Queensland beehives in recent years (unpublished data, C.H.M.), Australian cetoniines were not known to be harmful to agriculture (Moeseneder *et al.*, 2019). The larvae of most species live in and feed on decaying wood and function as organic recyclers, within standing or fallen trees. The larvae of a several species are found freely in soil (Moeseneder & Hutchinson, 2016).

The Australian cetoniine fauna is relatively depauperate, comprising 141 species (Moeseneder *et al.*, 2019), or 3% of the world fauna on 5% of the land surface. Sixty-eight percent of genera and 90% of species are endemic to the continent. Three of the twelve cetoniine tribes are represented in Australia: Schizorhinini, Cetoniini and Valgini. The Schizorhinini evolved *in situ* (Krikken, 1984) and is the continent's most speciose tribe, with 111 described species. The majority of Schizorhinini, however, occur in the Malay Archipelago and Melanesia.

While attractive and common species are better known, those with unusual characteristics and those that occur in remote regions were often lumped into unnatural genera, primarily in *Schizorhina* Kirby, 1825 (e.g. Macleay, 1863; 1861), *Diaphonia* Newman, 1840 (e.g. Janson, 1873; 1874; 1889) and *Pseudoclithria* van de Poll, 1886 (e.g. Macleay, 1871). Such oddities have been the focus of the authors' (C.H.M. and P.M.H.) past work (detailed in Moeseneder *et al.,* 2019), describing four new genera and seven new species. Based on the published literature and our own observations, we suspected cryptic species to be present in *Diaphonia*, *Pseudoclithria*, *Bisallardiana* Antoine, 2003, *Chondropyga* Kraatz, 1880*, Chlorobapta* Kraatz, 1880 and *Glycyphana* Burmeister, 1842. This DNA barcoding study is a first step towards resolving these taxonomic issues.

DNA barcoding is a widely used tool in taxonomy, with almost 4,000 papers published since its inception 16 years ago. It has overcome many of its early controversies as methods have matured and its utility in taxonomy, ecology and conservation has become widely appreciated (DeSalle & Goldstein, 2019). Lepidopterists, in particular, have embraced DNA barcoding as the very large datasets that have been developed for the faunas of North America, Europe and Australia facilitate routine species identification and aid taxonomic research in many families. Much progress is being made with beetles as well, e.g. Hendrich *et al.* (2015) published 16,000 barcodes for 3,500 European species.

The few DNA-based studies to date that have included the Cetoniinae are summarized in Table 1 (refer to Methods for the search methodology used). Most studies were either higher-level phylogenies which used a single sample per species (e.g. Ahrens *et al.*, 2008; 2011; Gunter *et al.,* 2016; McKenna *et al.,* 2015; Sipek *et al.*, 2016) or studies of a single genus. A notable exception was the DNA barcoding study of Hendrich *et al.* (2015) which included 70 samples from 14 European cetoniine species. A search of the BOLD Public Data Portal for "Cetoniinae Australia" (on 19 March 2020) revealed only 29 barcodes of Australian cetoniines. Fifteen of these were from Gunter (2016) and identified only to genus level. Twelve of the remaining barcodes represent three common species which we sampled in this study. Approximately half of these were identified using the BOLD Identification Engine (IDE). Of the remaining two barcodes, one species had been misidentified at the generic level. None of these 29 sequences were included in our study as they did not add any diversity to our study, and we did not have access to the specimens.

Before our submissions, a search in GenBank (10 October 2019) for "Cetoniinae" and ("Cytochrome c oxidase subunit I" or "COI" or "CO1" or "COX1") found 1260 records, of which 47% are from two genera, *Osmoderma* Lepeltier & Serville 1828 and *Protaetia* Burmeister 1842. Our submitted records increase this number by 23%.

The goals of this study are to build a foundational DNA barcode library for Australian Cetoniinae with the purpose of aiding the discovery of Australian species, anchoring the process of revising their taxonomy and facilitating identification of larvae.

Materials & Methods

Insect specimens and taxonomy

Our study covers Australia, including its external territories, although of these, cetoniines are known to be present only on Christmas Island and the Cocos (Keeling) Islands. Collecting permits were provided by the Queensland Department of Environment and Science (permit numbers WITF18701717, WITK15549915, WITK10612112, WITK05498008, TWB/02/2015, TWB/03B/2012, TWB/04A/2010, TWB/27B/2010, TWB/27A/2008 and TWB/26/2008), the

NSW National Parks and Wildlife Service (permit number SL100610) and the Western Australia Department of Biodiversity, Conservation and Attractions (permit numbers F025000050, 08-000563-2 and SF008817).

Images of specimens were taken with a Nikon D5100 camera, a Micro Nikkor 105 mm macro lens and four 3-Watt LED lights. The camera was controlled by Nikon Camera Control Pro 2 version 2.28.2 from a laptop computer. Focus-stacking was performed with a unit built by C.H.M. (Moeseneder, 2017).

Male genitalia were removed by separating the abdomen from the thorax by sliding Dumont #5 tweezers in the gap between abdomen and thorax at several points, usually requiring the metatibia to be forced slightly away from the abdomen. The aedeagus was then extracted from the abdomen and the abdomen re-attached with cyanoacrylate glue. The aedeagus was mounted with a micro pin into a small foam piece which was pinned on the same pin as the specimen. A small amount of cyanoacrylate glue was applied where the micro pin pierced the aedeagus to keep it from rotating or being lost. The method allowed rapid extraction without externally visible damage, storage of the aedeagus with the specimen, and three-dimensional inspection of the aedeagus at any time without obscuring any part. For species identification, all collections which are listed in the abbreviations were used.

Abbreviations of institutions and museums:

AFBRC -	Australian Flower Beetle Research Collection, Redland Bay, Qld, Australia
AMS -	Australian Museum, Sydney, NSW, Australia
ANIC -	Australian National Insect Collection, Canberra, Australia
CSIRO-	Commonwealth Scientific and Industrial Research Organisation
DAF -	Department of Agriculture and Fisheries, Queensland
NHML-	Natural History Museum, London, United Kingdom
PMH -	Paul Hutchinson, Beckenham, Perth, WA, Australia
QM -	Queensland Museum, Brisbane, Qld, Australia
SAM -	South Australian Museum, Adelaide, SA, Australia
WAM -	Western Australian Museum, Perth, WA, Australia

Further abbreviations: ACT - Australian Capital Territory, NSW - New South Wales, NT - Northern Territory, Qld - Queensland, SA - South Australia, Vic. - Victoria, WA - Western Australia.

Collection data and images of each specimen were uploaded to the Barcode of Life Data System (BOLD; Ratnasingham & Hebert, 2007) as public project *AUCET, Australian Cetoniinae*.

Where potentially undescribed species are mentioned in this work, they are identified by a code in the format sp_xxx_chm where the ‘xxx’ is a unique code. Their taxonomy will be resolved in future studies.

To find previous taxonomic and phylogenetic studies that produced DNA data for Cetoniinae we: 1) searched for the keywords “Cetoniinae and DNA or molecular” in Web of Science (https://apps.webofknowledge.com), 2) performed a Google search with the same keywords, 3) performed a search of GenBank for Cetoniinae COI sequences, and a subsequent search of Google Scholar for the studies that produced those sequences, and 4) consulted the reference list in each paper that was found.

We use the term ”well-known species” for our subjective measure of those Australian cetoniine species 1) with high numbers of specimens in collections, 2) which have more often been used to represent the subfamily, for example in literature and displays, 3) with larger numbers of records in The Atlas of Living Australia (http://www.ala.org.au), and 4) which are seen by the public in backyards and parks, and hence reported to museums, mentioned in digital media posts (e.g. Flickr, http://www.flickr.com) and citizen science projects (e.g. QuestaGame, http://www.questagame.com).

DNA barcoding

An initial trial round of sampling from both, archival specimens and those collected within the last approximately 10 years, produced a high rate of unsuccessful DNA extractions. Thereafter, the standardized sampling procedure described below was implemented which increased the success rate of DNA sequencing to 98%.

We sampled one to 12 specimens per species (mean = 3.34, median = 3) maximizing geographic coverage where possible. Live adult specimens were collected directly into laboratory-grade ethanol and samples for DNA extraction were taken immediately after death. Sampling was performed by removing the rear left leg with forceps, which were sterilized between samples by wiping with a clean tissue, dipping in 100% ethanol and flaming. A new, sterile surgical blade was used to cut the femur at both apices to exclude the joints. The central part of the femur was cut into two or more fragments to expose muscle tissue. Approximately equal-sized samples were used across all taxa to obtain comparable DNA concentrations. Samples were transferred to a tissue sample plate with sterilized forceps. During this process, all neighbouring wells were kept covered to reduce the chance of contamination. The sampling plate was stored in a freezer at approximately -12°C. Exceptions to this sampling protocol were: 1) the 18 specimens collected in flight intercept traps which were killed in a mixture of propylene glycol and water, and transferred to ethanol after approximately 2-4 weeks, and 2) *Microvalgus* Kraatz, 1883 specimens, where the entire specimen was macerated and used for sequencing. In these cases, the samples were one of a series of specimens collected at the same time, on the same tree and morphologically identical. In each case the series of specimens was kept as reference material.

Since specimen age ranged from 1-22 years, we used the PCR primers and amplification strategy developed by Mitchell (2015) for decades-old insect specimens. In summary, an attempt was made to PCR-amplify a 667-bp fragment of COI. If this was unsuccessful, two shorter overlapping PCR fragments, each approximately 300 bp were amplified, and subsequently reamplified using an internally nested primer on one end. When aligned, the two short fragments yielded 559 bp of contiguous COI sequence within the DNA barcode region.

Sequence trace files were assembled, PCR primers were trimmed, and consensus sequences aligned using Geneious 9.1.8 (Kearse *et al.*, 2012). Trace files and consensus sequences were uploaded to BOLD (http://boldsystems.org/) and are available as public project *Australian Cetoniinae*, project code *AUCET* (http://www.boldsystems.org/index.php/MAS_Management_DataConsole?codes=AUCET). Sequences were also submitted to GenBank as accession numbers XXXXXXXX – XXXXXXXX [Note: sequences have been submitted to GenBank by BOLD staff, but at the time of submission accession numbers were not yet available].

The BOLD platform was used for barcode-specific analyses, including the calculation of intraspecific and within-genus interspecific K2P distances, barcode gap analysis and BIN discordance analysis, i.e., comparison of morphology-based species identifications with Barcode Index Numbers (BINs) which are operational taxonomic units (OTUs) derived using RESL clustering (Ratnasingham & Hebert, 2013). We note, however, that sequences that do not meet all quality criteria, including for length, are not assigned to BINs. Therefore, for a more complete comparison of OTUs based on RESL clustering versus morphospecies, we also performed RESL clustering on all sequences using the "cluster sequences" function on BOLD. Finally, we tested for possible isolation by distance within every species, using the Geographic Distance Correlation tool on the BOLD platform, which calculates a Mantel correlation coefficient for geographic distance between sample localities versus K2P distance, and provides a Mantel test P value.

FaBox v. 1.4.2 (Villesen, 2007) was used to edit sequence names. Phylogenetic analyses were performed on the online science gateway CIPRES v. 3.3 (Miller *et al.*, 2010). Partitionfinder v.2 (Lanfear *et al.*, 2016) was used to select a partitioning scheme and to select the most appropriate models, which, in all cases, was a single data partition and the General Time Reversible model with Gamma-distributed rates and Invariable sites (GTR+G+I). Phylogenetic analyses were performed by Bayesian Inference (BI) using MrBayes v. 3.2.6 (Ronquist *et al.*, 2012) and under maximum likelihood using RAxML v. 8.2.10 (Stamatakis, 2014). The MrBayes analysis was set to run for 20 million generations, with a sample frequency of 1,000, using 2 runs, setting the number of chains to 4. The stopping rule was used to end the analysis when the average standard deviation of split frequencies dropped below 0.01, indicating convergence of the chains. The burnin fraction was set to 0.25. RAxML analysis used the hill climbing algorithm with 1,000 rapid bootstrap replicates. All trees were rooted on the node separating *Microvalgus* from the remaining taxa.

Results

We obtained DNA barcode data from 284 specimens, of which 256 were adults (90%) and 28 were larvae. We sampled 68 described species and up to 27 putative undescribed species at an average of 3 specimens per species. Our total of 68 described species includes an unidentified species of *Microvalgus* which is likely to be a described species.

Two hundred and forty-five sequences (86%) are BARCODE standard compliant, defined as > 486 bp in length, with two or fewer ambiguous bases and with at least two high-quality sequence trace files uploaded. Only six sequences were less than 300 bp in length.

Mean specimen age at DNA extraction was 4.2 years, although for the first batch of 94 samples the mean age was 7.4 years. The oldest sample to yield barcode-standard compliant data was 22.6 years old.

Bayesian Inference was completed after 18,625,000 generations when the average standard deviation of split frequencies reached 0.009997. The structure of the BI tree is summarized in Fig. 1, with strongly supported branches (posterior probabilities (PP) ≥ 0.99 and bootstrap percentages (BP) from the RAxML ≥ 95%) indicated by asterisks. The full BI tree is provided as supplementary Fig. S1, and the full RAxML tree is provided as supplementary Fig. S2.

Eleven genera were represented by a single species in our data set, including seven monotypic genera (*Phyllopodium* Schoch, 1895, *Octocollis* Moeseneder & Hutchinson, 2012*, Lenosoma* Kraatz, 1880*, Stenopisthes* Moser, 1913, *Hemipharis* Burmeister, 1842, *Neoclithria* van de Poll, 1886, *Micropoecila* Kraatz, 1880) and four additional genera (*Mycterophallus* van de Poll, 1886, *Poecilopharis* Kraatz, 1880, *Evanides* Thomson, 1880, *Storeyus* Hasenpusch & Moeseneder, 2009). In all six cases where these species had multiple samples, the species were recovered as monophyletic and distinct from other species.

Of the remaining 22 genera, for which multiple species were sampled, half were recovered as monophyletic. These are listed with the number of species sampled and number of specimens (n) sampled in parentheses: *Microvalgus* (4 spp., n = 13), *Ischiopsopha* Gestro, 1874 (2 spp., n = 8), *Lomaptera* Gory & Percheron, 1833 (2 spp., n = 5), *Schizorhina* (2 spp., n = 6), *Navigator* Moeseneder & Hutchinson, 2016 (2 spp., n = 5), *Lyraphora* Kraatz, 1880 (3 spp., n = 11), *Tapinoschema* Thomson, 1880 (3 spp., n = 12), *Bisallardiana* (10 spp., n = 31), *Neorrhina* Thomson, 1878 (2 spp., n = 11), *Chlorobapta* (3 spp., n = 11) and *Metallesthes* Kraatz, 1880 (4 spp., n = 16).

RESL cluster analysis grouped sequences into 100 OTUs, with 32 of these being singletons. There were 21 singleton species, and the remaining 11 singleton OTUs represented divergent lineages within species. RESL clustering split 13 species, some of them into as many as 4 OTUs, as summarised in Table 2.

Only two BINs contained multiple species: the BIN containing *Hemichnoodes mniszechi* (Janson, 1873), *H. parryi* (Janson, 1873) and *Diaphonia* sp_dnul_chm, and the BIN containing both *Glycyphana (Caloglycyphana) papua* (Wallace, 1867) and *G. (Caloglycyphana) pulchra* (Macleay, 1871) (maximum within-OTU distance = 1.57%).

Discussion

This preliminary study reports DNA barcode data for 68 described species from 33 genera, representing 48% of currently known Australian species and 83% of the genera (141 described species in 40 genera; Moeseneder *et al.*, 2019; Hutchinson & Moeseneder, 2019). Our goal is a comprehensive DNA barcode dataset, and complementary nuclear gene and morphological data, to address both species-level and higher-level relationships of the Australian cetoniines, facilitating integrative revisionary taxonomy. Here we recognise likely undescribed species and note cases of likely generic misassignment of species but refrain from making taxonomic decisions, as that would require generic revisions, which are beyond the scope of the current study.

In general, there was concordance between morphology-based identifications and barcode-based clustering. This concordance is not obvious since RESL clustering split many species and produced 100 OTUs. However, our preliminary morphological investigations suggest that in addition to the 68 described species we sampled, the 100 OTUs include up to 27 undescribed species.

Of the 27 possible undescribed species, five were known to us previously and are easily distinguished morphologically, ~~6~~ were suspected but with some uncertainty due to their similarity to described species, and 16 were completely unexpected (potential “cryptic species”) and were only revealed by their DNA barcodes. Their morphological similarity to described species is striking, and further work, including analysis of nuclear genes and male genitalia from a larger series of specimens, is needed to rigorously assess their taxonomic status. The number of undescribed species hence may represent a potential increase to the size of the Australian fauna of 12-19%.

There was one OTU that contained more than one species: *Glycyphana (Caloglycyphana) pulchra* plus *G. (Caloglycyphana) papua*, with 1.57% distance between them.

While Barcode Index Numbers (BINs) are calculated by BOLD using the RESL clustering algorithm, sequences on BOLD must meet criteria such as minimum sequence length and quality to be included in a BIN, thus only 252 sequences were placed into BINs. We therefore also performed a separate RESL clustering analysis on the complete 284 sequence dataset to obtain OTUs. The only differences between these analyses were that 1) both species of *Hemichnoodes* Kraatz, 1880, plus *Diaphonia* “sp dnul chm” were assigned to a single BIN (*D. luteola* (Janson, 1873) was not assigned to any BIN), while the cluster analysis split these four taxa into separate OTUs corresponding to their morphological identification, 2) *Dilochrosis balteata* was placed in a single BIN but split into two OTUs with 2.15% distance between them, and 3) *Eupoecila australasiae* was divided into ~~2~~ BINs with 2.71% distance between them, but comprised a single OTU.

The Geographic Distance Correlation test was significant ($p \leq 0.05$) for only ~~3~~ species (*Ischiopsopha wallacei* (Thomson, 1857), *Metallesthes anneliesae* Moeseneder & Hutchinson, 2014, *Glycyphana stolata*) and highly significant ($p \leq 0.01$) for a single species, *Chondropyga dorsalis*. The 12 specimens of *C. dorsalis* were collected within approximately 70 km of each other in Southeast Queensland in varying habitat types. An attempt at finding unique, easily visible characters for each group is ongoing.

Our trees were rooted on Valgini (*Microvalgus*) since the most comprehensive molecular phylogeny of the subfamily to date (Sipek *et al*., 2016) placed Valgini, and Trichiini in part, as sister-group to the remaining 10 tribes that they sampled.

While we do not expect a small and rapidly evolving fragment of a single mitochondrial gene to yield a robust phylogeny of the Cetoniinae, phylogenetic analysis of DNA barcodes is likely to give a good indication of relationships among closely related species, to provide a guide to where undescribed taxa should be placed, and suggest where further evidence is needed on supraspecific relationships. The discussion below is meant in that context, acknowledging the limited deeper phylogenetic utility of DNA barcodes.

In the most well-known Australian cetoniine species, *Eupoecila australasiae*, *Neorrhina punctatum*, *Glycyphana stolata*, *Chondropyga dorsalis* and *Bisallardiana gymnopleura*, we found high levels of DNA diversity. While this is not unusual for DNA barcoding studies, e.g. in Elateridae (Oba *et al*., 2015) and stemborer moths (Lee *et al*., 2019), our preliminary morphological examination of the species implies that these high levels of COI diversity are for the most part correlated with morphological diversity. This suggests that many of these OTUs may in fact represent undescribed species. Further cases of discordance between prior expectations based on current taxonomy and DNA barcoding results are detailed below.

Trichaulax (4 spp., $n = 11$) was rendered paraphyletic by the insertion of *Lenosoma fulgens* (1 spp., $n = 3$) (Fig. 2).

Chondropyga (4 spp., $n = 20$) was rendered paraphyletic by the insertion of *Pseudoclithria hirticeps* (Macleay, 1871) (1 sp., $n = 1$) (Fig. 3). *Pseudoclithria hirticeps*, the type species of *Pseudoclithria*, is placed incorrectly and likely belongs in genus *Chondropyga*. However, as we sampled only a single specimen of *P. hirticeps* this result requires confirmation with data from further specimens and genes.

The lineage containing *Dilochrosis* (4 spp., $n = 27$) had *Glycyphana pulchra*/*G. papua* (2 spp., $n = 4$) embedded within it in the Bayesian tree. The RAxML tree was similar, except that *Protaetia (Protaetia) fusca* (Herbst, 1790) ($n = 6$) was also embedded with *Dilochrosis*, as sister group to the two *Glycyphana* species. However, based on morphological evidence, the length of the branch subtending *G. pulchra*/*G. papua* (Fig. S1) and the instability of these nodes when analysed by maximum likelihood methods, it appears unlikely that these placements reflect true phylogenetic affinities, and further evidence is needed to resolve these questions.

Neoclithria (1 sp., n = 3) is embedded within *Clithria* (3 spp., n = 7) (Fig. 4), and *Micropoecila* (1 sp., n = 2) is embedded within *Eupoecila* (3 spp., n = 12). Thus, both *Neoclithria* and *Micropoecila* may need to be synonymised with the genera they are placed within.

Glycyphana was consistently split into two distantly related groups, one containing the closely related *G. (Caloglycyphana) pulchra* and *G. (Caloglycyphana) papua*, merged into a single BIN, and the other containing *G. (Glycyphaniola) brunnipes* (Kirby, 1818) and *G. (Glycyphaniola) stolata* (Fig. 5). *Glycyphana brunnipes* is split into two BINs while *G. stolata* is split into four BINs. Bacchus (1974) split *G. stolata* into three forms. Substantial further integrative taxonomic work is required to reassess species boundaries in these species complexes.

Relationships among *Diaphonia*, *Aphanesthes* Kraatz, 1880*, Hemichnoodes*, *Pseudoclithria* and *Metallesthes* were complex. There was moderate to strong support, a Bayesian posterior probabilities (PP) of 0.99 and maximum likelihood bootstrap percentage (BP) of 65%, for a clade including *Aphanesthes succinea* (Hope, 1844) (n = 4), *Diaphonia* (3 spp., n = 6) and *Hemichnoodes* (2 spp., n = 6). There was weaker support (PP of 0.95, BP < 50%) for the sister-group to the above clade, comprising *A. pullata* (Janson, 1873) (n = 3), "*A.* sp_aisa chm" (a possible undescribed species, n = 1), *Pseudoclithria* (5 spp., n = 13) excluding *P. hirticeps* (mentioned above) and *Metallesthes*. While the sampled species in this group are well defined, extensive further work is required to understand the relationships between these species and the status of the five current genera.

The remaining seven genera that constitute the Australian cetoniine fauna were not sampled because no recent material was available for DNA sequencing. These are *Aurum* Hutchinson & Moeseneder, 2019, *Axillonia* Krikken, 2018, *Grandaustralis* Hutchinson & Moeseneder, 2013, *Macrotina* Strand, 1934, *Territonia* Krikken, 2018, *Chalcopharis* Heller, 1903 and *Charitovalgus* Kolbe, 1904. The first four of these genera are monospecific and the last two are represented in Australia by a single species each.

Conclusions

We produced a DNA barcode dataset for Australian flower beetles that includes approximately half of the country's species. We found that DNA barcodes provide species-level resolution in almost all cases. The high levels of DNA diversity were unexpected within many species, and preliminary morphological investigations suggest that there may be as many as 27 undescribed species in our dataset. Further integrative taxonomic work, incorporating COI-based DNA barcoding, nuclear gene data and detailed morphological investigations, are needed to better understand the diversity of Australian Cetoniinae and to document and describe numerous undescribed species.

Acknowledgements

We are very grateful to the following people for their assistance. Jack Hasenpusch (Australian Insect Farm, Innisfail, Qld) for donating many specimens to the AFBRC collection. Dr. Christine Lambkin and Susan Wright (QM) for guidance on various topics, collecting permits, loans and assistance with specimens. Dr. Geoff Monteith (QM) for ecological information and flight intercept traps. For specimen loans: Cate Lemann and Tom Weir (ANIC), Derek Smith (AM), Justin Bartlett (DAF), Dr Nikolai Tatarnic and Brian Hanich (WAM), Andras Szito (DAFWA), Peter Hudson (SAM), Dr. Max Barclay and Beulah Garner (NHML). For fieldwork and donations: Adam and Matthew Yates (Alice Springs, NT), Allen Sundholm (Turrella, NSW), Mark Hura (Parafield Gardens, SA), Geoff Walker (Murrumbeena, Vic.) and David Baume (Cairns, Qld). Glen Smith (Duaringa, Qld) and Lindsay Popple (Cairns, Qld) for ecological observations. Melissa Syme (DAF) for access to parks. For access to their private properties: Tracey Rosser (Numinbah, NSW), Michael Corke (Numinbah, NSW) and Selwyn Podlich (Coulson, Qld). Sabine Moeseneder (Redland Bay, Qld) for assistance with fieldwork. Dr. Nicole Gunter (Cleveland Museum of Natural History, Cleveland, Ohio, USA) for additional information on specimens. Richard Zietek (Cabalaba, Qld) for specimens and loans. Dr. Eva Plaganyi-Lloyd (CSIRO) for supporting excursions via awards.

References

Bacchus ME (1974) A revision of the Australian species of the genus *Glycyphana* Burmeister (Coleoptera: Scarabaeidae). *Journal of the Australian Entomological Society* 13: 111–128.

DeSalle R, Goldstein P (2019) Review and interpretation of trends in DNA barcoding. *Frontiers in Ecology and Evolution* 7: 302.

Gunter NL, Weir TA, Slipinski A, Bocak L, Cameron SL (2016) If dung beetles (Scarabaeidae: Scarabaeinae) arose in association with dinosaurs, did they also suffer a mass co-extinction at the K-PG boundary? *PLoS ONE* 11: e0153570.

Hasenpusch J, Moeseneder C (2009) *Storeyus*, a new genus and *Storeyus pseudodipterus*, a new species of Cetoniinae (Coleoptera: Scarabaeidae) from Australia, with a redescription of *Lenosoma* Kraatz. *Australian Entomologist* 36: 221–234.

Hutchinson P, Moeseneder C (2019) *Aurum*, a new monotypic genus of Cetoniinae (Coleoptera: Scarabaeidae) endemic to Western Australia, Australia. *Zootaxa* 4683: 277–285.

Hutchinson P, Moeseneder C (2013) *Chondropyga insignicosta*, a new species of Cetoniinae (Coleoptera Scarabaeidae) from Queensland, Australia. *Zootaxa* 3710: 281–288.

Hutchinson P, Moeseneder C (2013) *Grandaustralis*, a new genus and *Grandaustralis boomerang*, a new species of Cetoniinae (Coleoptera Scarabaeidae) from Western Australia. *Zootaxa* 3669: 17–26.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.

Krajčík M (2012) Checklist of the world Scarabaeoidea. *Animma.X* 5: 1–278.

Krikken J (1984) A new key to the suprageneric taxa in the beetle family Cetoniidae, with annotated lists of the known genera. *Zoologische Verhandelingen* 210: 1–75.

Krikken J (2018) Two new genera of Australian flower chafers, with an annotated overview of their relatives (Coleoptera: Scarabaeidae: Cetoniinae). *Haroldius* 5: 1–36.

Lanfear R, Frandsen PB, Wright AM, Senfeld T, Calcott B (2016) PartitionFinder 2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analyses. *Molecular Biology and Evolution* 34: 772–773.

Lee TRC, Anderson SJ, Tran-Nguyen LTT, Sallam N, Le Ru BP, Conlong D, Powell K, Ward A, Mitchell A (2019) Towards a global DNA barcode reference library for quarantine identifications of lepidopteran stemborers, with an emphasis on sugarcane pests. *Scientific Reports* 9: 7039.

Miller MA, Pfeiffer W, Schwartz T (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. Gateway Computing Environments Workshop (GCE), New Orleans, LA, USA. IEEE, 1–8.

Mitchell A (2015) Collecting in collections: a PCR strategy and primer set for DNA barcoding of decades-old dried museum specimens. *Molecular Ecology Resources* 15: 1102–1111.

Moeseneder C (2017) Focus-stacking macro photography system for around $1,200. *Scarabs* 83: 1–7.

Moeseneder C, Weir T, Lemann C, Hutchinson P (2019) 34. Scarabaeidae: Cetoniinae Leach, 1815. In Slipinski A & Lawrence J (eds.) *Australian Beetles Volume 2 Archostemata, Myxophaga, Adephaga, Polyphaga (part).* CSIRO Publishing. 531-553. pl. 73–76.

Moeseneder C, Hutchinson PM (2016) *Navigator*, a new endemic genus of Cetoniinae (Coleoptera: Scarabaeidae) from Australia with descriptions of two new species and behavioural studies. *Zootaxa* 4173: 530–556.

Moeseneder C, Hutchinson P, Lambkin C (2014) Revision of the genus *Metallesthes* Kraatz and description of *Metallesthes anneliesae*, a new species of Cetoniinae (Coleoptera: Scarabaeidae) from Queensland and New South Wales, Australia. *Zootaxa* 3881: 301–327.

Moeseneder C, Hutchinson P (2012) *Octocollis*, a new genus and *Octocollis setosus*, a new species of Cetoniinae (Coleoptera: Scarabaeidae) from Queensland, Australia. *Zootaxa* 3557: 40–48.

Monteith G (2010) *The butterfly man of Kuranda: Frederick Parkhurst Dodd*. Queensland Museum, South Brisbane, Australia. 46 p.

Moore B (1987) *A Guide to the Beetles of South-Eastern Australia*. Australian Entomological Press, Greenwich, NSW.

Oba Y, Ôhira H, Murase Y, Moriyama A, Kumazawa Y (2015) DNA Barcoding of Japanese Click Beetles (Coleoptera, Elateridae). *PLoS ONE* 10: e0116612.

Ratnasingham S, Hebert PDN (2007) BOLD: The Barcode of Life Data System (www.barcodinglife.org). *Molecular Ecology Notes* 7: 355–364.

Ratnasingham S, Hebert PDN (2013) A DNA-Based Registry for All Animal Species: The Barcode Index Number (BIN) System. *PLoS ONE* 8: e66213.

Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539–542.

Šípek P, Fabrizi S, Eberle J, Ahrens D (2016) A molecular phylogeny of rose chafers Coleoptera: Scarabaeidae: Cetoniinae) reveals a complex and concerted morphological evolution related to their flight mode. *Molecular Phylogenetics and Evolution* 101: 163–175.

Stamatakis A (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30: 1312–1313.

Svensson GP, Oleksa A, Gawroński R, Lassance J, Larsson MC (2009) Enantiomeric conservation of the male-produced sex pheromone facilitates monitoring of threatened European hermit beetles (*Osmoderma* spp.). *Entomologia Experimentalis et Applicata* 133: 276–282.

Villesen P (2007) FaBox: an online toolbox for FASTA sequences. *Molecular Ecology Notes* 7: 965–968.

Vondracek D, Fuchsova A, Ahrens D, Kral D, Sipek P (2018) Phylogeography and DNA-based species delimitation provide insight into the taxonomy of the polymorphic rose chafer *Protaetia (Potosia) cuprea* species complex (Coleoptera: Scarabaeidae: Cetoniinae) in the Western Palearctic. *PLoS ONE* 13: e0192349.

Williams GA, Adam P (1998) Pollen loads collected from large insects in Australian subtropical rainforests. *Proceedings of the Linnean Society of New South Wales* 120: 49–67.

Zauli A, Carpaneto GM, Chiari S, Mancini E, Nyabuga FN, Redolfi De Zan L, Romiti F, Sabbani S, Audisio PA, Hedenström E, Bologna MA, Svensson GP (2016) Assessing the taxonomic status of Osmoderma cristinae (Coleoptera: Scarabaeidae), endemic to Sicily, by genetic, morphological and pheromonal analyses. *Journal of Zoological Systematics and Evolutionary Research* 54: 206–214.

Figure 1

Bayesian phylogenetic tree for all data.

Branches are collapsed to illustrate genus-level relationships. *Microvalgus* was treated as the outgroup. Asterisks indicate nodes with strong support from both Bayesian posterior probabilities (PP ≥ 0.99) and maximum likelihood bootstrap percentage (BP ≥ 95). Closed circles indicate nodes with strong support under only one of these methods.

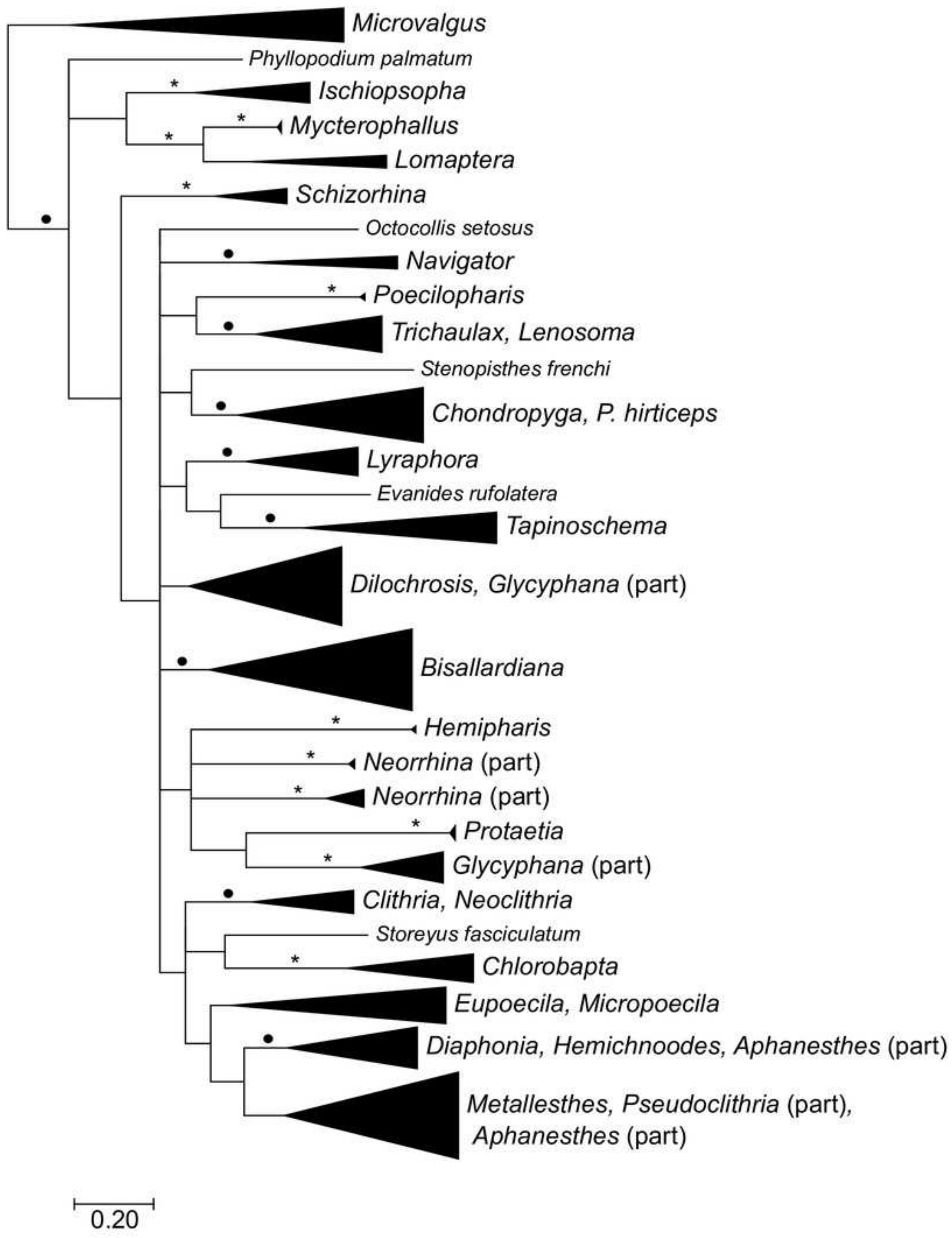
Microvalgus
Phyllopodium palmatum
Ischiopsopha
Mycterophallus
Lomaptera
Schizorhina
Octocollis setosus
Navigator
Poecilopharis
Trichaulax, Lenosoma
Stenopisthes frenchi
Chondropyga, P. hirticeps
Lyraphora
Evanides rufolatera
Tapinoschema
Dilochrosis, Glycyphana (part)
Bisallardiana
Hemipharis
Neorrhina (part)
Neorrhina (part)
Protaetia
Glycyphana (part)
Clithria, Neoclithria
Storeyus fasciculatum
Chlorobapta
Eupoecila, Micropoecila
Diaphonia, Hemichnoodes, Aphanesthes (part)
Metallesthes, Pseudoclithria (part), Aphanesthes (part)
0.20

Figure 2

Extract of Bayesian phylogenetic tree for *Trichaulax* and *Lenosoma*.

Asterisks indicate nodes with strong support from both Bayesian posterior probabilities (PP ≥ 0.99) and maximum likelihood bootstrap percentage (BP ≥ 95). Closed circles indicate nodes with strong support under only one of these methods.

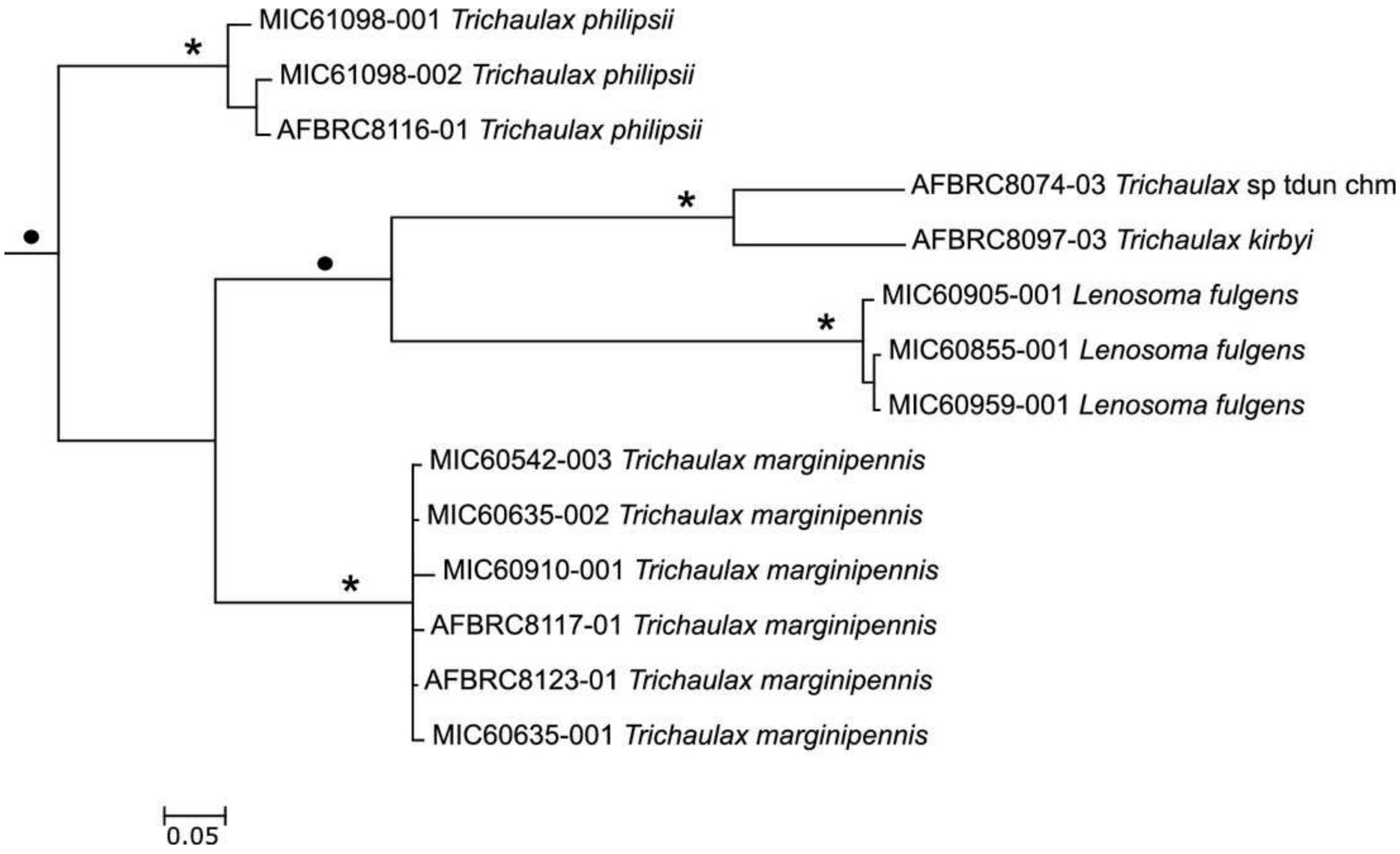

Figure 3

Extract of Bayesian phylogenetic tree for *Chondropyga* and *Pseudoclithria* (in part).

Asterisks indicate nodes with strong support from both Bayesian posterior probabilities (PP ≥ 0.99) and maximum likelihood bootstrap percentage (BP ≥ 95). Closed circles indicate nodes with strong support under only one of these methods.

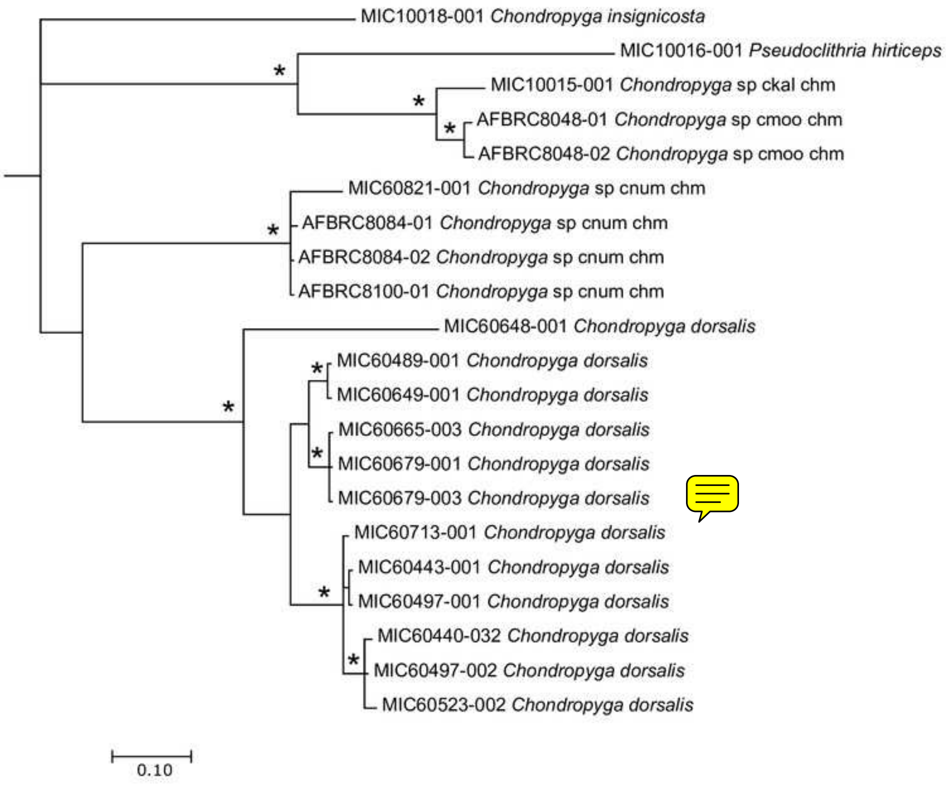
MIC10018-001 *Chondropyga insignicosta*
MIC10016-001 *Pseudoclithria hirticeps*
MIC10015-001 *Chondropyga* sp ckal chm
AFBRC8048-01 *Chondropyga* sp cmoo chm
AFBRC8048-02 *Chondropyga* sp cmoo chm
MIC60821-001 *Chondropyga* sp cnum chm
AFBRC8084-01 *Chondropyga* sp cnum chm
AFBRC8084-02 *Chondropyga* sp cnum chm
AFBRC8100-01 *Chondropyga* sp cnum chm
MIC60648-001 *Chondropyga dorsalis*
MIC60489-001 *Chondropyga dorsalis*
MIC60649-001 *Chondropyga dorsalis*
MIC60665-003 *Chondropyga dorsalis*
MIC60679-001 *Chondropyga dorsalis*
MIC60679-003 *Chondropyga dorsalis*
MIC60713-001 *Chondropyga dorsalis*
MIC60443-001 *Chondropyga dorsalis*
MIC60497-001 *Chondropyga dorsalis*
MIC60440-032 *Chondropyga dorsalis*
MIC60497-002 *Chondropyga dorsalis*
MIC60523-002 *Chondropyga dorsalis*
0.10

Figure 4

Extract of Bayesian phylogenetic tree for *Clithria* and *Neoclithria*.

Asterisks indicate nodes with strong support from both Bayesian posterior probabilities (PP ≥ 0.99) and maximum likelihood bootstrap percentage (BP ≥ 95). Closed circles indicate nodes with strong support under only one of these methods.

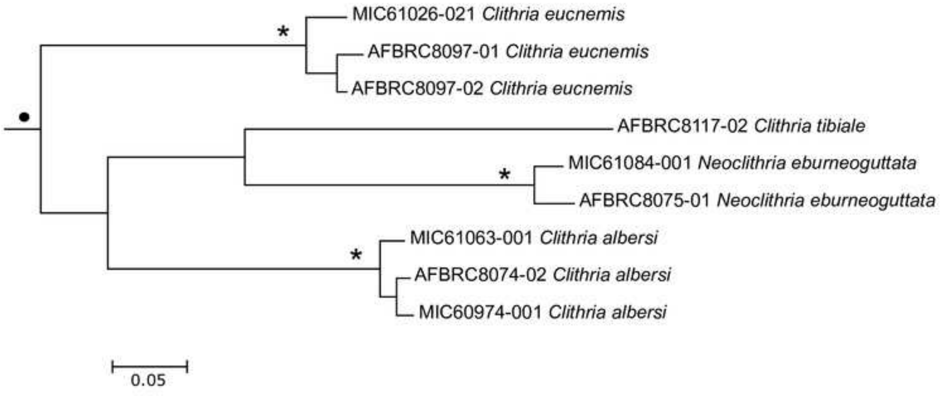
*
MIC61026-021 *Clithria eucnemis*
AFBRC8097-01 *Clithria eucnemis*
AFBRC8097-02 *Clithria eucnemis*
●
AFBRC8117-02 *Clithria tibiale*
*
MIC61084-001 *Neoclithria eburneoguttata*
AFBRC8075-01 *Neoclithria eburneoguttata*
*
MIC61063-001 *Clithria albersi*
AFBRC8074-02 *Clithria albersi*
MIC60974-001 *Clithria albersi*
0.05

Figure 5

Extract of Bayesian phylogenetic tree for *Glycyphana* (in part).

Asterisks indicate nodes with strong support from both Bayesian posterior probabilities (PP ≥ 0.99) and maximum likelihood bootstrap percentage (BP ≥ 95). Closed circles indicate nodes with strong support under only one of these methods.

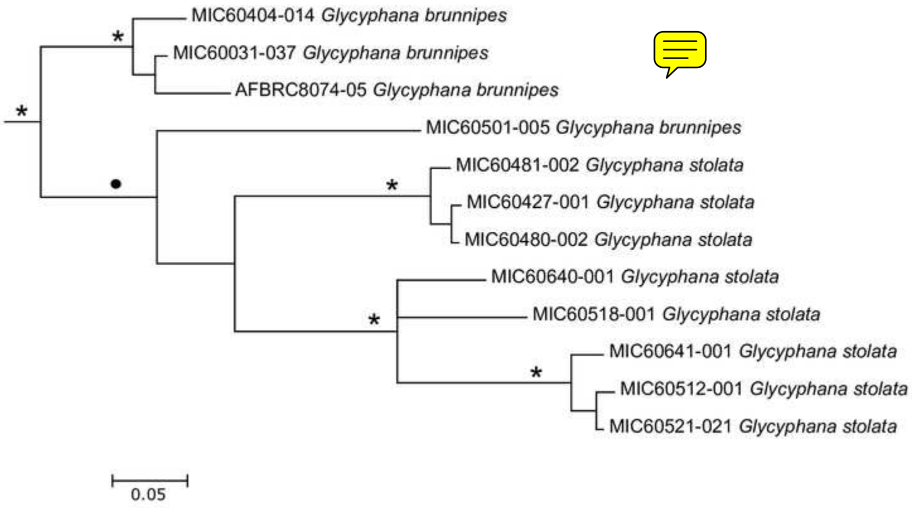
MIC60404-014 *Glycyphana brunnipes*
MIC60031-037 *Glycyphana brunnipes*
AFBRC8074-05 *Glycyphana brunnipes*
MIC60501-005 *Glycyphana brunnipes*
MIC60481-002 *Glycyphana stolata*
MIC60427-001 *Glycyphana stolata*
MIC60480-002 *Glycyphana stolata*
MIC60640-001 *Glycyphana stolata*
MIC60518-001 *Glycyphana stolata*
MIC60641-001 *Glycyphana stolata*
MIC60512-001 *Glycyphana stolata*
MIC60521-021 *Glycyphana stolata*
*
●
0.05

Table 1(on next page)

Previous DNA-based studies of Cetoniinae.

Table 1 Previous DNA-based studies of Cetoniinae.

Reference	**Number of samples**	**Number of species sampled (Australian)**	**Gene regions sampled**	**Taxon focus**	**Study purpose**
Ahrens *et al.* (2007)	11	2 (0)	COI, 16S, 28S	Scarabaeidae	Larval-adult association
Ahrens *et al.* (2008)	8	8 (1)	COI, 16S, 28S	Sericini	Higher-level phylogeny
Ahrens *et al.* (2011)	7	7 (0)	COI, 16S, 28S	Hopliini	Higher-level phylogeny
Ahrens *et al.* (2013)	230	1 (0)	COI, ITS1	*Cetonia aurata* complex	Phylogeography, species-level taxonomy
Audisio *et al.* (2008)	26	1 (0)	COI	*Osmoderma*	Species -level taxonomy
Audisio *et al.* (2009)	26	5 (0)	COI	*Osmoderma*	Species -level taxonomy
Gunter *et al.* (2016)	15	15 (15)	COI, 16S, 12S, 28S	Scarabaeinae	Higher-level phylogeny
Han *et al.* (2017)	16	3 (0)	COI	*Osmoderma*	Species-level taxonomy
Hendrich *et al.* (2015)	70	14 (0)	COI	European Coleoptera	DNA barcoding
Kim *et al.* (2013)	1	1 (0)	mitogenome	*Protaetia brevitarsis*	Genomics
Landvik *et al.* (2013)	7	1 (0)	COI	*Osmoderma*	Species identification
Lee *et al.* (2015)	50	5 (0)	COI, 16S	*Dicronocephalus*	Species-level phylogeny
McKenna *et al.* (2015)	5	5 (1?)	28S, CAD	Staphyliniformia, Scarabaeiformia	Higher-level phylogeny
Philips *et al.* (2016)	12	11 (0)	COI, 28S	*Trichiotinus*	Species-level phylogeny
Seidel (2018)	29	5 (0)	COI	*Eudicella*	Species-level taxonomy
Sipek *et al.* (2016)	130	125 (2)	COI, 16S, 28S	Cetoniinae	Higher-level phylogeny
Song *et al.* (2018)	4	4 (0)	5 mitogenomes	Scarabaeidae	Genomics, higher-level phylogeny
Svensson *et al.* (2009)	38	5 (0)	COI	*Osmoderma*	Species identification
Vondracek *et al.* (2018)	65	15 (0)	COI, CytB	*Potosia*	Species-level taxonomy
Zauli *et al.* (2016)	27	1 (0)	COI	*Osmoderma*	Species-level taxonomy

Table 2(on next page)

Results of RESL clustering for species showing >2% maximum uncorrected intraspecific distance.

Table 2

Results of RESL clustering for species showing >2% maximum uncorrected intraspecific distance.

Species	**Maximum uncorrected intraspecific distance**	**Number of RESL OTUs**	**Number of BOLD BINs**
Aphanestes pullata	6.20%	2	2
Chondropyga dorsalis	6.09%	4	4
Glycyphana (Glycyphaniola) brunnipes	5.61%	2	2
Glycyphana (Glycyphaniola) stolata	5.52%	4	4
Neorrhina punctatum	3.27%	2	2
Micropoecila cincta	4.84%	2	2
Dilochrosis brownii	4.20%	2	2
Lyraphora obliquata	3.76%	2	2
Aphanesthes succinea	2.51%	2	2
Chondropyga "sp_cmoo_chm"	2.87%	2	2
Eupoecila australasiae	2.71%	1	2
Metallesthes anneliesae	2.24%	2	2
Dilochrosis balteata	2.15%	2	1