CSIRO PUBLISHING

RESEARCH PAPER
https://doi.org/10.1071/MF21291

MARINE & FRESHWATER RESEARCH

DNA barcoding and metabarcoding of highly diverse aquatic mites (Acarina) can improve their use in routine biological monitoring

Melissa E. Carew[A,B,*], Wen Kyle Yow[A], Katie L. Robinson[A], Rhys A. Coleman[C] and Ary A. Hoffmann[A]

For full list of author affiliations and declarations see end of paper

*Correspondence to:
Melissa E. Carew
Pest and Environmental Adaptation Research Group (PEARG), School of BioSciences, Bio21 Institute, 30 Flemington Road, The University of Melbourne, Vic. 3010, Australia
Email: mecarew@unimelb.edu.au

Handling Editor:
Donald Baird

ABSTRACT

Context. Acarina are commonly collected in macroinvertebrate surveys used to monitor freshwater ecosystems. However, they can be difficult to identify morphologically requiring considerable taxonomic skill for identification to finer taxonomic levels. Therefore, in biomonitoring they are identified to subclass despite high species diversity and varied environmental responses. DNA barcoding individuals and DNA metabarcoding of bulk samples enables species to be accurately and routinely identified. However, poor DNA barcode coverage of Australian aquatic mites has hampered their use in DNA studies. **Aims.** Here, we aim to generate DNA barcodes for mites from Greater Melbourne, Australia. **Key results.** For many specimens, we link DNA barcodes to genus-level morphological identifications using genetic analysis of DNA barcodes to understand biodiversity. We then test if new DNA barcodes can improve identification of mites in samples processed with DNA metabarcoding. We found Australian aquatic mites showed high diversity with many DNA barcodes represented by single specimens. **Conclusions.** Increased mite DNA barcode library coverage improved their detection using DNA metabarcoding. **Implications.** Given high species diversity, much effort will be required to improve DNA barcode coverage for aquatic mites in Australia and integrate barcodes with species level taxonomy, allowing Acarina to be better incorporated into DNA-based biological monitoring.

Keywords: Australia, biodiversity, freshwater, Halacaroidea, Hydracarina, macroinvertebrates, Mesostigmata, Oribatida, species identification.

Introduction

Globally, 'water mites' or Hydracarina are a highly diverse group of invertebrates inhabiting a wide range of freshwater environments (Di Sabatino *et al.* 2007; Goldschmidt 2016). The Hydracarina (also referred to as Hydrachnellae, Hydrachnidia and Hydrachnida) are members of the large suborder Prostigmata within the order Trombidiformes (Harvey 1998). Over 6000 species of Hydracarina from 57 families have been described worldwide, while in Australia ~780 species have been recorded from over 30 families (Harvey 1998; Di Sabatino *et al.* 2007). The Hydracarina are the most species-rich and abundant mites in freshwater habitats, having successfully colonised most freshwater environments around the world (Di Sabatino *et al.* 2007). Other mites, such as the Halacaroidea, the Oribatida (order Sarcoptiformes), Mesostigmata and other Trombidiformes (such as the Trombidioidea) are also commonly found in freshwater macroinvertebrate samples (Harvey 1998; Schatz and Behan-Pelletier 2008; Proctor *et al.* 2015). Although some of these are considered aquatic or sub aquatic, others are terrestrial and captured due to their proximity to waterways (Schatz and Behan-Pelletier 2008; Walter and Proctor 2013).

The Hydracarina, along with other mite groups, are often underutilised in routine biomonitoring as they are not identified to finer taxonomic levels (Proctor 2007). In Australia, they are identified only at subclass level (Acarina) for bioassessment (e.g. Chessman 1995; Simpson and Norris 2000). Identification of mites at lower taxonomic

Received: 6 October 2021
Accepted: 7 April 2022
Published: 17 May 2022

Cite this:
Carew ME *et al.* (2022)
Marine and Freshwater Research, **73**(7), 900–914.
doi:10.1071/MF21291

levels can be difficult due to their small size and high diversity, often requiring a compound microscope and considerable taxonomic expertise for sorting to family, genus or species levels (Harvey 1998). However, they can be highly responsive to environmental change in freshwater environments when considered at finer taxonomic levels (e.g. Miccoli *et al.* 2013; Goldschmidt *et al.* 2016; Zawal *et al.* 2017), including in Australian rivers (e.g. Growns 2001).

DNA barcoding offers a reliable means to routinely identify taxonomically challenging species and has revealed previously unknown cryptic biodiversity in different invertebrate groups (Hebert *et al.* 2003; Jackson *et al.* 2014; Carew and Hoffmann 2015; Young *et al.* 2019). In water mites, DNA barcoding has revealed high diversity, including many previously unrecognised (cryptic) species (e.g. Stålstedt *et al.* 2013; Vasquez *et al.* 2017; Blattner *et al.* 2019; Montes-Ortiz and Elías-Gutiérrez 2020). DNA barcodes can also enable easier identification of males and females in sexually dimorphic species where only one sex has been taxonomically described, and can help link undescribed juveniles (larvae and nymphs) to adults (e.g. Glowska *et al.* 2014; Więcek *et al.* 2021). However, DNA barcode coverage of water mites in most regions of the world, including Australia, is poor. Prior to beginning this study there were only ~30 DNA barcodes from six families available on the public Barcode of Life Database (BOLD) for water mites from eastern Australia, whereas there were a further 289 DNA barcodes from two freshwater mite families from one location in Perth, Western Australia (http://www.boldsystems.org/index.php/Public_BINSearch?searchtype=records, accessed 4 November 2020). Given the high species diversity in water mites found in studies carried out elsewhere (e.g. Young *et al.* 2019), improving DNA barcoding coverage of water mites in Australia will better enable their routine identification in DNA-based studies and biomonitoring in freshwater environments. This is particularly important as 'DNA metabarcoding' of macroinvertebrate or environmental samples is becoming increasingly used in providing cost-effective species level biodiversity assessments (Hajibabaei *et al.* 2011; Yu *et al.* 2012; Porter and Hajibabaei 2018; Carew *et al.* 2021). Taxa can only be reliably identified using DNA metabarcoding if reference DNA barcode libraries are available with no large taxonomic gaps (Weigand *et al.* 2019). DNA sequences from DNA metabarcoding are typically assigned to species based on greater than 97% sequence matches to reference DNA barcodes (Carew *et al.* 2018*b*; Elbrecht and Steinke 2019) but this is rarely possible in Australia for water mites. Therefore, improving DNA barcode coverage of poorly represented groups, like water mites, facilitates their inclusion in DNA metabarcoding-based biodiversity assessments.

In this study, we generate DNA barcodes for mites found in macroinvertebrate samples collected from Greater Melbourne, Australia. We aim to improve the current understanding of local freshwater mite biodiversity and increase DNA barcode coverage to facilitate the inclusion of mites in environmental monitoring using DNA methods. We include photographed specimens to provide a morphological record of specimens to accompany DNA barcodes. We use a new high-throughput DNA sequencing protocol to provide DNA barcodes for individual mites, and using DNA metabarcoding, test the ability of these new DNA barcodes to improve detection and identification of mites in mixed macroinvertebrate samples.

Materials and methods

Sample collection and study overview

Macroinvertebrate samples containing Acarina were collected from a total of 50 sites (34 sites for both individual DNA barcoding and DNA metabarcoding and an additional 16 sites for individual DNA barcoding only) in Greater Melbourne, Australia (Fig. 1, Supplementary Table S1). Samples were taken from edge habitats using standard rapid bioassessment methods used in Australia (Chessman 1995). All sampling involved sweeping (edge habitats) over ~10 m^2 of benthic zone from wadable areas for ~5–10 min into a 250-μm mesh net. All bulk net contents (including sampling debris) from each sample were preserved on-site in absolute ethanol for subsampling in the laboratory. Absolute ethanol was changed after 24–48 h and samples were stored at 4°C. Samples were then randomly subsampled in the laboratory by initially taking a 10% subsample, and if less than 300 individuals were found, additional subsampling was taken until a count of 300 individuals was reached (Walsh 1997). For a small number of sites used to provide mites for DNA barcoding only, macroinvertebrates were live picked on-site for 30 min according to the Guidelines for Environmental Management – Water (EPA Victoria 2021) (Table S1). These samples were also stored in absolute ethanol at 4°C and absolute ethanol was changed after 24–48 h.

The 34 macroinvertebrate samples that were processed with DNA metabarcoding were collected by Ecology Australia and AQUEST on behalf of Melbourne Water Corporation as part of a previous study (Carew *et al.* 2019). The sites were selected so that the three major terrestrial bioregions in Greater Melbourne were included and sites had a spread of values for mean annual runoff depth, attenuated imperviousness (Walsh and Kunapo 2009) and attenuated forest cover (Walsh and Webb 2014). These samples were first sorted morphologically to mostly family level (i.e. Chessman 1995) but with Acarina identified to subclass (Fig. 2, step 1). Macroinvertebrate samples were then bulk processed non-destructively using DNA metabarcoding (see below) (Fig. 2, step 2). Amplicon sequence variants (ASVs – see below) resulting from the bioinformatic analysis of the DNA metabarcoding data were searched against existing DNA barcode reference libraries for freshwater invertebrates using the megablast algorithm. Any ASVs

Fig. 1. Locations around Greater Melbourne, Australia, where macroinvertebrate samples containing Acarina were collected. Large circles indicate sites used for DNA barcoding and metabarcoding, whereas small circles indicate sites only used for individual DNA barcoding.

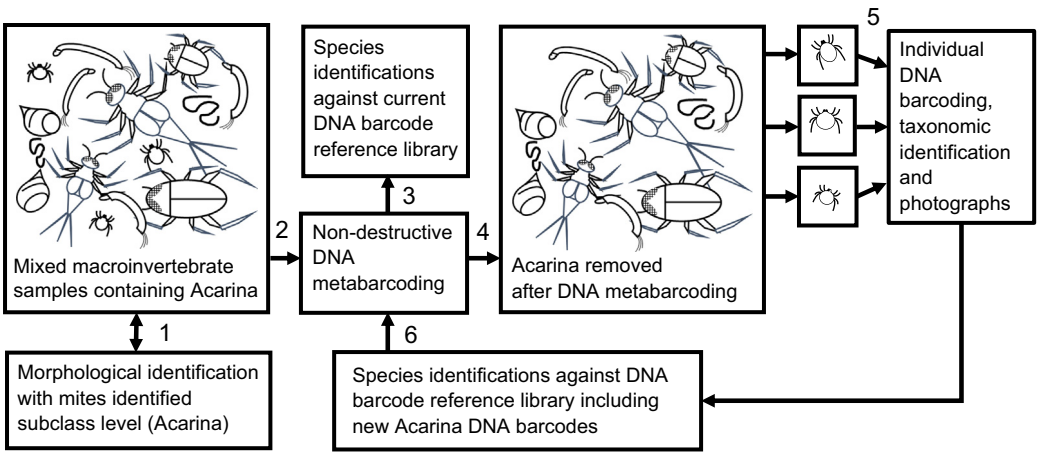

Fig. 2. DNA metabarcoding experimental design used in this study. Numbers indicate the chronological order in which different steps of the study were conducted. Double headed arrow indicates mites were removed for identification and then added back into the sample.

with a highest percentage identity match with query coverage over 80% to an Acarina sequence were recorded. Mites were then removed from bulk macroinvertebrate samples after DNA metabarcoding for individual DNA barcoding, photographing and morphological identification to the lowest possible taxonomic level (Fig. 2, steps 4, 5). We also included water mites (not used for DNA metabarcoding) from additional sites in greater Melbourne (Fig. 1, Table S1) to expand the reference DNA barcode library coverage and to improve taxon coverage for analysis. These additional mites were from samples collected over different years, site types and some with different sorting methods (see Table S1) including some from previously published studies (Carew *et al.* 2018*a*, 2018*b*, 2021). Acarina DNA barcodes generated were searched using the BOLD system identification engine (http://www.boldsystems.org/index.php/IDS_OpenIdEngine, accessed 6 June 2021) to confirm they were of arachnid origin and were analysed to identify

possible species clusters (see below). They were then added to the existing DNA barcode reference library to determine if the new DNA barcodes improved identification of Acarina in bulk macroinvertebrate samples processed with DNA metabarcoding (Fig. 2, step 6).

Macroinvertebrate sample DNA metabarcoding

After morphological identification of macroinvertebrates, the 34 samples used for DNA metabarcoding were bulk processed using a non-destructive DNA extraction protocol (see Carew *et al.* 2018*a*). In brief, this involved dissecting a leg or piece of tissue from large taxa into a single tube, whereas small taxa, including all mites, were co-immersed in a DNA extraction buffer (180 μL of the T1 buffer and 25 μL of proteinase K) from the Nucleospin tissue kit (Macherey-Nagel Inc.) and incubated for 2–3 h at 56°C to release DNA. Specimens were then transferred back to >95% ethanol where they could be used for individual DNA barcoding and taxonomic examination, and the DNA extraction buffer containing the macroinvertebrate DNA was processed using a Nucleospin tissue kit (Macherey-Nagel Inc.) following the manufacturer's instructions to isolate DNA for DNA metabarcoding.

DNA metabarcoding was performed using a two-step PCR process. The first PCR involved amplifying part of DNA barcode region (Hebert *et al.* 2003) using three overlapping PCR primer sets (Fig. 3). The primer sets included BF2/BR2 (Elbrecht and Leese 2017), B/E (Hajibabaei *et al.* 2012) and miCOIintF (Leray *et al.* 2013)/HCO2198-LepR1 (Folmer *et al.* 1994; Hebert *et al.* 2004). Primers were selected based on their ability to amplify a broad range of macroinvertebrate taxa, thereby maximising species detection, and mitigating amplification biases of individual primer sets. Two technical replicates were undertaken per sample.

First round PCR primers contained (5′–3′) a universal adaptor (Illumina Nextera transposase sequence; Illumina Corporation, San Diego, CA, USA) followed by the gene-specific primer sequence (above). First round PCR reactions contained 2 μL of DNA template (1:10 dilution), 8.3 μL of molecular biology grade water, 12.5 μL of KAPA3G PCR buffer with $MgCl_2$ (KAPA Biosystems), 1 μL of $MgCl_2$ (25 mM), 0.5 μL of forward primer (10 μM), 0.5 μL of reverse primer (10 μM), and 0.2 μL of KAPA3G polymerase (5 U mL^{-1}) (KAPA Biosystems) in a total volume of 25 μL and were amplified using the PCR conditions from Carew *et al.* (2021). The three sets of PCR amplicons were pooled for each sample in a ratio 1:1, except that 30% more of the longer (BF2/BR2) amplicon was added to allow for length-based biases in amplification in second round PCRs. Pooled amplicons were cleaned using ExoSAP-IT (Applied Biosystems, Waltham, MA, USA) according to the manufacturer's instructions.

The cleaned pooled amplicons were then used as templates for second round PCRs. Second round primers contained (5′–3′) Illumina p5/p7 adaptor sequence, a unique 8 bp index sequence and part of the universal adaptor sequence (Illumina Corporation). Reactions used 3 μL of the pooled first-round amplicons, 12.5 μL of MiFi mix (Bioline, London, UK), 2.5 μL of forward p5 index primer (10 μM), and 2.5 μL of reverse p7 primer (10 μM). PCR conditions were as follows: 94°C for 5 min followed by 12 cycles of 94°C for 30 s, 55°C for 30 s, 72°C for 30 s, then one cycle of 72°C for 5 min. Amplicons were pooled in equal amounts and the library was gel purified using a PureLink Quick Gel Extraction and PCR Purification Combo Kit (Invitrogen, Waltham, MA, USA). Illumina MiSeq sequencing was performed by the Australian Genome Research Facility Ltd (AGRF) using a 600-cycle flow cell MiSeq sequencing kit V3 (300 bp × 2) (Illumina Corporation).

DNA metabarcoding included negative controls from first round PCR amplifications; a macroinvertebrate control sample of known composition (to verify continuity of species detection between MiSeq runs); and a sample containing DNA from *Scaptodrosophila xanthorrhoeae* – an invertebrate restricted to north Queensland, Australia and

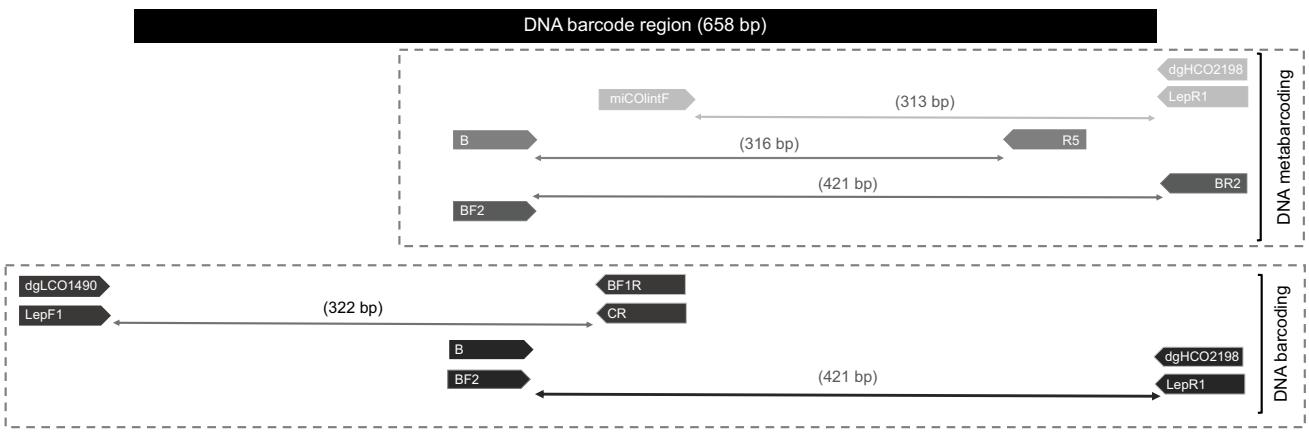

Fig. 3. A map of the primer combinations used for DNA metabarcoding of bulk macroinvertebrate samples (three amplicons) and DNA barcoding of individual Acarina (two amplicons).

not expected to occur in aquatic invertebrate samples – to check for sample cross-contamination and index switching. DNA from this species should not appear in macroinvertebrate samples if there is no cross-contamination and index switching.

Read pre-processing and sample demultiplexing was performed by AGRF. The DNA metabarcoding data was then analysed using a custom pipeline. Reads were trimmed to remove primer sequences and sorted into groups corresponding to the different amplicon regions (see Fig. 3) using Cutadapt (ver. 1.16, see https://cutadapt.readthedocs.org/; Martin 2011), standard Unix bash commands and the filter_fasta.py script from QIIME (ver. 1.9.0, see http://qiime.org/; Caporaso *et al*. 2010). Quality filtering, chimaera removal and read clustering were performed with QIIME2 (ver. 2019.4, see https://qiime2.org/; Bolyen *et al*. 2019) using the DADA2 plugin (Callahan *et al*. 2016). All resulting ASVs were filtered at 0.001% to remove singletons and low frequency ASVs where ASVs were removed from a sample if they made up less than 0.001% of the reads in that sample (see Elbrecht and Steinke 2019). The remaining ASVs were then searched against freshwater invertebrate DNA barcodes from the BOLD Systems (ver. 4, http://www.boldsystems.org/, accessed 6 June 2021) database and our private freshwater DNA barcode database (see Carew *et al*. 2021) using the megablast algorithm, within Geneious prime (ver. 2021.1, see https://www.geneious.com, accessed 6 June 2021). The BLAST search was then repeated after adding new individual Acarina DNA barcodes to our existing freshwater invertebrate DNA barcode reference library. The overall detection of Acarina in macroinvertebrate samples and the BLAST search results were compared between the old and new DNA barcode reference libraries.

Individual DNA barcoding and identification of Acarina

Mites either were retrieved from macroinvertebrate samples after DNA metabarcoding or separated from other macroinvertebrates (for the DNA barcoding only samples) and placed in individual 0.5-mL tubes containing 100% ethanol. Non-destructive Chelex extractions following Carew *et al*. (2018*a*) were used to obtain DNA for PCR and subsequent DNA barcoding. In brief, this method involved placing an individual mite into a 0.5-mL tube with 150 μL of 5% Chelex solution and 2 μL of proteinase K (22 mg mL^{-1}) and then incubating at 56°C for 30 min. Mites were then carefully retrieved from the tube using forceps with the assistance of a Leica L2 dissecting microscope (Leica Microsystems, Wetzlar, Germany) and placed back into 0.5-mL tubes containing 100% ethanol for morphological examination. After mites were removed, Chelex extractions were incubated at 90°C for 10 min to inactivate the proteinase K. Chelex extractions were centrifuged at 15 871*g* for 2 min and the supernatant was used as a source of DNA for PCR and subsequent DNA barcoding.

DNA barcoding was conducted mostly using Illumina MiSeq sequencing using a protocol modified from Shokralla *et al*. (2015). In brief, two overlapping amplicons which cover the entire DNA barcode region were amplified using degenerate primer pair cocktails (Fig. 3). The primer sets included BF2 – B (forward) with dgHCO2198 – LepR1 (reverse) (Folmer *et al*. 1994; Hebert *et al*. 2004) to amplify the 5′ end of the DNA barcode region and dgLCO1490–LepF1 (forward) (Folmer *et al*. 1994; Hebert *et al*. 2004) with C (CR) and BR1 (BF1R) (reverse) (Hajibabaei *et al*. 2012; Elbrecht and Leese 2017) to amplify the 3′ end of the DNA barcode region. Amplifications used the same workflow, PCR conditions and profiles as DNA metabarcoding (above) except all reaction volumes were reduced by 50% and a single replicate was used. Samples for individual DNA barcoding were sequenced by AGRF on an Illumina MiSeq platform, using a 600-cycle flow cell MiSeq sequencing kit V3 (300 bp × 2) (Illumina Corporation).

Read pre-processing and sample demultiplexing were also performed by AGRF. The DNA barcoding data were then analysed using customised automated workflows created using Geneious prime (https://www.geneious.com). A different customised automated workflow was used for each amplicon. The workflow was configured to merge paired reads, trim primers, *de novo* assemble reads into contigs and then search the contig consensus sequences against the same DNA database of freshwater invertebrates used for DNA metabarcoding (described above) using the megablast algorithm. After inspecting the BLAST results for contigs likely to be from Acarina, the two amplicons were assembled (using the overlapping region between the two amplicons) with the *de novo* assemble function in Geneious prime to produce a full-length DNA barcode. Assembled DNA barcodes obtained from mites were then searched using the BOLD systems identification engine (see http://www.boldsystems.org/index.php/IDS_OpenIdEngine, accessed 6 June 2021) and percentage identity to publicly available data were recorded. DNA barcoding of a small number of individuals (5) were taken from Carew *et al*. (2018*a*) where mites were Sanger DNA sequenced.

Major Acarina groups (Hydracarina, Oribatida, Halacaroidea, Mesostigmata, Trombidioidea) were identified using a Leica L2 dissecting microscope (Leica Microsystems) with the online keys from the Centre for Freshwater Ecosystems (formerly the Murray–Darling Freshwater Research Centre), La Trobe University (see https://www.mdfrc.org.au/bugguide/display.asp?class=16&subclass=&order=&Couplet=0&Type=2, accessed 6 June 2021). Hydracarina were identified using a high magnification dissecting Leica m80 microscope (Leica Microsystems) and, for some genera, using a Biotic B2 compound microscope (Motic, Hong Kong, PR China) to finer taxonomic levels with keys from Harvey (1996, 1998), Smit (2010), and Viets (1978). Pezidae from the Halacaroidea were identified using Harvey (1989). A small number of identifications from the other groups were

assigned from BOLD if DNA barcodes closely matched (>97%) sequences with taxonomic identifications in the BOLD systems database.

Analysis of DNA barcodes

All Acarina DNA barcodes were aligned using Clustal Omega (Thompson *et al.* 1999) in MEGA X (ver. 11, see www.megasoftware.net, accessed 6 June 2021). A Kimura-2-parameter neighbour-joining tree with 1000 bootstraps was used to examine the clustering pattern of DNA barcodes, for comparison to morphological data.

The generalised mixed Yule coalescent (GMYC) model was used to delineate species (Pons *et al.* 2006). Two datasets were constructed for GMYC analysis. One dataset contained full length DNA barcodes (658 bp) from our samples. To include a greater number of samples we constructed a second dataset by using truncated DNA barcodes (of 362 bp) from our dataset. DNA barcodes were excluded from the GYMC analysis if they belonged to clades with conflicting morphology (identified through the NJ tree). We conducted single threshold GMYC species delineation (Monaghan *et al.* 2009). Ultrametric gene trees for the GMYC approach were reconstructed under a strict molecular clock model with BEAST (ver. 2.6.3, see https://beast.community/index.html; Drummond *et al.* 2012). The BEAST input files were generated with BEAUti (ver. 2.6.3, see https://beast.community/beauti; Drummond *et al.* 2012) and were based on the HKY gamma model – the optimum evolutionary model under the Akaike Information Criterion (AIC). Mean substitution rate was set to one, the base frequencies were estimated from the data, six gamma categories were used, and the substitution model, the rate heterogeneity and the base frequencies were unlinked across partitions. All other parameters were set to the default values. Ten independent MCMC chains were run for 10 million generations and sampled every 1000 generations. Run convergence was visualised using Tracer (ver. 1.7.1, A. Rambaut and A. Drummond, see http://beast.bio.ed.ac.uk/Tracer). To account for burn-in, the first 5000 trees were discarded from each run and the independent log and tree files were combined with LogCombiner (ver. 2.6.3, see https://beast.community/logcombiner). Combined runs were completed once estimated sample sizes (ESS) >200 were reached – this parameter has been used in similar studies using single GMYC thresholds (Vuataz *et al.* 2011; Puillandre *et al.* 2012). TreeAnnotator (ver. 2.6.3, see https://www.beast2.org/treeannotator/) was run to produce a single tree using the maximum clade credibility tree with all other options set to default. For each dataset, single GMYC models were applied to the DNA barcode tree using the script available within the SPLITS package (see http://r-forge.r-project.org/projects/splits/, accessed 6 June 2021) for R. The GMYC putative species groups from each dataset were compared for concordance and node support.

All DNA barcodes, geographical information and photographs were submitted to the Barcode of Life BOLD systems V4 database (see http://www.boldsystems.org/index.php/databases, accessed 6 June 2021) and DNA barcode sequences were submitted to GenBank (Table S3). Barcode Index Numbers (BINs) were assigned in the Barcode of Life BOLD systems V4 after submission using refined single-linkage cluster (RESL) analysis (Ratnasingham and Hebert 2013). BINs were then compared for concordance to GMYC groupings.

Results

Morphological analysis of individual Acarina

A total of 152 mites were removed from macroinvertebrate samples from 50 sites. After Chelex extraction, we recovered 148 Acarina specimens for morphological analysis as we were unable to locate four specimens. Mites recovered could be identified to the major group levels Hydracarina, Oribatida, Halacaroidea, Mesostigmata, and Trombidioidea using the Centre for Freshwater Ecosystems keys (Fig. 4). The

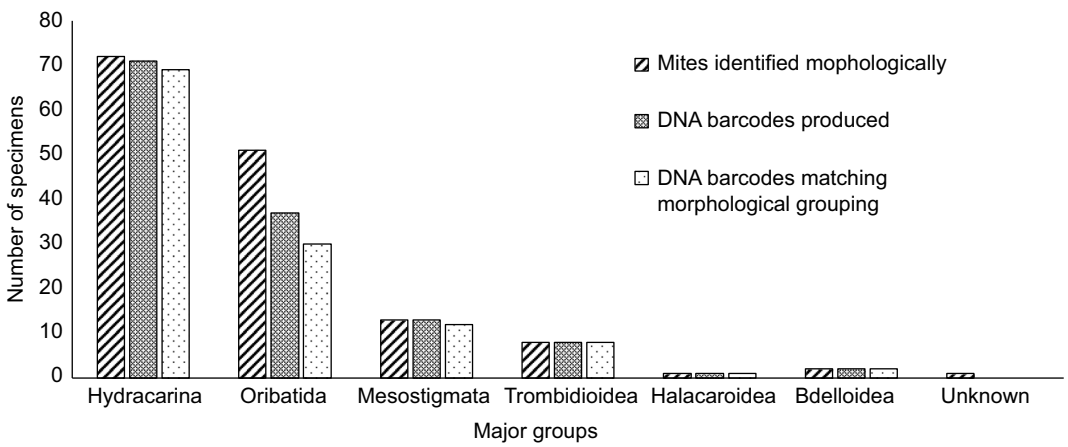

Fig. 4. Individual DNA barcoding of Acarina from macroinvertebrate samples by major taxonomic group.

Hydracarina were the most abundant group (72 specimens) followed by the Oribatida (51), Mesostigmata (13), Trombidioidea (8) and Halacaroidea (1). We also found two mites from the Bdelloidea and one which we could not identify and was possibly a terrestrial species.

We were able to identify a total of 20 families, 21 genera and 3 species. Most identified specimens belonged to the Hydracarina with 12 families and 18 genera identified (Table 1). We were able to identify four families and one species in the Oribatida, and one freshwater family and genus in the Halacaroidea. Some identifications in the Mesostigmata (2) and Oribatida (1) were made based on species-level DNA barcode matches (i.e. >97%) to DNA barcode sequences on the BOLD systems database (Table 1).

Table 1. Morphological identification of Acarina found in freshwater macroinvertebrate samples.

Minor group	Family	Lowest identification	Number of individuals	Habitat
Bdelloidea	Bdellidae	Bdellidae	2	Terrestrial
Halacaroidea	Pezidae	*Peza*	1	Freshwater
Hydracarina	Arrenuridae	*Arrenurus*	1	Freshwater
	Aturidae	*Austraturus*	1	Freshwater
	Aturidae	Aturidae	1	Freshwater
	Hydrachnidae	*Hydrachna*	2	Freshwater
	Hydrodromidae	*Hydrodroma*	1	Freshwater
	Hydryphantidae	*Diplodontus*	2	Freshwater
	Hygrobatidae	*Aspidiobates*	3	Freshwater
	Hygrobatidae	*Australiobates*	8	Freshwater
	Hygrobatidae	*Australorivacarus*	1	Freshwater
	Hygrobatidae	*Caenobates*	14	Freshwater
	Hygrobatidae	*Hygrobates*	3 (2)	Freshwater
	Hygrobatidae	*Procorticacarus*	4	Freshwater
	Hygrobatidae	*Rhynchaustrobates*	4	Freshwater
	Hygrobatidae	Hygrobatidae	1 (1)	Freshwater
	Limnesiidae	*Limnesia*	12	Freshwater
	Limocharidae	*Austrolimnochares*	2	Freshwater
	Mideopsidae	Mideopsidae	2	Freshwater
	Oxidae	*Oxus*	5	Freshwater
	Pionidae	*Acercella*	1	Freshwater
	Pionidae	*Piona*	2	Freshwater
	Pionidae	Pionidae	1	Freshwater
	Unionicolidae	*Neumania*	2	Freshwater
Mesostigmata	Macrochelidae	Macrochelidae[A]	2	Terrestrial
	Parasitidae	*Pergamasus crassipes*[A]	1	Terrestrial
	Parasitidae	*Pergamasus quisquiliarum*[A]	1	Terrestrial
		Mesostigmata	9 (1)	Terrestrial
Oribatida	Humerobatidae	Humerobatidae[A]	1	Terrestrial
	Hydrozetidae	Hydrozetidae	9	Freshwater
	Oribatulidae	*Zygoribatula undulata*[A]	1	Terrestrial
	Neotrichozetidae	Neotrichozetidae	2	Terrestrial
		Oribatida	36 (8)	Uncertain
Trombidioidea		Trombidioidea	9	Uncertain
Unknown			1 (1)	Terrestrial

The number of specimens mismatched between morphology and DNA barcodes are shown in parentheses.
[A]Finer level identifications made through species matches (>97%) to the BOLD systems database.

Individual Acarina DNA barcoding

Acarina DNA barcodes were obtained from 135 mites; this included 117 full length DNA barcodes and 18 partial DNA barcodes (12 of 422 bp, 3 of 322 bp, 1 of 655 bp from MiSeq DNA sequencing; 1 of 362 bp, 1of 580 bp from Sanger DNA sequencing). This dataset also included our previously published DNA barcodes (GenBank Accession number MG976102, MG976100, M976203, MW051403, KX198797, KX198766). An additional six DNA barcodes were found by BLAST searches to be from other macroinvertebrates (two Baetidae, two Leptophlebiidae, one Chironomidae and one Paramelitidae) that had been previously co-stored and extracted with the Acarina specimens and were omitted from further analysis. DNA amplification and barcoding success across Acarina groups varied. Notably, we failed to produce DNA barcodes for 27% of the Oribatida (Fig. 4). In comparison, 98% of specimens from other mite groups produced either partial or full-length DNA barcodes.

To examine the clustering between the DNA barcodes, a neighbour-joining tree was used to display distance-based grouping of all 135 Acarina DNA barcodes (Fig. 5). Terminal grouping of specimens mostly aligned with their morphology. However, we did find that the DNA barcodes of 13 specimens did not group with specimen morphology (Table 1, Fig. 4). This tended to occur mostly in the Oribatida, where eight specimens had DNA barcodes that grouped in our neighbour-joining tree (Fig. 5) with (mostly co-stored) Trombidiformes specimens. There were three instances where a Hydracarina specimen's DNA barcode grouped to the wrong Hydracarina genus or family. A Mesostigmata specimen had a DNA barcode which incorrectly grouped to the Trombidioidea and the single unidentified mite specimen had a DNA barcode which incorrectly grouped to *Piona* (Pionidae).

Species delineation using GMYC analysis of 117 full-length DNA barcodes (658 bp) produced 24 GMYC groups using 76 DNA barcodes, with the remaining 41 DNA barcodes represented by singletons (Fig. 6, Table S2). An additional GMYC group was found when examining 131 truncated DNA barcodes (362 bp) using GMYC analysis. We found that 25 GMYC groups were formed from 87 DNA barcodes and 44 DNA barcodes were represented by singletons. Overall, the GMYC groups were the same for both datasets. The greatest number of GMYC groups were found in the Hydracarina with 16 groups found (Fig. 7, Table S2). High support was found for all Hydracarina GMYC groupings, with node support of 1. There were 16 singletons in the Hydracarina when considering DNA barcodes that matched morphology, and no extra GMYC groups were found when considering DNA barcodes from the BOLD systems database. The Oribatida formed five GMYC groups (Fig. 5). One GMYC group was formed to due species matches of 100% to a Humerobatidae specimen from New Zealand (BOLD:ACP5818) on the BOLD systems database. Two other groups were formed by specimens from our study belonging to the Neotrichozetidae and an unidentified family from the Plenty River site (Table S1). The Hydrozetidae formed two putative species groups although support for this split was low for the truncated dataset (0.58) but higher for the full-length DNA barcodes (0.78). In the Mesostigmata where matches of 100% to *Pergamasus quisquiliarum* and *Pergamasus crassipes* from the BOLD systems database were found (Fig. 5), forming two GMYC groups. A well supported GMYC group was seen in the Trombidioidea with the full-length dataset and using the truncated dataset an extra GMYC group was found.

The RESL analysis, which assigned BIN codes to data, was fully concordant with the groupings from the GMYC analysis. There were 61 different BINs found in the dataset for specimens with DNA barcodes with a length of greater than 450 bp (Table S2). The Hydrozetidae formed two GYMC groups and were assigned two different BINs.

Many novel DNA barcodes were found, with only 13% of GMYC groups or singleton DNA barcodes matching at greater than 97% sequence identity to DNA barcodes on the BOLD systems database. Most DNA barcodes matched at less than 85% sequence identity (Fig. 6), showing that most DNA barcodes from Australian Acarina were substantially different to those on the BOLD systems database. GenBank Accession numbers for DNA barcodes and, BOLD systems taxon codes and BINs can be found in Table S2.

DNA metabarcoding

MiSeq sequencing of samples from 34 sites yielded 2 474 765 high quality reads (Table S3) with greater than 12 800 reads (average 36 260 reads; s.d. ±10 766 reads) per replicate and totalling greater than 45 000 reads (average 72 521 reads; s.d. ±15 390 reads) per macroinvertebrate sample. Raw reads from DNA metabarcoding were uploaded to the National Centre for Biotechnology Sequence Read Archive under BioProject PRJNA764021 (Table S3). We found no sequences from the *S. xanthorrhoeae* control sample in the other samples, suggesting that sample cross-contamination and index switching were non-existent or at a very low level. The detection of species in the control sample of known taxonomic composition was consistent with previous MiSeq runs. We found a small number of reads (<10 reads) in the PCR control.

We were able to assign greater than 85% of the total ASVs found to species level using the original DNA barcode reference library. Macroinvertebrate family and species diversity varied between the 34 samples (Table S3). Including Acarina, a total of 112 families and 551 species were identified. Technical replicates showed some variability with 62% of the ASVs present in both replicates. However, less variation was found when ASVs were identified to species with 86% of species found in both replicates. We found 106 ASVs had a top BLAST match with Acarina sequences from the original DNA

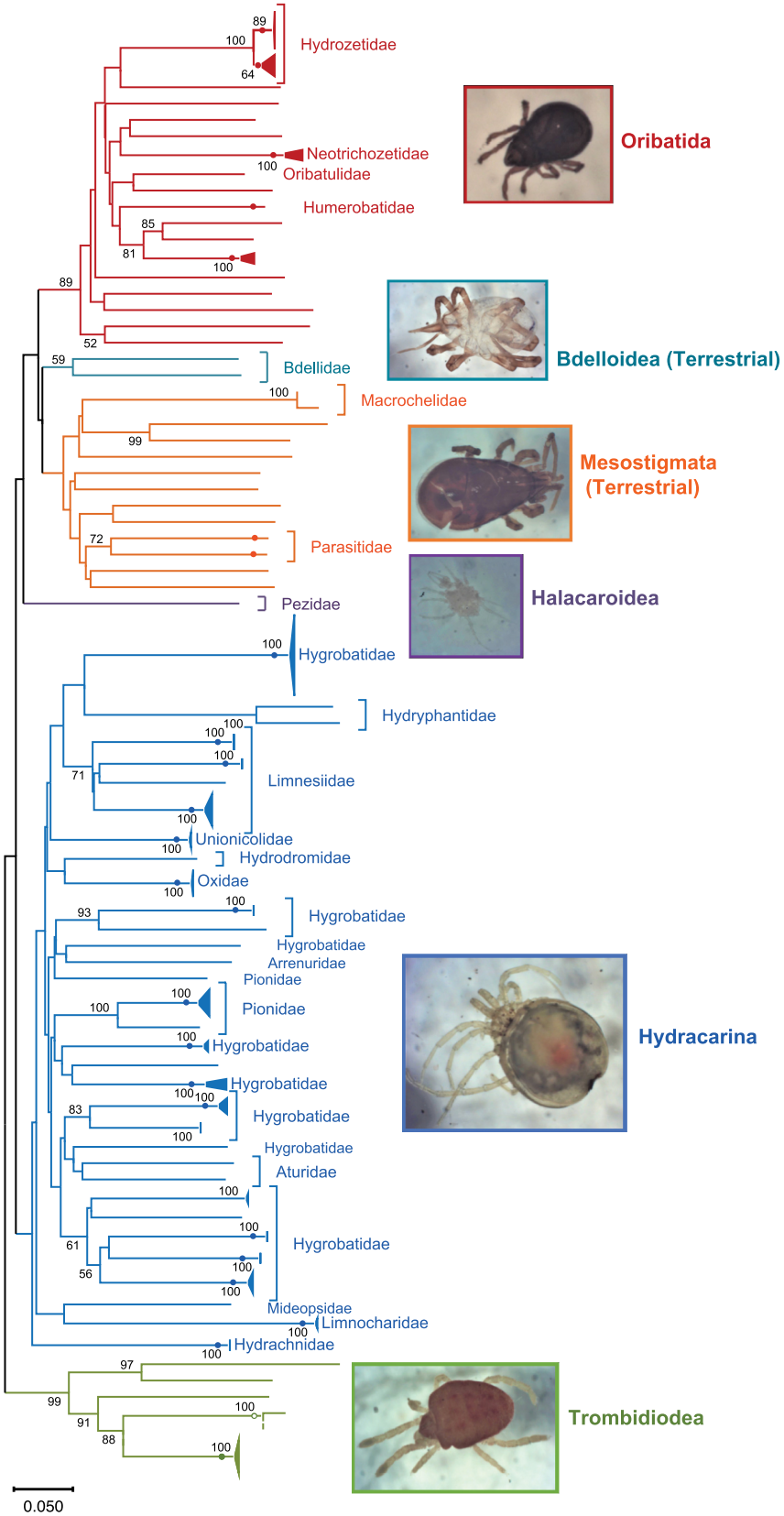

Fig. 5. Neighbour-joining tree of 135 mite DNA barcodes found in this study showing the diversity of minor taxonomic groups found in macroinvertebrate samples. Each major group is colour coded and the coloured circles on terminal branches represent GMYC groups and BOLD BINs. One GMYC grouping with an open circle represents the additional group in the Trombidioidea found using with analyses of the truncated dataset. Family level identification are provided for Oribatida and Mesostigmata if they matched taxa using the BOLD system identification engine. Bootstraps are based on 1000 replicates, with bootstrap values below 50 not shown. Scale bar is the substitutions per nucleotide position.

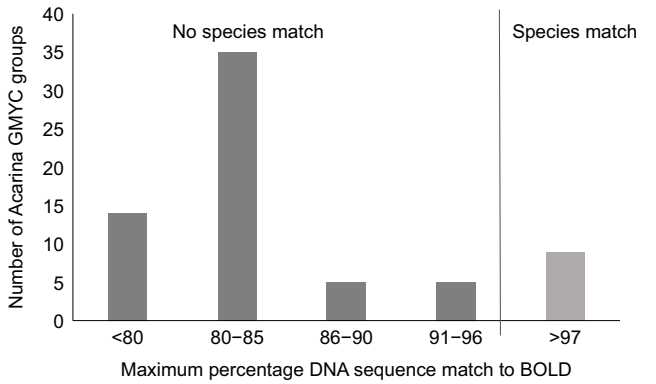

Fig. 6. Acarina DNA barcode matches to the BOLD systems database. Matches of >97% are considered species matches. Most 'no species match' sequences matched Acarina or Arachnida at finer levels.

barcode reference library. These ASVs were found among 30 of the 34 samples that contained Acarina. However, 70% of ASVs that had a mite reference sequence as the top-scoring BLAST hit shared <97% identity with the match, which is below the threshold required to infer a likely species match and provide confidence in taxonomic assignment (Fig. 8). After adding new mite DNA barcodes to the DNA barcode reference library, we found an extra 14 ASVs that matched Acarina in the DNA metabarcoding dataset. This also increased the number of samples in which we detected Acarina to 32 out of 34 macroinvertebrate samples with 0.93% of the total reads (22 853 reads) identified as Acarina. As expected, we found that many of the Acarina detected (88%) had greater than 97% matches to Acarina DNA barcodes in the new reference library (Fig. 8) with finer-level taxonomic identification resulting from our morphologically identified voucher collection. However, there were still 18 ASVs with a top BLAST match to Acarina that were not in the new reference library (Table S3). The detection of DNA barcoded mites compared with those detected using DNA metabarcoding from each site was mostly concordant, with 58% of taxa detected using both methods. Most of the difference between the methods (70%) could be attributed to a lack detection of terrestrial taxa with DNA metabarcoding, whereas 30% of aquatic taxa were not detected. This resulted in an overall concordance between DNA barcodes and metabarcoding of 72% when only freshwater taxa were considered. We also found 6% of taxa were only detected with DNA metabarcoding.

Discussion

Here we generated DNA barcodes for 135 freshwater mites from south-eastern Australia, which included 52 BIN's (OTU's) that are new to the BOLD and GenBank databases. This new resource will facilitate future biodiversity assessments and research within the Australian context and is an important step in remedying the inadequate representation of freshwater mites within reference sequence libraries. For our DNA metabarcoding dataset, these new barcodes allowed a greater number of mite taxa to be identified and improved the taxonomic resolution of mite identifications.

We found substantial gaps in existing DNA barcode library coverage for Australian mites, with few mite DNA barcodes generated in this study having species level matches (>97%) in publicly available DNA databases. This impedes confident identification of Acarina with DNA metabarcoding as megaBLAST matches of Acarina ASVs to current databases were often less than 85%, and Acarina ASVs were often misidentified with low level BLAST matches to other macroinvertebrate taxa. For example, DNA barcodes for the Pezidae (*Peza*) and Limnocharidae (*Austrolimnochares*) showed low percentage matches with barcodes from other invertebrates on the BOLD systems database. Not surprisingly, both the percentage match and Acarina detection substantially improved after targeted individual DNA barcoding of Acarina. Hence, our study highlights the importance of comprehensive DNA barcode libraries for the accurate identification species with DNA metabarcoding, as examined previously by Weigand *et al.* (2019).

Identification of mites detected with DNA metabarcoding compared to those recovered from samples for individual DNA barcoding were mostly concordant. However, we did find that terrestrial taxa, such as many Oribatida and Mesostigmata, were more likely to be missed by DNA metabarcoding compared to aquatic taxa, such as the Hydracarina and Hydrozetidae. This could be attributed to the smaller size of many terrestrial species, potential DNA degradation if animals were dead in samples before being captured, and hard exoskeletons (mainly of the Oribatida) that impeded the release of DNA during the non-destructive DNA extraction process used for DNA metabarcoding (see Carew *et al.* 2018*a*). However, we also failed to detect some Hydracarina in some samples, which was often associated with individuals being present as singletons or juvenile life stages. Compared to larger macroinvertebrates, mites contributed smaller quantities of DNA for DNA metabarcoding, which was reflected by the overall low percentage of reads (<1%) that were recovered for mites. We would suggest using a greater sequencing depth when metabarcoding to increase the opportunity to detect Acarina, as many taxa in this study were detected with fewer than 20 reads, with missing taxa possibly below our detection limit. We also detected some additional taxa only with DNA metabarcoding, which may be due to a failure to recover some individuals for DNA barcoding or differences in the amplification success between the primer sets used for DNA barcoding versus metabarcoding. It is not clear if these extra mite detections represent false positives.

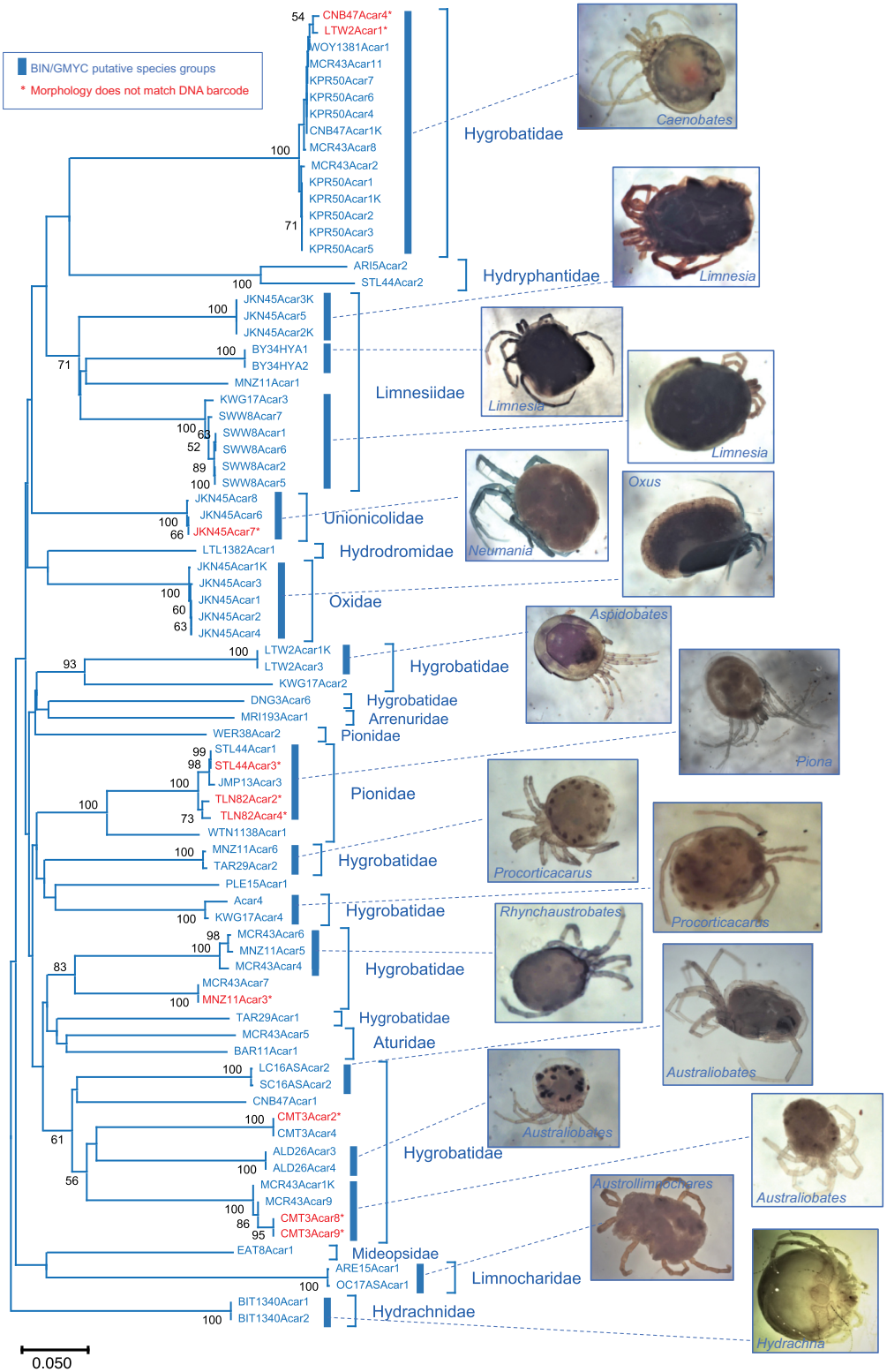

Fig. 7. Neighbour joining tree of Hydracarina DNA barcodes found in this study showing species delineation from BIN and GMYC analysis including DNA barcodes (in red) that did not match the specimen's morphology (contamination). Bootstraps are based on 1000 replicates, with bootstrap values below 50 not shown. Scale bar is the substitutions per nucleotide position.

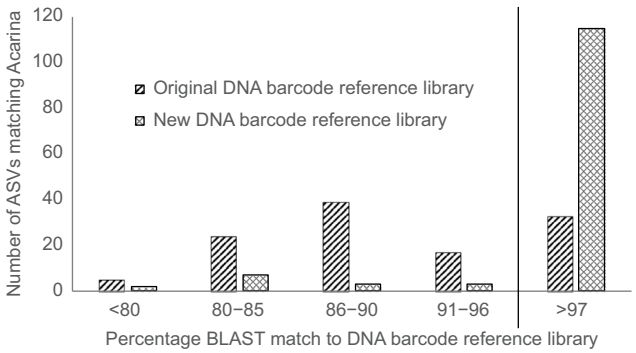

Fig. 8. DNA metabarcoding ASV BLAST matches before and after targeted individual DNA barcoding of Acarina. Matches of greater than 97% are considered a species match.

Our study has shown that grouping Acarina to the sub-class level misses the opportunity to differentiate between terrestrial and aquatic species. Many taxa detected such as those from the Mesostigmata and Bdelloidea (and some Oribatida and Trombidioidea) represent terrestrial species (Schatz and Behan-Pelletier 2008; Walter and Proctor 2013). However, these can be easily distinguished from the fully aquatic Hydracarina when finer taxonomic identification is used, which can be easily facilitated by using DNA barcodes linked to taxonomic identified specimens. Furthermore, using DNA barcodes to provide finer level identification can improve information on family, genus and species environmental responses that can be used in biomonitoring (Baird and Hajibabaei 2012). We were able to provide genus level identifications for most Hydracarina using currently available taxonomic keys (e.g. Harvey 1998), and with the involvement of experienced taxonomists many of the genera we found could be further identified to species. The use of finer level identifications of Hydracarina has been shown to be useful for biomonitoring of European lotic environments (Miccoli *et al.* 2013) and in water mites in Australia (Growns 2001).

The non-destructive DNA extraction techniques outlined in Carew *et al.* (2018*a*) proved useful for targeting Acarina for individual DNA barcoding, as most specimens could be identified morphologically after using non-destructive DNA extraction for DNA metabarcoding followed by a second non-destructive DNA extraction for individual DNA barcoding. Although most specimens were recovered and remained fully intact, some tissue was dissolved with proteinase K during the extraction process which meant that some species became hyaline or had missing eye spots, which could make identification of a small number of taxa more challenging. Unfortunately, the Oribatida were the least suited to this approach as many specimens were damaged or did not produce DNA barcodes. This may have been due to the small size and the hard exoskeleton of Oribatid mites affecting the penetration of proteinase K during the non-destructive Chelex DNA extraction. These mites were also more fragile and easily damaged when being removed from the Chelex DNA extraction compared to softer bodied specimens. Despite this, the extraction process did not appear to affect setae for most other minor groups, which is important because these are often key characters for specimen identification (Harvey 1998).

Most of the DNA barcodes in our study were generated using MiSeq DNA sequencing, which we found to be successful for producing DNA barcodes for Acarina. Shokralla *et al.* (2015) noted a vast improvement in the success of DNA barcoding for Trombidiformes using MiSeq DNA sequencing compared to Sanger DNA sequencing and our study supports this finding as our earlier attempts to sequence Acarina barcodes using Sanger sequencing often failed (data not shown). However, it is still important to verify that DNA barcodes recovered from a specimen also match morphology. We found that some DNA barcodes, especially those isolated from the Oribatida, did not match morphology but represented other species including other mites. This often occurred when mites from different orders were found in the same sample. We suspect this could have occurred by two processes; either the mites were feeding on other species before their capture or DNA leaching occurred during co-storage or bulk DNA extraction. The latter is the most likely culprit due to the high amount of DNA released during the non-destructive extraction process. We suspect that this issue is most problematic for mite species that are poorly identified through the current DNA barcoding primer sets developed for insects (e.g. Folmer *et al.* 1994). With the Oribatida having the greatest number of failed PCR amplifications, DNA barcodes from this group may need to be generated with other primer sets. Alternatively, Oribatids may be less suited to this barcoding approach due to their small size and hard exoskeleton, or may have shown a tendency towards low quality DNA yields because their DNA had been previously extracted for DNA metabarcoding.

We were able to identify the aquatic Oribatida family, Hydrozetidae, based on morphology and species matches (>97%) to the BOLD systems database. To date, a single species with a cosmopolitan distribution, *Hydrozetes lemnae*, has been recorded in Australia, although the possibly of more species was noted by Colloff and Halliday (1998). The Hydrozetidae species found in this study matched DNA barcodes from Canada, suggesting the species is not endemic to Australia. Similarly, we also found DNA barcodes that had species matches (>97%) to European and Canadian Mesostigmata mite samples from the Parasitidae, including *Pergamasus quisquiliarum* and *Pergamasus crassipes*.

Despite extensive DNA barcode coverage of mites from Canada (Young *et al.* 2012, 2019) and increasing coverage from Europe (Weigand *et al.* 2019), there were no matches of these DNA barcodes to Australian Hydracarina, which shows the importance of DNA barcoding local fauna in this

region and that local fauna are often not cosmopolitan. Although we were largely successful in providing finer-scale identification for Hydracarina, we struggled to add taxonomic information for other minor Acarina groups. For example, multiple species of Trombidioidea were likely to have been present, some of which may have been aquatic or sub-aquatic. However, we were unable to find keys to assist with identification of this group.

Conclusion

In conclusion, water mites from a region in south-east Australia showed high diversity and endemism, with many new DNA barcode BIN's, and several unique DNA barcodes being represented by single specimens. Further expanding DNA barcoding of aquatic mites throughout Australia will undoubtedly provide new insights into species diversity and distributions. A coordinated effort with taxonomists is needed to link species level taxonomy to DNA barcodes to provide a national comprehensive species-level DNA barcode library for aquatic mites. This DNA barcode library will provide the confidence needed to identify Acarina both as individuals and in bulk samples or environmental DNA samples analysed with DNA metabarcoding. This combined with environmental data can improve our understanding of the species-specific responses for use in biomonitoring.

Supplementary material

Supplementary material is available online.

References

Baird DJ, Hajibabaei M (2012) Biomonitoring 2.0: a new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. *Molecular Ecology* **21**, 2039–2044. doi:10.1111/j.1365-294X.2012.05519.x

Blattner L, Gerecke R, von Fumetti S (2019) Hidden biodiversity revealed by integrated morphology and genetic species delimitation of spring dwelling water mite species (Acari, Parasitengona: Hydrachnidia). *Parasites & Vectors* **12**, 492. doi:10.1186/s13071-019-3750-y

Bolyen E, Rideout JR, Dillon MR, *et al.* (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**, 852–857. doi:10.1038/s41587-019-0209-9

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**, 581–583. doi:10.1038/nmeth.3869

Caporaso JG, Kuczynski J, Stombaugh J, *et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature Methods* **7**, 335–336. doi:10.1038/nmeth.f.303

Carew ME, Hoffmann AA (2015) Delineating closely related species with DNA barcodes for routine biological monitoring. *Freshwater Biology* **60**, 1545–1560. doi:10.1111/fwb.12587

Carew ME, Coleman RA, Hoffmann AA (2018*a*) Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ* **6**, e4980. doi:10.7717/peerj.4980

Carew ME, Kellar CR, Pettigrove VJ, Hoffmann AA (2018*b*) Can high-throughput sequencing detect macroinvertebrate diversity for routine monitoring of an urban river? *Ecological Indicators* **85**, 440–450. doi:10.1016/j.ecolind.2017.11.002

Carew ME, Hoffmann AA, Kellar CL, Stevenson K, Walsh CJ, Danger A, Coleman RA (2019) Integrating DNA metabarcoding into Melbourne Water's macroinvertebrate monitoring program. Melbourne Water Corporation, Australia.

Carew ME, Coleman RA, Robinson KL, Hoffmann AA (2021) Using unsorted sweep-net samples to rapidly assess macroinvertebrate biodiversity. *Freshwater Science* **40**, 551–565. doi:10.1086/716215

Chessman BC (1995) Rapid assessment of rivers using macroinvertebrates: a procedure based on habitat-specific sampling, family level identification and a biotic index. *Austral Ecology* **20**, 122–129. doi:10.1111/j.1442-9993.1995.tb00526.x

Colloff MJ, Halliday RB (1998) 'Oribatid mites: a catalogue of Australian genera and species.' (CSIRO Publishing: Melbourne, Vic., Australia)

Di Sabatino A, Smit H, Gerecke R, Goldschmidt T, Matsumoto N, Cicolani B (2007) Global diversity of water mites (Acari, Hydrachnidia; Arachnida) in freshwater. *Hydrobiologia* **595**, 303–315. doi:10.1007/s10750-007-9025-1

Drummond AJ, Suchard MA, Xie D, Rambaut A (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**, 1969–1973. doi:10.1093/molbev/mss075

Elbrecht V, Leese F (2017) Validation and development of COI metabarcoding primers for freshwater macroinvertebrate bioassessment. *Frontiers in Environmental Science* **5**, 11. doi:10.3389/fenvs.2017.00011

Elbrecht V, Steinke D (2019) Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology* **64**(2), 380–387. doi:10.1111/fwb.13220

EPA Victoria (2021) Guidelines for environmental management; water. Available at https://www.epa.vic.gov.au/about-epa/publications/604-2 [Verified 9 October 2021]

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.

Glowska E, Dragun-Damian A, Broda L, Dabert J, Dabert M (2014) DNA barcodes reveal female dimorphism in syringophilid mites (Actinotrichida: Prostigmata: Cheyletoidea): *Stibarokris phoeniconaias* and *Ciconichenophilus phoeniconaias* are conspecific. *Folia Parasitologica* **61**(3), 272–276. doi:10.14411/fp.2014.030

Goldschmidt T (2016) Water mites (Acari, Hydrachnidia): powerful but widely neglected bioindicators – a review. *Neotropical Biodiversity* **2**, 12–25. doi:10.1080/23766808.2016.1144359

Goldschmidt T, Helson JE, Williams DD (2016) Ecology of water mite assemblages in Panama – first data on water mites (Acari, Hydrachnidia) as bioindicators in the assessment of biological integrity of neotropical streams. *Limnologica* **59**, 63–77. doi:10.1016/j.limno.2016.03.007

Growns JE (2001) Aquatic mites as bioindicators, with an Australian example. In 'Acarology: proceedings of the 10th international congress', 5–10 July 1998, Canberra, ACT, Australia. (Eds DE Walter, RB Halliday, HC Proctor, RA Norton, MJ Colloff) pp. 136. (CSIRO Publishing: Melbourne, Vic., Australia)

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS ONE* **6**, e17497. doi:10.1371/journal.pone.0017497

Hajibabaei M, Spall JL, Shokralla S, van Konynenburg S (2012) Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology* **12**, 28. doi:10.1186/1472-6785-12-28

Harvey MS (1989) Pezidae, a new freshwater mite family from Australia (Acarina: Halacaroidea). *Invertebrate Systematics* **3**, 771–781. doi:10.1071/IT9890771

Harvey MS (1996) A review of the water mite family Pionidae in Australia (Acarina: Hygrobatoidea). *Records of the Western Australian Museum* **17**, 361–393.

Harvey MS (1998) 'Australian water mites: a guide to families and genera.' (CSIRO Publishing: Melbourne, Vic., Australia)

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal*

Society of London – B. Biological Sciences **270**, 313–321. doi:10.1098/rspb.2002.2218
Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 14812–14817. doi:10.1073/pnas.0406166101
Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science* **33**, 312–324. doi:10.1086/675225
Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. *Frontiers in Zoology* **10**, 34. doi:10.1186/1742-9994-10-34
Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12. doi:10.14806/ej.17.1.200
Miccoli FP, Lombardo P, Cicolani B (2013) Indicator value of lotic water mites (Acari: Hydrachnidia) and their use in macroinvertebrate-based indices for water quality assessment purposes. *Knowledge and Management of Aquatic Ecosystems* **411**, 08. doi:10.1051/kmae/2013075
Monaghan MT, Wild R, Elliot M, *et al.* (2009) Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Systematic Biology* **58**, 298–311. doi:10.1093/sysbio/syp027
Montes-Ortiz L, Elías-Gutiérrez M (2020) Water mite diversity (Acariformes: Prostigmata: Parasitengonina: Hydrachnidiae) from Karst ecosystems in southern of Mexico: a barcoding approach. *Diversity* **12**, 329. doi:10.3390/d12090329
Pons J, Barraclough TG, Gomez-Zurita J, Cardoso A, Duran DP, Hazell S, Kamoun S, Sumlin WD, Vogler AP (2006) Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* **55**, 595–609. doi:10.1080/10635150600852011
Porter TM, Hajibabaei M (2018) Scaling up: a guide to high-throughput genomic approaches for biodiversity analysis. *Molecular Ecology* **27**, 313–338. doi:10.1111/mec.14478
Proctor HC (2007) Aquatic mites in assessment of stream invertebrate diversity. In 'Acarology XI: proceedings of the international congress', 8–13 September 2002, Coyoacán, Mexico City. (Eds JB Morales-Malacara, V Behan-Pelletier, E Ueckermann, TM Peréz, EG Estrada-Venegas, M Badii) pp. 105–117. (Instituto de Biología and Facultad de Ciencias, Universidad Nacional Autónoma de México; Sociedad Latinoamericana de Acarología)
Proctor HC, Smith IM, Cook DR, Smith BP (2015) Chapter 25 – Subphylum Chelicerata, class Arachnida. In 'Thorp and Covich's freshwater invertebrates', 4th edn. (Eds JH Thorp, DC Rogers) pp. 599–660. (Academic Press: Boston, MA, USA)
Puillandre N, Modica MV, Zhang Y, Sirovich L, Boisselier M-C, Cruaud C, Holford M, Samadi S (2012) Large-scale species delimitation method for hyperdiverse groups. *Molecular Ecology* **21**, 2671–2691. doi:10.1111/j.1365-294X.2012.05559.x
Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: the barcode index number (BIN) system. *PLoS ONE* **8**, e66213. doi:10.1371/journal.pone.0066213
Schatz H, Behan-Pelletier V (2008) Global diversity of oribatids (Oribatida: Acari: Arachnida). In 'Freshwater animal diversity assessment'. (Eds EV Balian, C Lévêque, H Segers, K Martens) pp. 323–328. (Springer Netherlands: Dordrecht, Netherlands)
Shokralla S, Porter TM, Gibson JF, Dobosz R, Janzen DH, Hallwachs W, Golding GB, Hajibabaei M (2015) Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports* **5**, 9687. doi:10.1038/srep09687
Simpson JC, Norris RH (2000) Biological assessment of river quality: development of AUSRIVAS models and outputs. In 'Assessing the biological quality of fresh waters RIVPACS and other techniques'. (Eds JF Wright, DW Sutcliffe, MT Furse) pp. 125–142. (Freshwater Biological Association, The Ferry House: Far Sawrey, Ambleside, UK)
Smit H (2010) Australian water mites of the subfamily Notoaturinae Besch (Acari: Hydrachnidia: Aturidae), with the description of 24 new species. *International Journal of Acarology* **36**, 101–146. doi:10.1080/01647950903512807
Stålstedt J, Bergsten J, Ronquist F (2013) "Forms" of water mites (Acari: Hydrachnidia): intraspecific variation or valid species? *Ecology and Evolution* **3**, 3415–3435. doi:10.1002/ece3.704
Thompson JD, Plewniak F, Poch O (1999) A comprehensive comparison of multiple sequence alignment programs. *Nucleic Acids Research* **27**, 2682–2690. doi:10.1093/nar/27.13.2682
Vasquez AA, Qazazi MS, Fisher JR, Failla AJ, Rama S, Ram JL (2017) New molecular barcodes of water mites (Trombidiformes: Hydrachnidiae) from the Toledo Harbor region of Western Lake Erie, USA, with first barcodes for *Krendowskia* (Krendowskiidae) and *Koenikea* (Unionicolidae). *International Journal of Acarology* **43**, 494–498. doi:10.1080/01647954.2017.1349178
Viets KO (1978) New water mites (Hydrachnellae: Acari) from Australia. *Marine and Freshwater Research* **29**, 77–92. doi:10.1071/mf9780077
Vuataz L, Sartori M, Wagner A, Monaghan MT (2011) Toward a DNA taxonomy of alpine *Rhithrogena* (Ephemeroptera: Heptageniidae) using a mixed Yule-coalescent analysis of mitochondrial and nuclear DNA. *PLoS ONE* **6**, e19728. doi:10.1371/journal.pone.0019728
Walsh CJ (1997) A multivariate method for determining optimal subsample size in the analysis of macroinvertebrate samples. *Marine and Freshwater Research* **48**, 241–248. doi:10.1071/MF96096
Walsh CJ, Kunapo J (2009) The importance of upland flow paths in determining urban effects on stream ecosystems. *Journal of the North American Benthological Society* **28**, 977–990. doi:10.1899/08-161.1
Walsh CJ, Webb JA (2014) Spatial weighting of land use and temporal weighting of antecedent discharge improves prediction of stream condition. *Landscape Ecology* **29**, 1171–1185. doi:10.1007/s10980-014-0050-y
Walter DE, Proctor HC (2013) Acari underwater, or, why did mites take the plunge? In 'Mites: ecology, evolution and behaviour'. (Eds DE Walter, HC Proctor) pp. 229–280. (Springer Netherlands: Dordrecht, Netherlands)
Weigand H, Beermann AJ, Čiampor F, *et al.* (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. *Science of the Total Environment* **678**, 499–524. doi:10.1016/j.scitotenv.2019.04.247
Więcek M, Broda Ł, Proctor H, Dabert M, Smith BP, Dabert J (2021) Species boundaries among extremely diverse and sexually dimorphic, *Arrenurus*, water mites (Acariformes: Hydrachnidiae: Arrenuridae). *bioRxiv* 2021.04.04.438411. doi:10.1101/2021.04.04.438411
Young MR, Behan-Pelletier VM, Hebert PDN (2012) Revealing the hyperdiverse mite fauna of subarctic Canada through DNA barcoding. *PLoS ONE* **7**, e48755. doi:10.1371/journal.pone.0048755
Young MR, Proctor HC, deWaard JR, Hebert PDN (2019) DNA barcodes expose unexpected diversity in Canadian mites. *Molecular Ecology* **28**, 5347–5359. doi:10.1111/mec.15292
Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z (2012) Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution* **3**, 613–623. doi:10.1111/j.2041-210X.2012.00198.x
Zawal A, Stryjecki R, Stępień E, Buczyńska E, Buczyński P, Czachorowski S, Pakulnicka J, Śmietana P (2017) The influence of environmental factors on water mite assemblages (Acari, Hydrachnidia) in a small lowland river: an analysis at different levels of organization of the environment. *Limnology* **18**, 333–343. doi:10.1007/s10201-016-0510-y

Data availability. All DNA sequence data have been submitted to the National Center for Biotechnology Information and the BOLD systems database under the accession numbers found in Tables S2 and S3.

Conflicts of interest. The authors declare that they have no conflicts of interest.

Declaration of funding. This study was funded by the Melbourne Water Corporation, the Native Australian Animals Trust and the Australian Research Council through a Linkage grant (LP200100381).

Acknowledgements. The authors thank Chris Walsh and Yung En Chee for providing comments on this manuscript. We also thank Heather Proctor and an anonymous reviewer for providing comments on this manuscript.

Author affiliations

[A]Pest and Environmental Adaptation Research Group (PEARG), School of BioSciences, Bio21 Institute, 30 Flemington Road, The University of Melbourne, Vic. 3010, Australia.

[B]Waterway Ecosystem Research Group (WERG), School of Ecosystem and Forestry Sciences, 500 Yarra Boulevard, Richmond, Vic. 3121, Australia.

[C]Applied Research, Melbourne Water, 990 La Trobe Street, Docklands, Vic. 3008, Australia.