

Difficulties barcoding in the dark: the case of crustacean stygofauna from eastern Australia

Maria G. Asmyhr^{A,D} and Steven J. B. Cooper^{B,C}

^ADepartment of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia.

^BEvolutionary Biology Unit, South Australian Museum, North Terrace, Adelaide, SA 5000, Australia.

^CAustralian Centre for Evolutionary Biology and Biodiversity, The University of Adelaide, SA 5005, Australia.

^DCorresponding author. Email: maria.asmyhr@mq.edu.au

Abstract. The eastern Australian aquifers remain mostly unexplored; however, recent surveys suggest that there could be substantial levels of subterranean biodiversity hidden in these aquifers. Groundwater fauna (stygofauna) is often characterised by short-range endemism. Furthermore, high levels of cryptic species, and lack of formal taxonomic descriptions and taxonomic expertise for many of the groups demand innovative approaches for assessing subterranean biodiversity. Here we evaluate the potential of using DNA barcoding as a rapid biodiversity assessment tool for the subterranean groundwater fauna of New South Wales, Australia. We experienced low amplification success using universal and more taxon-specific primers for PCR amplification of the barcoding gene (COI) in a range of crustacean stygofauna. Sequence comparisons of the most commonly used COI universal primers in selected crustacean taxa revealed high levels of variability. Our results suggest that successful amplification of the COI region from crustacean stygofauna is not straightforward using the standard ‘universal’ primers. We propose that the development of a multiprimer (taxon specific) and multigene approach for DNA barcode analyses, using next-generation sequencing methodologies, will help to overcome many of the technical problems reported here and provide a basis for using DNA barcoding for rapid biodiversity assessments of subterranean aquatic ecosystems.

Received 26 April 2012, accepted 19 September 2012, published online 19 December 2012

Introduction

Until recently Australian aquifers were unexplored and considered to contain highly impoverished subterranean fauna attributed to the continent’s aridity and absence of glaciation (Humphreys 2008). However, over the last two decades, intensive research focusing on subterranean habitats, particularly in Western Australia, has revealed unique ecosystems comprising higher taxa new to the Southern Hemisphere as well as ancient lineages previously known only from the fossil record (see Humphreys 2008; Guzik *et al.* 2010 for reviews). In Western Australia (WA) surveys of subterranean fauna are now a compulsory feature of environmental impact assessment (EIA) reports, with the result that most research in this field has been focussed on areas of extensive mining such as the Pilbara and Yilgarn regions of WA (Eberhard *et al.* 2005; Allford *et al.* 2008; Leys and Watts 2008). Less effort has been focussed on sampling eastern Australian aquifers (but see Thurgate *et al.* 2001a, 2001b), but results from surveys of five alluvial aquifers in New South Wales (NSW) and Queensland indicate that there could also be substantial levels of subterranean biodiversity hidden in the aquifers of eastern Australia (Hancock and Boulton 2008; Korb and Hose 2011; Lategan *et al.* 2012; Cook *et al.* 2012).

Tiny crustaceans (Copepoda, Amphipoda, Syncarida, Ostracoda and Isopoda, referred to here as stygofauna:

Humphreys 2008) typically dominate the invertebrate fauna of subterranean groundwater ecosystems. Adaptations to the subterranean environment have resulted in fauna from various taxa developing similar morphological characters (Culver *et al.* 1995), thus it is often difficult to identify and classify species on the basis of morphology alone (Bradford *et al.* 2010). Moreover, the isolated nature of many subterranean habitats occupied by stygofauna, combined with the limited dispersal abilities of stygofauna, have resulted in distributions being often limited to the single aquifer that was originally colonised by ancestral taxa (e.g. Cooper *et al.* 2002; Leys *et al.* 2003). Consequently, many Australian stygofaunal groups consist of short-range endemics (i.e. ranges <10 000 km²: Harvey 2002) confined to a single aquifer (Cooper *et al.* 2007, 2008; Guzik *et al.* 2008).

Guzik *et al.* (2010) estimated that more than 80% of Australia’s subterranean fauna remains undiscovered. The high level of cryptic species (see for example Guzik *et al.* 2008), and lack of formal taxonomic descriptions and taxonomic expertise for many of the taxonomic groups, demands innovative approaches for assessing subterranean biodiversity (Proudlove and Wood 2003). There is also a need for development of rapid biodiversity assessment tools to improve the efficiency of EIAs, and ongoing environmental monitoring, of groundwater ecosystems. One such approach for assessing species diversity

that provides a basis for these tools is DNA barcoding (Hebert *et al.* 2003).

The use of a standardised mitochondrial gene (cytochrome oxidase 1, COI) was proposed by Hebert *et al.* (2003) to serve as a global identification tool for all animal species. This approach has been successfully applied for large-scale biodiversity studies of butterflies (Hebert *et al.* 2004), fish (Ward *et al.* 2005), birds (Johnsen *et al.* 2010) and crustaceans (Costa *et al.* 2007).

The ideal DNA barcode marker should fulfill several criteria (Hebert *et al.* 2003). First, priming sites needs to be sufficiently conserved to permit the development of universal primers for reliable PCR amplification across taxa. Second, the gene needs to be variable enough to be informative at the interspecific level, but sufficiently conserved at the intraspecific level. Finally, its amplification and sequencing should be as robust as possible, also under variable laboratory conditions and protocols.

It has been suggested that universal primers (LCO1490/HCO2198; Folmer *et al.* 1994) may enable amplification of COI for representatives of most animal phyla (Hebert *et al.* 2003). However, several authors have revealed that these standard primers fail or have proven to be unreliable for PCR-amplification of COI in a variety of taxa (e.g. Amphibia: Vences *et al.* 2005; scale insects: Kondo *et al.* 2008; syncarids: Camacho *et al.* 2011; Copepoda: Karanovic and Cooper 2011a, 2012). These studies suggest that either additional modified primers may need to be developed for COI in certain taxonomic groups or

alternative DNA barcode markers (e.g. 16S rRNA and 18S rDNA) may need to be utilised (see, for example, Corse *et al.* 2010; Hardy *et al.* 2010). It is, therefore, important that there is documentation of both the failed and successful PCR-amplification results when using the above standard universal COI primers.

Here we test the utility of using the standard primers of Folmer *et al.* (1994) for PCR-amplification of COI within the understudied east Australian crustacean stygofauna and evaluate the potential of using DNA barcoding as a rapid biodiversity assessment tool.

Materials and methods

Sampling method

Stygofauna was collected from several geologically different aquifers (sandstone, sandbeds and alluvium) in eastern Australia between 2008 and 2011: Kangaloon, Bylong, Hunter, Pages, the Central coast and Wellington in New South Wales, Australia (Fig. 1). Access to the aquifers relied on the availability of boreholes developed for groundwater monitoring (NSW water). Bores were sampled using a motorised inertia pump and/or a plankton net, depending on the depth of the bore (>30 m were netted). When pumping, 150–300 L of water was removed from the bore and subsequently passed through a 63- μ m sieve. When using the plankton net approach, the net was hauled

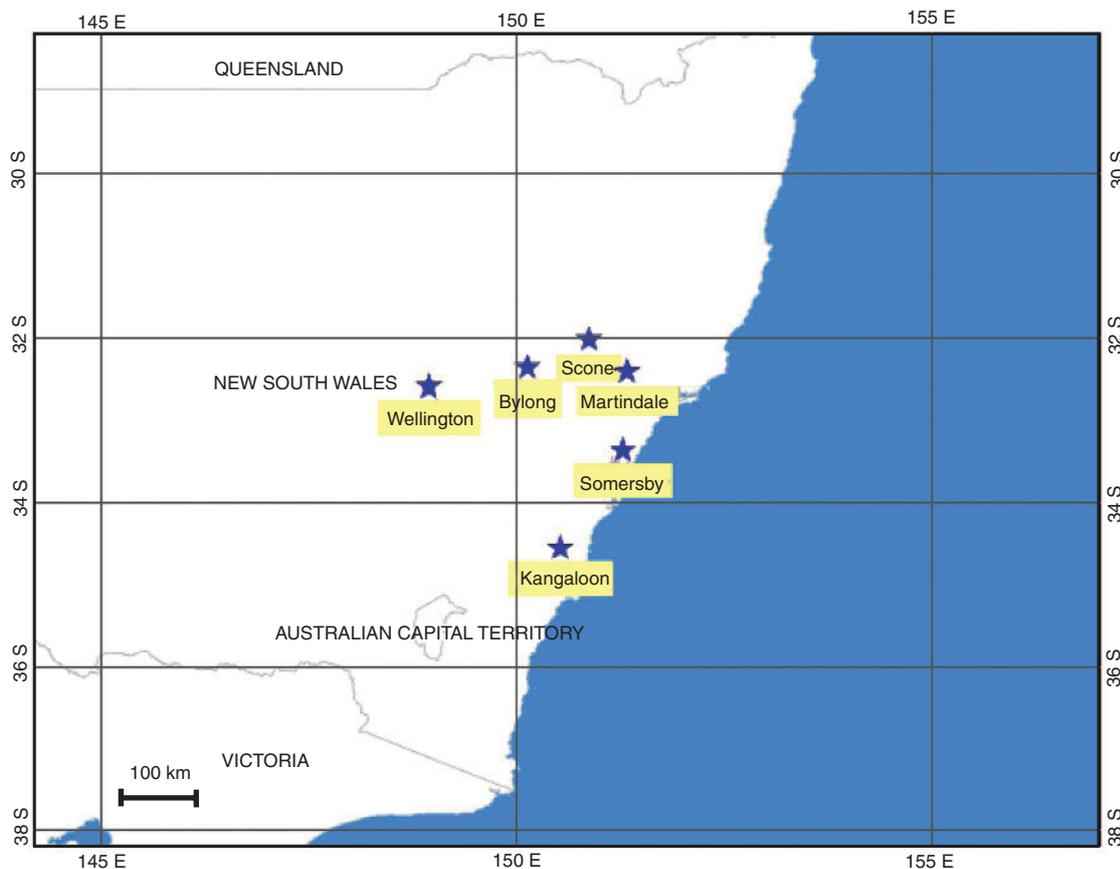


Fig. 1. The location of the study sites in eastern Australia.

through the water column in the boreholes multiple times. First, three hauls were made with a fine (63 µm) mesh net followed by three hauls with a coarser (100 µm) mesh net. For each haul, the net was lowered to the bottom of the bore, bounced up and down to stir up sediment, and then slowly retrieved so that it filtered both stirred up sediment and the water column. In the end, all hauls were combined and passed through the 63-µm sieve. The invertebrates retained in the sieve were preserved in 100% ethanol until DNA extraction. The stored samples were sorted under a light microscope and divided into morphotypes: Amphipoda, Copepoda (Cyclopoida and Harpacticoida), Syncarida (Bathynellidae and Parabathynellidae), Psammaspidae, Isopoda and Ostracoda.

DNA extraction and PCR amplification

DNA was extracted from legs and from whole animals (<2 mm in size) using Bioline Genomic DNA extraction kit (Bioline). A total of 125 samples were used for the amplification tests described below (8 isopods, 15 amphipods, 10 ostracods, 35 cyclopoids, 20 harpacticoids, 14 bathynellids, 3 psammaspides and 20 parabathynellids). We attempted to amplify a region of COI for all higher taxa using the 'universal' primer pair HCO2198 and LCO1490 (Folmer *et al.* 1994). Amplification success rate

was very low for all taxa (<20%) leading to concerns about: (1) the DNA quality, and (2) primer fit. To evaluate whether lack of PCR amplification was due to low DNA quality/quantity, we tested a range of additional DNA extraction methods: 10% Chelex, Genereleaser with Proteinase K (Schizas *et al.* 1997), High-salt (Sunnucks and Hales 1996), NucleoSpin (Macherey-Nagel) and the QIAamp DNA Mini Kit using carrier RNA (Qiagen).

DNA was subsequently amplified using a universal primer pair for a region of 18S rDNA. The primer pair 18SF and 18SR amplifies a fragment between 200 and 450 bp, and amplification success rate was 90% for all taxa. In order to evaluate the potential use of this marker as an alternative/complementary barcode marker for stygofauna, we generated 18S rRNA sequences for a subset of the samples from which we had obtained COI sequences.

In addition, we tested the 'universal' DNA mini-barcode primer pair Uni-MinibarF1 and UniMinibarR1 (Meusnier *et al.* 2008) that amplifies ~150 bp of COI. The success rate resulting from this trial was low, indicating that the problem was likely to be due to unsuitable primers for COI amplification. We therefore went on to test a wider range of 'universal' and more taxon-specific primers for COI, including a set of primers designed for

Table 1. COI primers tested in this study

Primer	PCR	Primer sequence	Reference
LCO1490	1	GGTCAACAAATCATAAAGATATTGG	Folmer <i>et al.</i> 1994
HCO2198	1	TGATTTTTGGTCACCCTGAAGTTTA	Folmer <i>et al.</i> 1994
C1-J-1718	2 ^B	GGAGGATTGGAAAATTGATTAGTTCC	Simon <i>et al.</i> 1994
C1-J-2329	2 ^B	ATCGTAAATATATGATGAGCTCA	Simon <i>et al.</i> 1994
C1-J-1709	2 ^B	AAATGGWGGWTTYGGAAAYTG	Simon <i>et al.</i> 1994
C1-J2195	2 ^B	TGATTCTTTGGWCACCCWGAAGT	Simon <i>et al.</i> 2006
N2353	2 ^B	GCTCGTGTATCAACGTCTATWCC	Simon <i>et al.</i> 2006
Uni-minibarF	2	GAAAATCATAATGAAGGCATGAGC	Meusnier <i>et al.</i> 2008
Uni-minibarR	2	TCCACTAATCACAAARGATATTGGTAC	Meusnier <i>et al.</i> 2008
tRWF1	3	AAACTAATARCCTTCAAAG	Park <i>et al.</i> 2010
tRWF2	3	AAACTAATAATYTTCAAATTA	Park <i>et al.</i> 2010
M1322 ^A		TCAAAATARRTYGTGRTAWARHAC	Karanovic and Cooper 2011a
M1321 ^A		TRRNGAYGAYCARRTTTATAATGT	Karanovic and Cooper 2011a
M1323 ^A		GAYGAYCARRTTTATAATGT	Karanovic and Cooper 2011a
M479	2	TTTATTTTAGGDGCMGTGATC	Cooper <i>et al.</i> 2007
M480	2	AATGADGTRTTTARRTTTCG	Cooper <i>et al.</i> 2007
CpF1	1 and 2	RYARTATRGTRATAGTCCWGCTA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
CpR1	1 and 2	AAYGWKTWGTVACRGCTCA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
CpF2	1 and 2	GTWAYTARMACTGMYCARGCAA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
CpR2	1 and 2	GTWRTWGTGTHACDGCTCA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
CpF3	1 and 2	GCTCCWGTCAAWACWGGTA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
CpF4	1 and 2	YATDGTTRATTAGCHCCWGCTA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
CpF5	1 and 2	ARMACWGMYCAGDCAAA	S. M. H. Smith, M. G. Asmyhr, G. Hose, A. Stow, and J. Ross, J., unpubl. data
M876	2	CARTTYCCAAAMCCNCCAATTA	Cooper <i>et al.</i> 2008
M877	2	AATTTTTATTACAACRTRATHAATATRCG	Cooper <i>et al.</i> 2008

^ANested PCR on LCO1490/HCO2198 template.

^BCycling conditions when different from previously published protocols.

a pilot study (S. M. H. Smith, M. G. A. Asmyhr, G. Hose, A. Stow and J. Ross, unpubl. data of harpacticoid copepods (Table 1).

PCR and DNA sequencing

The primers summarised in Table 1 were tested in various combinations. All primer combinations were tested for a subset of 28 samples (5 amphipods, 5 isopods, 5 cyclopoids, 5 harpacticoids, 5 syncarids and 3 ostracods). Taxon-specific primers, such as, for example, M479/480 for amphipods, were tested on subsets comprising 5–15 samples, depending on the taxon. Optimised PCRs were carried out in 25- μ L volumes containing 1–3 μ L of DNA template, 1 μ M of each primer, 2–4 mM MgCl₂, 2–4 μ g BSA, 3 μ L 5 \times Go Taq FlexiBuffer (Promega), 800 μ M of dNTPs, and 1U Taq Polymerase (Promega). As a starting point, primers were tested using the PCR protocols recommended by the primer reference paper (listed in Table 1). Gradient PCRs were run for each primer combination where no such information was available in order to find the optimal annealing temperature. For most primer combinations, the application of touchdown protocols resulted in the highest amplification success. The following protocols were therefore used in subsequent experiments: (1) Initial denaturation for 3 min at 94°C followed by five cycles of (94°C, 30 s; 45°C, 30 s; 1 min; 72°C, 1 min) and 35 cycles of (94°C, 30 s; 52°C, 30 s; 1 min; 72°C, 1 min) with a final extension of 10 min at 72°C. (2) Initial denaturation for 3 min at 94°C, followed by six cycles of denaturation (94°C, 30 s), primers annealing (55°C, 53°C, 51°C, 49°C, 47°C, 45°C, 45 s) and an extension (72°C, 1 min). Following this, 33 cycles of (94°C, 30 s; 55°C, 45 s; 72°C, 1 min) and a final extension (72°C, 10 min). (3) Initial denaturation for 3 min at 94°C followed by five cycles of (94°C, 40 s; 45°C, 40 s; 1 min; 72°C, 70 s) and 35 cycles of (94°C, 40 s; 51°C, 40 s; 72°C, 70 s) with a final extension of 5 min at 72°C. PCR product was purified with ExoSAP-IT according to the manufacturer's instructions (USB Corporation). Sequencing was performed on an ABI PRISM310, 3130 automated sequencer by the Macquarie University Sequencing Facility. All DNA sequences >200 bp are available on GenBank as accessions JX948792–JX948818

16S rRNA

Some authors have recommended the use of the 16S rRNA gene for barcoding of some animal groups (Vences *et al.* 2005) because of more reliable and universal amplifications than for COI. We therefore tested the 'universal' primer pair 16Sar and 16Sbr (Palumbi *et al.* 1991) for amplification of a fragment of the mitochondrial 16S gene in a subset (32) of samples mentioned above. Master mix was prepared as described above and cycling conditions consisted of: 95°C for 3 min followed by 35 cycles (95°C, 30 s; 51°C, 45 s; 72°C, 45 s) and a final extension at 72°C for 10 min.

Priming site analysis

In order to further investigate the cause of amplification failure in COI, we did a survey of priming site variability for the most commonly used COI universal primer pair LCO1490 and HCO2198 (Folmer *et al.* 1994). This analysis was carried out

using five complete crustacean mitochondrial sequences from GenBank. We selected the mitochondrial genomes of surface-living relatives of the most common stygofaunal taxa (Cyclopoida NC012455.1, Harpacticoida DQ913891.2, Isopoda DQ442914.1 and Ostracoda AB114300.1). In addition, the mitochondrial genome of a subterranean amphipod (AM944817.1) was included. Orders Anaspidacea and Bathynellacea (Syncarida) were not represented with a complete mitochondrial sequence in GenBank. For LCO1490 we also included the sequence of an unknown stygobite amphipod from Wellington, NSW. Sequences were aligned using MEGA 5 (Tamura *et al.* 2011) and the number of mutations for each priming site quantified. Given the low amplification success for 16Sar and 16Sbr, priming site variability was also analysed for these primers.

Sequence analysis

Partial mitochondrial COI and nuclear 18S rRNA sequences were imported into MEGA 5 (Tamura *et al.* 2011), manually edited by eye and aligned using the default in CLUSTAL_X ver. 2.0 (Larkin *et al.* 2007). The Kimura two-parameter model (K2P)

Table 2. The different primer combinations tested in this study

Taxon	Primer pair	Amplification success (%)	Confirmed by sequencing (%)
Amphipoda	LCO1490/HCO218	12	100
	UniMinibarF1/R1	0	
	M479/M480	0	
	M479/HCO2198	0	
	LCO1490/M480	10	0
	tRWF1 and tRWF2/HCO2198	70	50
Syncarida	LCO1490/HCO218	25	10
	UniMinibarF1/R1	10	100
	C1-J1718/HCO2198	50	
	C1-J-1718/N2329	40	30
	C1-J-1718/N2353	90	50
	C1-1709/HCO2198	0	
Copepoda (harpacticoids and cyclopoids)	C1-J1709/N2352	50	30
	LCO1490/HCO2198	26	70
	UniMinibarF1/R1	10	100
	CpF1/R1	10	100
	F1/R2	0	
	F2/R1	0	
	F2/R2	0	
	F4/R1	0	
	F4/R2	0	
	F5/R1	0	
F5/R2	0		
Nested PCR (LCO1490 +HCO2198)	M1322/1321	50	25
	M1323/HCO2198	0	
Isopoda	LCO1490/HCO2198	10	10
	UniMinibarF1/R1	6	
	M876/M877	20	
Ostracoda	LCO1490/HCO2198	0	
	UniMinibarF1/R1	0	

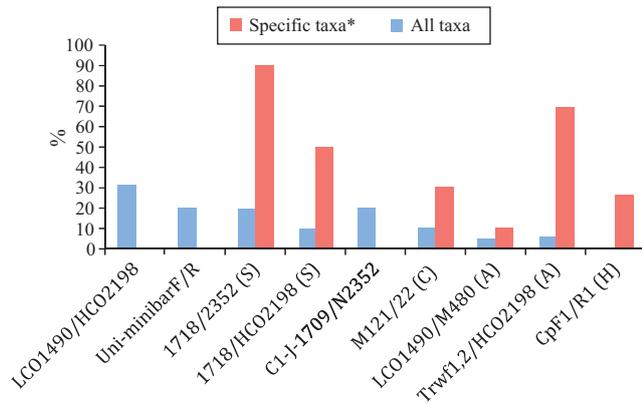


Fig. 2. Summary of the amplification success of different primer combinations. All taxa = amplification success when samples from all taxa are combined and * Specific taxa shows the amplification success within S = Syncarida, C = Cyclopoida, A = Amphipoda, H = Harpacticoida. See Table 1 for a list of primer codes.

was used to calculate nucleotide divergences (Kimura 1980). The comparisons of sequence divergence between the COI and 18S rRNA markers were based on amphipod sequences ($n = 10$, 452 and 223 bp respectively) and cyclopoid copepods ($n = 5$, 468 bp and 103 bp).

Results

COI amplification

Table 2 summarises the amplification and sequencing success of the different primer combinations for each taxon. For the ostracods, no amplification product could be obtained. Although some primer combinations resulted in amplification success rate as high as 70–90%, the resulting sequencing success rate was much lower. For some of the universal primers such as C1-J-1718/ N2353 we observed that human DNA was amplified and sequenced. There was also a problem with very low signal when sequencing (i.e. very weak PCR amplification), resulting in sequences that were impossible to interpret with certainty. For the nested PCR using primer pair M1322/1321 we observed double sequences for half of the samples, making them impossible to interpret. Of the primer pairs developed for the harpacticoid copepods (CpF1, F2, F3, F4, F5, R1, R2) in this study only F1/R1 resulted in amplification products. These primers were then used to generate sequences for six samples that failed to amplify using the Folmer primers. However, when these primers were tested on a subset of copepod samples collected from a different aquifer, PCR-amplification failed. Some of the primer pairs were tested on samples representing all taxa and the results of this trial are summarised in Fig. 2. We failed to find a set of primers that would consistently amplify DNA from samples from all taxa.

16S rRNA amplification

PCR amplifications using the universal primers 16Sar and 16Sbr resulted in 8 (28%) samples amplifying 1 isopod, 2 syncarids and 5 harpacticoids. Sequencing of the successful amplicons revealed that only four of the sequences (isopod, 2 syncarids and 1

Table 3. Variability of COI priming sites for selected crustacean taxa

Taxa	Sequence
Primer LCO1490	GGTCAACAAATCATAAAGATATTGG
Ostracoda	ACTCCACAAATCACAAAGATATTGG
Harpacticoida	GGGCTACGAACCACAAAGATATTGG
Isopoda	ATTCTACCAACCATAAAGATATTGG
Cyclopoida	ATTCTATGTAACCACAAAGATATTGG
Subterranean Amphipoda	TTTCAACTAACCATAAAGACATTGG
Amphipoda (Wellington, current study)	GATCAACAAATAAAGGCATTCCGT
HCO2198	TGATTTTTTGGTACCCTGAAGTTTA
Ostracoda	TGATTTTTTGGTACCCTGAAGTTTA
Harpacticoida	TGATTTTTTGGACATCCTGAGGTGTA
Isopoda	TGGTTTTTGGTCCATCCAGAGGTGTA
Cyclopoida	TGATTTTTTGGTCCATCCGAGGTATA
Subterranean Amphipoda	TGATTTTTTGGCCATCCAGAAGTCTA

harpacticoid) originated from crustacean species, whereas the remaining sequences were identical to human COI (4 harpacticoid copepods).

Priming site analysis

We found extremely high variability in the priming site of primer LCO1490 (Table 3a) with 4–12 mutations for all taxa. With the exception of the ostracod sequence, which showed no site variants, the priming site for HCO2198 (Table 3b) contained 3–4 mutations for all taxa with the majority of the mutations in the third codon position. The subsequent analysis of priming site variability in 16S revealed an even greater level of variability in the 16S priming sites, with 16Sar containing 5–12 substitutions and 16Sbr containing 1–8 substitutions (Table 4a, b).

Comparison of the COI and 18S rRNA marker

Where to draw the boundary between two species based on COI sequences varies between taxa (e.g. >7.1%: Abrams *et al.* 2012; >11%: Guzik *et al.* 2010; >17%: Costa *et al.* 2007) and without formal taxonomic description of the species in our study we had to use this as an approximation only. For the amphipod sequences analysed in this study we found that the mean sequence divergence between COI sequences was 23.5%. In contrast, the 18S sequences differed by a mean of 1.6%. For the

Table 4. Variability of 16S priming sites for selected crustacean taxa

Primer	Sequence
16Sar	CGCCTGTTTATCAAAAAACAT
Ostracoda	TGCTTATCCAACAAAAACCT
Harpacticoida	GGTCTGAACCTCAAATCATGT
Isopoda	CTGTAGTTTATIAAAAAACGT
Cyclopoida	TGACGGATTATCAAAATACAC
Subterranean Amphipoda	CACTTTTTTATTTAAAAGAT
16Sbr	CCGGTCTGAACCTCAGATCACGT
Ostracoda	CCGGTCTGAACCTCAGATCATGT
Harpacticoida	TGGGTATAAACTCAAATTTTGT
Isopoda	CCGGTCTGAACCTCAAATCATGT
Cyclopoida	TCGATCTTAACTCAAATCATGT
Subterranean Amphipoda	CCGGTCTGAACCTCAAATCATGT

cyclopoid copepod sequences there was a mean sequence divergence of 29.8% and 5.5% for COI and 18S respectively. We found that for pairs of sequences with a COI divergence of >7%, the 18S sequences differed by >0.5%, thus indicating that this marker could potentially be used as an alternative barcode marker for the taxa collected in this study.

Discussion

The high levels of endemism of stygofauna and the ever-increasing use of Australian aquifers for various anthropogenic activities (e.g. mining, agriculture) threaten the survival of many groundwater invertebrate species (Hancock and Boulton 2008). Our research aimed to build up a sequence library for stygofaunal species, which, following morphological assessment of the taxa, could be used as a basis for rapid biodiversity assessments (RBAs) of groundwater ecosystems and EIAs in the future. With the exception of Abrams *et al.* (2012), the current study is the first attempt to amplify DNA from stygofauna from NSW.

Universal primers and rapid assessment of subterranean biodiversity

The amplification success of the most commonly used COI primers (Folmer *et al.* 1994) was low in all taxa.

An explanation for the low amplification success rate is likely to be the high level of mismatch between the primer sequences and target DNA. We found high variability in the priming sites of both primer LCO1490 (Table 2a) and HCO2198, with 4–7 mutations (Table 3b). This finding is similar to that observed by Vences *et al.* (2005), who reported extremely low amplification success using LCO1490/HCO2198 in class Amphibia.

Our second approach to PCR-amplify COI was to utilise or develop a pair of primers for each higher taxon. We trialed primers previously tested in stygofauna of WA (Cooper *et al.* 2007; Guzik *et al.* 2008), and although this did result in an increased amplification success rate for syncarida (90% in the trial PCR), amplification success in the remaining taxa did not increase above 50% and was mostly well below that.

Using COI sequences from a few populations of harpacticoid copepods generated in this project and data from GenBank we designed several primers (both specific and degenerate) (Hose 2008; Smith *et al.* in prep). These primers performed well for individuals collected within these specific populations, but did not amplify DNA in samples from other populations (primer pair CpF1/CpR1) (Fig. 1). Similarly, Karanovic and Cooper (2011a, 2011b, 2012) used a nested PCR-amplification approach (primers M1321–3) (Table 1) designed specifically for copepods; however, the PCR success rate remained at ~50% and several species (e.g. *Kinnecaris esbe*: Karanovic and Cooper 2011a) failed to amplify despite repeated attempts and modifications to protocols. In addition, contaminating human DNA was PCR amplified and sequenced for five specimens (Cooper, unpublished results), suggesting that the PCR primers did not work optimally for some of the copepod species.

The amplification success using 16S universal primers was no higher than what we experienced when amplifying COI. The sequencing success was even lower than for COI (50% of the amplicons being human DNA), suggesting that it may be very

difficult to develop a ‘universal’ approach for amplification of any of these genes in the crustacean stygofauna.

In contrast, universal primers for the nuclear 18S rRNA gene amplified in most samples from all taxa. This difference in amplification success between nuclear 18S and mitochondrial COI and 16S markers could be caused by the relative ratio of mitochondrial to nuclear rRNA gene copies present in different samples, with the latter potentially being more abundant, and hence more likely to PCR-amplify (Rogers and Bendich 1987; Veltri *et al.* 1990). Moreover, the 18S rRNA region is a slowly evolving marker, thus the primer regions will vary less between taxa, facilitating the development and use of universal primers across highly divergent taxa. Although the analysis of slowly evolving markers such as 18S rRNA is more frequently used to detect higher taxa (e.g. at the Family and Genus level), several recent studies have reported its success as an alternative barcode (e.g. Hardy *et al.* 2010; Corse *et al.* 2010). For the amphipod and cyclopoid copepod samples analysed in this study, we found that although 18S rRNA sequence divergence for different species (species boundary based on COI sequence divergence level) was considerably lower than what was observed for COI, there was sufficient variation to delineate between individuals originating from geographically close populations. Therefore, for the purpose of this project (sequence database for RBAs), our results indicate that the 18S rRNA marker could be used as an alternative barcode marker.

A case for development of a multiprimer and multigene barcoding approach

There is considerable scope in the future to use next-generation sequencing methods (e.g. Illumina, FLX 454, Ion Torrent 2nd generation sequencing and, recently, PacBio and Nanopore 3rd generation sequencing) as a basis for rapidly surveying biodiversity using amplicon sequencing or ‘environmental genomic’ approaches (e.g. Bybee *et al.* 2011). The advantage of these methods is that thousands of samples and multiple loci can be pooled and simultaneously sequenced, and bioinformatic approaches can be used to sort the sequences and BLAST search a library of ‘species-specific DNA barcode’ sequences to identify the taxa present. To make this approach feasible as a robust RBA tool it is important that PCR products can be reliably amplified from each target species. Our study and others (Vences *et al.* 2005; Kondo *et al.* 2008; Camacho *et al.* 2011; Karanovic and Cooper 2011a, 2011b, 2012) suggest that the use of COI only is often technically problematic and that there is a need to commence building databases of additional loci that can be reliably PCR-amplified and sequenced. In addition to the technical problems we report here, amplification of nuclear copies of mitochondrial DNA (NUMTs) can also be a major problem, as recently noted for crustaceans (Buhay 2009), which can lead to overestimation of levels of species diversity. One way to help circumvent the NUMT problem is to sequence additional mtDNA markers and determine whether there is concordance in gene trees produced from the independently sequenced markers; concordance is expected if both markers are from a single mtDNA locus.

Given that the next-generation sequencing methods are not greatly limited by the number of pooled loci (or different primer

combinations used for the same locus), it makes sense that a multiprimer and multigene approach is used to PCR-amplify 'DNA barcodes' for species identification. Alternatively, capture array methods (see Lemmon *et al.* 2012) may also be used to enrich for target DNA from specific loci, before sequencing. The choice of additional DNA barcoding loci requires further debate, but we suggest that, in the short term, in addition to the use of COI, 16S rRNA, 18S and the internal transcribed spacer (*ITS*) of the nuclear rRNA gene family be further developed, the latter recently being proposed as the DNA barcode for fungi (Schoch *et al.* 2012).

We further suggest that cocktails of taxon-specific primers be developed for genes such as COI and 16S rRNA by modification of the primer region originally developed by Folmer *et al.* (1994) and Palumbi *et al.* (1991) respectively, as these regions have proven utility across a broad range of taxa, despite some of the shortcomings we report here. These developments would be facilitated by obtaining complete mitochondrial genome sequence data from target taxa, which is now relatively straightforward and cost effective by next-generation sequencing analyses of genomic DNA (e.g. Gardner *et al.* 2011).

Conclusion

One of the major applications of DNA barcoding in the future is to enable the use of RBA tools that can be used for EIAs and long-term monitoring of the environment.

This application is particularly relevant to subterranean aquatic habitats and stygofauna of Australia, which are often found in regions rich in mineral resources or of high agricultural value. It is, however, important to note that although molecular methods can be used as a rapid initial screening tool, further identification based on morphology is necessary in order to fully characterise and understand these ecosystems.

We propose that the recent advances in sequencing technology now make it feasible to use a multigene approach for DNA barcode analyses, and that this approach will help overcome many of the technical problems reported in our study of Australian east coast stygofauna.

Acknowledgements

This work was supported by the Australian Research Council (ARC: DP1095200), the Department of Biological Sciences, Macquarie University, Gosford Council and the Linnean Society of New South Wales. Special thanks go to Grant Hose, Adam Stow and Shannon Smith (MQ) for providing valuable assistance and technical support with sampling and analysis. We are grateful to Phil Cranidge (Gosford Council), John Paul Williams, Ben Mulhearn, Rod Gleeson, Ross Beasley, Andrew Pearce (NSW water), Paul Watson (Wyang Council), Peter Graham and Andy Baker (UNSW), Alexander Michie and Michael Strack (MQ) for facilitating and helping out with the collection of invertebrate samples. SJBC's research is supported by an ARC Linkage grant (LP 100200494).

References

- Abrams, K. M., Guzik, M. T., Cooper, S. J. B., Humphreys, W. F., King, R. A., Cho, J.-L., and Austin, A. D. (2012). What lies beneath: molecular phylogenetics and ancestral state reconstruction of the ancient subterranean Australian Parabathynellidae (Syncarida, Crustacea). *Molecular Phylogenetics and Evolution* **64**, 130–144. doi:10.1016/j.ympev.2012.03.010
- Allford, A., Cooper, S. J. B., Humphreys, W. F., and Austin, A. D. (2008). Diversity and distribution of groundwater fauna in a calcrete aquifer: does sampling method influence the story? *Invertebrate Systematics* **22**, 127–138. doi:10.1071/IS07058
- Bradford, T., Adams, M., Humphreys, W. F., Austin, A. D., and Cooper, S. J. B. (2010). DNA barcoding of stygofauna uncovers cryptic amphipod diversity in a calcrete aquifer in Western Australia's arid zone. *Molecular Ecology Resources* **10**, 41–50. doi:10.1111/j.1755-0998.2009.02706.x
- Bybee, S. M., Bracken-Grissom, H., Haynes, B. D., Hermansen, R. A., Byers, R. L., Clement, M. J., Udall, J. A., Wilcox, E. R., and Crandall, K. A. (2011). Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. *Genome Biology and Evolution* **3**, 1312–1323. doi:10.1093/gbe/evr106
- Buhay, J. E. (2009). "COI-like" sequences are becoming problematic in molecular systematic and DNA barcoding studies. *Journal of Crustacean Biology* **29**, 96–110.
- Camacho, A. I., Dorda, B. A., and Rey, I. (2011). Identifying cryptic speciation across groundwater populations: first *co1* sequences of bathynellidae (crustacea, syncarida). *Graellsia* **67**, 7–12. doi:10.3989/graeellsia.2011.v67.031
- Cook, B. D., Abrams, K. M., Marshall, J., Perna, C. N., Choy, S., Guzik, M. T., and Cooper, S. J. B. (2012). Species diversity and genetic differentiation of stygofauna (Syncarida: Bathynellacea) across an alluvial aquifer in north-eastern Australia. *Australian Journal of Zoology* **60**, 152–158. , –
- Cooper, S. J. B., Hinze, S., Leys, R., Watts, C. H. S., and Humphreys, W. F. (2002). Islands under the desert: molecular systematics and evolutionary origins of stygobitic water beetles (Coleoptera: Dytiscidae) from central Western Australia. *Invertebrate Systematics* **16**, 589–590. doi:10.1071/IT01039
- Cooper, S. J. B., Bradbury, J. H., Saint, K. M., Leys, R., Austin, A. D., and Humphreys, W. F. (2007). Subterranean archipelago in the Australian arid zone: mitochondrial DNA phylogeography of amphipods from central Western Australia. *Molecular Ecology* **16**, 1533–1544. doi:10.1111/j.1365-294X.2007.03261.x
- Cooper, S. J. B., Saint, K. M., Taiti, S., Austin, A. D., and Humphreys, W. F. (2008). Subterranean archipelago: mitochondrial DNA phylogeography of stygobitic isopods (Oniscidea: *Haloniscus*) from the Yilgarn region of Western Australia. *Invertebrate Systematics* **22**, 195–203. doi:10.1071/IS07039
- Corse, E., Costedoat, C., Chappaz, M. I., Pech, N., Martin, J.-F., and Gilles, A. (2010). A PCR-based method for diet analysis in freshwater organisms using 18S rDNA barcoding on faeces. *Molecular Ecology* **10**, 96–108. doi:10.1111/j.1755-0998.2009.02795.x
- Costa, F. O., deWaard, J. R., Boutillier, J., Ratnasingham, S., Dooh, R. T., Hajibabaei, M., and Hebert, P. D. (2007). Biological identifications through DNA barcodes: the case of the Crustacea. *Canadian Journal of Fisheries and Aquatic Sciences* **64**, 272–295. doi:10.1139/f07-008
- Culver, D. C., Kane, T., and Fong, D. W. (1995). 'Adaptation and Natural Selection in Caves: the Evolution of *Gammarus minus*.' (Harvard University Press: Cambridge, MA.)
- Eberhard, S. M., Halse, S. A., and Humphreys, W. F. (2005). Stygofauna in the Pilbara region, north-west Western Australia: a review. *Journal of the Royal Society of Western Australia* **88**, 167–176.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., and 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.
- Gardner, M. G., Fitch, A. J., Bertozzi, T., and Lowe, A. J. (2011). Rise of the machines – recommendations for ecologists when using next generation sequencing for microsatellite development. *Molecular Ecology Resources* **11**, 1093–1101. doi:10.1111/j.1755-0998.2011.03037.x

- Guzik, M. T., Abrams, K. M., Cooper, S. J. B., Humphreys, W. F., Cho, J.-L., and Austin, A. D. (2008). Phylogeography of the ancient Parabathynellidae (Crustacea: Bathynellacea) from the Yilgarn region of Western Australia. *Invertebrate Systematics* **22**, 205–216. doi:10.1071/IS07040
- Guzik, M. T., Austin, A. D., Cooper, S. J. B., Harvey, M. S., Humphreys, W. F., Bradford, T., Eberhard, S. M., King, R. A., Leys, R., Muirhead, K. A., and Tomlinson, M. (2010). Is the Australian subterranean fauna uniquely diverse? *Invertebrate Systematics* **24**, 407–418. doi:10.1071/IS10038
- Hancock, P. J., and Boulton, A. J. (2008). Stygofauna biodiversity and endemism in four alluvial aquifers in eastern Australia. *Invertebrate Systematics* **22**, 117–126. doi:10.1071/IS07023
- Hardy, C. M., Krull, E. S., Hartley, D. M., and Oliver, R. L. (2010). Carbon source accounting for fish using combined DNA and stable isotope analyses in a regulated lowland river weir pool. *Molecular Ecology* **19**, 197–212. doi:10.1111/j.1365-294X.2009.04411.x
- Harvey, M. S. (2002). Short-range endemism amongst the Australian fauna: some examples from non-marine environments. *Invertebrate Systematics* **16**, 555–570. doi:10.1071/IS02009
- Hebert, P. D. N., Cywinska, A., Ball, S. L., and deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings. Biological Sciences* **270**(1512), 313–321. doi:10.1098/rspb.2002.2218
- Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H., and 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
- Hose, G. C. (2008). Stygofauna baseline assessment for Kangaloon Borefield Investigations – 354 Southern Highlands, NSW. Report to Sydney Catchment Authority, Access Macquarie Ltd, Sydney.
- Humphreys, W. F. (2008). Rising from Down Under: developments in subterranean biodiversity in Australia from a groundwater fauna perspective. *Invertebrate Systematics* **22**, 85–101. doi:10.1071/IS07016
- Johnsen, A., Rindal, E., Ericson, P. G. P., Zuccon, D., Kerr, K. C. R., Stoeckle, M. Y., and Lifjeld, J. T. (2010). DNA barcoding of Scandinavian birds reveals divergent lineages in trans-Atlantic species. *Journal für Ornithologie* **151**, 565–578. doi:10.1007/s10336-009-0490-3
- Karanovic, T., and Cooper, S. J. B. (2011a). Molecular and morphological evidence for short range endemism in the *Kinnecaris solitaria* complex (Copepoda, Parastenocarididae), with descriptions of seven new species. *Zootaxa* **3026**, 1–64.
- Karanovic, T., and Cooper, S. J. B. (2011b). Third genus of parastenocarid copepods from Australia supported by molecular evidence (Harpacticoida: Parastenocarididae). In ‘Studies on Freshwater Copepoda: a Volume in Honour of Bernard Dussart’. Crustaceana Monographs 16. (Eds D. Defaye, J. C. von Vaupel Klein, and G. Wyngaard.) pp 283–326. (Brill: Leiden.)
- Karanovic, T., and Cooper, S. J. B. (2012). Explosive radiation of the genus *Schizopera* Sars (Copepoda: Harpacticoida) in a small subterranean island in Western Australia: unravelling the cases of cryptic speciation, size differentiation, and multiple invasions. *Invertebrate Systematics* **26**, 115–192. doi:10.1071/IS11027
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution* **16**, 111–120. doi:10.1007/BF01731581
- Kondo, T., Gullan, P. J., and Williams, D. J. (2008). Coccidology: the study of scale insects (Hemiptera: Sternorrhynca: Coccoidea). *Revista Corpoica – Ciencia y Tecnológica Agropecuaria* **9**, 55–61.
- Korbel, K. L., and Hose, G. C. (2011). A tiered framework for assessing groundwater ecosystem health. *Hydrobiologia* **661**, 329–349. doi:10.1007/s10750-010-0541-z
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **23**, 2947–2948. doi:10.1093/bioinformatics/btm404
- Lategan, M. J., Torpy, F. R., Newby, S., Stephenson, S., and Hose, G. C. (2012). Fungal diversity of shallow aquifers in southeastern Australia. *Geomicrobiology Journal* **29**, 352–361. doi:10.1080/01490451.2011.559306
- Lemmon, A. R., Emme, S., and Lemmon, E. M. (2012). Anchored hybrid enrichment for massively high throughput phylogenomics. *Systematic Biology* **61**, 727–744. doi:10.1093/sysbio/sys049
- Leys, R., and Watts, C. H. (2008). Systematics and evolution of the Australian subterranean hydroporine diving beetles (Dytiscidae), with notes on *Carabhydrus*. *Invertebrate Systematics* **22**, 217–233. doi:10.1071/IS07034
- Leys, R., Watts, C. H. S., Cooper, S. J. B., and Humphreys, W. F. (2003). Evolution of subterranean diving beetles (Coleoptera: Dytiscidae, Hydroporini, Bidessini) in the arid zone of Australia. *Evolution* **57**, 2819–2834. doi:10.1111/j.0014-3820.2003.tb01523.x
- Meusnier, I., Singer, G. A. C., Landry, J.-F., Hickey, D. A., Hebert, P. D. N., and Hajibabaei, M. (2008). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics* **9**, 214. doi:10.1186/1471-2164-9-214
- Palumbi, S., Martin, A., Romano, S., McMillan, W. O., Stice, L., and Grabowski, G. (1991). The simple fool’s guide to PCR. Version 2.0. Department of Zoology and Kewalo Marine Laboratory, University of Hawaii, Honolulu, HI.
- Park, D.-S., Suh, S.-J., Oh, H.-W., and Hebert, P. D. N. (2010). Recovery of the mitochondrial COI barcode region in diverse Hexapoda through tRNA-based primers. *BMC Genomics* **11**, 423. doi:10.1186/1471-2164-11-423
- Proudlove, G., and Wood, P. J. (2003). The blind leading the blind: cryptic subterranean species and DNA taxonomy. *Trends in Ecology & Evolution* **18**, 272–273. doi:10.1016/S0169-5347(03)00095-8
- Rogers, S. O., and Bendich, A. J. (1987). Ribosomal RNA genes in plants: variability in copy number and in the intergenic spacer. *Plant Molecular Biology* **9**, 509–520. doi:10.1007/BF00015882
- Schizas, N. V., Street, G. T., Coull, B. C., Chandler, G. T., and Quattro, J. M. (1997). An efficient DNA extraction method for small metazoans. *Molecular Marine Biology and Biotechnology* **6**, 381–383.
- Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spougea, J. L., Levesque, C. A., and Chen, W. Fungal Barcoding Consortium (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 6241–6246. doi:10.1073/pnas.1117018109
- Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., and Flook, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of the Entomological Society of America* **87**, 651–701.
- Simon, C., Buckley, T. R., Frati, F., Stewart, J. B., and Beckenbach, A. T. (2006). Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annual Review of Ecology Evolution and Systematics* **37**, 545–579. doi:10.1146/annurev.ecolsys.37.091305.110018
- Sunnucks, P., and Hales, D. (1996). Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and Evolution* **13**, 510–524. doi:10.1093/oxfordjournals.molbev.a025612

- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* **28**, 2731–2739. doi:[10.1093/molbev/msr121](https://doi.org/10.1093/molbev/msr121)
- Thurgate, M. E., Gough, J. S., Spate, A., and Eberhard, S. (2001a). Subterranean biodiversity in New South Wales: from rags to riches. *Records of the Western Australian Museum* **64**(Supplement), 37–47.
- Thurgate, M. E., Gough, J. S., Clarke, A. K., Serov, P., and Spate, A. (2001b). Stygofauna diversity and distribution in eastern Australian cave and karst areas. *Records of the Western Australian Museum* **64**(Supplement), 49–62.
- Veltri, K. L., Espiritu, M., and Singh, G. (1990). Distinct genomic copy number in mitochondria of different mammalian organs. *Journal of Cellular Physiology* **143**, 160–164. doi:[10.1002/jcp.1041430122](https://doi.org/10.1002/jcp.1041430122)
- Vences, M., Thomas, M., van der Meijden, A., Chiari, Y., and Vieites, D. R. (2005). Comparative performance of the 16S rRNA gene in DNA barcoding of amphibians. *Frontiers in Zoology* **2**, 5. doi:[10.1186/1742-9994-2-5](https://doi.org/10.1186/1742-9994-2-5)
- Ward, R. D., Zemlak, T. S., Innes, B. H., Last, P. R., and Hebert, P. D. N. (2005). DNA barcoding Australia's fish species. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **360**(1462), 1847–1857. doi:[10.1098/rstb.2005.1716](https://doi.org/10.1098/rstb.2005.1716)