CSIRO PUBLISHING

Australian Journal of Zoology, 2017, **65**, 283–291
https://doi.org/10.1071/ZO17076

The first complete mitochondrial genomes of subterranean dytiscid diving beetles (*Limbodessus* and *Paroster*) from calcrete aquifers of Western Australia

Josephine Hyde[A,F], *Steven J. B. Cooper*[A,B], *Pablo Munguia*[C], *William F. Humphreys*[D,E] *and Andrew D. Austin*[A]

[A]Australian Centre for Evolutionary Biology and Biodiversity, School of Biological Sciences, University of Adelaide, Adelaide, SA 5005, Australia.
[B]Evolutionary Biology Unit, South Australian Museum, North Terrace, Adelaide, SA 5000, Australia.
[C]RMIT Studios, RMIT University, Melbourne, Vic. 3000, Australia.
[D]Western Australian Museum, Welshpool DC, WA 6986, Australia.
[E]School of Animal Biology, University of Western Australia, Crawley, WA 6009, Australia.
[F]Corresponding author. Email: josephine.hyde@adelaide.edu.au

Abstract. Comparative analyses of mitochondrial (mt) genomes may provide insights into the genetic changes, associated with metabolism, that occur when surface species adapt to living in underground habitats. Such analyses require comparisons among multiple independently evolved subterranean species, with the dytiscid beetle fauna from the calcrete archipelago of central Western Australia providing an outstanding model system to do this. Here, we present the first whole mt genomes from four subterranean dytiscid beetle species of the genera *Limbodessus* (*L. palmulaoides*) and *Paroster* (*P. macrosturtensis*, *P. mesosturtensis* and *P. microsturtensis*) and compare genome sequences with those from surface dytiscid species. The mt genomes were sequenced using a next-generation sequencing approach employing the Illumina Miseq system and assembled *de novo*. All four mt genomes are circular, ranging in size from 16 504 to 16 868 bp, and encode 37 genes and a control region. The overall structure (gene number, orientation and order) of the mt genomes is the same as that found in eight sequenced surface species, but with genome size variation resulting from length variation of intergenic regions and the control region . Our results provide a basis for future investigations of adaptive evolutionary changes that may occur in mt genes when species move underground.

Received 12 November 2017, accepted 7 March 2018, published online 16 April 2018

Introduction

Stygobionts, invertebrates that complete their life cycle entirely in subterranean groundwater habitats, are highly adapted to their environment (Galassi 2001). These habitats are characterised by darkness, a reduction in primary production, a lack of oxygen and, often, fragmented microhabitats (Hüppop 2000; Culver *et al*. 2005). Stygobionts are characteristically blind and depigmented, and show adaptations to living underground that include elongated appendages, reduced metabolism and reproduction rates, loss of wings, as well as an extended life span (Jeffery 2001; Di Lorenzo *et al*. 2015). Understanding the evolution of mitochondrial (mt) genomes of these animals could provide insights into a key part of their metabolic processes. However, there are currently very few mt genomes sequenced from subterranean animals, making such evolutionary analyses difficult to perform.

The arid Yilgarn region in central Western Australia is a biodiversity hotspot for subterranean invertebrates (Guzik *et al.* 2011). The region has hundreds of isolated calcrete (carbonate) bodies, with those examined having their own array of endemic stygofaunal species. Of particular note are the aquatic diving beetles (Dytiscidae), of which ~100 stygobitic species have been described from two tribes, Bidessini and Hydroporini (Balke and Ribera 2004; Leys and Watts 2008; Watts and Humphreys 2009). In each calcrete with stygobitic beetles present, there are generally between two and four species from non-overlapping size classes. In several there are sympatric sister species, suggesting that they have evolved from a stygobitic ancestral species within the calcrete (Cooper *et al*. 2002; Leys *et al*. 2003; Leijs *et al*. 2012). However, most species have evolved independently from surface ancestors, providing a powerful system for exploring the adaptive and regressive changes that occur during the evolution of subterranean animals.

Advances in DNA sequencing technology in recent years now make it possible to obtain whole mt genomes of diverse animal groups. Most of the genomes that have been published to

Journal compilation © CSIRO 2017

date have relied on long-range PCR, which can be both challenging and time-consuming (Hahn *et al*. 2013) because it requires high-molecular-weight DNA, and available primers may not work on the target animal group. With next-generation sequencing of total genomic DNA, some of these issues can be resolved. Due to the small size of the mt genome and high copy number, only relatively shallow sequencing is required to reconstruct the complete genome (Cameron 2014; Kocher *et al*. 2014; Linard *et al*. 2016).

In this study, we present four new mt genomes from subterranean diving beetle species from the genera *Limbodessus* and *Paroster*. Additionally, we compare the overall structure (gene content, order, orientation and size) of these genomes with those of eight epigean (surface) dytiscid species.

Materials and methods

Specimen collection

The study sites included two calcretes found in the Yilgarn region of Western Australia, at Laverton Downs and Sturt Meadows pastoral stations. Thirteen specimens from four species were sequenced in this study (collection details are listed in Table 1). Species included *Limbodessus palmulaoides*, the largest beetle species found in the Laverton calcrete (Watts and Humphreys 2009), and *Paroster macrosturtensis*, *P. mesosturtensis* and *P. microsturtensis*, three sympatric sister species from the Sturt Meadows calcrete (Guzik *et al*. 2009; Watts and Humphreys 2009). Adult beetles were identified on the basis of morphological characters, and larval beetles by *COI* barcoding and BLAST comparison with *COI* data from GenBank. All specimens, except two, were preserved by snap freezing in liquid nitrogen and stored at −80°C. An additional specimen (25545) was preserved in 100% ethanol and then stored at −20°C, while the other (25542) was killed with 100% ethanol just prior to DNA extraction.

DNA extraction and sequencing

Genomic DNA was extracted from whole specimens using a modified Gentra Pure-Gene DNA purification kit protocol (Gentra Systems, Minneapolis, MN, USA). All extractions were performed inside an Aura PCR cabinet (EuroClone, Pero, Italy). Genomic DNA was sent for library construction, and sequencing at the Australian Genome Research Facility; libraries were prepared using a Nextera DNA library prep kit (Caruccio 2011). Each library contained a single specimen and three Illumina Miseq runs were performed. The first Miseq run (300 bp paired end sequencing) included five libraries, four of which were included in the current study. A second Miseq run (150 bp paired end sequencing) included six libraries, four of which were included in the current study. A final Miseq run (300 bp paired end sequencing) had five libraries that were all included in the current study.

Analysis and annotation

Raw sequences were initially analysed to filter out low-quality sequences and those that contained unknown nucleotides (N bases) using Trimmomatic (Bolger *et al*. 2014) and Prinseq (Schmieder and Edwards 2011). A reference database of all complete beetle mt genomes on GenBank ($n = 172$, March 2017) was assembled. The dytiscid sequence data were mapped to the reference file using bowtie2, and default parameter settings (Langmead and Salzberg 2012), to separate the mt genome sequences from the rest of the sequence data. The mt genomes were assembled using a combination of MIRA4 (using the default settings for mirabait) (Chevreux *et al*. 1999) and the circular genome assembly tool in Geneious 8.1.9 (Kearse *et al*. 2012). The *De Novo* assembly tool used in Geneious was set to custom sensitivity with minimum overlap set to 100 bp, minimum overlap ID 95%, word length 50 bp and maximum mismatch set to 5%. The mt genomes were annotated using MITOS (Bernt *et al*. 2013), applying the invertebrate mt genetic code. Any tRNA genes (tRNAs) not found using MITOS were checked against the beetle mt genomes from Linard *et al*. (2016). The protein-coding genes (PCGs) and rRNA genes (rRNAs) were verified by using Blast+ searches (Camacho *et al*. 2009) and then the 5′ and 3′ ends of the genes were refined by comparing the sequences against beetles in the

Table 1. Collection information for the beetles sequenced in this study

All collection localities are in Western Australia. Collectors: KKJ, K. K. Jones; SJBC, S. J. B. Cooper; BL, B. Langille; WFH, W. F. Humphreys; JH, J. Hyde; AA, A. Allford

Genus	Species	Life stage	Location	Collection date	Collector	BPA catalogue no.[A]
Limbodessus	*palmulaoides*	Adult	Mount Windarra	Sep. 2015	KKJ, SJBC, BL	102.100.100/25542
Limbodessus	*palmulaoides*	Adult	Mount Windarra	Apr. 2015	WFH, SJBC, JH	102.100.100/27821
Limbodessus	*palmulaoides*	Larva	Mount Windarra	Apr. 2015	WFH, SJBC, JH	102.100.100/28086
Limbodessus	*palmulaoides*	Larva	Mount Windarra	Apr. 2015	WFH, SJBC, JH	102.100.100/28088
Paroster	*macrosturtensis*	Adult	Sturt Meadows	Sep. 2015	KKJ, SJBC, BL	102.100.100/25544
Paroster	*macrosturtensis*	Adult	Sturt Meadows	Nov. 2006	WFH, SJBC, AA	102.100.100/25545
Paroster	*macrosturtensis*	Adult	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/27822
Paroster	*macrosturtensis*	Larva	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/28084
Paroster	*macrosturtensis*	Larva	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/28085
Paroster	*macrosturtensis*	Larva	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/28089
Paroster	*mesosturtensis*	Adult	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/27823
Paroster	*microsturtensis*	Larva	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/25543
Paroster	*microsturtensis*	Adult	Sturt Meadows	Apr. 2015	WFH, SJBC, JH	102.100.100/27824

[A]These catalogue numbers are part of a larger stygofauna diversity joint project between BioPlatforms Australia (BPA) and The University of Adelaide and sequence data are available on request.

suborder Adephaga (available on GenBank, March 2017) (Table 2) and refining the regions by eye.

Sequence analysis

For comparison to published genomes, eight complete mtDNA genomes from surface dytiscid species were downloaded from GenBank (21/02/2017) (Table 2). The AT Skew of each of the four completed genomes, and those for the additional eight dytiscids were calculated using (A–T)/(A+T) and the GC skew was calculated using (G–C)/(G+C); both skew calculations were based on the majority strand sequence (Grigoriev 1998). AT skews were calculated for the whole genome, the PCGs, tRNAs, rRNAs and the control region (CR). AT percentage ratios were also calculated for the coding region, the CR and the rRNAs. The relative synonymous codon usage was also calculated using MEGA6 (Tamura *et al*. 2013).

Results and discussion

Genome organisation

Complete mt genomes were obtained for *L. palmulaoides* (16 868 bp), *P. macrosturtensis* (16 676 bp) (Fig. 1), *P. mesosturtensis* (16 663 bp), and *P. microsturtensis* (16 504 bp) (Table 3). The mean coverage for all the combined mt genomes by the raw sequence data was between 66.8× (*P. mesosturtensis*) and 2612.5× (*L. palmulaoides*) and every nucleotide was covered a minimum of 2× (*P. microsturtensis*) to 23× (*P. macrosturtensis*) (Table S1, Fig. S1, Supplementary Material). None of the sequences had a Q30 score below 81.5 and the mean confidence for the genomes was 1116.61 (Table S2, Supplementary Material).

Previous studies have suggested that the coding region of the coleopteran mt genome is reasonably stable at an average of 14 700 bp in length (Sheffield *et al*. 2008). However, the mean coding region length for the four new mt genomes is 15 370.8 bp (with a standard deviation of 284.2 bp) and this only decreases to a mean of 15 145.5 bp (with a standard deviation of 334.1 bp) when the eight other genomes from the surface species are included (Linard *et al*. 2016). There is little variation in length among the three *Paroster* sister species with a standard deviation between the three genomes of 5.6 bp compared with the two epigean *Hydroporus* spp., which had a standard deviation of 81.5 bp and the three epigean *Hygrotus* sp., which had a standard deviation of 220.3 bp. The gene content in the genomes of all four subterranean species are typical of insect genomes previously reported, with 13 PCGs, 22 tRNA genes, two mt rRNA genes, and a single CR. The orientation and the order of the genes in *Limbodessus* and *Paroster* are identical to those of other beetle mt genomes previously reported and to the ancestral insect mt genome (Boore *et al*. 1998; Hwang *et al*. 2001). This lack of alteration to the gene order and the absence of any additional genes suggest that the larger sizes of the dytiscid mt genomes are likely due to an increase in the length or number of intergenic regions (IGRs) or the length of the PCGs, compared with other beetle families. When comparing the length of coding regions and the 37 mt genes from the family Dytiscidae to those in beetles from other families (Sheffield *et al*. 2008), there is evidence that the increased length may result from an increase in length of both IGRs and PCGs. Of particular note is the IGR between the tRNAs tRNA-Ile and tRNA-Gln genes, which is considerably expanded in the dytiscid beetles and is likely to account for most of the difference in length. These features are further explored in comparative analyses given below.

Overlapping genes, intergenic regions and skewness

The evolution of the mt genome favours a reduction in size (Andersson and Kurland 1998) and, from an evolutionary perspective, it makes sense that there would be a reduction in the number of IGRs, potentially even to the point of gene overlap. However, gene overlap appear to be the exception rather than the rule as it is rarely the case that the end of one gene is a useful part of the next, plus overlapping genes can lead to post-transcriptional complications (Burger *et al*. 2003;

Table 2. List of complete mt genomes of species in the suborder Adephaga used in this study and their genome sizes

Species	Family	Size (bp)	Accession no.	Reference
Trachypachus holmbergi	Trachypachidae	15 722	NC_011329	Sheffield *et al*. (2008)
Aspidytes niobe	Aspidytidae	14 257	NC_012139	Pons *et al*. (2010)
Macrogyrus oblongus	Gyrinidae	16 643	NC_013249	Cameron *et al*. (2009)
Damaster mirabilissimus mirabilissimus	Carabidae	16 823	NC_016469	Wan *et al*. (2012)
Calosoma sp. BYU-CO241	Carabidae	16 462	NC_018339	Song *et al*. (2010)
Abax parallelepipedus	Carabidae	17 701	NC_030592	Linard *et al*. (2016)
Hygrobia hermanni	Hygrobiidae	16 336	NC_030593	Linard *et al*. (2016)
Acilius sp.	Dytiscidae	20 689	KT876878	Linard *et al*. (2016)
Hygrotus sp.	Dytiscidae	16 730	KT876899	Linard *et al*. (2016)
Hygrotus sp.	Dytiscidae	17 968	KT876900	Linard *et al*. (2016)
Hygrotus sp.	Dytiscidae	17 071	KT876901	Linard *et al*. (2016)
Colymbetes sp.	Dytiscidae	16 211	KT876885	Linard *et al*. (2016)
Liopterus sp.	Dytiscidae	16 541	KT876902	Linard *et al*. (2016)
Hydroporus sp.	Dytiscidae	23 380	KT876896	Linard *et al*. (2016)
Hydroporus sp.	Dytiscidae	17 698	KT876897	Linard *et al*. (2016)
Limbodessus palmulaoides	Dytiscidae	16 868	MG912994	This study
Paroster macrosturtensis	Dytiscidae	16 676	MG912995	This study
Paroster mesosturtensis	Dytiscidae	16 663	MG912996	This study
Paroster microsturtensis	Dytiscidae	16 504	MG912997	This study

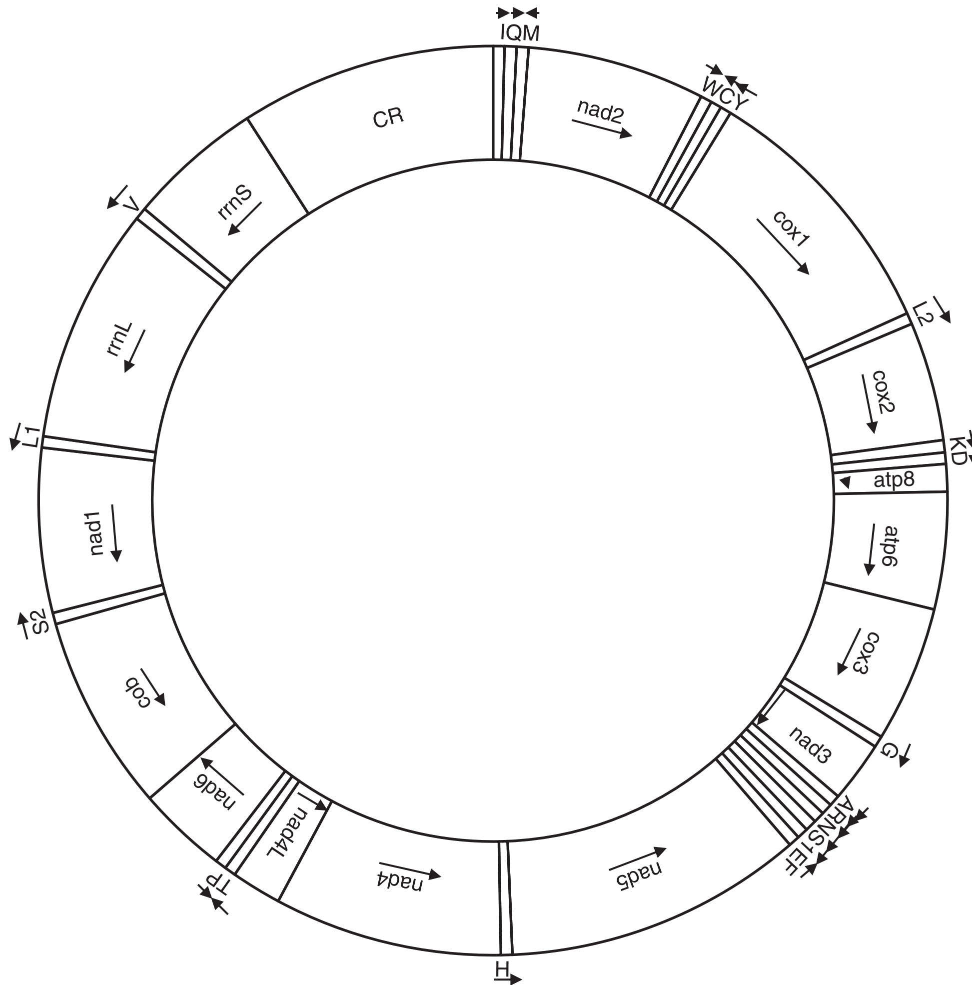

Fig. 1. Map of the mt genome of *Paroster macrosturtensis.* The tRNAs are labelled according to IUPAC-IUB. One-letter symbols – S1, S2, L1, L2 – denote the codons *tRNA-Ser*(AGN), *tRNA-Ser* (UCN), *tRNA-Leu*(CUN) and *tRNA-Leu*(UUR), respectively. The arrow direction indicates if the gene is on the majority or minority strand. An identical mt genome structure was found for all the dytiscid species that were sequenced in the current study.

Sheffield *et al*. 2008). In all four genomes sequenced here, several PCGs (*cox1*, *nad3*, *and cob*) overlap tRNA genes (Table 3). Additionally, in all three *Paroster* species *nad2* overlaps by 1 bp with the tRNA-trp gene.

The mt genome of *L. palmulaoides* contains 1163 bp of IGRs spread over six regions, ranging from 1101 to 2 bp; the longest region is between tRNAs tRNA-Ile and tRNA-Gln. The mt genomes of the three *Paroster* species contain IGRs of 482 bp (*P. macrosturtensis*), 480 bp (*P. mesosturtensis*), and 304 bp (*P. microsturtensis*). The longest region (418, 416, 241 bp) in each species (respectively) occurs between tRNAs tRNA-Ile and tRNA-Gln. In all three species, the shortest IGR is 1 bp. Except for KT876896 (*Hydroporus* sp.) and KT876902 (*Liopterus* sp.), all other dytiscid mt genomes sequenced have their longest region between tRNAs tRNA-Ile and tRNA-Gln and range between 1563 and 100 bp (Linard *et al*. 2016). This region contained no tandem repeats and did not return any significant blast results. Additionally, it did not fold like tRNAs or contain any open reading frames, suggesting that the region is non-coding and non-functional. While most IGRs are unique to each species, there is one well known IGR common to Coleoptera and other insect orders. It is a small IGR between tRNA-Ser(UCN) and *nad1*, with a 5-bp conserved region (TACTA) (Cameron and Whiting 2008; Sheffield *et al*. 2008). All 12 sequenced dytiscids have a 16-bp conserved region including a 5-bp TACTA motif between tRNA-Ser(UCN) and *nad1* with only two species having a single T → A point mutation.

The overall AT content of the genomes of the four subterranean beetles ranged from 75.8% to 77.03% (Table S3, Supplementary Material), which is within the range of the eight surface species' genomes (75.1–81.2%). It is also within the ranges that have been reported previously for other Coleoptera (65.6–78.2%) (Sheffield *et al*. 2008). Three of the previously sequenced mt genomes (KT876878 *Acilius* sp., KT876896 *Hydroporus* sp., and KT876897 *Hydroporus* sp.) have overall AT contents that are above the range previously reported for Coleoptera and all the dytiscid beetles sequenced are at the top end of the range. The four new genomes, like the eight previously

Table 3. The annotation and gene organisation of the four mt genomes sequenced

Gene	Direction	Location	Length	Anticodon	Codon start	Codon stop	Intergenic
L. palmulaoides							
Transfer RNA-Ile	F	1–68	68	gat	–	–	0
Transfer RNA-Gln	R	1170–1238	69	ttg	–	–	1101
Transfer RNA-Met	F	1238–1306	69	cat	–	–	–1
nad2	F	1307–2332	1026	–	ATT	TAA	0
Transfer RNA-Trp	F	2333–2396	64	tca	–	–	0
Transfer RNA-Cys	R	2428–2492	65	gca	–	–	31
Transfer RNA-Tyr	R	2493–2557	65	gta	–	–	0
cox1	F	2550–4094	1545	–	ATT	TAA	–8
Transfer RNA-Leu(UUR)	F	4090–4154	65	taa	–	–	–5
cox2	F	4155–4842	688	–	ATG	T	0
Transfer RNA-Lys	F	4843–4913	71	ctt	–	–	0
Transfer RNA-Asp	F	4914–4979	66	gtc	–	–	0
atp8	F	4980–5138	159	–	ATT	TAA	0
atp6	F	5132–5806	675	–	ATG	TAA	–7
cox3	F	5806–6594	789	–	ATG	TAA	–1
Transfer RNA-Gly	F	6594–6659	66	tcc	–	–	–1
nad3	F	6660–7025	366	–	ATT	TAA	0
Transfer RNA-Ala	F	7012–7075	64	tgc	–	–	–14
Transfer RNA-Arg	F	7076–7139	64	tcg	–	–	0
Transfer RNA-Asn	F	7137–7203	67	gtt	–	–	–3
Transfer RNA-Ser(AGN)	F	7204–7270	67	gct	–	–	0
Transfer RNA-Glu	F	7271–7336	66	ttc	–	–	0
Transfer RNA-Phe	R	7335–7401	67	gaa	–	–	–2
nad5	R	7401–9133	1733	–	ATT	TA	–1
Transfer RNA-His	R	9131–9195	65	gtg	–	–	0
nad4	R	9196–10516	1321	–	ATA	T	–1
nad4l	R	10528–10818	291	–	ATT	TAA	11
Transfer RNA-Thr	F	10821–10886	66	tgt	–	–	2
Transfer RNA-Pro	R	10887–10953	67	tgg	–	–	0
nad6	F	10956–11471	516	–	ATT	TAA	2
cob	F	11471–12607	1137	–	ATG	TAG	–1
Transfer RNA-Ser(UCN)	F	12606–12670	65	tga	–	–	–2
nad1	R	12687–13637	954	–	TTG	TAG	16
Transfer RNA-Leu(CUN)	R	13638–13700	63	tag	–	–	0
rrnL	R	13666–15003	1338	–	–	–	–13
Transfer RNA-Val	R	15002–15072	71	tac	–	–	–1
rrnS	R	15072–15859	788		–	–	–1
CR	–	15859–16868	1009	–	–	–	–
P. macrosturtensis							
Transfer RNA-Ile	F	1–66	66	gat	–	–	0
Transfer RNA-Gln	R	485–553	69	ttg	–	–	418
Transfer RNA-Met	F	561–629	69	cat	–	–	–1
nad2	F	630–1658	1029	–	ATT	TAA	0
Transfer RNA-Trp	F	1658–1723	66	tca	–	–	–1
Transfer RNA-Cys	R	1748–1809	62	gca	–	–	24
Transfer RNA-Tyr	R	1810–1873	64	gta	–	–	0
cox1	F	1866–3410	1545	–	ATT	TAA	–8
Transfer RNA-Leu(UUR)	F	3406–3471	66	taa	–	–	–5
cox2	F	3473–4160	688	–	ATG	T	1
Transfer RNA-Lys	F	4161–4231	71	ctt	–	–	0
Transfer RNA-Asp	F	4232–4297	66	gtc	–	–	0
atp8	F	4298–4456	159	–	ATG	TAG	0
atp6	F	4450–5127	678	–	ATG	TAA	–7
cox3	F	5127–5915	789	–	ATG	TAA	–1
Transfer RNA-Gly	F	5915–5980	66	tcc	–	–	–1
nad3	F	5981–6334	354	–	ATC	TAG	0
Transfer RNA-Ala	F	6333–6397	65	tgc	–	–	–2
Transfer RNA-Arg	F	6398–6462	65	tcg	–	–	0

(*continued next page*)

Table 3. (*continued*)

Gene	Direction	Location	Length	Anticodon	Codon start	Codon stop	Intergenic
Transfer RNA-Asn	F	6463–6526	64	gtt	–	–	0
Transfer RNA-Ser(AGN)	F	6527–6593	67	gct	–	–	0
Transfer RNA-Glu	F	6594–6657	64	ttc	–	–	0
Transfer RNA-Phe	R	6656–6720	65	gaa	–	–	–2
nad5	R	6720–8453	1734	–	ATT	TAA	–1
Transfer RNA-His	R	8451–8516	66	gtg	–	–	0
nad4	R	8517–9850	1334	–	ATA	T	–1
nad4l	R	9849–10139	291	–	ATT	TAA	11
Transfer RNA-Thr	F	10142–10206	65	tgt	–	–	2
Transfer RNA-Pro	R	10207–10272	66	tgg	–	–	0
nad6	F	10274–10795	522	–	ATC	TAA	1
Cob	F	10795–11931	1137	–	ATG	TAG	–1
Transfer RNA-Ser(UCN)	F	11930–11995	66	tga	–	–	–2
nad1	R	12012–12959	1184	–	TTG	TAG	16
Transfer RNA-Leu(CUN)	R	12960–13024	65	tag	–	–	0
rrnL	R	12990–14340	1351	–	–	–	–13
Transfer RNA-Val	R	14339–14409	71	tac	–	–	–1
rrnS	R	14411–15193	783		–	–	–1
CR	–	15194–16676	1482	–	–	–	–
P. mesosturtensis							
Transfer RNA-Ile	F	1–65	65	Gat	–	–	0
Transfer RNA-Gln	R	482–550	69	Ttg	–	–	416
Transfer RNA-Met	F	558–626	69	Cat	–	–	7
nad2	F	627–1655	1029	–	ATT	TAA	0
Transfer RNA-Trp	F	1655–1721	67	Tca	–	–	–1
Transfer RNA-Cys	R	1745–1806	62	Gca	–	–	23
Transfer RNA-Tyr	R	1807–1871	65	Gta	–	–	0
cox1	F	1864–3408	1545	–	ATT	TAA	–8
Transfer RNA-Leu(UUR)	F	3404–3469	66	Taa	–	–	–5
cox2	F	3472–4159	688	–	ATG	T	2
Transfer RNA-Lys	F	4160–4230	71	Ctt	–	–	0
Transfer RNA-Asp	F	4231–4295	65	Gtc	–	–	0
atp8	F	4296–4454	159	–	ATG	TAA	0
atp6	F	4448–5125	678	–	ATG	TAA	–7
cox3	F	5125–5913	789	–	ATG	TAA	–1
Transfer RNA-Gly	F	5913–5978	66	Tcc	–	–	–1
nad3	F	5979–6332	354	–	ATC	TAG	0
Transfer RNA-Ala	F	6331–6395	65	Tgc	–	–	–2
Transfer RNA-Arg	F	6395–6458	64	Tcg	–	–	–1
Transfer RNA-Asn	F	6459–6522	64	Gtt	–	–	0
Transfer RNA-Ser(AGN)	F	6523–6589	67	Gct	–	–	0
Transfer RNA-Glu	F	6590–6653	64	Ttc	–	–	0
Transfer RNA-Phe	R	6652–6716	65	Gaa	–	–	–2
nad5	R	6716–8449	1734	–	ATT	TAA	–1
Transfer RNA-His	R	8447–8511	65	Gtg	–	–	0
nad4	R	8512–9832	1334	–	ATG	T	–1
nad4l	R	9844–10134	291	–	ATT	TAA	11
Transfer RNA-Thr	F	10137–10201	65	Tgt	–	–	2
Transfer RNA-Pro	R	10202–10269	68	Tgg	–	–	0
nad6	F	10271–10792	522	–	ATT	TAA	1
cob	F	10792–11928	1137	–	ATG	TAG	–1
Transfer RNA-Ser(UCN)	F	11927–11993	67	Tga	–	–	–2
nad1	R	12010–12957	948	–	ttg	TAG	16
Transfer RNA-Leu(CUN)	R	12958–13022	65	Tag	–	–	0
rrnL	R	12988–14333	1346	–	–	–	–13
Transfer RNA-Val	R	14332–14402	71	Tac	–	–	1
rrnS	R	14404–15183	780		–	–	1
CR	–	15184–16663	1479	–	–	–	–

(*continued next page*)

Table 3. (*continued*)

Gene	Direction	Location	Length	Anticodon	Codon start	Codon stop	Intergenic
P. microsturtensis							
Transfer RNA-Ile	F	1–65	65	gat	–	–	0
Transfer RNA-Gln	R	307–375	69	ttg	–	–	241
Transfer RNA-Met	F	383–451	69	cat	–	–	7
nad2	F	452–1480	1029	–	ATT	TAA	0
Transfer RNA-Trp	F	1480–1547	68	tca	–	–	–1
Transfer RNA-Cys	R	1570–1631	62	gca	–	–	22
Transfer RNA-Tyr	R	1632–1698	67	gta	–	–	0
cox1	F	1691–3235	1545	–	ATT	TAA	–8
Transfer RNA-Leu(UUR)	F	3231–3296	66	taa	–	–	–5
cox2	F	3299–3986	688	–	ATG	T	2
Transfer RNA-Lys	F	3987–4057	71	ctt	–	–	0
Transfer RNA-Asp	F	4058–4123	66	gtc	–	–	0
atp8	F	4124–4282	159	–	ATG	TAA	0
atp6	F	4276–4953	678	–	ATG	TAA	–7
cox3	F	4953–5741	789	–	ATG	TAA	–1
Transfer RNA-Gly	F	5741–5806	66	tcc	–	–	–1
nad3	F	5807–6172	366	–	ATC	TAG	0
Transfer RNA-Ala	F	6159–6223	65	tgc	–	–	–14
Transfer RNA-Arg	F	6224–6289	66	tcg	–	–	0
Transfer RNA-Asn	F	6290–6353	64	gtt	–	–	0
Transfer RNA-Ser(AGN)	F	6354–6420	67	gct	–	–	0
Transfer RNA-Glu	F	6421–6484	64	ttc	–	–	0
Transfer RNA-Phe	R	6483–6547	65	gaa	–	–	–2
nad5	R	6547–8280	1734	–	ATT	TAA	–1
Transfer RNA-His	R	8278–8343	66	gtg	–	–	0
nad4	R	8344–9677	1334	–	ATA	T	–1
nad4l	R	9676–9966	291	–	ATT	TAA	11
Transfer RNA-Thr	F	9969–10033	65	tgt	–	–	2
Transfer RNA-Pro	R	10034–10099	66	tgg	–	–	0
nad6	F	10101–10622	522	–	ATT	TAA	1
cob	F	10622–11758	1137	–	ATG	TAG	–1
Transfer RNA-Ser(UCN)	F	11757–11822	66	tga	–	–	–2
nad1	R	11839–12786	948	–	ttg	TAG	16
Transfer RNA-Leu(CUN)	R	12787–12850	64	tag	–	–	0
rrnL	R	12817–14168	1352	–	–	–	–12
Transfer RNA-Val	R	14167–14237	71	tac	–	–	1
rrnS	R	14239–15018	780		–	–	1
CR	–	15019–16504	1485	–	–	–	–

reported, have a weak positive AT skew and a negative GC skew (Table S3, Supplementary Material).

Protein-coding genes

The 13 PCGs in all four mt genomes presented use standard start codons and both complete and incomplete stop codons with the exception of *nad1*, which uses the atypical start codon TTG (Table 3). They also have incomplete stop codons to terminate *cox2* (T) and *nad4* (T). Additionally, *L. palmulaoides* has an incomplete stop codon at the end of *nad5* (TA). Other beetle mt genomes, including those of dytiscids, have been found to include both the atypical start codon and incomplete stop codons (Sheffield *et al*. 2008; Linard *et al*. 2016). Relative synonymous codon usage values of each of the four mt genomes are summarised in Table S4 (Supplementary Material); in all three of the *Paroster* genomes, all of the codons are present. In *L. palmulaoides* only the codons CGC and CGG, which both code for Arg, are not represented in the coding sequence. The most frequent amino acids are leucine 2 (Leu(UUR) amino acid present 401–433 times), isoleucine (Ile amino acid present 379–393 times) and phenylalanine (Phe amino acid present 369–385 times); these amino acids are also abundant in the other dytiscid beetles as well as other insects (Sheffield *et al*. 2008; Dai *et al*. 2017). These frequencies are consistent with the range observed in the surface dytiscid beetles sequenced (Fig. S3, Supplementary Material). The average AT content of the 13 PCGs for *L. palmulaoides* is 73.91%, and in the three *Paroster* species it ranges from 72.30% (*P. macrosturtensis*) to 74.79% (*P. mesosturtensis*). The AT skew in PCGs of all four subterranean dytiscids was slightly negative, indicating a higher content of T than of A (Table S1, Supplementary Material).

The GC skew in *P. macrosturtensis* and *P. mesosturtensis* was also slightly negative, showing a higher content of C than G present in the PCGs. Similar results were found in the other epigean dytiscids investigated. However, the GC skew of both *P. microsturtensis* and *L. palmulaoides* is zero, indicating that an equal number of Gs and Cs are present in the 13 PCGs overall (Table S3, Supplementary Material).

Transfer RNAs

It was found that for the structure of the tRNA genes in the subterranean species, 14 are encoded on the major strand, and the remaining eight are encoded on the minor strand. The total length of the tRNAs of *L. palmulaoides* is 1460 bps, while the three *Paroster* species total tRNA lengths range from 1454 bp (*P. macrosturtensis*) to 1458 bp (*P. microsturtensis*). The maximum tRNA length of all subterranean species is 71 bp. The minimum tRNA length is 62 bp in three species while in *P. mesosturtensis* the shortest tRNA is 63 bp. The tRNAs AT content for *L. palmulaoides* is 77.67%, and for the three *Paroster* species it is between 79.00% (*P. microsturtensis*) and 79.59% (*P. mesosturtensis*). All four AT (0.01–0.04) and GC (0.15–0.19) skews were slightly positive, indicating more A and G occurrences compared with T and C nucleotides.

Ribosomal RNAs

The two rRNA genes (*rrnS* and *rrnL*) in all four mt genomes investigated are located between tRNA-Leu(CUN) and tRNA-Val, and tRNA-Val and the CR, respectively. The lengths of *rrnL* range from 1315 bp (*L. palmulaoides*) to 1327 bp (*P. microsturtensis*) and the lengths of *rrnS* range from 780 bp (*P. mesosturtensis*, *P. microsturtensis*) to 788 bp (*L. palmulaoides*). The AT content of the two ribosomal genes is very similar, ranging between 79.25% (*L. palmulaoides*) and 80.21% (*P. microsturtensis*). The rRNA AT skew for all four species was slightly negative (–0.04 to –0.05), indicating that there were more T nucleotides than A. The GC skew was positive (0.38 to 0.40), indicating that there were more G nucleotides than C, as found in other beetle mt genomes (Friedrich and Muqim 2003; Sheffield *et al*. 2008).

Control region

While coding regions are, to a large degree, constrained in their length, for the genes to function properly, the AT-rich CR has considerable length variation as it is the non-coding region of the mt genome and so is relatively free from these restraints (Fenn *et al*. 2007). While the size of CRs varies significantly across different beetle lineages, from less than 300 bp to over 6500 bp (Sheffield *et al*. 2008), in the Dytiscidae, especially, there appears to be a significant amount of variation in the length of the CR. It ranges in size from 1009 bp in *L. palmulaoides* to 8648 bp in one of the *Hydroporus* sp. (Linard *et al*. 2016), the latter being larger than any other coleopteran CR previously reported (Sheffield *et al*. 2008). The AT content of the CRs of the four mt genomes are remarkably consistent, ranging from 82.4% (*P. microsturtensis*) to 84.9% (*P. macrosturtensis*), considering the highly variable nature of the region. The AT content of surface dytiscid beetle CRs range from 84.9% (KT876899) to 94% (KT876897) (Table S3, Supplementary Material).

Conclusion

The mt genomes of the four subterranean diving beetles sequenced are well conserved, with no significant differences in the overall structure (number, order and orientation) and nucleotide composition compared with those of surface dytiscids. However, further analyses need to target the adaptive variation in individual mt genes, which would require a phylogenetic framework and additional contrasts among related surface and independently evolved subterranean taxa. These comparisons would enable tests of positive selection at the amino acid/nucleotide level in PCGs, to determine whether any metabolic changes in mt genes have evolved during the adaptation of species to subterranean life. The mt genomes reported here provide a basis for these future comparative analyses to be conducted on the ~100 subterranean dytiscids in the genera *Paroster* and *Limbodessus* from the calcrete archipelago of central Western Australia.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

We thank Flora, Peter and Paul Axford of the Sturt Meadows Station for providing both accommodation and access to their property. This research was funded by Australian Research Council linkage grants (LP100200494 and LP140100555) to ADA, SJBC, WFH, and PM, with industry partners Minara Resources Limited, the Western Australian Museum, the South Australian Museum, Biota Environmental Sciences, Bennelongia Pty Ltd and the Department of Biodiversity, Conservation and Attractions (WA), and by funding from Bioplatforms Australia and the School of Biological Sciences, the University of Adelaide.

References

Andersson, S. G., and Kurland, C. G. (1998). Reductive evolution of resident genomes. *Trends in Microbiology* **6**, 263–268. doi:10.1016/S0966-842X(98)01312-2

Balke, M., and Ribera, I. (2004). Jumping across Wallace's line: *Allodessus* Guignot and *Limbodessus* Guignot revisited (Coleoptera: Dytiscidae, Bidessini) based on molecular–phylogenetic and morphological data. *Australian Journal of Entomology* **43**, 114–128. doi:10.1111/j.1440-6055.2004.00415.x

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., Pütz, J., Middendorf, M., and Stadler, P. F. (2013). MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69**, 313–319. doi:10.1016/j.ympev.2012.08.023

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120. doi:10.1093/bioinformatics/btu170

Boore, J. L., Lavrov, D. V., and Brown, W. M. (1998). Gene translocation links insects and crustaceans. *Nature* **392**, 667–668. doi:10.1038/33577

Burger, G., Gray, M. W., and Lang, B. F. (2003). Mitochondrial genomes: anything goes. *Trends in Genetics* **19**, 709–716. doi:10.1016/j.tig.2003.10.012

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T. L. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421. doi:10.1186/1471-2105-10-421

Cameron, S. L. (2014). How to sequence and annotate insect mitochondrial genomes for systematic and comparative genomics research. *Systematic Entomology* **39**, 400–411. doi:10.1111/syen.12071

Cameron, S. L., and Whiting, M. F. (2008). The complete mitochondrial genome of the tobacco hornworm, *Manduca sexta* (Insecta: Lepidoptera: Sphingidae), and an examination of mitochondrial gene variability within butterflies and moths. *Gene* **408**, 112–123. doi:10.1016/j.gene.2007.10.023

Cameron, S. L., Sullivan, J., Song, H., Miller, K. B., and Whiting, M. F. (2009). A mitochondrial genome phylogeny of the Neuropterida (lace-wings, alderflies and snakeflies) and their relationship to the other holometabolous insect orders. *Zoologica Scripta* **38**, 575–590. doi:10.1111/j.1463-6409.2009.00392.x

Caruccio, N. (2011). Preparation of next-generation sequencing libraries using Nextera™ technology: simultaneous DNA fragmentation and adaptor tagging by *in vitro* transposition. In 'High-throughput Next Generation Sequencing: Methods and Applications'. (Eds Y. M. Kwon and S. C. Ricke.) pp. 241–255. (Humana Press: Totowa, NJ.)

Chevreux, B., Wetter, T., and Suhai, S. (1999). Genome sequence assembly using trace signals and additional sequence information. In 'German Conference on Bioinformatics'. pp. 45–56. (Heidelberg.)

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

Culver, D. C., White, W. B., and White, W. B. (2005). 'Encyclopedia of Caves.' (Elsevier: Amsterdam.)

Dai, L.-S., Li, S., Yu, H.-M., Wei, G., Wang, L., Qian, C., Zhang, C.-F., Li, J., Sun, Y., and Zhao, Y. (2017). Mitochondrial genome of the sweet potato hornworm, *Agrius convolvuli* (Lepidoptera: Sphingidae), and comparison with other Lepidoptera species. *Genome* **60**, 128–138. doi:10.1139/gen-2016-0058

Di Lorenzo, T., Di Marzio, W. D., Spigoli, D., Baratti, M., Messana, G., Cannicci, S., and Galassi, D. M. P. (2015). Metabolic rates of a hypogean and an epigean species of copepod in an alluvial aquifer. *Freshwater Biology* **60**, 426–435. doi:10.1111/fwb.12509

Fenn, J., Cameron, S., and Whiting, M. (2007). The complete mitochondrial genome sequence of the Mormon cricket (*Anabrus simplex*: Tettigoniidae: Orthoptera) and an analysis of control region variability. *Insect Molecular Biology* **16**, 239–252. doi:10.1111/j.1365-2583.2006.00721.x

Friedrich, M., and Muqim, N. (2003). Sequence and phylogenetic analysis of the complete mitochondrial genome of the flour beetle *Tribolium castanaeum. Molecular Phylogenetics and Evolution* **26**, 502–512. doi:10.1016/S1055-7903(02)00335-4

Galassi, D. M. P. (2001). Groundwater copepods: diversity patterns over ecological and evolutionary scales. *Hydrobiologia* **453**, 227–253. doi:10.1023/A:1013100924948

Grigoriev, A. (1998). Analyzing genomes with cumulative skew diagrams. *Nucleic Acids Research* **26**, 2286–2290. doi:10.1093/nar/26.10.2286

Guzik, M. T., Cooper, S. J. B., Humphreys, W. F., and Austin, A. D. (2009). Fine-scale comparative phylogeography of a sympatric sister species triplet of subterranean diving beetles from a single calcrete aquifer in Western Australia. *Molecular Ecology* **18**, 3683–3698. doi:10.1111/j.1365-294X.2009.04296.x

Guzik, M. T., Austin, A. D., Cooper, S. J., Harvey, M. S., Humphreys, W. F., Bradford, T, Eberhard, S. M., King, R. A., Leys, R, and Muirhead, K. A. (2011). Is the Australian subterranean fauna uniquely diverse? *Invertebrate Systematics* **24**, 407–418.

Hahn, C., Bachmann, L., and Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads – a baiting and iterative mapping approach. *Nucleic Acids Research* **41**, e129. doi:10.1093/nar/gkt371

Hüppop, K. (2000). How do cave animals cope with the food scarcity in caves? In 'Subterranean Ecosystems: Ecosystems of the world'. Vol. 30. (Eds H. Wilkins, D. C. Culver, and W. F. Humphreys.) pp. 159–188. (Elsevier: Amsterdam.)

Hwang, U. W., Park, C. J., Yong, T. S., and Kim, W. (2001). One-step PCR amplification of complete arthropod mitochondrial genomes. *Molecular Phylogenetics and Evolution* **19**, 345–352. doi:10.1006/mpev.2001.0940

Jeffery, W. R. (2001). Cavefish as a model system in evolutionary developmental biology. *Developmental Biology* **231**, 1–12. doi:10.1006/dbio.2000.0121

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., and Duran, C. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649. doi:10.1093/bioinformatics/bts199

Kocher, A., Kamilari, M., Lhuillier, E., Coissac, E., Péneau, J., Chave, J., and Murienne, J. (2014). Shotgun assembly of the assassin bug *Brontostoma colossus* mitochondrial genome (Heteroptera, Reduviidae). *Gene* **552**, 184–194. doi:10.1016/j.gene.2014.09.033

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357–359. doi:10.1038/nmeth.1923

Leijs, R., Van Nes, E. H., Watts, C. H. S., Cooper, S. J. B., Humphreys, W. F., and Hogendoorn, K. (2012). Evolution of blind beetles in isolated aquifers: a test of alternative modes of speciation. *PLoS One* **7**, e34260. doi:10.1371/journal.pone.0034260

Leys, R., and Watts, C. H. S. (2008). Systematics and evolution of the Australian subterranean hydroporine diving beetles (Dytiscidae), with notes on *Carabhydrus. Invertebrate Systematics* **22**, 217–225. 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:10.1554/03-165

Linard, B., Arribas, P., Andújar, C., Crampton-Platt, A., and Vogler, A. (2016). Lessons from genome skimming of arthropod-preserving ethanol. *Molecular Ecology Resources* **16**, 1365–1377. doi:10.1111/1755-0998.12539

Pons, J., Ribera, I., Bertranpetit, J., and Balke, M. (2010). Nucleotide substitution rates for the full set of mitochondrial protein-coding genes in Coleoptera. *Molecular Phylogenetics and Evolution* **56**, 796–807. doi:10.1016/j.ympev.2010.02.007

Schmieder, R., and Edwards, R. (2011). Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864. doi:10.1093/bioinformatics/btr026

Sheffield, N., Song, H., Cameron, S., and Whiting, M. (2008). A comparative analysis of mitochondrial genomes in Coleoptera (Arthropoda: Insecta) and genome descriptions of six new beetles. *Molecular Biology and Evolution* **25**, 2499–2509. doi:10.1093/molbev/msn198

Song, H., Sheffield, N. C., Cameron, S. L., Miller, K. B., and Whiting, M. F. (2010). When phylogenetic assumptions are violated: base compositional heterogeneity and among-site rate variation in beetle mitochondrial phylogenomics. *Systematic Entomology* **35**, 429–448. doi:10.1111/j.1365-3113.2009.00517.x

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–2729. doi:10.1093/molbev/mst197

Wan, X., Hong, M. Y., Liao, A., Kim, M. I., Kim, K. G., Han, Y. S., and Kim, I. (2012). Complete mitochondrial genome of a carabid beetle, Damaster mirabilissimus mirabilissim (Coleoptera: Carabidae). *Entomological Research* **42**, 44–54. doi:10.1111/j.1748-5967.2011.00355.x

Watts, C. H. S., and Humphreys, W. F. (2009). Fourteen new Dytiscidae (Coleoptera) of the genera *Limbodessus* Guignot, *Paroster* Sharp, and *Exocelina* Broun from underground waters in Australia. *Transactions of the Royal Society of South Australia* **133**, 62–107. doi:10.1080/03721426.2009.10887112

Handling Editor: Paul Cooper