10.1071/WR21052
Wildlife Research

Supplementary Material

Diverse moth prey identified in the diet of the critically endangered southern bent-wing bat (*Miniopterus orianae bassanii*) using DNA metabarcoding of scats

Johanna G. Kuhne[A,*], *Jeremy J. Austin*[B], *Terry B. Reardon*[C] and *Thomas A. A. Prowse*[D]

[A]School of Biological Sciences, The University of Adelaide, Adelaide, SA 5005, Australia

[B]Australian Centre for Ancient DNA, School of Biological Sciences, University of Adelaide, Darling Building, North Terrace Campus, SA 5005, Australia

[C]South Australian Museum, North Terrace, Adelaide, SA 5000, Australia

[D]School of Mathematical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia

*Correspondence to: Email: johanna.kuhne@adelaide.edu.au

Diverse moth prey identified in the diet of the critically endangered southern bent-wing bat (Miniopterus orianae bassanii) using DNA metabarcoding of scats.

Supplementary Materials

Supplementary Methods

This file contains a complete, detailed description of the methods.

Study sites and sample collection

In total we collected 66 samples of scats or guano from seven caves - 39 sheet samples of scats (20 from Naracoorte Bat Cave and 19 from Glencoe West Cave in South Australia) and 27 floor samples of guano (five from Naracoorte Bat Cave, six from Glencoe West Cave and 16 from five Victorian caves) as described below. SBWB are the only species of bat to have been recorded in any of these caves (TBR personal observations) so we are confident that all scats and guano collected came from SBWBs.

SBWB scat samples and guano were collected between February and April 2019. Over this period the majority of SBWBs gather at the two maternity caves, though non-breeding caves are still used by small groups. Two maternity caves and five non-breeding caves were sampled (Fig. 1). Samples from South Australian caves were collected on 17/04/2019 under Scientific Research Permit (no. Y26828-2) issued by the Department for Environment and Water, South Australia, while sampling from Victorian caves occurred from 17/02/2019 to 21/02/2019 and was performed under Wildlife Act Scientific Permit (no. 10008640) issued by the Department of Environment, Land, Water and Planning, Victoria.

SBWB scat samples were collected at the two South Australian caves using plastic sheets (hereafter ‘sheet samples’), in an effort to reduce contamination by the DNA of non-diet species such as cave fauna, and to ensure freshness of scat material. Numbered plastic sheets measuring approximately 1 × 1.5 m were positioned below bat roosts and in prominent flyways. Stakes were used to elevate the sheets from the cave floor to minimise access by cave invertebrates. Twenty sheets were placed in Naracoorte Bat Cave and 19 were placed in Glencoe West Cave. Sheets were left *in situ* for 24 hours to allow for a full foraging/resting cycle while minimising time for contamination from crawling and flying arthropods living in the cave. All

scats deposited on sheets were collected using sterile forceps or plastic spoons and stored in sterile jars at -20°C.

A second collection method sampled guano directly from the top of guano piles located below bat roosts on the cave floors (hereafter 'floor samples'). Floor samples consisted of approximately 150-200 g of scat material which was collected using a sterile plastic spoon, placed into sterile zip-lock bags and stored at -20 °C. Eleven floor samples were collected at the two South Australian caves (Glencoe West Cave, six samples; Naracoorte Bat Cave; five samples). An additional 16 floor samples were opportunistically collected from five Victorian caves, with sample sizes ranging from two to five per cave (see Supp. Fig. 1 for further details).

DNA extraction

For both floor and sheet samples, ten to twelve scats were randomly selected from each sample for extraction. Scats were approximately 3-4 mm long and 2-3 mm wide. For samples with ten or more scats, the weight of material used for extraction ranged from 0.09 g to 0.34 g. Seven sheet samples contained less than ten scats, and in these cases the extracted material weighed between 0.02 g to 0.09 g. Extraction was completed using the NucleoSpin®Soil DNA extraction kit (Machery Nagel) as per the manufacturer's instructions. Following an initial optimisation trial on a subset of five samples we elected to use the SL1 lysis buffer with enhancer. Mechanical sample lysis was completed using a Bead Ruptor 24 (Omni International Inc), set to a speed of $5ms^{-1}$ for 30 seconds. Extractions were completed in batches of 8-13 samples, with each batch including an extraction blank (EB) to control for environmental, laboratory or reagent contamination. There were eight extraction blanks in total. DNA extracts were quantified using Qubit Fluorometer (Thermo Fisher Scientific) with the Broad Range Assay. Extracted DNA was stored at -20°C until further analysis.

Amplicon library preparation and sequencing

To control for possible taxonomic PCR amplification bias, we generated amplicon libraries for a small section of the mtDNA Cytochrome Oxidase subunit 1 (COI) barcoding region using two different set of

primers – ANML (Jusino *et al.* 2019) and ZBJ (Zeale *et al.* 2011). Both primers were insect-specific and have been used previously in bat dietary and/or insect metabarcoding studies (Zeale *et al.* 2011; Swift *et al.* 2018; Jusino *et al.* 2019). The regions targeted by ANML and ZBJ were largely overlapping, with the ZBJ barcode extending 16 bp and three bp beyond the 5' ends of the ANML forward and reverse primers respectively (Table 1). We used a two-step PCR protocol: (1) to amplify the target region; and (2) to attach dual 8bp indexes and Illumina adapters to each sample (Table 1). Negative controls were included for each set of extractions, initial PCRs and indexing PCRs to identify laboratory or reagent contamination. Heterogeneity spacers of 1-3 bp were included on some PCR primers to increase sequence heterogeneity in the sequencing run (Table 1).

1st step gene-specific PCRs

The initial PCRs using the ANML and ZBJ primers were done in 12.5µl final volumes containing 1xMRT buffer (1x ImmoBuffer, 1.5mM $MgCl_2$, 0.2mM each dNTP, 0.1mg/ml BSA), 0.4µM forward and reverse primers, 0.25 U IMMOLASE DNA polymerase (Bioline) and 1µl DNA extract.

To amplify with the ANML primers, thermocycling consisted of initial denaturing at 95°C for 10 minutes, followed by 35 cycles of denaturation at 94°C for 30s, annealing at 51°C for 30s, and extension at 72°C for 30s, ending with a final extension at 72°C for 10 min.

For the ZBJ primers, thermocycling used a touch-down procedure consisting of initial denaturation at 95°C for 10 minutes, followed by 16 cycles of denaturation at 94°C for 30s, annealing at 61°C for 30s (decreasing 0.5°C every 30s), and extension at 72°C for 30s, then an additional 19 cycles of denaturation at 94°C for 30s, annealing at 53°C for 30s, and extension at 72°C for 30s and ending with a final extension at 72°C for 10 min.

Sixty-six samples, eight extraction blanks and one PCR blank were amplified in triplicate using the ANML and ZBJ primers. Triplicates were used to broaden diversity detection and to overcome PCR bias resulting

from stochastic amplification of low concentrations of mixed DNA , and were pooled for each sample/primer combination after the first step PCR.

2nd step Indexing PCRs

A unique combination of P5 and P7 indexes were used for each of the 75 ANML PCR products and 75 ZBJ PCR products, plus an indexing PCR blank for both ANML and ZBJ (see Supp. Table 3 for P5/P7 indexes). Dual indexing was used to reduce chimeric molecules that can occur when only single indexes are used per sample (Kircher *et al.* 2012).

Indexing PCR (iPCR) was conducted in a 13.5µl volume containing 1xMRT buffer (1x ImmoBuffer, 1.5mM $MgCl_2$, 0.2mM each dNTP, 0.1mg/ml BSA), 0.4µM forward and reverse indexing primer, 0.25 U IMMOLASE DNA polymerase (Bioline) and 2µl neat PCR product. The iPCR protocol consisted of initial denaturation at 95°C for 10 minutes, followed by eight cycles of denaturation at 94°C for 30s, annealing at 55°C for 30s, and extension at 72°C for 30s, ending with a final extension at 72°C for 10 min.

Indexed PCR products were purified using 1:1 ratio Axyprep magnetic beads (Axygen) using established methods (Rohland and Reich 2012) and quantified using a Qubit Fluorometer (Thermo Fisher Scientific) with the High Sensitivity Assay. The amplicons were then pooled in approximately equimolar concentrations to create a single library pool. A further purification step using Axyprep beads was completed before analysing the pool on a Fragment Analyzer (Advanced Analytical Technologies) to ensure no small DNA fragments remained and DNA concentration was sufficient. The final concentration of the sequencing library was 4nM. Metabarcoding libraries were sequenced on two 2x250 bp paired end read sequencing runs on the Illumina MiSeq platform at the ACRF Cancer Genomics Facility, SA Pathology, Adelaide.

DNA sequence processing and taxonomic assignment

Sequences were initially demultiplexed by SA Pathology using the P5/P7 indexes. All bioinformatics was then performed using *Mothur* software (Schloss *et al.* 2009). Briefly, forward and reverse sequences from

each primer/sample set were paired, and sequences that could not be joined or produced ambiguous base calls were discarded. Primer sequences and sequences with lengths > 10 bp different to that of the target amplicon were removed. Identical sequences (PCR duplicates) were collapsed to a single consensus sequence for identification, but sequence counts were retained for relative abundance analysis. Any unique sequences that were present in a single sample only and were observed less than five times in that sample were discarded, since such sequences are unlikely to represent key diet species and could potentially result from PCR and/or sequencing error.

To facilitate the classification of amplicon sequences to taxonomic groups, a reference library for all available Australian arthropod COI sequences was compiled from the Barcode of Life Data System (BOLD, accessed 02.10.2019) (Ratnasingham and Hebert 2007). Three *Miniopterus* (SBWB genus) sequences were also included from BOLD to rule out contamination of scats by bat DNA. The reference library was filtered by removing sequences with missing or incomplete taxonomic information, and sequences with ambiguous base calls. The retained sequences were then aligned (using *Mothur*) to a publicly available reference alignment of high-quality eukaryotic COI sequences (Machida *et al.* 2017), and the resulting alignment was trimmed to retain the amplified ANML or ZBJ sequence within it. Finally, sequences which did not cover the entire ZBJ and ANML amplicon were removed. These steps resulted in a final classification library of 78,652 high-quality reference COI sequences.

Each unique sequence was assigned to the lowest possible taxonomic level by comparison to the reference database, using the *Mothur* function ‘classify.seqs’ (Schloss *et al.* 2009) with default settings. This produced taxonomic classifications ranging from the Kingdom to Species level, along with a measure of classification certainty based on a bootstrapping algorithm.

Statistical Analysis

A binomial linear mixed-effects model was used to examine the effect of collection method (sheet or floor) and primers (ANML or ZBJ) on the proportion of sequences identified to species level. For this analysis, cave was treated as a random effect on the model intercept.

To identify insect Orders that were regularly observed, the prevalence for each Order was calculated (i.e. the proportion of samples that contained sequences of that Order) by cave, primers and method. To examine the effect of collection method and primers on the observed prevalence of Diptera a binomial linear mixed-effects model was used. However, prevalence data for the other insect Orders were not amenable to analysis in this way, due to prevalence values of zero or one in many cells of the design. To allow comparison across samples with different sequencing depths and to account for PCR amplification biases due to variation in primer efficiencies between taxa we calculated mean relative abundance of each insect Order per sample, for each cave × primers × collection method combination.

Since SBWB scat samples from sheets (with less potential for contamination from cave fauna) were dominated by Lepidoptera sequences (see Results), a subsequent analysis for Lepidopteran species was performed. Lepidoptera sequences that were classified to species level with >97% bootstrap support were accepted for species richness and prevalence analysis. Species richness was calculated for each sample as the total number of Lepidoptera species identified. These data were analysed with a Poisson regression model that included the fixed effects of cave, collection method, primers and their interactions. To account for the possible correlation between species richness and sequencing depth, the total number of sequences obtained for each sample was also included as a continuous covariate.

To investigate whether the probability of occurrence for different Lepidoptera species differed spatially, data for 32 Lepidoptera species with a mean relative abundance >0.05% were extracted. Binomial regression models were used to examine the effect of cave, collection method, primers, and sample sequence count on the probability of species presence. As recent research suggests the Bat Cave and Glencoe West Cave comprise a subpopulation distinct from the Victorian caves (E. van Harten & L. Lumsden, unpublished data), for each prey species, the probability of presence was compared between Victorian and South Australian caves with a planned contrast. To investigate the relationship between these 32 lepidopteran species and agriculture we examined published literature to determine which species have been recorded as eating agricultural plants.

References

Jusino, MA, Banik, MT, Palmer, JM, Wray, AK, Xiao, L, Pelton, E, Barber, JR, Kawahara, AY, Gratton, C, Peery, MZ, Lindner, DL (2019) An improved method for utilizing high-throughput amplicon sequencing to determine the diets of insectivorous animals. *Molecular Ecology Resources* **19**, 176-190.

Kircher, M, Sawyer, S, Meyer, M (2012) Double indexing overcomes inaccuracies in multiplex sequencing on the Illumina platform. *Nucleic Acids Research* **40**, e3.

Machida, RJ, Leray, M, Ho, S-L, Knowlton, N (2017) Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data* **4**, 1-7.

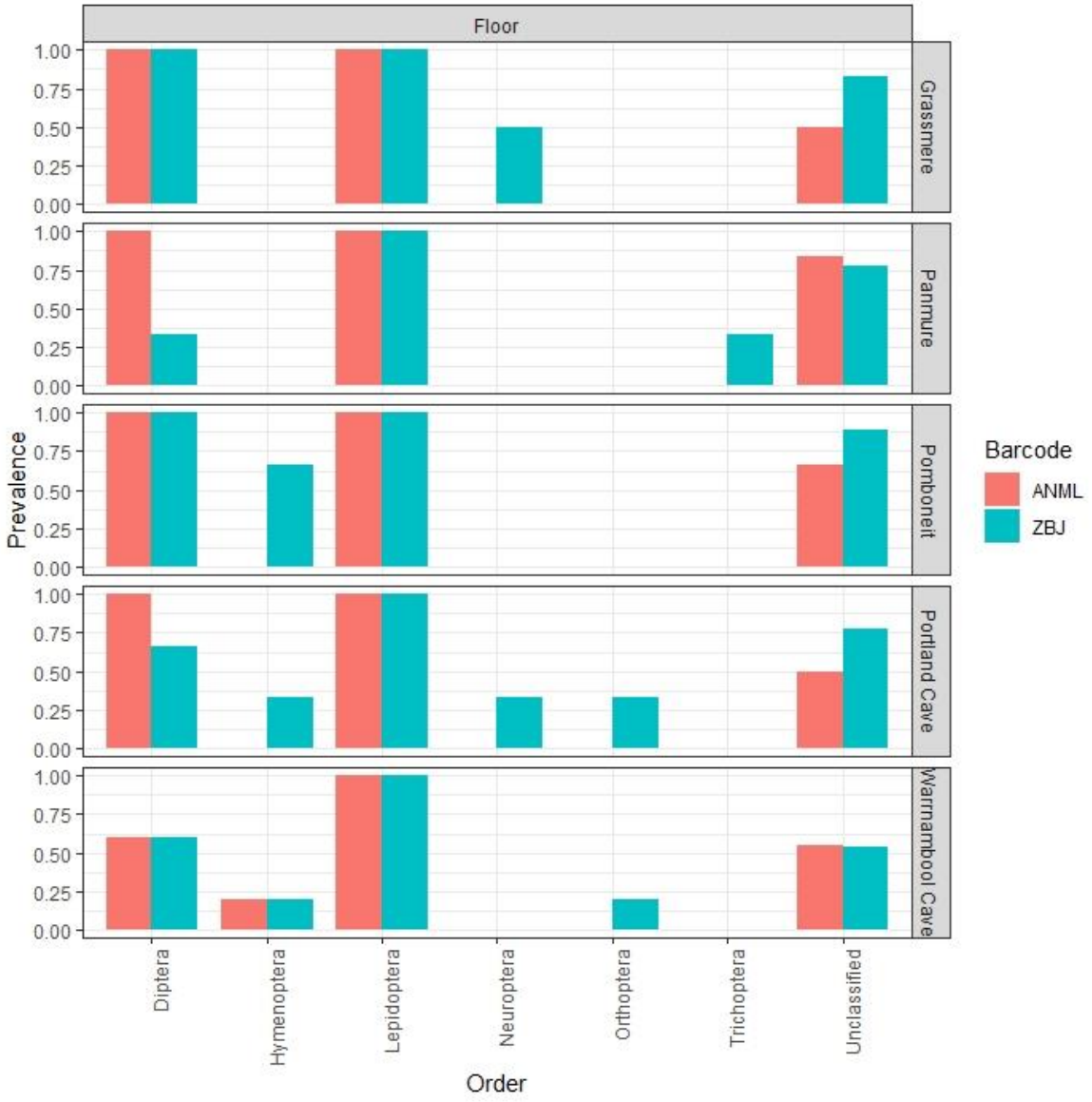
Ratnasingham, S, Hebert, PDN (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). *Molecular Ecology Notes* **7**, 355-364.

Rohland, N, Reich, D (2012) Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research* **22**, 939-946.

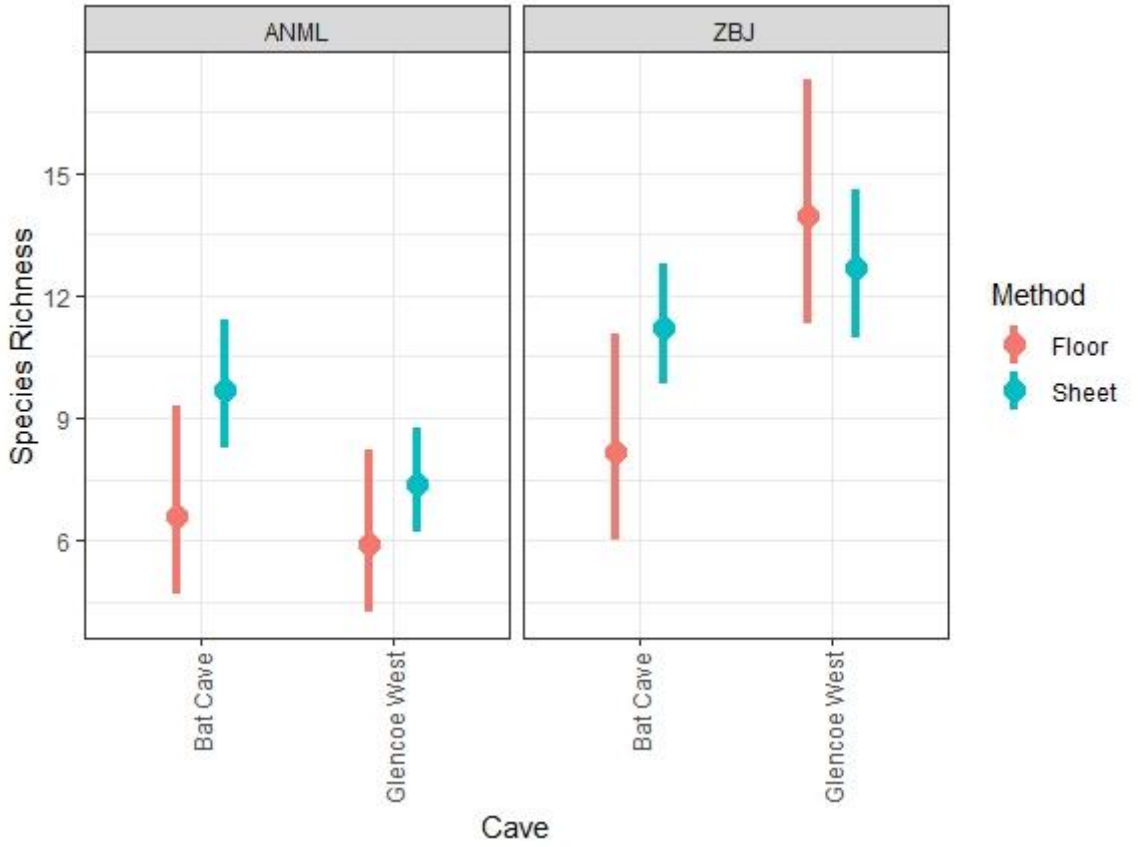
Schloss, PD, Westcott, SL, Ryabin, T, Hall, JR, Hartmann, M, Hollister, EB, Lesniewski, RA, Oakley, BB, Parks, DH, Robinson, CJ, Sahl, JW, Stres, B, Thallinger, GG, Van Horn, DJ, Weber, CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology* **75**, 7537-7541.

Swift, JF, Lance, RF, Guan, X, Britzke, ER, Lindsay, DL, Edwards, CE (2018) Multifaceted DNA metabarcoding: Validation of a noninvasive, next-generation approach to studying bat populations. *Evolutionary Applications* **11**, 1120-1138.

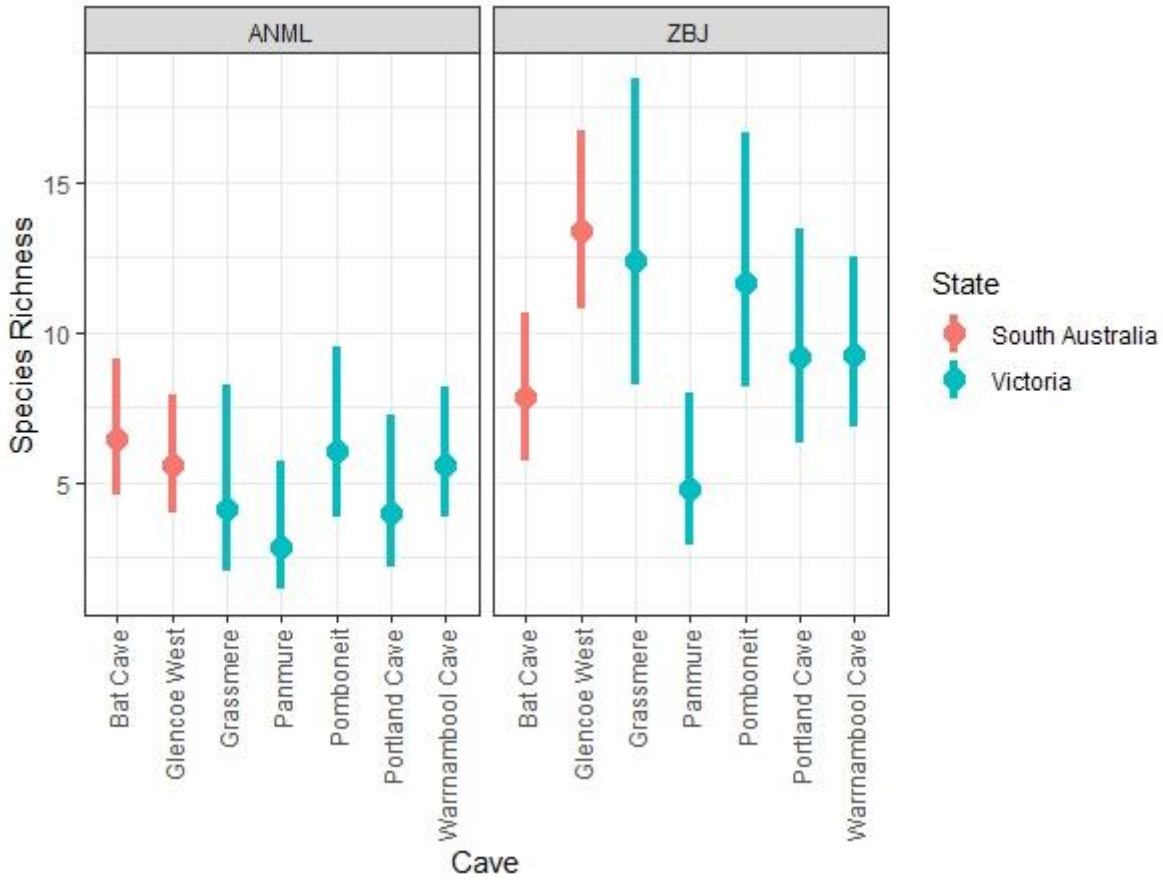
Zeale, MRK, Butlin, RK, Barker, GLA, Lees, DC, Jones, G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources* **11**, 236-244.



Supplementary Figure 1: Prevalence (i.e., the proportion of positive samples) of Insecta Orders identified from 16 samples of southern bent-wing bat scats taken from five caves in Victoria, using two COI barcodes. Sample sizes from top to bottom were n=2, n=3, n=3, n=5, n=3.



Supplementary Figure 2: Estimated species richness per sample (± 95 % confidence intervals) of Lepidoptera species identified from 50 samples of Southern bent-wing bat scats taken from two caves in South Australia, using two COI barcodes and two collection methods. Sample sizes were: n=6 and 19 for Glencoe West floor and sheet samples, respectively; and n=5 and 20 for Bat Cave floor and sheet samples, respectively. Victorian samples (n=16) were not included in this figure but can be seen in Supplementary Fig. 3.



Supplementary Figure 3: Estimated species richness per sample (± 95 % confidence intervals) of Lepidoptera species identified from 66 samples of Southern bent-wing bat scats taken from the floors of seven caves in Southern Australia and Victoria using two COI barcodes.

Supplementary Table 1: Number of insect COI sequences identified at each taxonomic level by method and barcode used. 'Species (%)' shows proportion of all sequences per cave/method/barcode identified to species level.

Cave	Method	Barcode	Kingdom	Phylum	Class	Order	Family	Genus	Species	Species (%)
Bat Cave	Sheet	ANML	413688	398145	398008	396265	247376	234116	222862	**53.87**
		ZBJ	965197	957822	909864	902420	621696	607220	587341	**60.85**
	Floor	ANML	220517	132552	131265	129677	114263	105224	68896	**31.24**
		ZBJ	278691	55516	53558	52724	50076	49972	48132	**17.27**
Glencoe West	Sheet	ANML	613198	542138	537941	535967	465165	449765	431949	**70.44**
		ZBJ	1335626	1269237	1147082	1139290	995149	937655	933280	**69.88**
	Floor	ANML	330968	50090	47644	47004	46831	46327	45976	**13.89**
		ZBJ	319915	286173	279075	276843	254428	247819	246828	**77.15**

Supplementary Table 2: Mean number of diet taxa identified per sample at each taxonomic level by cave, method and barcode for 66 Southern bent-wing bat scat samples.

Cave	Method	Barcode	Class	Order	Family	Genus	Species
Bat Cave	Sheet	ANML	1	2.0	7.7	9.9	9.6
		ZBJ	1	1.5	8.3	13.0	11.8
	Floor	ANML	1	2.0	6.2	8.8	7.2
		ZBJ	1	2.8	8.2	10.4	9.4
Glencoe West Cave	Sheet	ANML	1	1.9	5.1	7.7	7.6
		ZBJ	1	1.7	7.2	14.8	14.1
	Floor	ANML	1	1.7	4.0	6.3	6.3
		ZBJ	1	3.3	8.2	16.2	15.8
Warrnambool Cave	Floor	ANML	1	1.8	4.8	7.0	6.8
		ZBJ	1	2.0	6.2	9.0	9.2
Grassmere Cave	Floor	ANML	1	2.0	2.5	4.0	5.0
		ZBJ	1	2.5	9.0	14.0	14.0
Pomboneit Cave	Floor	ANML	1	2.0	4.7	7.0	8.0
		ZBJ	1	2.7	7.7	11.3	11.7
Panmure Cave	Floor	ANML	1	2.0	3.7	4.3	3.7
		ZBJ	1	1.7	5.0	6.3	5.7
Portland Cave	Floor	ANML	1	2.0	3.7	5.3	4.3
		ZBJ	1	2.7	6.7	10.3	10.3

Supplementary Table 3: P5 and P7 indexes used in unique combinations for indexing 66 Southern bent-wing bat samples.

P5 Index#	P5 Index Seq	P5 Primer ID	P5 Primer Sequence
1	TAGATCGC	P5_index1	AATGATACGGCGACCACCGAGATCTACACTAGATCGCTCGTCGGCAGCGTC
2	CTCTCTAT	P5_index2	AATGATACGGCGACCACCGAGATCTACACCTCTCTATTCGTCGGCAGCGTC
3	TATCCTCT	P5_index3	AATGATACGGCGACCACCGAGATCTACACTATCCTCTTCGTCGGCAGCGTC
4	AGAGTAGA	P5_index4	AATGATACGGCGACCACCGAGATCTACACAGAGTAGATCGTCGGCAGCGTC
5	GTAAGGAG	P5_index5	AATGATACGGCGACCACCGAGATCTACACGTAAGGAGTCGTCGGCAGCGTC
6	ACTGCATA	P5_index6	AATGATACGGCGACCACCGAGATCTACACACTGCATATCGTCGGCAGCGTC
7	AAGGAGTA	P5_index7	AATGATACGGCGACCACCGAGATCTACACAAGGAGTATCGTCGGCAGCGTC
8	CTAAGCCT	P5_index8	AATGATACGGCGACCACCGAGATCTACACCTAAGCCTTCGTCGGCAGCGTC
9	CGTCTAAT	P5_index9	AATGATACGGCGACCACCGAGATCTACACCGTCTAATTCGTCGGCAGCGTC
10	TCTCTCCG	P5_index10	AATGATACGGCGACCACCGAGATCTACACTCTCTCCGTCGTCGGCAGCGTC
11	TCGACTAG	P5_index11	AATGATACGGCGACCACCGAGATCTACACTCGACTAGTCGTCGGCAGCGTC
12	TTCTAGCT	P5_index12	AATGATACGGCGACCACCGAGATCTACACTTCTAGCTTCGTCGGCAGCGTC
13	CCTAGAGT	P5_index13	AATGATACGGCGACCACCGAGATCTACACCCTAGAGTTCGTCGGCAGCGTC
14	GCGTAAGA	P5_index14	AATGATACGGCGACCACCGAGATCTACACGCGTAAGATCGTCGGCAGCGTC
15	CTATTAAG	P5_index15	AATGATACGGCGACCACCGAGATCTACACCTATTAAGTCGTCGGCAGCGTC
16	AAGGCTAT	P5_index16	AATGATACGGCGACCACCGAGATCTACACAAGGCTATTCGTCGGCAGCGTC
17	GAGCCTTA	P5_index17	AATGATACGGCGACCACCGAGATCTACACGAGCCTTATCGTCGGCAGCGTC
18	TTATGCGA	P5_index18	AATGATACGGCGACCACCGAGATCTACACTTATGCGATCGTCGGCAGCGTC
P7 Index#	**P7 Index Seq**	**P7 Primer ID**	**P7 Primer Sequence**
1	TCGCCTTA	P7_index1	CAAGCAGAAGACGGCATACGAGATTCGCCTTAGTCTCGTGGGCTCGG
2	CTAGTACG	P7_index2	CAAGCAGAAGACGGCATACGAGATCTAGTACGGTCTCGTGGGCTCGG
3	TTCTGCCT	P7_index3	CAAGCAGAAGACGGCATACGAGATTTCTGCCTGTCTCGTGGGCTCGG
4	GCTCAGGA	P7_index4	CAAGCAGAAGACGGCATACGAGATGCTCAGGAGTCTCGTGGGCTCGG
5	AGGAGTCC	P7_index5	CAAGCAGAAGACGGCATACGAGATAGGAGTCCGTCTCGTGGGCTCGG
6	CATGCCTA	P7_index6	CAAGCAGAAGACGGCATACGAGATCATGCCTAGTCTCGTGGGCTCGG
7	GTAGAGAG	P7_index7	CAAGCAGAAGACGGCATACGAGATGTAGAGAGGTCTCGTGGGCTCGG
8	CCTCTCTG	P7_index8	CAAGCAGAAGACGGCATACGAGATCCTCTCTGGTCTCGTGGGCTCGG
9	AGCGTAGC	P7_index9	CAAGCAGAAGACGGCATACGAGATAGCGTAGCGTCTCGTGGGCTCGG
10	CAGCCTCG	P7_index10	CAAGCAGAAGACGGCATACGAGATCAGCCTCGGTCTCGTGGGCTCGG
11	TGCCTCTT	P7_index11	CAAGCAGAAGACGGCATACGAGATTGCCTCTTGTCTCGTGGGCTCGG
12	TCCTCTAC	P7_index12	CAAGCAGAAGACGGCATACGAGATTCCTCTACGTCTCGTGGGCTCGG
13	TCATGAGC	P7_index13	CAAGCAGAAGACGGCATACGAGATTCATGAGCGTCTCGTGGGCTCGG
14	CCTGAGAT	P7_index14	CAAGCAGAAGACGGCATACGAGATCCTGAGATGTCTCGTGGGCTCGG
15	TAGCGAGT	P7_index15	CAAGCAGAAGACGGCATACGAGATTAGCGAGTGTCTCGTGGGCTCGG
16	GTAGCTCC	P7_index16	CAAGCAGAAGACGGCATACGAGATGTAGCTCCGTCTCGTGGGCTCGG
17	TACTACGC	P7_index17	CAAGCAGAAGACGGCATACGAGATTACTACGCGTCTCGTGGGCTCGG
18	AGGCTCCG	P7_index18	CAAGCAGAAGACGGCATACGAGATAGGCTCCGGTCTCGTGGGCTCGG
19	GCAGCGTA	P7_index19	CAAGCAGAAGACGGCATACGAGATGCAGCGTAGTCTCGTGGGCTCGG
20	CTGCGCAT	P7_index20	CAAGCAGAAGACGGCATACGAGATCTGCGCATGTCTCGTGGGCTCGG
21	GAGCGCTA	P7_index21	CAAGCAGAAGACGGCATACGAGATGAGCGCTAGTCTCGTGGGCTCGG
22	CGCTCAGT	P7_index22	CAAGCAGAAGACGGCATACGAGATCGCTCAGTGTCTCGTGGGCTCGG
23	GTCTTAGG	P7_index23	CAAGCAGAAGACGGCATACGAGATGTCTTAGGGTCTCGTGGGCTCGG
24	ACTGATCG	P7_index24	CAAGCAGAAGACGGCATACGAGATACTGATCGGTCTCGTGGGCTCGG
25	TAGCTGCA	P7_index25	CAAGCAGAAGACGGCATACGAGATTAGCTGCAGTCTCGTGGGCTCGG
26	GACGTCGA	P7_index26	CAAGCAGAAGACGGCATACGAGATGACGTCGAGTCTCGTGGGCTCGG