ARPHA Preprints

PENSOFT

PREPRINT

Author-formatted, not peer-reviewed document posted on 29/05/2023

DOI: https://doi.org/10.3897/arphapreprints.e106924

DNA metabarcoding data from faecal samples of the lesser (*Myotis blythii*) and the greater (*Myotis myotis*) mouse-eared bats from Bulgaria

Antoniya Hubancheva, Vedran Bozicevic, Jerome Moriniere, Holger Goerlitz

ARPHA Preprints *Author-formatted, not peer-reviewed document posted on 29/05/2023*. DOI: https://doi.org/10.3897/arphapreprints.e106924

DNA metabarcoding data from faecal samples of the lesser (*Myotis blythii*) and the greater (*Myotis myotis*) mouse-eared bats from Bulgaria

Antoniya Hubancheva[1,2], Vedran Bozicevic[3], Jérôme Morinière[3]*, Holger R. Goerlitz[1]*

1 Acoustic and Functional Ecology, Max Planck Institute for Biological Intelligence, Seewiesen, Germany

2 Department of Animal Diversity and Resources, Institute of Biodiversity and Ecosystem Research, Bulgarian Academy of Sciences, Sofia, Bulgaria

3 AIM - Advanced Identification Methods GmbH, Leipzig, Germany

Corresponding author: Antoniya Hubancheva (antonia.hubancheva@gmail.com)

*Shared last authorship

Abstract

A comprehensive understanding of trophic interactions in terrestrial ecosystems is crucial for ecological research and conservation. Recent advances in non-invasive methods, such as DNA metabarcoding, have enabled researchers to collect vast amounts of data on wild animal diets. However, sharing this data and metadata effectively and transparently presents new challenges. To address this, a new type of scholarly journal publication has emerged that aims to describe datasets rather than report research investigations. In this paper, we present a dataset of consumed prey species and parasites based on the metabarcoding of 113 faecal samples from the greater and lesser mouse-eared bats (*Myotis myotis* and *Myotis blythii*), along with a detailed description of the data sampling, laboratory analysis, and bioinformatics pipeline. Our dataset comprises 1018 unique Barcode Index Numbers (BINs) from 12 Classes and 43 Orders. In addition, we provide interactive Krona charts to visually summarize the taxonomic relationships and relative read abundance of the consumed prey species and parasites. This data can be used for meta-analysis, exploring new predator-prey and host-parasite interactions, studying inter and intraspecific ecological interactions, and informing protected area management, among other applications. By sharing this dataset, we hope to encourage other researchers to use it to answer additional ecological questions and advance our understanding of trophic interactions in terrestrial ecosystems.

Overview and background

Bats play a crucial role in terrestrial ecosystems worldwide by occupying various ecological niches and exploiting a range of food sources including insects, vertebrates, blood, nectar, pollen and fruit (Simmons, 2005, Kunz et al., 2011). Due to their ecological abundance, bats could be used as bioindicator species that provide quantitative information on the quality of ecosystems, enabling the tracking of environmental alterations (Russo et al 2021). Therefore,

high-quality and detailed information on the diets of bat species in areas of conservation importance is essential. Historically, bat diets have been studied through the morphological analysis of invasively collected gut content or of noninvasively collected faeces (Whitaker et al. 2009). These morphological methods are time-consuming, require specialized entomological knowledge, and often can only identify prey down to order level. In recent years, advances in DNA barcoding and metabarcoding using high-throughput sequencing revolutionized the study of animal diets by providing a powerful, accurate, and time- and cost-efficient tool that can often identify prey down to the species level. As a consequence, after 2012 metabarcoding became increasingly popular (Ando et al. 2020). However, the methodology is still developing, and mistakes can occur at any level from sample collection, through PCR amplification and the bioinformatics pipeline, to data interpretation (Alberdi et al. 2018, O'Rourke et al. 2020). Thus, the sharing of transparent, detailed and open protocols and data is crucial for promoting good practices and avoiding errors. Here, we provide the full methods for collecting, metabarcoding and analysing a dataset of faecal samples collected from the greater and the lesser mouse-eared bats (*Myotis myotis* and *Myotis blythii*) in Bulgaria.

Myotis myotis (Borkhausen, 1797) and *Myotis blythii* s.l. (Tomes, 1857; for summary and discussion on taxonomy and phylogeny of the species see Ruedi 2020), family Vespertilionidae, are closely related sibling bat species that coexist in parts of Europe and the Middle East. While both are listed as Least Concerned in IUCN (Coroiu et al 2016, Juste & Paunović 2016), in Europe they are protected under the Agreement on the Conservation of Populations of European Bats (EUROBATS). In Bulgaria, both species are listed as Near Threatened in the Red List of Protected Species of the country (Golemanski et al 2015). On the Balkan Peninsula, both bat species live almost exclusively in caves, or on rare occasions in mines, where they form mixed maternity and hibernation colonies. Interspecific competition is avoided mainly by habitat selection and different foraging strategies (Arlettaz 1999). *M. myotis* predominantly forages in forests and agricultural land with open, accessible ground (Arlettaz 1999, Stidsholt et al 2023). M. *blythii*, in contrast, tends to forage in steppe-like habitats with dense grass and small shrubs (Arlettaz 1999). Both species use echolocation for aerial hunting as well as passive listening for gleaning prey off the ground and vegetation (Arlettaz et al. 2001; Siemers and Güttinger 2006, Stidsholt et al. 2023). However, when gleaning, *M. myotis* listens for the rustling sounds of large walking prey (Siemers and Güttinger 2006) while *M. blythii* eavesdrops on the mating song of bushcrickets (Jones et al. 2011).

The diet of *M. myotis* is extensively studied throughout its range with morphological methods (Audet 1990, Beck 1995, Arlettaz et al. 1997a, Arlettaz et al. 1997b, Arlettaz et al. 1999, Pereira et al. 2002, Zahn et al 2006, Steck and Güttinger 2006, Graclik and Wasielewski 2012), while less is known about the diet of *M. blythii* (Arlettaz et al. 1997a, Arlettaz et al. 1997b, Arlettaz 1999). These studies cover well the geographical and ecological variation in the diet of the two bat species, but they have a low taxonomic resolution and mostly identify prey only down to the order. Recent metabarcoding studies provided higher taxonomic resolution, however, only two studies investigated the diet of *M. myotis* (Galan et al 2018, Alberdi et al 2020), and only one study the diet of *M. blythii* (Mata et al. 2021), which additionally sampled and analyzed both species as one species complex due to methodological limitations.

Here, we provide a detailed description of the metabarcoding analysis of the faeces and of the diet of 113 individual bats (60 *M. myotis* and 53 *M. blythii*) collected in an area with high biodiversity value (Cimatti et al. 2021). The dataset contains 1018 Barcode Index Numbers

(BIN) species from 12 Classes and 43 Orders, ranging from prey species to various ecto- and endo-parasites (including acars, fleas, tapeworms, roundworms, and others). Interestingly, the presence of Molluscan and Anneline species (such as *Pomatias rivulare*, *Lumbricus rubellus*, and *Eisenia fetida*) suggests that the data could contain information on further trophic interactions from prey species of predatory carabid beetles or other arthropods, that in turn have been consumed by the bats. The re-using potential of our data set varies from meta-analysis of the diets of insectivorous bats to investigations of predator-prey or host-parasite interactions and interspecific food webs and ecological interactions, and to the management of protected areas.

In summary, the dataset we present in this paper is a valuable resource that can aid in advancing ecological research and conservation efforts. We hope that by sharing our data, we can contribute to a more collaborative and transparent research environment that will lead to more effective conservation and management of terrestrial ecosystems.

Methods

Sampling

Geographic Coverage

Faecal samples were collected from individual bats at the entrance of the Orlova Chucka cave, Pepelina, Dve Mogili District, Bulgaria (N 43.593240 E 25.960108). The cave is inhabited by 15 bat species all year round. In summer, however, it is predominantly occupied by mixed maternity colonies of *M. myotis* and *M. blythii*, as well as *Rhinolophus euryale* and *Rh. mehelyi* (Borissov 2010). Mouse-eared bats are highly mobile with a hunting range of about 23 km around the cave (Egert-Berg et al 2018, Stidsholt et al. 2023). While we collected the faecal samples at the cave entrance, our effective sampling area thus matches the foraging area of the bats, covering an area of approximately 1600 km^2 (Fig. 1A). Notably, a large proportion of the foraging grounds of the bats are in protected areas including NATURA 2000 sites and the Rusenski Lom Natural Park (Borissov 2009). The preferred foraging sites of the mouse-eared bats in the study area consist of small-scale agricultural areas, forests, open grasslands, karstic areas and riverine habitats (Fig. 1B, C, D).

ARPHA Preprints *Author-formatted, not peer-reviewed document posted on 29/05/2023*. DOI: https://doi.org/10.3897/arphapreprints.e106924

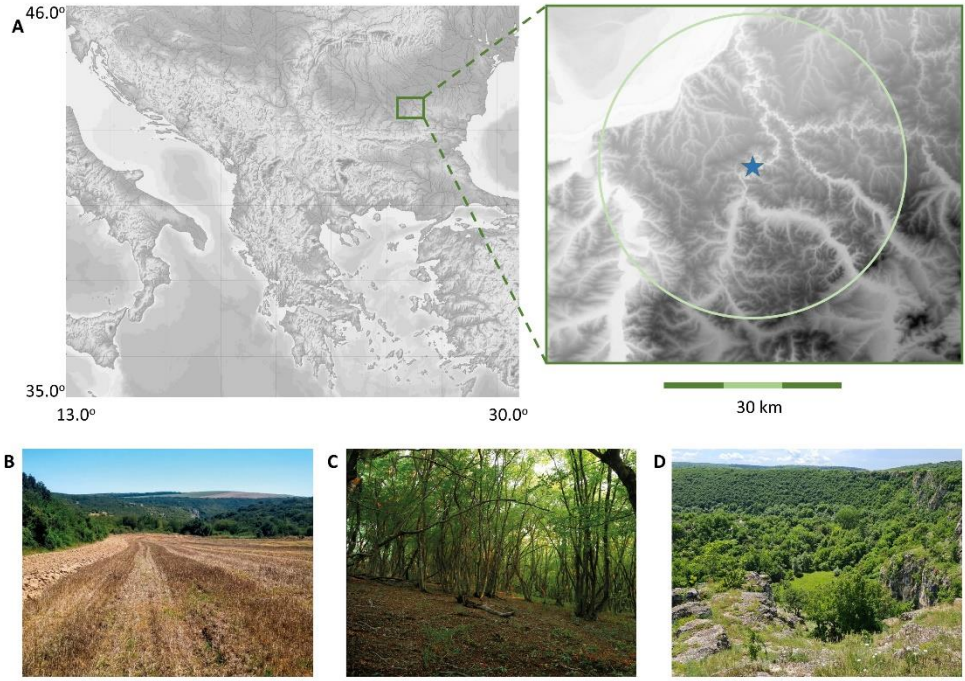


Fig. 1 Geographic coverage, sampling area and typical hunting habitats. A) Topographical map showing the study area with the sampling site marked by a star and the approximate foraging range of the bats represented by a circle. B-D) Representative hunting habitats of the lesser and greater mouse-eared bats, including small-scale agricultural fields, forests, karstic areas, and riverine habitats (modified after Stidsholt et al. 2023).

Temporal Coverage

Samples were collected from June to August 2017 and 2018. This period covers the lactation and post-lactation period of the female bats, during which they have to forage more actively to provide enough nutrition to both themselves and the pup.

Sampling Methodology

Bats were captured in the morning (when returning to the roost after foraging) with a harp trap placed in front of the cave entrance. We emptied the trap every 5 to 10 minutes to minimise defecation in the trap, and thus potential cross-contamination between individualsby faeces attached to the fur. However, we could not fully prevent bats from defecating in the trap, therefore, a small proportion of cross-contamination between the different individuals might have occurred. After being removed from the trap, bats were placed in individual cotton bags until they defecated. Prior to data collection, the bags were brushed from previous guano and washed at 90°C with bleach. After the bats had defecated in the bags, they were measured, sexed and identified to species level following Dietz et al 2004. To avoid misidentification, however, we used a conservative approach and only sampled individuals that could be clearly

identified based on morphological measurements, identifying individuals as *M. myotis* if the length of the upper jaw (i.e., from the canine to the third molar, CM^3) was >9.4 mm and the forearm length (FA) >61 mm, and as *M. blythii* for CM^3 <9.0 mm and FA <59 mm. The guano pellets were placed in 2 ml Eppendorf tubes with 98% ethanol, which were subsequently stored in a freezer at -18°C until further treatment. Bat catching and sample collection were performed under permit granted by the Ministry of Environment and Waters, Bulgaria and under the control of the Regional Environment and Water Inspection Ruse, permit number 696/19.01.2017.

Laboratory procedures

DNA extraction, amplification, and metabarcoding

DNA metabarcoding was conducted at the AIM Lab (AIM—Advanced IdentificationMethods GmbH, Leipzig, Germany). Genomic data was extracted using the Quick-DNA Fecal/Soil Microbe 96 Kit (Zymo Research Corporation, Irvine CA, USA) and following the manufacturer's instructions. To control for artifacts arising from lab contamination, we ran 6 empty vials as negative control samples through the lab procedure: 2 before extraction, and 2 before each of the two rounds of PCR. These negative control samples were processed in the same way as the faecal samples. Further laboratory analyses were carried out as per the methods described in Uhler *et al.* (2022). High Throughput Sequencing (HTS) was performed on an Illumina MiSeq Illumina Inc., San Diego, USA) “v3 chemistry” (2 × 300 base pairs, 600 cycles, maximum of 25 million paired end reads).

Bioinformatics

Preprocessing of raw Illumina reads

From each sample, paired-end reads were merged using the *-fastq_mergepairs* utility of USEARCH v11.0.667 (Edgar, 2010) with the following parameters: *-fastq_maxdiffs* 99, *-fastq_pctid* 75, *-fastq_trunctail* 0. Next, adapter sequences were removed using CUTADAPT (Martin, 2011) (single-end mode, with default parameters). Reads that did not contain the appropriate adapter sequences were filtered out in this step using CUTADAPT's *--discard-untrimmed* option. The remaining pre-processing steps (quality filtering, dereplication, chimera filtering, and clustering) were carried out using the VSEARCH suite v2.9.1 (Rognes et al., 2016).

Quality filtering was performed using *--fastq_filter*, allowing a maximum of 1 expected error along the length of the sequence and a minimum read length of 300 bases (parameters: *--fastq_maxee* 1, *--minlen* 300). This was followed by dereplication on the sample level using *--derep_fulllength,* keeping only a single copy of each unique sequence (parameters: *--sizeout, --relabel Uniq*). Cleaned and dereplicated sample files were concatenated into one large FASTA file, which was then dereplicated again, and also filtered for sequences occurring only once in the entire dataset (singletons) with the parameters *--minuniquesize* 2, *--sizein* , *--sizeout* , *--fasta_width* 0.

To save processing power, a clustering step (at 98% identity) was employed before chimera filtering using the VSEARCH utility *--cluster_size* and the centroids algorithm (parameters: *--id* 0.98, *--strand plus*, *--sizein*, *--sizeout*, *--fasta_width* 0, *--centroids*). Chimeric sequences were then detected and filtered out from the resulting file using the VSEARCH *-- uchime_denovo*

utility (parameters: *--sizein*, *--sizeout*, *--fasta_width* 0, *--nonchimeras*). Next, a perl script obtained from the authors of VSEARCH (https://github.com/torognes/vsearch/wiki/VSEARCH-pipeline) was used to regenerate the concatenated FASTA file, but without the subsequently detected chimeric sequences. The resulting chimera-filtered file was then used to cluster the reads into operational taxonomic units (OTUs) using SWARM v.3.1.0 (Mahé *et al.* 2021; parameters: *-d*13 *-z*). The value for the *d* parameter was chosen based on the results for the mitochondrial cytochrome oxidase subunit I (COI) mini-barcode (Leray *et al.* 2013) from *in silico* experiments performed by Antich *et al.* (2021). The representative sequences of each OTU cluster were then sorted using VSEARCH (parameters: *--fasta_width* 0 *--sortbysize*). An OTU table was constructed from the resulting FASTA file using the VSEARCH utility *--usearch_global* (parameters: *--strand plus --sizein --sizeout --fasta_width* 0).

To reduce the risk of false positives, a cleaning step was employed that excluded read counts in the OTU table constituting <0.01% of the total number of reads in the sample. OTUs were additionally removed from the results based on negative control samples. If the number of reads for the OTU in any sample was less than the maximum for that OTU among negative controls, those reads were excluded from further analysis.

BLAST, reference database construction, and annotation

OTU representative sequences were blasted with the program Megablast (parameters: maximum hits: 1; scoring (match mismatch): 1-2; gap cost (open extend): linear; max E-value: 10; word size: 28; max target seqs 100) against (1) a custom database downloaded from GenBank (a local copy of the NCBI nucleotide database downloaded from ftp://ftp.ncbi.nlm.nih.gov/blast/db/), and (2) a custom database built from data downloaded from BOLD (www.boldsystems.org) (Ratnasingham & Hebert, 2007, 2013) including taxonomy and BIN information. BLAST searches were performed using the GUI software suite Geneious (v.10.2.5 – Biomatters, Auckland, New Zealand).

All available Animalia data was downloaded from the BOLD database on 29 July 2022 using the available public data API (http:// www.boldsystems.org/index.php/resources/api) in a combined TSV file format. The combined TSV file was then filtered to keep only the records that: (1) had a sequence (field 72, “nucleotides”); (2) had a sequence that did not hold exclusively one or more “-” (hyphens); had a sequence that did not contain non-IUPAC characters; (3) belonged to COI (the pattern “COI-5P” in either field 70 (“markercode”) or field 80 (“marker_codes”)); 5) had an available BIN (field 8, “bin_uri”). In (5), an exception was made in cases where the species belonging to that record did not occur with a BIN elsewhere in the dataset. In other words, “BIN-less” records were kept if their species were also completely BIN-less in the dataset.

The dataset was then filtered to include only Romanian and Bulgarian records by keeping only records that where (1) field 55 (“country”) contained the country names "Romania" and "Bulgaria"; or (2) latitude (field 47, “lat”) was between 40.4 and 49.2 and longitude (field 48, “lon”) was between 19 and 30.9. These latitude and longitude values were derived from the extreme north, south, east, and west points of both countries combined, adding a buffer of 100 km in all direction using the “Measure on Map” function of SunEarthTools.com.

Finally, a FASTA file annotated with (1) a Process ID (field 1, “processid”),(2) BIN (field 8), (3) taxonomy (fields 10, 12, 14, 16, 18, 20, 22 - “phylum_name”, “class_name”, “order_name”, “family_name”, “subfamily_name”, “genus_name”, “species_name”), (4)

geolocation data (fields 47, 48, 55), and (5) GenBank ID (field 71, “genbank_accession”) was created from the filtered combined TSV file. This FASTA file was then converted into a BLAST database using Geneious v10.2.6 (Biomatters, Auckland, New Zealand). The results were exported and further processed according to methods described by Uhler *et al.* (2022).

Briefly, the resulting CSV files containing BLAST results were exported from Geneious and combined with the OTU table generated by the bioinformatic pre-processing pipeline. The CSVs included: (1) OTU ID; (2) BOLD Process ID; (3) BIN; (4) Hit-%-ID value (the percentage of identical base pairs of the OTU query sequence with its closest counterpart in the reference database); (5) Grade-%-ID value (a value that combines query coverage, E-value and Hit-%-ID with weights of 0.5, 0.25 and 0.25 respectively); (6) length of the top BLAST hit sequence; (7) phylum, class, order, family, genus and species for each detected OTU.

As an additional measure of control other than BLAST, the OTUs were classified into taxa using the Ribosomal Database Project (RDP) naïve Bayesian classifier (Wang et al., 2007), which was trained on a cleaned COI dataset of Arthropods and Chordates (plus outgroups; see Porter & Hajibabei, 2018). OTUs were also annotated with the taxonomic information from the NCBI (downloaded from https://ftp.ncbi.nlm.nih.gov/pub/taxonomy/), followed by the creation of a taxonomic consensus between BOLD, NCBI and RDP to facilitate assessment of the resulting matches across the three reference databases. To create the taxonomic consensus, we first adjusted the taxonomic depths of each hit from the three reference databases based on its Grade-%-ID value (>97% for species, >95% for genus, >90% for family, >85% for order, >80% for class, and >75% for phylum). In cases where a taxonomically identical match was found in all three reference databases (BOLD BLAST, NCBI BLAST, and RDP classifier), the OTUs were assigned the taxonomic score "A". Where BOLD & NCBI agreed, but RDP disagreed, the OTUs were assigned the score "B". This was in most cases the result of certain taxa either missing or not being represented with sufficient numbers in the RDP classifier's training set. Finally, where NCBI & RDP agreed, but BOLD disagreed, the OTUs got the score "C". A score of "C" commonly occurs in cases where BOLD cannot resolve a species due to a phenomenon commonly referred to as "BIN sharing". For the purposes of constructing the consensus, in every case of a BIN that is shared between 2 or more species in the database, we disregarded the species-level information given by the BOLD BLAST result. In this way, we gave precedence to a species-level annotation with a score of "C" (by means of NCBI and RDP) over a hypothetical genus-level annotation with a score of "A". We treated cases of identifications to different taxonomic levels across the three references in the same way, i.e., a lower score (consensus level) was preferred if it meant an increase to the taxonomic resolution.

BOLD taxonomy was then used to create Krona charts (Fig. 2). These interactive HTML charts were created by means of KronaTools v2.7 (Ondov et al. 2011) (https://github.com/marbl/Krona/wiki/KronaTools). Krona charts are a variation of a sunburst diagram, a pie-chart-like visualization, which is commonly used to plot hierarchical data in a way that emphasizes their taxonomic relationships and relative abundance. A Krona chart shows hierarchy through a series of concentric rings, where each ring corresponds to a level in the hierarchy, and each ring is segmented proportionally to represent read abundance. Where multiple OTUs were identified to the same taxon, read counts were summed over all those OTUs. A set of

charts was created: one for each individual sample, one summed over all samples, as well as one each summed over *M. myotis*- or *M. blythii*-derived samples, respectively. First, a custom script was used to extract from the final Excel results table only the OTU table counts and associated taxonomic annotations. Then, intermediate sample count (.TAX) files for KronaTools were created using a bash script obtained from https://github.com/GenomicaMicrob/OTUsamples2krona. The charts were created by the same script using the command "ktImportText [SAMPLE.TAX] -n SAMPLE -o SAMPLE.html".

Results and Discussion

The presented dataset is a comprehensive collection of 1018 BIN species belonging to 12 Classes and 43 Orders. The interactive Krona charts, based on BOLD taxonomy, provide a useful tool for visualizing the dataset (Fig. 2). These charts present the proportions of the sequence counts of respective species/taxonomic groups relative to all counts. However, it is essential to acknowledge that the sequence counts obtained from the samples do not necessarily correspond to the biomass of the original sample due to inherent methodological limitations and considerable variability in species sizes (Elbrecht et al. 2017, Lamb et al. 2019). Therefore, the relative read abundance (RRA, Deagle et al. 2019) presented in the Krona charts should only be considered a visual guide to the taxonomic diversity and relative abundance of the reads, and not an indication of the actual consumed biomass of the respective species.

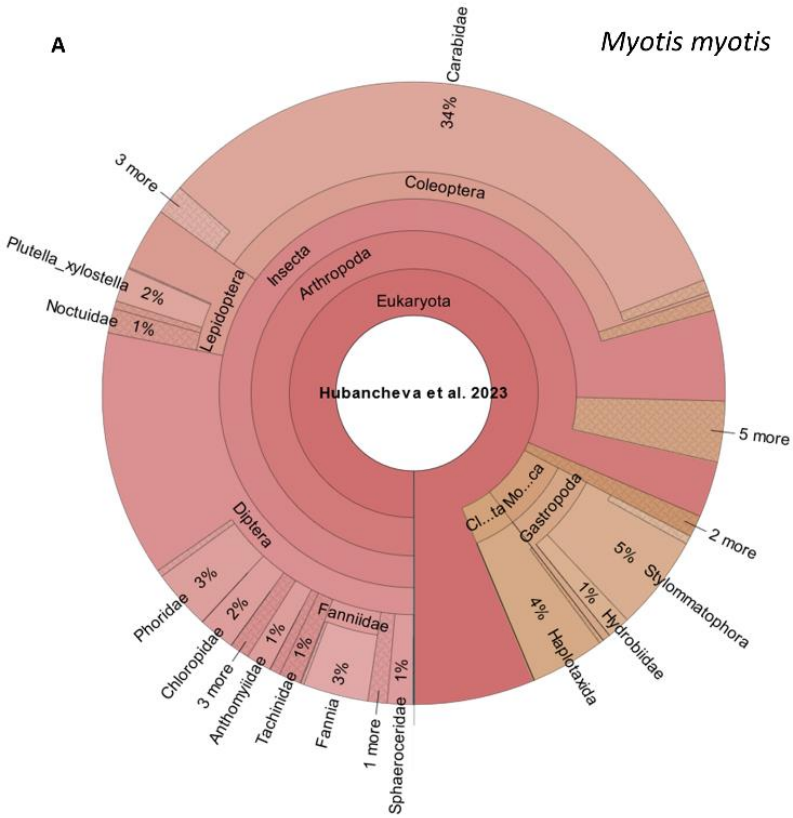
Insects made up the largest proportion of the detected species. The observed differences in the RRA from carabid beetles between *M. myotis* (34%) and *M. blythii (*3%) aligns with the differences in diet, foraging style and habitat of these species (Arlettaz et al. 1997a, Arlettaz et al. 1997b, Arlettaz 1999). However, the differences in the RRA of other taxa shown in the Krona charts were less pronounced between the two bat species. This might be due to the lower resolution power of the BOLD taxonomy for groups such as Orthoptera, which are a significant portion of the diet of *M. blythii* (Arlettaz et al. 1997a, Arlettaz 1999). Nevertheless, the NCBI Genbank and the Ribosomal Database Project Classifier algorithm revealed a diverse range of Orthopteran taxa, including *Tettigonia*, *Phaneroptera*, *Poecilimon*, and *Isophya* species. These findings emphasize the importance of adjusting the analysis approach to the provided dataset according to the research question. Additionally, we recommend complementing the use of this dataset with other research techniques, such as biologging (Stidsholt et al. 2023) and behavioral studies, to increase the reliability and the scope of the results.

Notably, in addition to the bats' prey species, the provided dataset also includes reads from various ecto- and endo-parasites, such as ticks (Ixodes), acarids (Mesostigmata and Sarcoptiformes), roundworms (Strongylida and Rhabditida), and other parasite species. Furthermore, we identified molluscs (Gastropoda) and worms (Annelida) in the samples, including *Pomatias rivulare*, *Lumbricus rubellus*, and *Eisenia fetida*, which were likely consumed by predatory carabid beetles or other arthropods that were then consumed by the bats. Moreover, the presence of species from the roundworm genus *Steinernema*, which are known to parasitize mole crickets and other bat prey, suggests that the dataset also contains parasites of the bats' prey species. This comprehensive dataset thus offers valuable insights into the diversity and abundance of the parasites, the prey and their associated species of the

ARPHA Preprints *Author-formatted, not peer-reviewed document posted on 29/05/2023*. DOI: https://doi.org/10.3897/arphapreprints.e106924

318 greater and the lesser mouse-eared bats.

319



320

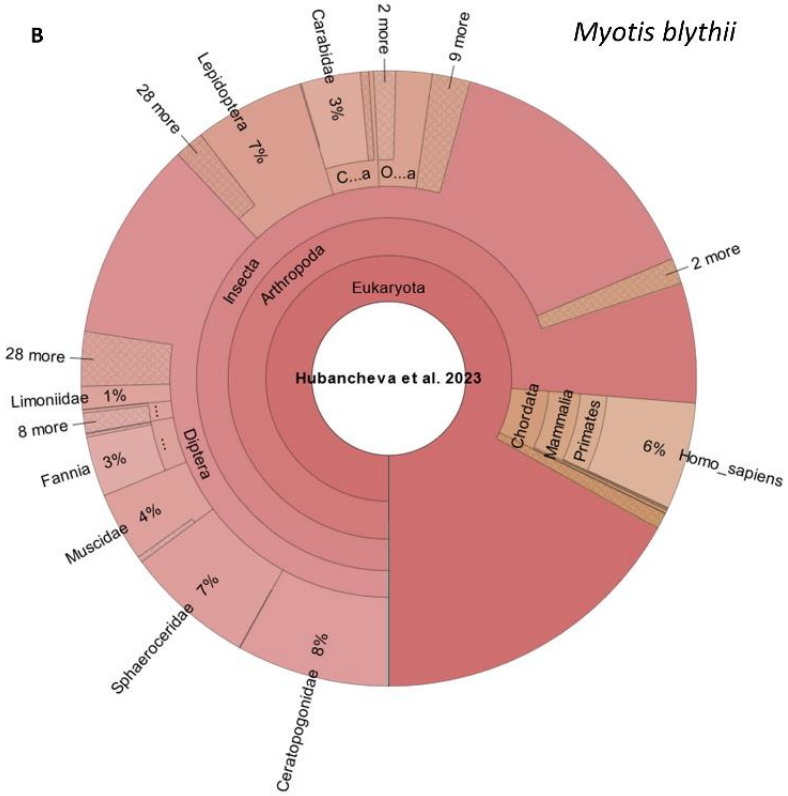


Fig. 2 Taxonomic relationships and relative abundance of prey and parasite species in faecal samples collected from **A)** 60 individuals of the greater mouse-eared bat, *Myotis myotis,* and **B)** 53 individuals of the lesser mouse-eared bat, *Myotis blythii*. The Krona charts presented in this figure exclude the reads from the two bat species. However, Krona charts with included bat reads can be found in the Supplementary Information. An interactive graph is also available in the online version of this publication, offering a more in-depth analysis of the data.

Acknowledgments

We thank Martin Georgiev, Theresa Hügel, Kathrin Dimitrova, and the entire field team of the Siemers Bat Research Station (field seasons 2017 and 2018) for their invaluable assistance with sample collection, methodological, logistical, and mental support. We acknowledge the Max Planck Institute for Ornithology and the National Museum of Natural History, Sofia for infrastructure and support. This research was supported by the Bulgarian Academy of Sciences to A.H (Grant No. DFNP-631 17-71/28.07.2017) and an Emmy Noether grant to H.R.G. by the Deutsche Forschungsgemeinschaft (grant no. 241711556).

Dataset description

The dataset comprises of two files: one Excel file and one HTML file. The Excel file, titled " mmyotis-mblythii-metabarcoding-data-bulgaria.xlsx", includes DNA metabarcoding data from the faecal samples of the lesser (*Myotis blythii*) and greater (*Myotis myotis*) mouse-eared bats from Bulgaria and a second sheet that contains descriptions of the columns in the metabarcoding dataset, along with their meanings. The second file is an interactive Krona chart, which provides a visual representation of the data, and is named "mmyotis-mblythii-

metabarcoding-data-bulgaria-krona.html".

Object name

mmyotis-mblythii-metabarcoding-data-bulgaria.xlsx

mmyotis-mblythii-metabarcoding-data-bulgaria-krona.html

Creation dates

07 April 2023

Dataset creators

Vedran Bozicevic and Antoniya Hubancheva

Dataset contributors

Vedran Bozicevic, Antoniya Hubancheva, Jérôme Morinière

Language

English

License

CC-BY

Repository location

Supplementary information, this paper

Re-use potential

The data could be used for meta-analysis of the diets of insectivorous bats, investigation of predator-prey and host-parasite interactions, interspecific ecological interactions, management of protected areas, etc.

Reference

Alberdi, A., Aizpurua, O., Gilbert, M.T.P. and Bohmann, K., 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution, 9(1), pp.134-147.

Alberdi, A., Razgour, O., Aizpurua, O., Novella-Fernandez, R., Aihartza, J., Budinski, I., Garin, I., Ibáñez, C., Izagirre, E., Rebelo, H. and Russo, D., 2020. DNA metabarcoding and spatial modelling link diet diversification with distribution homogeneity in European bats. Nature

Communications, 11(1), p.1154.

Ando, H., Mukai, H., Komura, T., Dewi, T., Ando, M. and Isagi, Y., 2020. Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. Environmental DNA, 2(4), pp.391-406.

Antich, A., Palacin, C., Wangensteen, O. S., & Turon, X. (2021). To denoise or to cluster, that is not the question: optimizing pipelines for COI metabarcoding and metaphylogeography. BMC bioinformatics, 22, 1-24.

Arlettaz, R., 1999. Habitat selection as a major resource partitioning mechanism between the two sympatric sibling bat species *Myotis myotis* and *Myotis blythii*. Journal of Animal Ecology, 68(3), pp.460-471.

Arlettaz, R., Jones, G. and Racey, P.A., 2001. Effect of acoustic clutter on prey detection by bats. Nature, 414(6865), pp.742-745.

Arlettaz, R., Perrin, N. and Hausser, J., 1997a. Trophic resource partitioning and competition between the two sibling bat species *Myotis myotis* and *Myotis blythii*. Journal of animal ecology, pp.897-911.

Arlettaz, R., Ruedi, M., Ibáñez, C., Palmeirim, J. and Hausser, J., 1997b. A new perspective on the zoogeography of the sibling mouse-eared bat species *Myotis myotis* and *Myotis blythii*: morphological, genetical and ecological evidence. Journal of Zoology, 242(1), pp.45-62.

Audet, D., 1990. Foraging behavior and habitat use by a gleaning bat, *Myotis myotis* (Chiroptera: Vespertilionidae). Journal of Mammalogy, 71(3), pp.420-427.

Beck, A., 1995. Fecal analyses of European bat species. Myotis, 32(33), pp.109-119.

Borissov I. M. 2010 Проучване на прилепите (Mammalia: Chiroptera) по долината на река Русенски Лом и притоците ѝ. Значение на подземните местообитания за прилепното съобщество. Софийски университет, София

Cimatti, M., Brooks, T.M. and Di Marco, M., 2021. Identifying science-policy consensus regions of high biodiversity value and institutional recognition. Global Ecology and Conservation, 32, p.e01938.

Coroiu, I., Juste, J. & Paunović, M. 2016. *Myotis myotis*. The IUCN Red List of Threatened Species 2016: e.T14133A22051759. https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T14133A22051759.en.

Deagle, B.E., Thomas, A.C., McInnes, J.C., Clarke, L.J., Vesterinen, E.J., Clare, E.L., Kartzinel, T.R. and Eveson, J.P., 2019. Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? Molecular ecology, 28(2), pp.391-406.

Dietz, C. and von Helversen, O., 2004. Illustrated identification key to the bats of Europe.

Edgar, R. (2010). Usearch. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Egert-Berg, K., Hurme, E.R., Greif, S., Goldstein, A., Harten, L., Flores-Martínez, J.J., Valdés, A.T., Johnston, D.S., Eitan, O., Borissov, I. and Shipley, J.R., 2018. Resource ephemerality drives social foraging in bats. Current Biology, 28(22), pp.3667-3673.

Elbrecht, V., Peinert, B. and Leese, F., 2017. Sorting things out: Assessing effects of unequal specimen biomass on DNA metabarcoding. Ecology and evolution, 7(17), pp.6918-6926.

Galan, M., Pons, J.B., Tournayre, O., Pierre, E., Leuchtmann, M., Pontier, D. and Charbonnel, N., 2018. Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis. Molecular ecology resources, 18(3), pp.474-489.

Golemanski, V. and Peev, D.R. eds., 2015. Red Book of the Republic Bulgaria. Volume 2. Animals. BAS & Ministry of Education, Sofia

Graclik, A. and Wasielewski, O., 2012. Diet composition of *Myotis myotis* (Chiroptera, Vespertilionidae) in western Poland: results of fecal analyses. Turkish Journal of Zoology, 36(2), pp.209-213.

Jones, P.L., Page, R.A., Hartbauer, M. and Siemers, B.M., 2011. Behavioral evidence for eavesdropping on prey song in two Palearctic sibling bat species. Behavioral Ecology and Sociobiology, 65, pp.333-340.

Kunz, T.H., Braun de Torrez, E., Bauer, D., Lobova, T. and Fleming, T.H., 2011. Ecosystem services provided by bats. Annals of the New York academy of sciences, 1223(1), pp.1-38.

Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer, S., Davies, R.G. and Taylor, M.I., 2019. How quantitative is metabarcoding: A meta-analytical approach. Molecular ecology, 28(2), pp.420-430.

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T. and Machida, R.J., 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in zoology, 10(1), pp.1-14.

Mata, V.A., da Silva, L.P., Veríssimo, J., Horta, P., Raposeira, H., McCracken, G.F., Rebelo, H. and Beja, P., 2021. Combining DNA metabarcoding and ecological modelling to assess the diet of the greater mouse-eared bat. Molecular Ecology, 30(9), pp.2185-2201.

Mahé, F., Czech, L., Stamatakis, A., Quince, C., de Vargas, C., Dunthorn, M., & Rognes, T. (2022). Swarm v3: towards tera-scale amplicon clustering. Bioinformatics, 38(1), 267-269.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal, 17(1), 10-12.

Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. PeerJ, 4, e2584.

Ondov, B. D., Bergman, N. H., & Phillippy, A. M. (2011). Interactive metagenomic visualization in a Web browser. BMC bioinformatics, 12(1), 1-10.

O'Rourke, D.R., Bokulich, N.A., Jusino, M.A., MacManes, M.D. and Foster, J.T., 2020. A total crapshoot? Evaluating bioinformatic decisions in animal diet metabarcoding analyses. Ecology and Evolution, 10(18), pp.9721-9739.

Porter, T. M., & Hajibabaei, M. (2018). Automated high throughput animal CO1 metabarcode classification. Scientific Reports, 8(1), 4226.

Juste, J. & Paunović, M. 2016. *Myotis blythii*. The IUCN Red List of Threatened Species 2016: e.T14124A22053297 https://dx.doi.org/10.2305/IUCN.UK.2016-2.RLTS.T14124A22053297.en.

Pereira, M.J.R., Rebelo, H., Rainho, A. and Palmeirim, J.M., 2002. Prey selection by *Myotis myotis* (Vespertilionidae) in a Mediterranean region. Acta Chiropterologica, 4(2), pp.183-193.

Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System (http://www. barcodinglife. org). Molecular ecology notes, 7(3), 355-364.

Ratnasingham, S., & Hebert, P. D. (2013). A DNA-based registry for all animal species: the Barcode Index Number (BIN) system. PloS one, 8(7), e66213.

Ruedi, M., 2020. Lesser Mouse-Eared Bat Myotis blythii (Tomes, 1857). Handbook of the Mammals of Europe, pp.1-25.

Russo, D., Salinas-Ramos, V.B., Cistrone, L., Smeraldo, S., Bosso, L. and Ancillotto, L., 2021. Do we need to use bats as bioindicators? Biology, 10(8), p.693.

Siemers, B.M. and Güttinger, R., 2006. Prey conspicuousness can explain apparent prey selectivity. Current biology, 16(5), pp. R157-R159.

Simmons, N.B., 2005. An Eocene big bang for bats. Science, 307(5709), pp.527-528.

Stidsholt L., Hubancheva A., Greif S., Goerlitz H. R., Johnson M., Yovel Y., Madsen P.T., 2023. Echolocating bats prefer a high risk-high gain foraging strategy to increase prey profitability, eLife, 12:e84190.

Uhler, J., Haase, P., Hoffmann, L., Hothorn, T., Schmidl, J., Stoll, S., Welti, E.A., Buse, J. and Müller, J., 2022. A comparison of different Malaise trap types. Insect Conservation and Diversity, 15(6), pp.666-672.

Wang, Q., Garrity, G. M., Tiedje, J. M., & Cole, J. R. (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and environmental microbiology, 73(16), 5261-5267.

Whitaker, J. O., McCracken, G. F. and Siemers, B. M. 2009 Foog habits analysis of insectivorous bats. In: Kunz, TH, and S. Parsons (eds.) Ecological and Behavioral Methods for the Study of Bats. Johns Hopkins University Press, Baltimore, Maryland, 901 pp. ISBN 978-0-8018-9147-2

Zahn, A., Rottenwallner, A. and Güttinger, R., 2006. Population density of the greater mouse-

539 eared bat (*Myotis myotis*), local diet composition and availability of foraging habitats. Journal of
540 Zoology, 269(4), pp.486-493.