SEASONAL SHIFT IN THE DIET OF THE NOTCHED-EARED BAT (*MYOTIS EMARGINATUS*): FROM FLIES TO SPIDERS

MAMMALIAN BIOLOGY

Nerea Vallejo[1], Joxerra Aihartza[1], Lander Olasagasti[1], Miren Aldasoro[1], Urtzi Goiti[1], Inazio Garin[1]

[1] Department of Zoology and Animal Cell Biology, Faculty of Science and Technology, University of the Basque Country UPV/EHU, Sarriena z.g., Leioa, The Basque Country

Corresponding author: Nerea Vallejo (nerea.vallejo@ehu.eus)

ORCID: 0000-0002-1478-1536

ONLINE RESOURCE 1

DNA Extraction

For each sample, 1-6 pellets were used for each sample (approximate weight 30-60 mg). Each sample was weighted, and the number of pellets were counted prior to extraction. We used DNeasy PowerSoil Kit and DNeasy PowerSoil Pro Kit (Qiagen). We had to change extraction kits because DNeasy Power Soil Kit was discontinued while DNA extractions were ongoing. Each extraction round included 23 samples and one negative extraction control. We followed manufacturer's protocol but did some modifications. We tried to apply similar modifications to the new DNeasy PowerSoil Pro Kit that had previously been successful for the old DNeasy PowerSoil Kit. All extractions were performed inside a biosafety cabinet.

The exact protocols are detailed below:

<u>For DNeasy PowerSoil</u>

1. Add each sample to the PowerBead Tubes containing beads and bead solution.

2. Add 60 µl of C1 solution to each tube and vortex briefly.
3. Incubate the tubes at 65 °C for 15 minutes.
4. Homogenize each sample at 6500 rpm for 2x20 seconds on a Precellys 24 Tissue Homogenizer (Bertin technologies).
5. Centrifuge the tubes at 13000g for 3 minutes.
6. Transfer the supernatant (400-500 µl, pipetted two times) to a new 1.5 ml tube.
7. Add 250 µl of C2 solution and vortex briefly.
8. Incubate at 4 °C for 10 minutes.
9. Centrifuge the tubes at 13000g for 1 minute.
10. Transfer the supernatant (600 µl) to a new 1.5 ml tube.
11. Add 200 µl of C3 solution and vortex briefly.
12. Incubate at 4 °C for 10 minutes.
13. Centrifuge the tubes at 13000g for 1 minute.
14. Transfer the supernatant (750 µl) to a new 2 ml tube.
15. Add 1100 µl of C4 solution and voltex briefly.
16. Load 650 µl of the mix to onto a spin column, gently and letting the liquid run through the walls of the tube.
17. Centrifule at 8000 g for 1 minute.
18. Discard the flowthrough and repeat steps 16 and 17 until all the mix is finished.
19. Add 500 µl of C5 solution to the spin column, gently and letting the liquid run through the walls of the tube.
20. Centrifuge at 13000g for 1 minute.
21. Discard the flowthrough and centrifuge again at 13000g for 1 minute.
22. Place the spin filter into a new 2ml collection tube and add 50 µl of C6.
23. Incubate the tubes for 15 minutes at 37 °C.
24. Centrifuge at 13000g for 1 minute and transfer the flowthrough to 1.5 ml safe-lock Eppendorf tubes.

25. Freeze the extracted DNA immediately.

For DNeasy PowerSoil Pro

Before starting:

- Preheat solutions CD1 and CD3 at 60 °C.
- Solution CD2 should be stored at 2-8 °C.

Procedure

1. Add each sample to the PowerBead Tubes containing beads.
2. Add 800 µl of C1 solution to each tube and vortex briefly.
3. Incubate the tubes at 65 °C for 15 minutes.
4. Homogenize each sample at 6500 rpm for 2x20 seconds on a Precellys 24 Tissue Homogenizer (Bertin technologies).
5. Centrifuge the tubes at 13000g for 3 minutes.
6. Transfer the supernatant (500 - 600 µl, pipetted two times) to a new 1.5 ml tube.
7. Add 200 µl of C2 solution and vortex briefly.
8. Centrifuge the tubes at 15000g for 1 minute.
9. Transfer the supernatant (700 µl) to a new 2 ml tube.
10. Add 600 µl of C3 solution and vortex briefly.
11. Load 650 µl of the mix to onto a spin column, gently and letting the liquid run through the walls of the tube.
12. Centrifuge at 13000 g for 1 minute.
13. Discard the flowthrough and repeat steps 11 and 12 until all the mix is finished.
14. Add 500 µl of EA solution to the spin column, gently and letting the liquid run through the walls of the tube.
15. Centrifuge at 13000g for 1 minute.
16. Discard the flowthrough, place the spin column into a new 1.5 ml tube and centrifuge again at 13000g for 1 minute.

17. Place the spin filter into a new 1.5 ml collection tube and add 50 µl of C6.
18. Incubate the tubes for 15 minutes at 37 °C.
19. Centrifuge at 13000g for 1 minute and transfer the flowthrough to 1.5 ml safe-lock Eppendorf tubes.
20. Freeze the extracted DNA immediately.

PCR amplification

We used a single primer set, FWH1 (Vamos et al. 2017) to amplify DNA from both potential prey items and bats. For the amplification process we used 10 µl of Qiagen multiplex PCR kit, 1 µl forward primer (10 µM), 1 µl reverse primer (10 µM), 7 µl H_2O and 1 µl of DNA. Cycling conditions started with 15 min at 95 °C for polymerase activation, followed by 40 cycles of 94 °C (30 s) - 45 °C (45 s) - 72 °C (2 m), and finished with 10 minutes at 72 °C. Amplification products were then purified using magnetic beads.

Library building and sequencing

Index sequences and Illumina adapters were attached using Nextera XT v2 sets A, and C following official Illumina protocols. A different combination of forward and reverse markers was used in each sample to allow their differentiation later on. Amplification success was checked by migrating the product in an agarose gel before purifying again and pooling all the samples at equal molarities for sequencing. Sequencing was performed in Illumina MiSeq, in four separate runs, together with more *M. emarginatus* samples from other colonies not included in this study. We used a 500 cycle v2 run kit.

Bioiformatical analysis and OTU building

Bioinformatic analyses were done with vsearch (Rognes et al. 2016) and Cutadapt (Martin 2011). The sequences from the four runs were analysed together. The steps are detailed below.

1. We merged paired-end reads using vsearch *-fastq_mergepairs* and the following options: *-fastq_truncquall 30 -fastq_minovlen 50 -fastq_maxdiffs 10 -fastqout*
2. We trimmed the primers in two steps using Cutadapt funcionts *-g* and *-a*. We specified that the primer sequence should be at the beginning or end of the sequence.
3. We selected sequences by length using Cutadapt functions *-m* and *-M*. As our amplicon has a length of 179 base pairs, sequence length was limited to 169-189.
4. All files were renamed according to their sample identifier using a custom python script. Then, every sequence inside each file was renamed using *vsearch -sortbysize* to utilise the option *-relabel*. Every sequence was uniquely relabelled according to their sample identifier, followed by a number
5. The fastq format files were pooled together into the same file.
6. Sequences were quality filtered using *vsearch -fastx_filter* with options *-fastq_maxee 1 --fastq_qmax 42*. All files were converted to fasta format with option *-fastaout*.
7. Singletons were removed, and all sequences were collapsed into uniques using *vsearch -derep_fulllength*, with options *-minuniquesize 2*. Information on the size of the sequences was added to sequence names using options *-sizein -sizeout.* Additionally, option *-uc* was used to write a file indexing every sequence to its unique representative.
8. The resulting fasta file was used to build OTUs at the 97% identity using *vsearch -cluster_size* with options *-id 0.97 -relabel -sizeout -sizein*. OTU centroids were outputted using option *-centroids*.
9. Chimeras were removed using *vsearch -uchime3_denovo*. The resulting fasta was considered the valid OTU list.
10. The list of sequences obtained in step 5 was mapped against the OTUs at the 97% identity, using *vsearch -usearch_global -id 0.97*. The output is a uc file indicating what

OTU each sequence is paired with. A custom script in R is then used to conver this into a sample-by-OTU table.

Taxonomic assignment of OTUs

1. Every OTU was assigned to their appropriate taxonomy using *blastn* function in BLAST+ (Camacho et al. 2009) to access the GenBank dataset, and Boldigger-cline (Buchner and Leese 2020) to access the BOLD dataset.

   For *blastn* we used options *-db nt -evalue 1e-20 -perc_identity 98 -remote* to remotely compare our sequences to the GenBank nucleotide database. Only matches over 98% similarity and e-value lower than 1e-20 were considered. Boldigger does not allow to filter matches by identity at this stage.
2. The results were curated using a custom script in R. We used package *taxizedb* (Chamberlain and Arendsee 2021) to add taxonomic information to each identification. Duplicate identifications for each OTU were collapsed into one, and the maximum identity value was assigned. Results from BoldSystems and GenBank were collapsed together.
3. The final curation was done manually, to assure that each OTU was assigned to one taxon only. If two species belonging to the same genus, family, etc. matched with the same OTU, the lowest common taxonomic level was chosen. The distribution of each species was checked in Fauna Europaea (de Jong et al. 2014) to establish whether they are present in the study area. Each assigned taxon was classified into one of the following categories: predator, prey, not in study area, environmental contamination.

Selection of samples and OTUs

Extraction and library blanks were examined to account for pervasive potential contamination events regarding potential prey items. We evaluated taxa whose ratio of

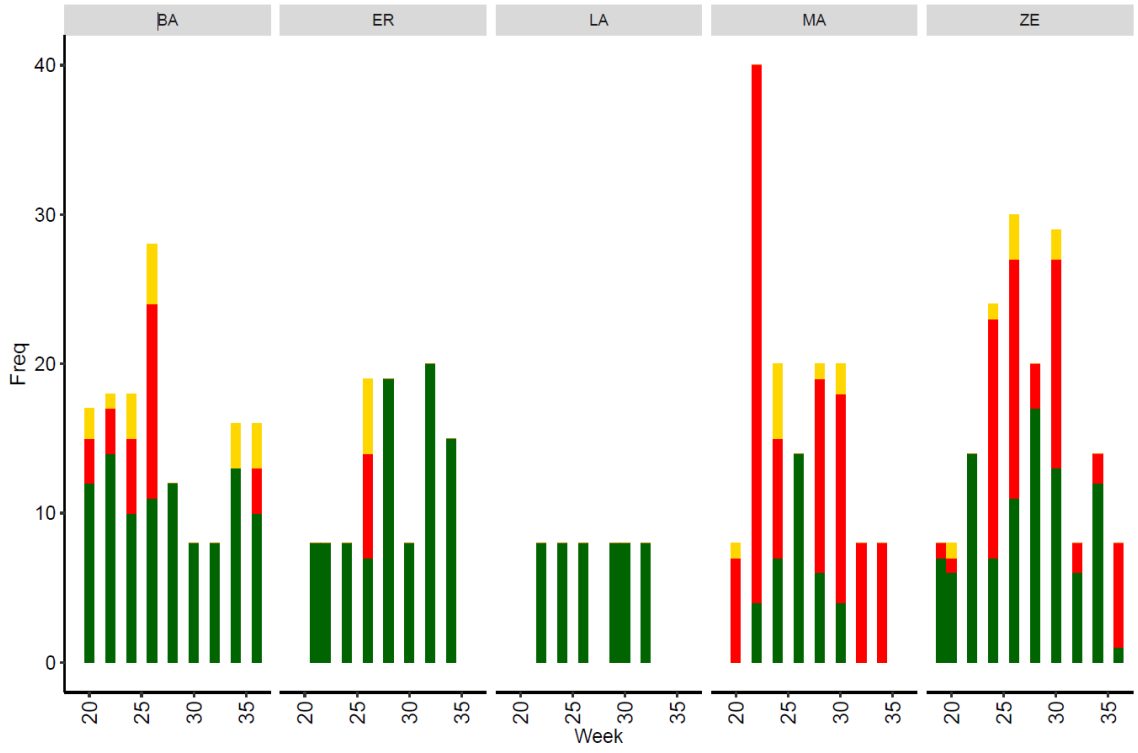
maximum read counts between blanks and samples was higher than 0.1. We found that these taxa could be classified according to:

1) Taxa that represent very common prey items, or prey items previously reported in the diet of *Myotis emarginatus*. They have high read counts in many samples, and one (or more) extraction blank. The appearance of the taxa in the extraction blank is most likely linked to tag jumps or punctual transfer from a sample to the blank, but not the other way around. We believe that these taxa are genuine prey items and removing these taxa from the analysis would seriously change the interpretation of the data.
2) Taxa that represent uncommon prey items. They appear in a single extraction blank and a few samples. However, the samples do not belong to the same batch as the extraction blank, so it cannot be attributed to a contamination event.

Therefore, due to the overall lack of correlation between reads appearing in blanks and samples from the same extraction blanks, and in fear of eliminating actual prey items with ecological significance to the study, no taxa were removed from the analysis at this point. Nonetheless, an abundance threshold was applied, so that in each sample, OTUs with less than 0.5% of reads were removed. While far from perfect, this method has proven to be quite effective in limiting contamination risk of multiple sources without eliminating too many rare prey taxa (Drake et al. 2021); therefore, the data was interpreted having this decision in mind.

Initially, between 8 and 10 samples were chosen from each combination of colony and date to be sequenced. We checked for prey specie coverage each time, and we aimed to reach 75% of the diet diversity per colony and day (Chao et al. 2014). We performed inter- and extrapolation of dietary richness using package iNEXT (Hsieh, Ma & Chao, 2020), and improved sampling effort in the colonies where it was necessary. To ensure that all samples considered in the analysis belonged to the target species, samples were discarded from the analysis if out of all the reads identified as bat species, more than 10% belonged to species other than *M. emarginatus*. Finally, from each combination of colony and date, between 8

and 40 samples were sequenced (SFig. 1). We managed to obtain at least eight to 12 samples eligible for analysis in most situations. Sampling size was greater wherever *M. emarginatus* shared roost with other bats (BA, MA and ZE) and when prey species coverage was well below 75% after a preliminary evaluation. Nevertheless, we could not retrieve DNA from *M. emarginatus* at samplings of MA in weeks 20, 32 and 34, and of ZE in week 36 and consequently removed these sample groups completely from the analysis.



Supplementary Figure 1 Frequency of samples processes in each colony and date (as indicated by the week number). Colour refers to sample quality in regard to amount of DNA corresponding to *Myotis emarginatus*: Green > 90%; Yellow 75-90%, Red < 75%. Only Green samples were chosen for the diet analysis

References

Buchner D, Leese F (2020) BOLDigger – a Python package to identify and organise sequences with the Barcode of Life Data systems. *Metabarcoding Metagenom.* 4:e53535. https://doi.org/10.3897/mbmg.4.53535

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

Chao A, Gotelli NJ, Hsieh TC, Sander EL, Ma KH, Colwell RK, Ellison AM (2014) Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monog*. 84(1):45-67. https://doi.org/10.1890/13-0133.1

Chamberlain S, Arendsee Z (2021) taxizedb: Tools for Working with Taxonomic' Databases. R package version 0.3.0. https://CRAN.R-project.org/package=taxizedb

Drake LE, Cuff JP, Young RE, Marchbank A, Chadwick EA, Symondson WOC (2022) An assessment of minimum sequence copy thresholds for identifying and reducing the prevalence of artefacts in dietary metabarcoding data. Methods Ecol. Evol. 13(3):694–710. https://doi.org/10.1111/2041-210X.13780

Hsieh TC, Ma KH, Chao A (2020) iNEXT: iNterpolation and EXTrapolation for species diversity. R package version 2.0.20 URL: http://chao.stat.nthu.edu.tw/wordpress/software-download/

de Jong Y, Verbeek M, Michelsen V, Bjørn PP, Los W, Steeman F, Bailly N, Basire C, Chylarecki P, Stloukal E, Hagedorn G, Wetzel FT, Glöckler F, Kroupa A, Korb G, Hoffmann A, Häuser C, Kohlbecker A, Müller A, Güntsch A, Stoev P, Oenev L (2014) Fauna Europaea - all European animal species on the web. *Biodivers. Data J.* 2:e4034. https://doi.org/10.3897/BDJ.2.e4034

Martin M (2011) Cutadapt removes adapter sequences from high-throughtput sequencing reads. *EMBnet J.* 17(1):10–12. https://doi.org/10.14806/ej.17.1.200

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. 4:e2584. https://doi.org/10.7717/peerj.2584

Vamos EE, Elbrecht V, Leese F (2017) Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding Metagenom.* 1:e14625. https://doi.org/10.3897/mbmg.1.14625