Appendix A for *Spatio-temporal patterns in arctic fox (*Vulpes alopex*) diets revealed by molecular analysis of scats from Northeast Greenland*.

Niels Martin Schmidt, Tomas Roslin, Lars Holst Hansen, Olivier Gilg, Johannes Lang, Benoit Sittler, Jannik Hansen, Loïc Bollache and Eero Vesterinen

Molecular analyses

DNA was extracted from 380 (4x95) Arctic fox scat samples using ZR 96 Kits (cat nr D6011, Zymo Research Corp., Irvine, CA, USA) following the manual provided by the manufacturer. A small piece of each scat was dissected and placed onto a disruption plate well. Based on individual weighing of a subset of single samples, the average weight of a single Arctic Fox scat sample in the DNA extraction was 94.13 mg (±19.53). Each plate included one extraction control. To amplify arctic fox vertebrate prey, we chose three primer pairs on two mitochondrial genes, of which two were used with a blocking primer. We designed two blocking primers for the current study, and added a *C3 spacer* modification to the 3'-end to decrease predator amplification in favour of the prey amplification (Vestheim and Jarman, 2008).

Mammal primers

The first primer pair "Fox-Mam" C1-J-1709 (5'-AATTGGWGGWTTYGGAAAYTG-3') and C1-N-1843d (5'-GMWARWGGWGGRTAWACWGTTCA-3') targets the mammalian mtDNA *cytochrome oxidase subunit I* (Simon et al. 2006), and was here used with a blocking primer, VulaR-blk-C3 (5'-ATA CAC GGT TCA CCC TGT TCC CGC GCC C-**C3 spacer** -3').

Bird primers

The second primer pair "Fox-Bird" CytbS-Bird-(5'-AATGGGATTTTGTCGCAGTC-3') and CytbS-Bird-R 5'-TCTCAGCCATCCCCTACATC-3') targets the bird mitochondrial *cytochrome b* gene (Pastor-Beviá et al., 2014).

Vertebrate primers

The third primer pair "COI-prey" was designed specifically for this study. We used all available full mitochondrial genomes for birds in the study area, and chose the candidate primers. We modified a forward primer COIPrey-FW from Pastor-Beviá et al. (2014) and designed a matching pair for that using Geneious R6 (Kearse et al., 2012). We then downloaded sequences of all the vertebrates regularly observed in Greenland, and removed duplicate sequences. We tested the new primer pair for all of these sequences and designed a blocking primer to bind solely on the arctic fox DNA. After testing, we chose the third primer set to be COIPrey-FW3 (5'-CGYGCAGARCTAGGCCARCC-3') and Bird_COI_303R2 (5'-CGTGGRAATGCTATGTCNGG-3'), targeting vertebrate prey mtDNA and used with a blocking primer Fox-Blk-R (5'-TGCTATGTCAGGAGCACCAATTATTAAGGGA-**C3 spacer** -3').

Primer-specific PCR conditions

To find the optimal PCR setup in the presence of possible inhibitors, the blocking primers and prey-specific primers were tested on gradient PCR's (with varying annealing temperature) and with various PCR enzymes. We also tested different blocker:primer ratios (0, 1:1, 1:10, and 1:20). We used a wolf scat DNA sample as "positive" control sample in our test protocols, as this sample had been successfully used in population genetic studies previously (see e.g. Harmoinen et al., 2021). Some of the tested chemicals were provided with PCR additives and enhancers, which were each tested in parallel. We also tried to increase PCR success by adding a common PCR stabiliser, BSA (bovine serum albumin). The final PCR and library preparation followed Vesterinen et al. (2018) and Vesterinen et al. (2016) with the following modifications. Fox-Mam PCR included 5 µL MyTaq Red Mix (2x), 200nM forward primer (C1-J-1709), 200nM reverse primer (C1-N-1843d), 2000nM blocking primer (VulaR-blk-C3), and 4 µL purified arctic fox DNA in a total 10 µL volume. PCR cycling profile for Fox-Mam started with initial denaturation at 95 °C for 5 min, then 40 cycles of denaturation for 30s at 95 °C, annealing for 30s at 55 °C, and extension for 60s at 72 °C, ending with final extension for 5 min at 72 °C. Fox-Bird PCR consisted of 5 µL MyTaq Red Mix (2x), 200nM forward primer (CytbS-Bird-Fm), 200nM reverse primer (CytbS-BirdR), 0.6 µL ultrapure water, and 4 µL purified arctic fox DNA in a total 10 µL volume. PCR cycling profile for Fox-Bird began with initial denaturation at 95 °C for 5 min, then 40 cycles of denaturation for 30s at 95 °C, annealing for 30s at 52 °C, and extension for 60s at 72 °C, ending with final extension for 5 min at 72 °C.

Library preparation PCR

All PCR reactions (Fox-Mam and Fox-Bird) were purified by adding 0.25 µL EXO, 0.25 µL SAP, and 3.0 µL of ultrapure water into each sample. The mixture was warmed at 37 °C for 10 min (enzyme activation), and then heated up to 85 °C for 5 min (enzyme destruction). The library PCR followed immediately and consisted of 6.25 µL KAPA HiFi HotStart ReadyMix (product nr KK2602, Roche Ltd, KAPA Biosystems, Wilmington, Massachusetts, USA), 300nm forward indexed i5-adapter, 300nm forward indexed i7-adapter, and 1.75 µL of DNA from the first PCR in a total volume of 12.5 µL. Fox-Bird PCR consisted of 5 µL MyTaq Red Mix (2x), 200nM forward primer (CytbS-Bird-Fm), 200nM reverse primer (CytbS-BirdR), 0.6 µL ultrapure water, and 4 µL purified arctic fox DNA in a total 10 µL volume. PCR cycling profile for Fox-Bird began with initial denaturation at 95 °C for 5 min, then 40 cycles of denaturation for 30s at 95 °C, annealing for 30s at 52 °C, and extension for 60s at 72 °C, ending with final extension for 5 min at 72 °C. All PCR reactions (Fox-Mam and Fox-Bird) were purified by adding 0.25 µL EXO, 0.25 µL SAP, and 3.0 µL of ultrapure water into each sample. The mixture was warmed at 37 °C for 10 min (enzyme activation), and then heated up to 85 °C for 5 min (enzyme destruction). The library PCR followed immediately and consisted of 6.25 µL KAPA HiFi HotStart ReadyMix, 300nm forward indexed i5-adapter, 300nm forward indexed i7-adapter, and 1.75 µL of DNA from the first PCR in a total volume of 12.5 µL. COI-Prey PCR was carried out in two parallel replicates for each sample, and consisted of 5 µL MyTaq Red Mix (2x), 200nM forward primer (COIPrey-FW3), 200nM reverse primer (Bird_COI_303R2), 2000nM blocking primer (Fox-Blk-R), and 4 µL purified arctic fox DNA in a total 10 µL volume. PCR cycling profile for COI-Prey began with initial denaturation at 95 °C for 5 min, then 40 cycles of denaturation for 30s at 95 °C, annealing for 30s at 52 °C, and extension for 60s at 72 °C, ending with final extension for 5 min at 72 °C. The library PCR followed immediately after the first PCR and consisted of 6.25 µL KAPA HiFi HotStart ReadyMix, 300nm forward indexed i5-adapter, 300nm forward indexed i7-adapter, and 1.75 µL of DNA from the first PCR in a total volume of 12.5 µL. For all samples, the library PCR profile

was: initial denaturation at 98 °C for 4 min, then 15 cycles of denaturation for 20s at 98 °C, annealing for 15s at 60 °C, and extension for 30s at 72 °C, ending with final extension for 3 min at 72 °C.

Pooling, purification, and sequencing

After indexing, 2.6 μl of each indexed sample was combined in four separate subset pools (Fox-Mam, Fox-Bird, COI-Prey1, and COI-Prey2) to maximize the diversity of the input template and to allow for scaling the input for each loci/replicate. These pools were purified using dual-SPRI method by Vesterinen et al. (2016) to remove primer-dimers, non-specific PCR products, and any impurities. The purified Fox-Mam and Fox-Bird pools were combined in equimolar ratios, and equally COI-Prey1 and COI-Prey2 were pooled together. Sequencing was performed on two distinct runs. First run included Fox-Mam+Fox-Bird pool (40% of the input DNA), and another arthropod-specific pool (60% of the input DNA), and was sequenced Macrogen Inc. (Seoul, South Korea) on one lane of Illumina HiSeq4000 flow cell using 2x150bp run length with TruSeq 3000 4000 SBS Kit v3 sequencing kit following HiSeq 3000 4000 System User Guide (Document # 15066496 v04 HCS 3.3.52). The second run included the COI-Prey1+2 pool on Illumina NextSeq500 platform at the Functional Genomics Unit of the University of Helsinki, Finland, using Mid Output with 300 cycles and 2*150 bp paired-end read length.

Bioinformatics for raw sequence data

The raw sequence output was uploaded to CSC servers (IT Center for Science, www.csc.fi) for trimming and further analysis. The HiSeq4000 run yielded 265,871,470 paired-end reads. These reads were assigned for samples, merged (*minlen=50, maxdiffs=50*), quality-filtered (*fastq_maxee=1*) using 64-bit *vsearch* version v2.14.2 on LINUX (Rognes et al., 2016), and trimmed for primers with cutadapt (minlen=100, errors=0.2%; Martin, 2011). Then, trimmed reads were dereplicated (*minuniquesize=2*), filtered for chimeras, clustered into OTUs (95% identity), and mapped back to the trimmed reads to construct an OTU table using *vsearch*. The NextSeq500 run yielded 31,899,334 paired-end reads. These reads were assigned for samples, merged (*minlen=50, maxdiffs=50, minmergelen=120*), quality-filtered (*fastq_maxee=1*) using 64-bit *vsearch* version v2.14.2 on LINUX (Rognes et al., 2016), and trimmed for primers with cutadapt (minlen=120, errors=0.2%; Martin 2011). Then, trimmed reads were dereplicated (*minuniquesize=2*), denoised into ZOTUs using 32-bit *usearch* v10 (Edgar, 2016), and mapped back to the trimmed reads to construct a zotutable using *vsearch*.

We identified 306 OTUs for Fox-Mam in 500,644 mapped reads in the OTU table, 526 OTUs for Fox-Bird in 8,419,225 mapped reads for OTU table, and 6392 ZOTUs for COI-Prey in 10,001,546 mapped reads. The prey identification criteria vary for different loci, and thus the prey assignation was initially done independently for each dataset, before merging the data. The OTUs from COI region were assigned to taxa 1) based on Barcode of Life Data systems API engine (Ratnasingham and Hebert, 2007), using custom-made scripts from Vesterinen et al. (2020), and 2) using BLAST against GenBank nt database (Altschul et al., 1990) and then importing results to software Megan for lowest common ancestor (LCA) analysis (Huson et al., 2007). CytB reads were assigned solely by BLAST plus LCA approach. The final dataset consisted of 500,644 reads in the Fox-Mam OTU table, 8,419,225 reads for Fox-Bird OTU table, and 11,955,842 reads for COI-Prey.

In the subsequent filtering for tag jumping (Illumina sample cross-talk) using *dplyr* (Wickham et al. 2020) in R, all OTUs were removed if the number of reads assigned for a sample pre OTU was lower than the maximum reads assigned to negative control sample. The whole OTUs were not removed even

though the same OTU was found in negative samples, since the most abundant OTUs in the whole data tend to "leak" into other samples. After this, the OTU was removed from the sample, if only one of the two replicates produced reads for that specific sample (COI-Prey dataset only). We then combined all the datasets into one, by summing all the assigned taxa per sample. We compared the results from each primer dataset, even though we did not expect similar results from different primers due to varying targets (Figure S1). We only retained matches for the target taxa: Birds, mammals, and fishes. We removed all the reads assigned to Canidae (Artic fox, potentially wolf *Canis lupus*), and our final dataset consisted of 3,361,848 prey reads assigned to prey taxa. For each site, the number of reads were divided as follows: ZAC: 2,193,246 reads (162 samples with data), KVP: 711,292 reads (43 samples with data), and HOC: 415,335 reads (32 samples with data). We were able to retrieve prey data from 303 out of the original 380 (~80%) samples. The rest of the samples only contained bacterial or other non-target reads.

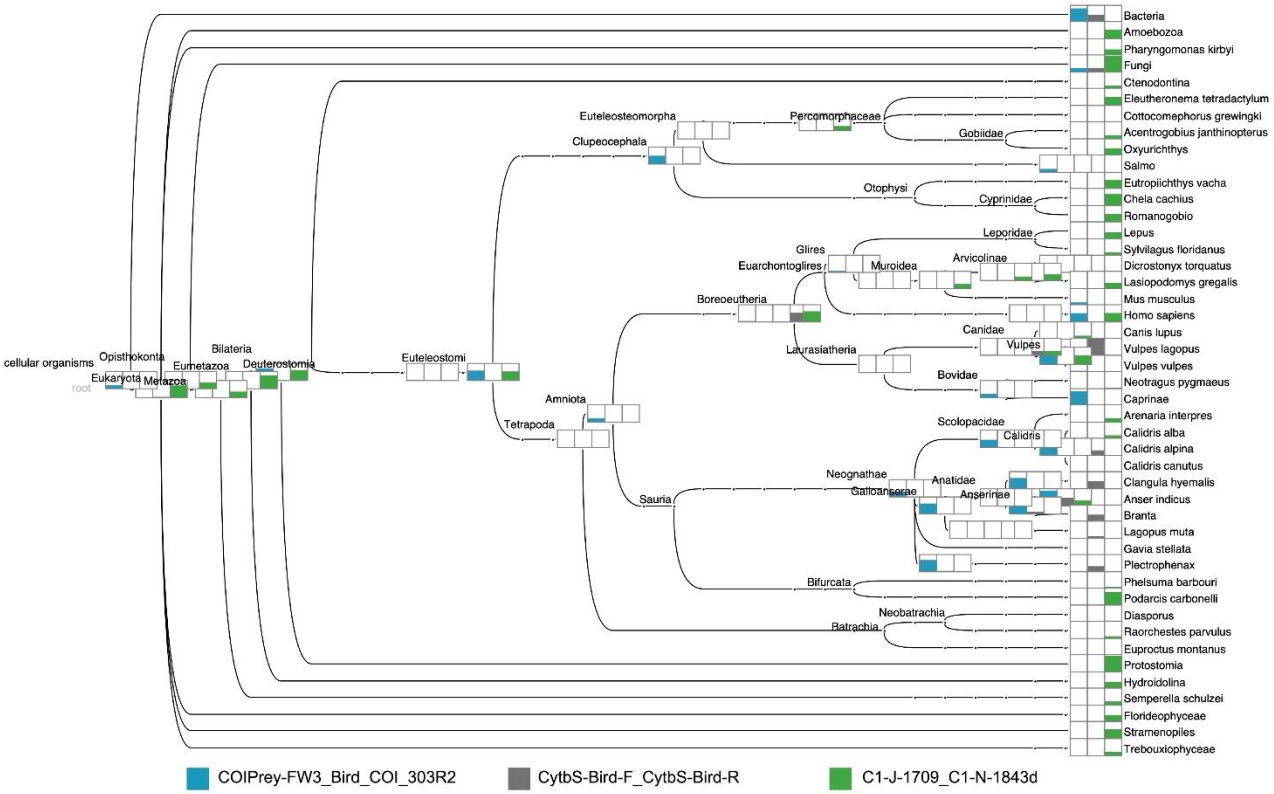

Figure S1. Assignation of the sequencing reads to taxa, as resolved per primer pair: mammalian prey as amplified by primers C1-J-1709+C1-N-1843d, bird prey amplified by primers CytbS-Bird+CytbS-Bird-R, and vertebrate prey amplified by primers COIPrey-FW3+Bird_COI_303R2. The figure shows which taxa were identified by which primer pair (see legend), and how consistent the detections and the proportions were for the different primer pairs. Thus, the three columns on the right show the three primer pairs as above, with the fill level of each bar representing the number of reads (in log scale) per primer pair assigned for the specified taxon. For each gene region and primer pair, the taxonomic assignation was done by BLASTing sequences against the GenBank nt database, and then importing BLAST output into software MEGAN to analyse the lowest common ancestor (LCA) for each sequence

References

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.

Edgar, R.C., 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. bioRxiv, 081257.

Harmoinen, J., Von Thaden, A., Aspi, J., Kvist, L., Cocchiararo, B., Jarausch, A., Gazzola, A., Sin, T., Lohi, H., Hytönen, M.K., Kojola, I., Stronen, A.V., Caniglia, R., Mattucci, F., Galaverni, M., Godinho, R., Ruiz-González, A., Randi, E., Muñoz-Fuentes, V., Nowak, C., 2021. Reliable wolf-dog hybrid detection in Europe using a reduced SNP panel developed for non-invasively collected samples. BMC Genomics 22.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377-386.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647-1649.

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

Pastor-Beviá, D., Ibáñez-Álamo, J.D., García-Mudarra, J.L., Juste, J., 2014. A molecular approach to the study of avian DNA in bat faeces. Acta Chiropterologica 16, 451-460.

Ratnasingham, S., Hebert, P.D., 2007. BOLD: The Barcode of Life Data System (http://www.barcodinglife.org). Mol. Ecol. Notes 7, 355-364.

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

Vesterinen, E.J., Kaunisto, K.M., Lilley, T.M., 2020. A global class reunion with multiple groups feasting on the declining insect smorgasbord. Sci. Rep. 10, 16595.

Vesterinen, E.J., Puisto, A.I.E., Blomberg, A.S., Lilley, T.M., 2018. Table for five, please: Dietary partitioning in boreal bats. Ecol. Evol. 8, 10914-10937.

Vesterinen, E.J., Ruokolainen, L., Wahlberg, N., Peña, C., Roslin, T., Laine, V.N., Vasko, V., Sääksjärvi, I.E., Norrdahl, K., Lilley, T.M., 2016. What you need is what you eat? Prey selection by the bat Myotis daubentonii. Mol. Ecol. 25, 1581-1594.

Vestheim, H., Jarman, S.N., 2008. Blocking primers to enhance PCR amplification of rare sequences in mixed samples – a case study on prey DNA in Antarctic krill stomachs. Frontiers in Zoology 5, 12.