Oct 24, 2024

LIFEPLAN Malaise sample metabarcoding

DOI
dx.doi.org/10.17504/protocols.io.5qpvokn3xl4o/v1

Jeremy deWaard[1], Stephanie deWaard[1], Maria Kuzmina[1], Jayme Sones[1], Kate Perez[1], Amy Thompson[1], Bess Hardwick[2], Brendan Furneaux[3], Tomas Roslin[4], Evgeny Zakharov[1]

[1]Centre for Biodiversity Genomics, University of Guelph; [2]University of Helsinki; [3]University of Jyväskylä; [4]Swedish University of Agricultural Sciences

Jeremy deWaard: present affiliation: Smithsonian National Museum of Natural History;

Bess Hardwick
University of Helsinki

DOI: https://dx.doi.org/10.17504/protocols.io.5qpvokn3xl4o/v1

Protocol Citation: Jeremy deWaard, Stephanie deWaard, Maria Kuzmina, Jayme Sones, Kate Perez, Amy Thompson, Bess Hardwick, Brendan Furneaux, Tomas Roslin, Evgeny Zakharov 2024. LIFEPLAN Malaise sample metabarcoding. **protocols.io https://dx.doi.org/10.17504/protocols.io.5qpvokn3xl4o/v1**

Protocol status: Working
We use this protocol and it's working

Created: March 28, 2024

Last Modified: October 24, 2024

Protocol Integer ID: 97488

Keywords: metabarcoding, COI, arthropods, malaise trapping, bulk arthropod samples to coi sequence, lifeplan malaise sample metabarcoding in project lifeplan, lifeplan malaise sample metabarcoding, bulk samples of arthropod, bulk arthropod sample, arthropod, project lifeplan, reads per bulk sample, dna extraction, species, coi sequence, bulk sample

Funders Acknowledgements:
Government of Canada New Frontiers in Research Fund
Grant ID: NFRFT-2020-00073
Large Scale Applied Research Program administered by Genome Canada and Ontario Genomics
Grant ID: OGI-208
Major Science Initiatives Fund administered by Canada Foundation for Innovation
Grant ID: Project number 42450
European Research Council
Grant ID: Grant 856506
Food from Thought Program at the University of Guelph, funded by the Canada First Research Excellence Fund
Grant ID: NA

Abstract

In Project LIFEPLAN, bulk samples of arthropods are collected with Malaise traps and the species are identified with a non-destructive COI metabarcoding process. This protocol describes the steps for lysis, DNA extraction, PCR amplification, library construction, sequencing and sequence analysis, going from bulk arthropod samples to COI sequences, aiming at 1 million reads per bulk sample.

Materials

Microfunnel 0.45 uM Supor membrane filters (Pall Laboratory)
Analytic and precision weighing balance (Sartorius)
6-Funnel manifold (Pall Laboratory)
Incubator
Shaker
500 µl 96-well deep well plates (Eppendorf)
Positive control: AMPtk synthetic, Palmer et al. 2018
96% ethanol
96-well microplates
3.0 µm Pall Supor Membrane glass fiber plates (Pall Laboratory)
Centrifuge
GuSCN
EDTA pH 8.0
Tris-HCl pH 8.0
0.5% Triton X-100
5% Tween20
Tris-HCl pH 6.4
4% Triton X-100
NaCl
Tris-HCl pH 7.4
forward primer BF3 and reverse primer BR2 (Elbrecht & Steinke 2018 ; Elbrecht et al., 2019)
PCR thermal cycler
Invitrogen e-Gel 2% Agarose gel
Biomek i7 Automated Liquid Handling Workstation (Beckman Coulter LIfe Sciences, Indianapolis)
5 – 15 mL tubes
Magnetic beads (Cytiva Sera-Mag™ Carboxylate-Modified Magnetic SpeedBeads)
Vortex shaker
1.5 mL Eppendorf LoBind tubes
Magnet for magnetic separation
Invitrogen Size-Select 2% agarose gel
Rainin pipette (Mettler Toledo, Mississauga, ON)
Invitrogen Qubit High Sensitivity kit
Bioanalyzer (Agilent Technology, Santa Clara)
Illumina NovaSeq 6000 platform with the S Prime (SP) reagent kit, SP flow cell, and 500-cycle configuration

Troubleshooting

Sample collection

1 As its foundation for insect sampling, the LIFEPLAN project adopted the basic protocols used in the Global Malaise Program.

Protocol

NAME

Global Malaise Trap Project and LIFEPLAN Malaise sampling

CREATED BY

Gaia Giedre Banelyte

Preview

Traps were serviced once per week, replacing the sample bottle and recording important collection information through the LIFEPLAN app (Android: https://play.google.com/store/apps/details?id=com.lifeplanapp and iOS: https://apps.apple.com/fi/app/lifeplan/id1533244844).

2 These samples were shipped to the Centre for Biodiversity Genomics (**https://biodiversitygenomics.net/**) for analysis, where they were preserved in fresh 96% ethanol and held at -20 °C until ethanol filtration and tissue lysis.

Comprehensive details of sample collection are provided by Sones et al. (2023).

Figure 1. Bulk samples of insect specimens collected from Malaise Traps, preserved in 96% ethanol, and stored at -20C pending analysis at the Centre of Biodiversity Genomics.

Ethanol filtration and tissue lysis

3 Ethanol was filtered from five bulk samples of insect specimens at a time using five Microfunnel 0.45 uM Supor membrane filters (Pall Laboratory), an analytic and precision weighing balance (Sartorius), and a 6-Funnel vacuum manifold (Pall Laboratory).

Figure 2. Five bulk samples of insect specimens prepared for ethanol filtration using a 6-funnel vacuum manifold.

4 First the weight of each of the five membrane filters was tared with source weights recorded in CBG's Collection Information Management System and then placed on five of the six funnels on the manifold.

5 The bulk sample of insect specimens were filtered through the membrane on the manifold until all specimens were transferred from the bulk bottle to the membrane and all ethanol was removed. The membrane and bulk sample contents were weighed using the balance with the initial tared weight subtracted, to measure the resulting wet arthropod biomass of the bulk sample.

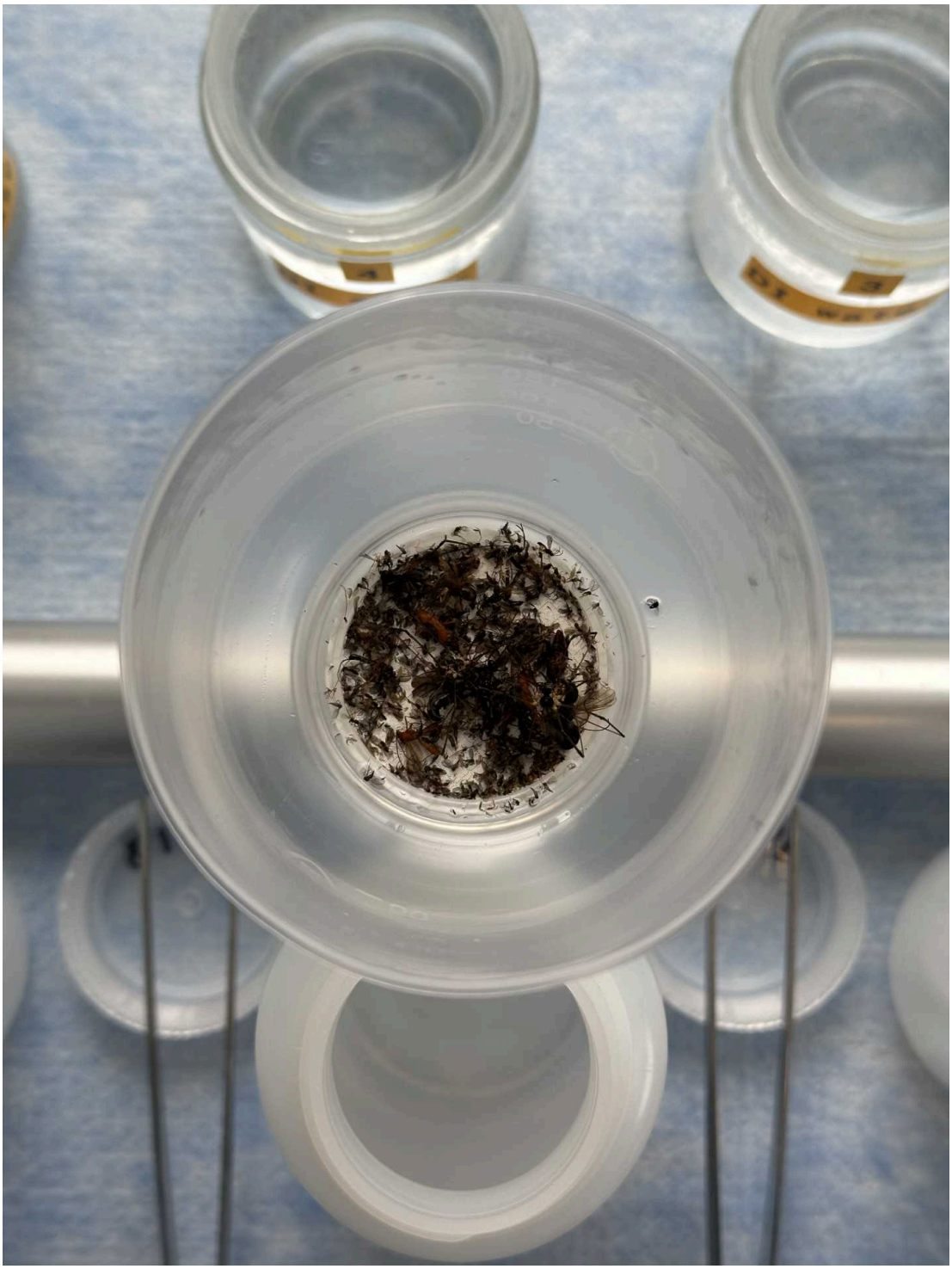


Figure 3. A bulk sample of insect specimens with ethanol filtered, ready to measure the wet arthropod biomass.

6 Lastly, the insect specimen contents and their membrane were transferred back to the bulk sample bottle for lysis.

7 The wet arthropod biomass was recorded for the remaining four bulk samples of insect specimens in the subset by repeating go to step #5 .

8 The amount of insect lysis buffer (go to step #12.1) was determined based on the wet arthropod biomass (see Table 1) and this volume of buffer was added to the bulk bottle containing the filtered bulk sample of insect specimens, ensuring that all specimens were submerged in the insect lysis buffer.

	A	B	C
	Minimum wet weight (g)	Maximum wet weight (g)	ILB (mL)
	0	9.9	50
	10	19.9	100
	20	29.9	200
	30	39.9	300
	40	49.9	400
	50	59.9	500
	60	69.9	600
	70	79.9	700
	80	100	800

Table 1. The amount of insect lysis buffer (ILB) required for lysis was determined based on a wet weight (g) to insect lysis buffer ratio (mL).

9 Bulk samples were incubated, rotating gently on a shaker (VWR), at 56 °C Overnight .

Figure 4. Bulk samples of insect specimen bottles containing lysis buffer and positioned on rotating shakers (VWR) for incubation at 56C.

After lysis, 3 technical replicates of 300 µL aliquots of lysate were taken from each sample in a set of 30 samples and transferred to a 500 µL 96-well deep well plate (Eppendorf).

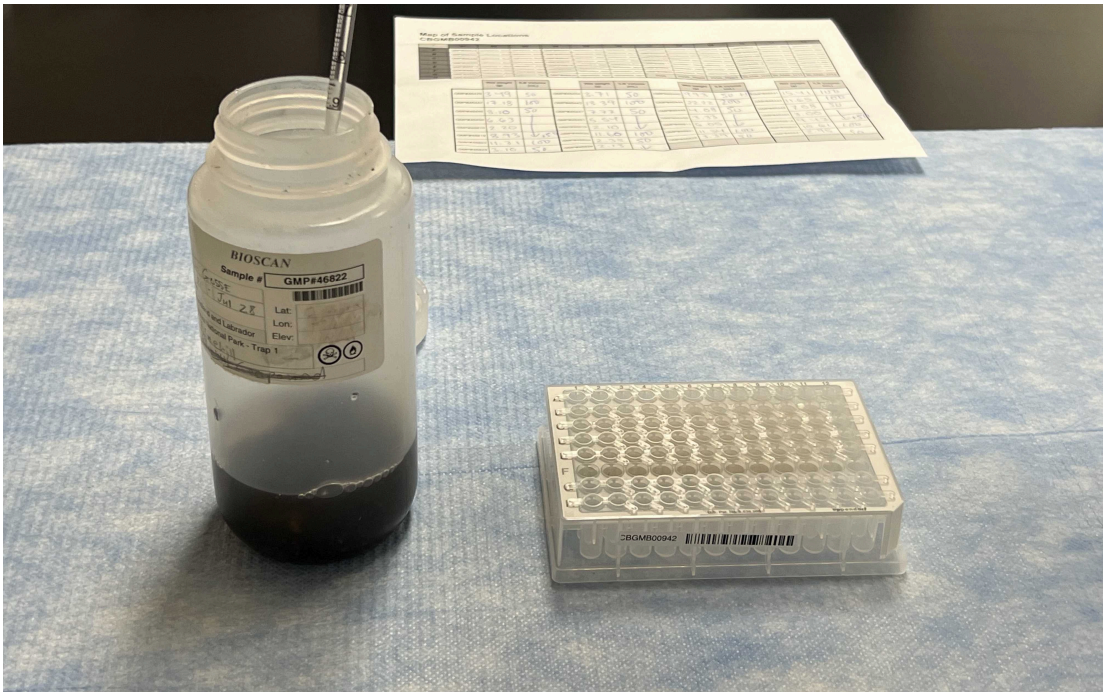


Figure 5. The sampling of 300uL of lysate from one bulk sample of insect specimens following lysis.

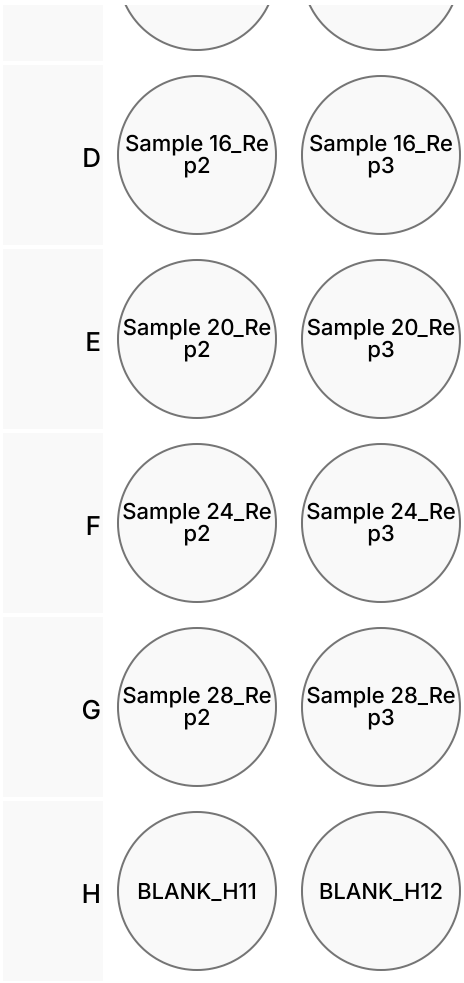
Six wells were designated for three negative (template-free) and three positive (AMPtk synthetic, Palmer et al. 2018) controls per block (see below for a microplate map of samples and controls).

	1	2	3	4	5
A	Sample 1	Sample 1_Rep 2	Sample 1_Rep 3	Sample 2	Sample 2_Rep 2
B	Sample 5	Sample 5_Rep 2	Sample 5_Rep 3	Sample 6	Sample 6_Rep 2
C	Sample 9	Sample 9_Rep 2	Sample 9_Rep 3	Sample 10	Sample 10_Re p2
D	Sample 13	Sample 13_Re p2	Sample 13_Re p3	Sample 14	Sample 14_Re p2
E	Sample 17	Sample 17_Rep 2	Sample 17_Rep 3	Sample 18	Sample 18_Re p2
F	Sample 21	Sample 21_Re p2	Sample 21_Re p3	Sample 22	Sample 22_Re p2
G	Sample 25	Sample 25_Re p2	Sample 25_Re p3	Sample 26	Sample 26_Re p2
H	Sample 29	Sample 29_Re p2	Sample 29_Re p3	Sample 30	Sample 30_Re p2

	6	7	8	9	10
A	Sample 2_Rep 3	Sample 3	Sample 3_Rep 2	Sample 3_Rep 3	Sample 4

B	Sample 6_Rep 3	Sample 7	Sample 7_Rep 2	Sample 7_Rep 3	Sample 8
C	Sample 10_Rep3	Sample 11	Sample 11_Rep 2	Sample 11_Rep 3	Sample 12
D	Sample 14_Rep3	Sample 15	Sample 15_Rep2	Sample 15_Rep3	Sample 16
E	Sample 18_Rep3	Sample 19	Sample 19_Rep2	Sample 19_Rep3	Sample 20
F	Sample 22_Rep3	Sample 23	Sample 23_Rep2	Sample 23_Rep3	Sample 24
G	Sample 26_Rep3	Sample 27	Sample 27_Rep2	Sample 27_Rep3	Sample 28
H	Sample 30_Rep3	POS	POS_Rep2	POS_Rep3	BLANK_H10

	11	12
A	Sample 4_Rep 2	Sample 4_Rep 3
B	Sample 8_Rep 2	Sample 8_Rep 3
C	Sample 12_Rep2	Sample 12_Rep3



16h

10 Subsequently, 50 µL of lysate for each well was transferred from the block to a microplate to proceed with DNA extraction and the remaining 250 µL of lysates in the block were archived for future analysis.

11 Lysate remaining in the bottles was filtered and disposed of, whereas the specimens were archived in 96% ethanol in labelled, heat-sealed bags for future use and stored at -20 °C .

Figure 6. Specimens within heat sealed bags containing 96% ethanol for archiving.

DNA extraction and PCR amplification

12 DNA extractions followed a membrane-based protocol (Ivanova et al. 2006). The components of each buffer used in tissue lysis and DNA extraction are the following (from Ivanova et al. 2006):

12.1 Insect lysis buffer – 700 mm GuSCN, 30 mm EDTA pH 8.0, 30 mm Tris-HCl pH 8.0, 0.5% Triton X-100, and 5% Tween20

12.2 Binding buffer (BB) – 6 M GuSCN, 20 mM EDTA pH 8.0, 10 mM Tris-HCl pH 6.4, and 4% Triton X-100, pre-warmed at 56 °C to dissolve

12.3 Binding mix – 50 mL of ethanol (96%) thoroughly mixed with 50mL of BB

12.4 Protein wash buffer – 70 mL of ethanol (96%) thoroughly mixed with 26 mL of BB

12.5 Wash buffer – ethanol (60%), 50 mm NaCl, 10 mm, Tris-HCl pH 7.4 and 0.5 mm EDTA pH 8.0

13 First, 100 µL of binding mix go to step #12.3 was added to the microplate containing 50 µL of lysate and the total volume of 150 µL was mixed and transferred to a 3.0µm Pall Supor Membrane glass fiber plate (Pall Laboratory) and 5000 x g, Room temperature, 00:05:00 . 5m

14 Then, 180 µL of protein wash buffer go to step #12.4 was added and centrifuged at 5000 x g, Room temperature, 00:02:00 followed by two additional washes with 600 µL of wash buffer go to step #12.5 with centrifugation 5000 x g, Room temperature, 00:05:00 for both washes. 12m

15 The 3.0µm Pall Supor Membrane glass fiber plate was placed onto a new microplate used for DNA storage and incubated at 56 °C for 00:30:00 . 30m

16 Finally, 60 µL of elution buffer was added, incubated at room temperature for 1 min, and centrifuged for 5 min at 5000 x g to elute the DNA.

17 A 418 bp region of cytochrome c oxidase subunit I (COI) was amplified from two rounds of PCR amplification using the forward primer BF3 and the reverse primer BR2 (Elbrecht & Steinke 2018; Elbrecht et al., 2019). The PCR1 reaction amplifies the BF3/BR2 fragment and the PCR2 reaction adds the P5/P7 Illumina adapters. DNA and PCR products were added to pre-made PCR plates using a Beckman Coulter Biomek FX^P automated workstation.

Figure 7. PCR1 dilution and transfer to PCR2 using a Beckman Coulter Biomek FX^{P} automated workstation.

Figure 8. Close up of 2ul of PCR1 product being transferred to a PCR2 plate.

18 PCR1 reactions were carried out with the following thermocycler settings: initial denaturation at 95 °C for 00:05:00, 30 cycles of denaturation at 94 °C for 00:00:30, annealing at 46 °C for 00:00:30, extension at 72 °C for 00:00:50, and lastly extension for 00:05:00 at 72 °C. A 2 µL volume of DNA template and AMPtk positive control were added to corresponding wells. The microplate of genomic DNA was heat-sealed and archived at -80 °C for future use. First-round PCR products were diluted by 1x before the second round of PCR. 11m 50s

19 PCR2 reactions employed fusion versions of BF3 and BR2 primers tailed with one of 24 6- to 10-base long in-line tags and P5 or P7 sequencing adaptors (Illumina) for a total of 1152 unique combinations of tag pairs. The total number of unique combinations factors in half of the samples being sequenced in reverse direction by swapping P5 and P7 tails on the indexed primers. PCR2 reactions were carried out with the following thermocycler settings: initial denaturation at 94 °C for 00:02:00, 15 cycles of denaturation at 94 °C for 00:00:40, annealing at 51 °C for 00:01:00, extension at 72 °C for 00:01:00, and lastly extension for 00:05:00 at 72 °C. A 2 µL volume of PCR1 template, including positive and negative controls, was added. 9m 40s

Library construction

20 Pooling PCR2 product

The PCR2 products labeled with P5/P7 Illumina indexing adapters were checked by running 4 µL of the product on the Invitrogen pre-cast E-Gel 2% agarose gel for 00:04:00. An equal volume of each sample (6 µL) was pooled into one 96-well microplate using an Biomek i7 Automated Liquid Handling Workstation (Beckman Coulter LIfe Sciences, Indianapolis) robot. The pooled PCR products were then transferred into a 5 – 15 mL tube.

Figure 9. Pooling of 12 96-well PCR2 plates into a single 96-well microplate using an Biomek i7 Automated Liquid Handling Workstation.

4m

21 Concentration of the DNA library

10m

The pooled DNA library was concentrated with a 0.8 – 1.2 ratio of magnetic beads (Cytiva Sera-Mag™ Carboxylate-Modified Magnetic SpeedBeads) to pooled PCR product. The DNA library was vortexed and spun. Four 1.5 mL Eppendorf LoBind tubes were prepared with pre-aliquoted 160 µL of magnetic beads. The pooled DNA library was vortexed and 200 µL of it was added to each tube. The standard recommendations for the library concentration were applied: 00:08:00 on the bench at Room temperature, 00:02:00 on the magnet, three washes with 1 mL of 80% freshly prepared ethanol, elution with 45 µL of water, and collecting 40 µL from each tube. The total volume of the concentrated DNA library was 160 µL.

22 Size selection

To target the desired length of amplicons, and thereby eliminate non-specific, especially long fragments of DNA from the DNA library, we used an Invitrogen Size-Select 2% agarose gel, according to the protocol recommended by the manufacturer. The final volume of the target DNA was approximately 200 µL.

23 Final purification and evaluation of the quality of the DNA library

The volume of the target DNA library was evaluated by using a Rainin pipette (Mettler Toledo, Mississauga, ON). The 0.8 ratio of Cytiva magnetic beads:pooled library was applied. For library concentration, we used the standard recommendations of the manufacturer, including 00:08:00 on the bench at Room temperature, 00:02:00 on the magnet, four washes with 1 mL of 80% freshly prepared ethanol, elution with 23 µL of water, and collecting 20 µL. To evaluate the concentration of the final DNA library, we used the Qubit High Sensitivity kit reagents and the Qubit 2.0 Fluorometer (Invitrogen). The normalized 1 ng/uL concentration of the final library was checked on the Bioanalyzer (Agilent Technology, Santa Clara) using the High Sensitivity kit in order to visualize the length of the target DNA.

Figure 10. The concentration of the final DNA library read using a Qubit 2.0 Fluorometer (Invitrogen)

10m

Sequencing

24 Sequencing was completed on an Illumina NovaSeq 6000 platform with the S Prime (SP) reagent kit, SP flow cell, and 500-cycle configuration. Each lane received its own library with 1152 dual-indexed amplicons. Because multiplexing employed inline tags and did not employ any native Illumina indices, indexing flows were redirected to sequencing of the insert for both reads, effectively making it 2×258bp paired-end sequencing. Up to 15% PhiX spike-in was used to ensure base diversity. Each sequencing run produced two sets of files, one for each lane, which included a single read 1 file and a single read 2 file. These files were subjected to the bioinformatics data flow as described below.

Sequence analysis

25 Reads from the three replicates for each sample were demultiplexed and batch-uploaded to mBRAVE (Multiplex Barcode Research And Visualization Environment; Ratnasingham 2019; **http://mbrave.net/**) for filtering, clustering and taxonomic assignment. For each NovaSeq lane, a new mBRAVE project was created (e.g., MBR-LPLAN21001 – 'LifePlan Malaise Trap Metabarcoding - December 2021') to analyze and archive the 1152 sample and control replicates. Prior to filtering, reads were trimmed 23 bp from both termini and primer masking was turned off. For inclusion in downstream analysis, reads were filtered with a minimum length >446 bp and the following three quality criteria: mean QV > 20; <6% positions with QV < 20; and <1% positions with QV < 10. Reads were then pre-clustered into OTUs (OTU threshold = 3%; minimum OTU size = 1 read) prior to querying against each reference library (with an ID distance threshold of 2% and a minimum overlap of 100bp).

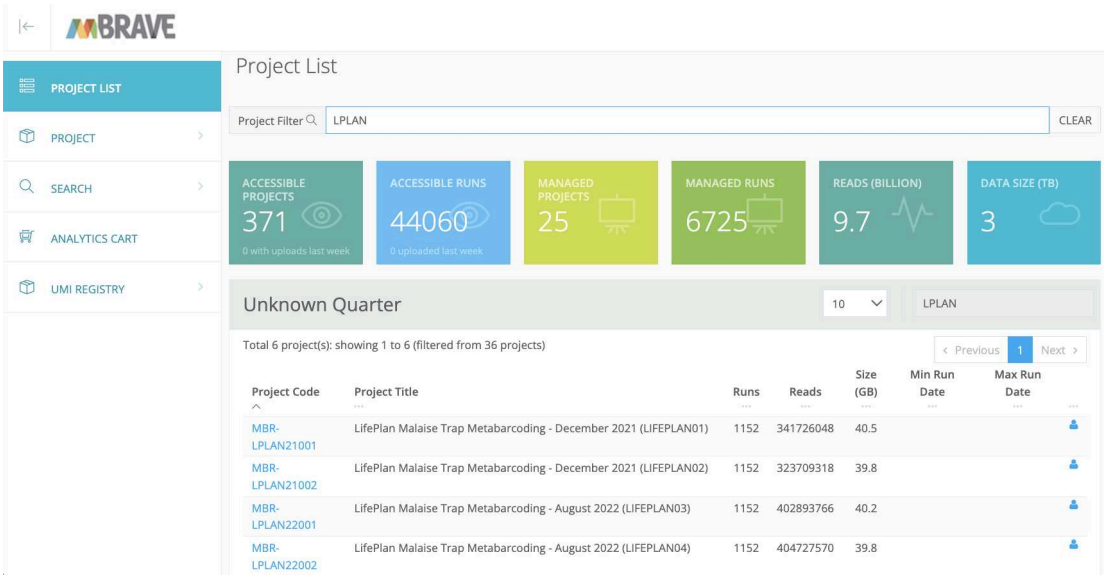


Figure 11. Screen capture from mBRAVE showing a list of LIFEPLAN projects.

26 All clusters were queried against ten system libraries from BOLD (Ratnasingham & Hebert 2007): Standard Contaminants Based on Reagent Production (SYS-MBRAVEC), Bacteria COI (SYS-CRLBACTERIA), Non-Arthropoda Invertebrates (SYS-CRLNONARTHINVERT), Non-Insect Arthropoda (SYS-CRLNONINSECTARTH), Insecta (SYS-CRLINSECTA), Chordata (SYS-CRLCHORDATA), Aves (SYS-CRLAVES), Protista (SYS-CRLPROTISTA), and Human Contamination Check (SYS-HUMC). These results were exported as tab-delimited text files (.tsv files). For more details on the data analysis using mBRAVE, see Steinke et al. (2022) and Liu et al. (2023).

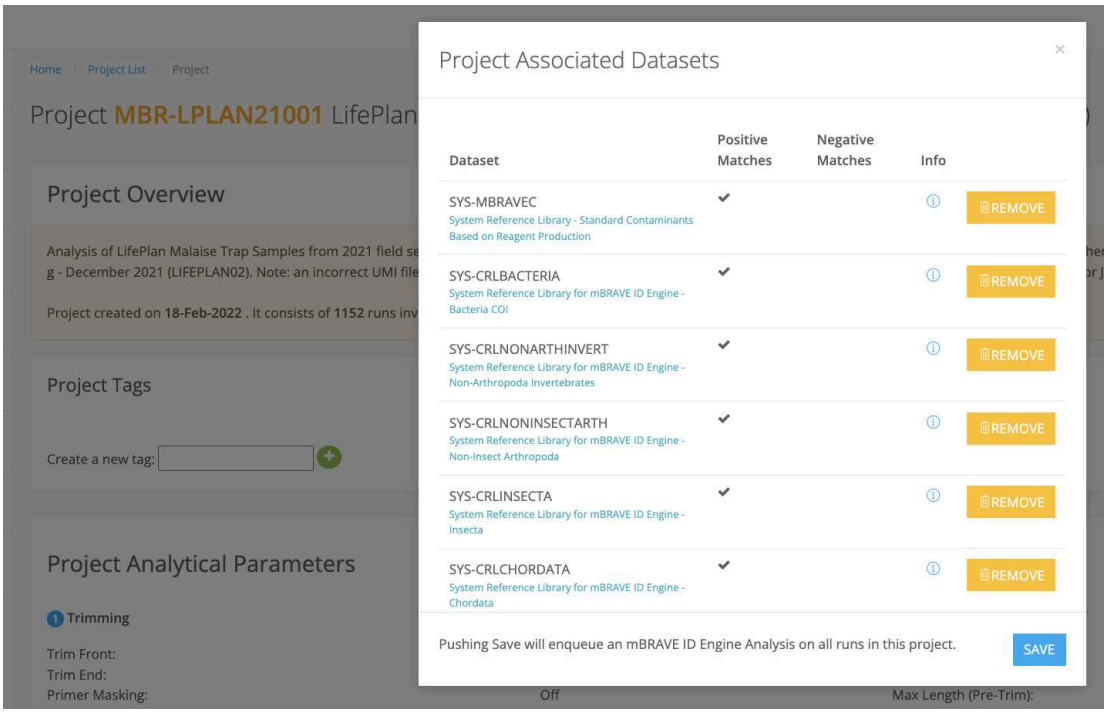


Figure 12. Screen capture from mBRAVE showing a list of system libraries from BOLD used to query clusters against.

27 In Excel, we generated BIN (Barcode Index Number; Ratnasingham & Hebert 2013) tables including all library queries for each individual plate (30 samples X 3 replicates, plus 3 negative and 3 positive controls for each plate). Read counts for any BINs recovered from the negative control on a plate were subtracted from the counts for the same BIN in the 90 non-control wells in the run. When this subtraction reduced the read count for a BIN to zero, its occurrence was removed. This step reduced the effects of rare tag switching on data integrity and reduced background contamination. Finally, all bacterial, protist and other contaminant reads were discarded from further analysis.

Protocol references

Braukmann TWA, Ivanova NV, Prosser SWJ, Elbrecht V, Steinke D, Ratnasingham S, deWaard JR, Sones JE, Zakharov EV, Hebert PDN. 2019. Metabarcoding a diverse arthropod mock community. Molecular Ecology Resources 19: 711–727.

Elbrecht V, Steinke D. 2018. Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. Freshwater Biology 64: 380–387.

Elbrecht V, Braukmann TWA, Ivanova NV, Prosser SWJ, Hajibabaei M, Wright M, Zakharov EV, Hebert PDN, Steinke D. 2019. Validation of COI metabarcoding primers for terrestrial arthropods. PeerJ 7: e7745.

Ivanova NV, deWaard JR and Hebert, PDN. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes 6: 998–1002.

Liu C, Ashfaq M, Yin Y, Zhu Y, Wang Z, Cheng H, Hebert P. 2023. Using DNA metabarcoding to assess insect diversity in citrus orchards. PeerJ 11: e15338.

Palmer JM, Jusino MA, Banik MT, Lindner DL. 2018. Non-biological synthetic spike-in controls and the AMPtk software pipeline improve mycobiome data. PeerJ 6: e4925.

Ratnasingham S. 2019. mBRAVE: The Multiplex Barcode Research And Visualization Environment. Biodiversity Information Science and Standards 3: e37986.

Ratnasingham S, Hebert PDN. 2013. A DNA-based registry for all animal species: the barcode index number (BIN) system. PLOS ONE 8: e66213.

Ratnasingham S, Hebert PDN. 2007. BOLD: The Barcode of Life Data system (www.barcodinglife.org). Molecular Ecology Notes 7: 355–364.

Sones, JE, Banelyte GG, deWaard JR, Farrell AM, Rogers HMK, Kerdraon D, Perez KHJ. 2023. Global Malaise Trap Project and LIFEPLAN Malaise sampling. protocols.io https://dx.doi.org/10.17504/protocols.io.kqdg3xkdqg25/v1

Steinke D, deWaard SL, Sones JE, Ivanova NV, Prosser SWJ, Perez K, Braukmann TWA, Milton M, Zakharov EV, deWaard JR, Ratnasingham S, Hebert PDN. 2022. Message in a Bottle—Metabarcoding enables biodiversity comparisons across ecoregions. GigaScience 11: giac040.