Pooling size sorted Malaise trap fractions to maximize taxon recovery with metabarcoding

Vasco Elbrecht[1,2], Sarah J. Bourlat[1], Thomas Hörren[3], Angie Lindner[1], Adriana Mordente[1], Niklas W. Noll[1], Livia Schäffler[1], Martin Sorg[3] and Vera M.A. Zizka[1]

[1] Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig, Bonn, Germany
[2] SimplexDNA AG, Winterthur, Switzerland
[3] Entomological Society Krefeld, Krefeld, Germany

Corresponding author
Vera M.A. Zizka,
v.zizka@leibniz-zfmk.de

ABSTRACT

Background: Small and rare specimens can remain undetected when metabarcoding is applied on bulk samples with a high specimen size heterogeneity. This is especially critical for Malaise trap samples, where most of the biodiversity is contributed by small taxa with low biomass. The separation of samples in different size fractions for downstream analysis is one possibility to increase detection of small and rare taxa. However, experiments systematically testing different size sorting approaches and subsequent proportional pooling of fractions are lacking, but would provide important information for the optimization of metabarcoding protocols. We set out to find a size sorting strategy for Malaise trap samples that maximizes taxonomic recovery but remains scalable and time efficient.

Methods: Three Malaise trap samples were sorted into four size classes using dry sieving. Each fraction was homogenized and lysed. The corresponding lysates were pooled to simulate unsorted samples. Pooling was additionally conducted in equal proportions and in four different proportions enriching the small size fraction of samples. DNA from the individual size classes as well as the pooled fractions was extracted and metabarcoded using the FwhF2 and Fol-degen-rev primer set. Additionally, alternative wet sieving strategies were explored.

Results: The small size fractions harboured the highest diversity and were best represented when pooling in favour of small specimens. Metabarcoding of unsorted samples decreases taxon recovery compared to size sorted samples. A size separation into only two fractions (below 4 mm and above) can double taxon recovery compared to not size sorting. However, increasing the sequencing depth 3- to 4-fold can also increase taxon recovery to levels comparable with size sorting, but remains biased towards biomass rich taxa in the sample.

Conclusion: We demonstrate that size fractionation of Malaise trap bulk samples can increase taxon recovery. While results show distinct patterns, the lack of statistical support due to the limited number of samples processed is a limitation. Due to increased speed and lower risk of cross-contamination as well as specimen damage we recommend wet sieving and proportional pooling of the lysates in favour of the small size fraction (80–90% volume). However, for large-scale projects with time constraints, increasing sequencing depth is an alternative solution.

INTRODUCTION

DNA metabarcoding is a useful approach to characterize arthropod communities. Instead of DNA barcoding individual specimens (*Hebert et al., 2003*), DNA is usually extracted from homogenized bulk samples (*Zizka, Geiger & Leese, 2020*; *Mata et al., 2020*; *Turunen et al., 2021*; *Hardulak et al., 2020*). A barcoding marker is amplified for the whole community and sequenced with high throughput sequencing. This allows for the identification of entire communities (*Ji et al., 2013*), often at species level resolution, depending on marker choice and reference database completeness (*Leray et al., 2019*; *Weigand et al., 2019*). However, DNA metabarcoding is affected by a large number of methodological biases preventing accurate specimen counts and often not all taxa in the sample are detected (*Elbrecht et al., 2017*; *Piper et al., 2019*). While there is a loose correlation between biomass and read abundance (*Lamb et al., 2019*), metabarcoding results are often skewed by primer bias (*Elbrecht & Leese, 2015*; *Piñol et al., 2015*; *Krehenwinkel et al., 2017*), mitochondrial copy number variation (*Braukmann et al., 2019*; *Deagle et al., 2019*), sequencing errors (*Singer et al., 2019*) and variation in specimen size, which can result in smaller taxa remaining undetected (*Elbrecht, Peinert & Leese, 2017*).

Arthropod bulk samples often show a wide range of specimen sizes (*Aylagas et al., 2016*). While imaginal stages of the same species have similar biomass, there can be substantial size variation between the life stages within species and of course between different species (*Stork & Blackburn, 1993*; *Siemann, Tilman & Haarstad, 1996*). Thus, when extracting DNA from bulk samples containing rare or small taxa with low biomass, those might be overshadowed by biomass rich taxa that contribute a large amount of DNA to the extraction. This natural variation can jeopardize the detection of small and rare specimens (*Deagle et al., 2018*), especially when taking PCR effects into account (*Kelly, Shelton & Gallego, 2019*). To maximize taxon detection in bulk samples showing substantial specimen biomass variation (*Aylagas et al., 2016*), size sorting might be required (*Liu et al., 2020*). Metabarcoding of marine arthropod bulk samples is commonly preceded by size fractioning of samples through sieving (*Cowart et al., 2015*; *Leray & Knowlton, 2015*; *Wangensteen & Turon, 2017*; *Wangensteen et al., 2018*). Size sorting has also been applied to freshwater macroinvertebrates (*Elbrecht, Peinert & Leese, 2017*; *Carew, Coleman & Hoffmann, 2018*; *Zizka et al., 2019*) and terrestrial arthropods (*Ji et al., 2013*; *Creedy, Ng & Vogler, 2019*; *Hausmann et al., 2020*). While the advantages of size sorting seem logical, the process can take a substantial amount of time if done manually (*Elbrecht, Peinert & Leese, 2017*) or even when using semi-automated methods like sieving. Additionally, while size sorting likely increases taxon recovery, it also generates sub samples that can either be pooled or sequenced individually. Regardless of the chosen size sorting and metabarcoding strategy, the amount of laboratory work and cost increases compared to the processing of a single metabarcoding bulk sample without size sorting. Thus, some authors suggest increasing sequencing depth instead of applying size sorting

strategies, especially when size variation is limited (*Creedy, Ng & Vogler, 2019*; *Elbrecht, Peinert & Leese, 2017*). In fact, most non-marine studies do metabarcode complete bulk samples without prior size sorting (*e.g. Hardulak et al., 2020*; *Hausmann et al., 2020*; *Steinke et al., 2020*; *Bush et al., 2020*).

In this study, we explore the effects of size sorting Malaise trap samples through sieving into four size fractions to maximize taxon recovery. We thereby want to circumvent time consuming and therefore expensive manual sorting, which is not feasible for large-scale approaches. Furthermore, manual sorting and determination of samples is difficult to standardize and depends on the researcher performing the sorting. We therefore focus on reproducible sieving approaches including a standardized mesh size and a defined sieving time and amplitude. Prior to DNA extraction, we specifically test the effect of pooling lysates of each size fraction in different proportions. Each size fraction is also extracted and metabarcoded individually, giving us the opportunity to further validate different size sorting scenarios. The goal of this study is to find a standardized and scalable DNA metabarcoding strategy for Malaise trap samples, which also captures the diversity of small and rare specimens.

MATERIALS AND METHODS

Samples and size sorting

Arthropods were collected on a meadow near Eschweiler, Germany (50.576708N, 6.730254E) using a Townes style Malaise trap (*Ssymank et al., 2018*). The trap was set up in the Natural Park "Hohes Venn" (calcareous grassland managed through grazing) with special permit from the nature conservation authority Euskirchen (reference number: 67.64.300/Pa). The trap was run with 80% denatured ethanol with 1% methyl ethyl ketone and samples were collected in 2017 from June 5th to June 26th (sample L1), from July 14th to August 11th (sample L2) and from May 16th to June 5th (sample L3). Samples were dried in an incubator (INCU Line ILS 6; VWR, Radnor, PA, USA) at 50 °C for 5 days. Samples were size sorted with a vibratory sieve shaker (AS 300 control; Retsch, Haan, Germany) using three stacked perforated plate sieves (305 mm in diameter) with 8, 4, and 2 mm diameter round holes respectively (Retsch, Haan, Germany). Sieving was performed for 5 min at 1.5 mm amplitude. Four size fractions (S < 2 mm, M > 2 mm, L > 4 mm, XL > 8 mm) were obtained. All specimens electrostatically attracted to the sieves were collected with tweezers and equipment was thoroughly cleaned with bleach between samples. Purified water was used for bleach removal and the equipment was subsequently UV irradiated for 5 min.

Size sorting of Malaise trap samples with the method outlined above takes 5–10 min, but removal of electrostatically charged specimens from the sieving equipment is time consuming and can take 2–3 h per sample. Thus, a second and third sieving method was explored (Fig. S1), but not sequenced due to time and financial constraints. Those alternative methods were tested on different samples than the dry sieving approach introduced in the previous section (L1–L3). Samples for the testing of alternative sieving approaches were collected in the nature protected area "Latumer Bruch", Lohbruchweg, Krefeld, Germany (51.326701N, 6.632973E) from 18th to 29th of May 2019

(dry sieving "shaker") and in a nature protected area near Vlotho, Germany (52.122705N, 8.785186E) from 19th of September to 3rd of October 2017 (wet sieving "lunch box"). For the second method, a dry sieving shaker was built. The shaker was constructed using one 500 ml and one 1 l sampling bottle (Kautex Textron GmbH & Co. KG, Bonn, Germany) which were connected with a metal mesh of either eight mm or four mm mesh size (Haver & Boecker OHG, Oelde, Germany) to obtain three size fractions (>8, >4, <4 mm). Dried specimens were separated into size classes by manually shaking the bottle for 2–3 min but subsequent specimen recovery was also affected by electrostatic charging from the plastic. As a third alternative size sorting approach, a four mm metal mesh (Haver & Boecker OHG, Oelde, Germany) was placed in a round glass container (IKEA365+; Ikea, Delft, Netherlands). The edges of the mesh were sealed with hot glue for better grip and stability within the glass dish. The sample was emptied onto the grid and fully immersed in 96% ethanol without prior drying. With repeated light shaking and pauses of 10 s, specimens could be separated into a small (<4 mm) and a large (>4 mm) size fraction. Specimens contained in both size fractions were dried and all specimens were measured and weighed to assess the effectiveness of this size sorting approach. Size measurements refer to the length of specimens as measured on a millimetre scale paper (excluding antennae and legs, Fig. S1).

Tissue homogenization, lysate pooling and DNA extraction

To reduce electrostatic charging after dry sieving, sieved insects were moistened with 96% ethanol before transferring them into 30 ml grinding tubes (Nalgene; Thermo Fisher Scientific, Waltham, MA, USA). Specimens were dried again for up to 5 days at 50 °C in the oven (INCU Line ILS 6, VWR, Radnor, PA, USA). Homogenization was performed in grinding tubes with five mm diameter steel beads (Retsch, Haan, Germany), using a mixer mill (MM400; Retsch, Haan, Germany) for 2 min at 30 hz. Ground tissue from each size fraction was transferred to either 50 ml falcon tubes or 500 ml bottles and weighed. For tissue lysis, 10 ml ATL lysis buffer (Qiagen, Hilden, Germany) supplemented with 2% Proteinase K (20 mg/ml) was added per g of tissue. Ground tissue with lysis buffer was incubated overnight at 56 °C in a shaking incubator (INCU Line ILS 6; VWR, Radnor, PA, USA) and centrifuged for 5 min at 4,149 g (Mega Star 1.6; VWR, Radnor, PA, USA). Samples were centrifuged to pelletize solid tissue at the bottom of the tubes and to facilitate the pooling of lysates by pipetting the respective proportions of supernatant. Thereby, lysates of the four size fractions (S < 2 mm, M > 2 mm, L > 4 mm, XL > 8 mm) were pooled in volumes ranging from 0.18 to 200 μl of supernatant (see Fig. 1 & Table S1 for details of lysate volumes pooled for the different ratios). In total, 200 μl of pooled (or single fraction) lysate was used for each DNA extraction. Since the amount of extraction buffer added was proportional to the weight of tissue in the respective size fraction, we assumed the DNA concentration in each lysate to be the same. Thus, pooling lysate volumes of the four size fractions proportional to the tissue weight simulates an unsorted Malaise trap sample (L1_gA, L2_gA, L3_gA). The other schemes (shown in
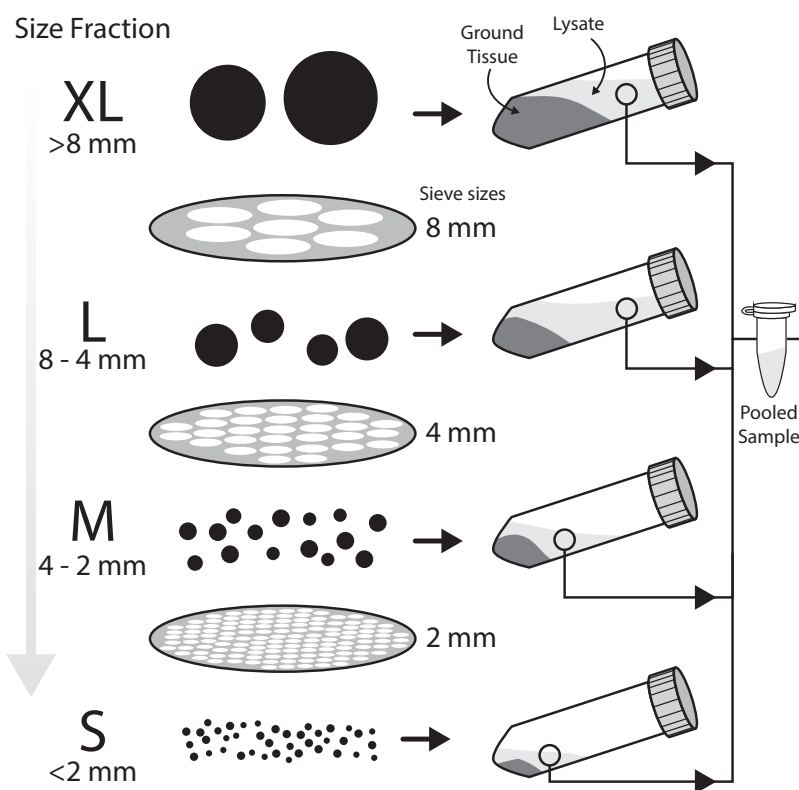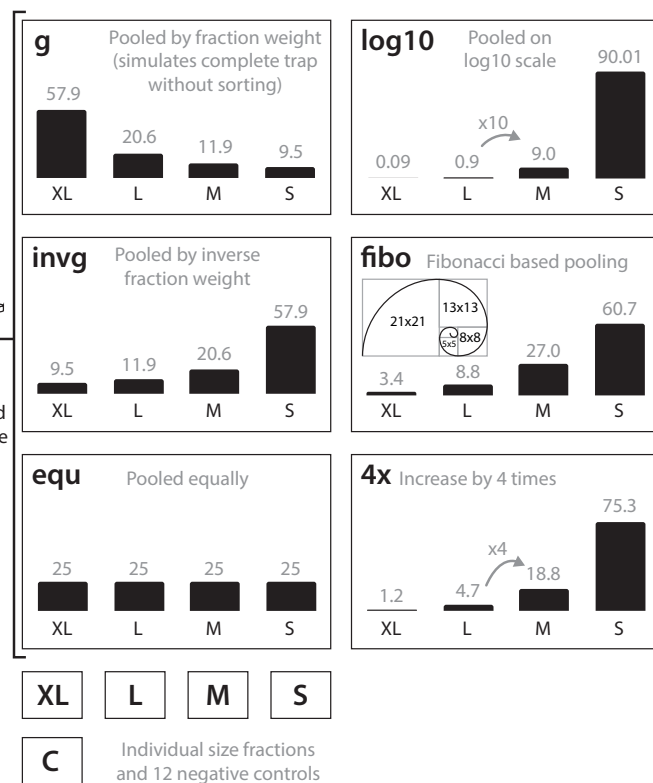
Figure 1 **Experimental design testing the effectiveness of size sorting Malaise trap samples and pooling the fractions in different proportions.** (A) Dried specimens are sieved into four different size classes, homogenized and lysed with a buffer amount corresponding to specimen weight in the respective fraction. (B) Lysate is incubated and pooled in different amounts. Grey numbers above the bars indicate the relative proportions (%). All six lysate pools were generated in duplicate, each individual size fraction (S, L, M, XL) was also extracted in duplicate, and nine negative controls were included.

Fig. 1, Table S1) simulate different pooling strategies of size fractions to test the effects on taxon recovery. In addition, individual size fractions (S, M, L, XL) of all three samples were also processed and sequenced individually and used for later *in silico* analysis.

DNA was extracted from the lysate using a Qiagen DNeasy 96 well Blood & Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol with only one adjustment (2% proteinase K). DNA was eluted with 200 µl buffer AE (Qiagen, Hilden, Germany). The 96 well plate map with sample locations is available as supporting information (Fig. S2), with 27 wells being used for primer tests as part of a separate project. In nine negative controls only ATL lysis buffer (Qiagen, Hilden, Germany) and Proteinase K without tissue were used for extraction (DNeasy 96 well Blood & Tissue Kit; Qiagen, Hilden, Germany). The eluates obtained from these negative controls were also included in downstream amplicon and indexing PCR. DNA quality was examined *via* gel electrophoresis on a 0.5% agarose gel with 0.1% GelRed (10,000×; Biotium Inc., Fremont, CA, USA) and 1× TBE. One µl of GeneRuler 100 bp plus DNA ladder (Thermo Fisher Scientific, Waltham, MA, USA) and three µl of each DNA extract were loaded onto the gel.

DNA metabarcoding

Sample metabarcoding was carried out for the six different pooling strategies of samples L1, L2 and L3 and also for each size fraction individually. An extraction replicate was included for each reaction. A two-step approach with standard illumina Nextera primers (San Diego, CA, USA) was used for dual tagging in PCR 2 (index PCR). All primers were obtained from Metabion International AG, Munich, Germany. PCR 1 (amplicon PCR) was carried out in a 96 well plate using reactions of 25 μl with one μl DNA (stock undiluted), 0.2 μM of each fwhF2 forward primer (*Vamos, Elbrecht & Leese, 2017*) and Fol-degen-rev reverse primer (*Yu et al., 2012*), 12.5 μl PCR Multiplex Plus buffer (Qiagen, Hilden, Germany) and 10.5 μl ddH$_2$O (Sigma–Aldrich, St. Louis, MO, USA Sigma–Aldrich). Primers included an universal Nextera tail to facilitate sample tagging in PCR 2. The thermal cycler 2,720 from ABI was used and set with the initial denaturation at 95 °C for 5 min; 30 cycles of: 30 sec at 95 °C, 30 sec at 50 °C and 50 sec at 72 °C; and a final extension of 5 min at 72 °C. One μl of PCR product was used as a template for the second PCR (no cleanup applied, *Elbrecht & Steinke, 2019*). In the second PCR 12 times 8 Nextera (Illumina, San Diego, CA, USA) based primer combinations were used to attach eight bp dual indexes to individual samples as well as the tails needed for Illumina sequencing (see Fig. S2 for further information on adapter combinations used). The 25 μl reactions in the second PCR were composed of one μl PCR1 (undiluted), 0.4 μM of each primer (Nextera, Illumina, San Diego, CA, USA), 12.5 μl PCR Multiplex Plus buffer (Qiagen, Hilden, Germany) and 10.5 μl ddH$_2$O. A higher concentration of primers (0.4 μM) was used in the second PCR than in the first PCR (0.2 μM) to achieve better DNA amplification. The same PCR cycling conditions as in the first PCR step were used, but the cycle number was reduced to 15. PCR success was checked on a 0.7% and 1% agarose gel for the first and second PCR, respectively. A higher agarose concentration allows a better separation of smaller bands *via* electrophoresis, thereby visualizing potentially problematic byproducts. Gels were composed as described above for the DNA quality check. PCR products were normalized using SequalPrep Normalization Plates (Thermo Fisher Scientific, Waltham, MA, USA; *Harris et al., 2010*) to obtain 25 ng product per reaction according to the manufacturer's protocol. Normalization success was verified on a 1% agarose gel as described above. A total of 10 μl of each normalised sample library was pooled, and the final library was cleaned up twice using a left sided size selection with 0.76× and 0.65× SPRIselect (Beckman Coulter, CA, USA) to completely remove primer dimers (including negative controls). The final library concentration (2.01 ng/μl) was measured using the QuantiFluor dsDNA system (E2670; Promega, Madison, WI, USA) and a Quantus fluorometer (Promega, Madison, WI, USA). Cleanup success was verified by running samples on a Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA) using the HS NGS Fragment Kit (Agilent Technologies, Santa Clara, CA, USA). The Fragment Analyzer data were evaluated using the software PROSize 2.0. Sequencing was carried out by Starseq (Mainz, Germany) using a 600 cycle Illumina MiSeq Reagent Kit v3 and 5% PhiX spike in.

Bioinformatic processing

Raw sequence data were delivered in blc format and converted to fastq files using bcl2fastq v2.20.0.422 and the option—create-fastq-for-index-reads. Sequence quality was verified using FastQC v0.11.9. Since there were 'no call' issues with the sequencing run, the base calls for each flow cell tile in each cycle were plotted and evaluated for problems using the PlotTiles function in JAMP v0.77 (github.com/VascoElbrecht/JAMP). No-call events were limited to the primer binding regions. Sequences were demultiplexed using deML setting exact matching (github commit dd87669 (*Renaud et al., 2015*)). Sequence data was processed using the JAMP pipeline that mostly relies on Vsearch v2.14.2 (*Rognes et al., 2016*) and Cutadapt v2.8 (*Martin, 2011*). Scripts are available in Script S1 ("JAMP_lysis_test_v3.R"). Paired-end merging of forward and reverse read was done using Vsearch with fastq_maxdiffpct=25 to maximize the number of reads merged. Primer sequences were removed from each sample using Cutadapt, allowing for 30% mismatches in the binding region, while retaining only sequences where both primers were successfully trimmed. Cutadapt was also used to remove sequences deviating more than 10 bp from expected amplicon length of 313 bp. Poor quality sequences with an expected error score above 1 (*Edgar & Flyvbjerg, 2015*) or containing N base calls were removed using Vsearch. All samples were pooled, dereplicated (minsize = 2) and clustered into Operational Taxonomic Units (OTUs) using Vsearch with cluster_smallmem and 97% similarity, followed by denovo chimera removal. Individual samples were dereplicated (retaining singletons) and mapped against the OTUs to generate an OTU table, using usearch_global with maxrejects = 256 and maxaccepts = 16. Potential false OTUs were flagged using LULU version 0.1.0 (*Frøslev et al., 2017*) and taxonomic assignment was conducted with https://www.gbif.org/tools/sequence-id. Subsequently, only flagged OTUs that shared the same taxonomy were merged. Flagged OTUs that were assigned to different taxonomies were analysed as separate molecular units for downstream analysis. Direct OTU clustering was used instead of denoising, to maximize the amount of sequences retained. Additionally, given the good coverage in the reference data, the removal of spurious OTUs with LULU gives a good confidence in OTU accuracy. For the comparison between size ratios OTU clustered data are sufficient and also represents species more closely than exact sequence variants (ESVs) which give a haplotype level resolution within species. Data were further cleaned up by subtracting the maximum read number of each OTU in the negative controls from all other samples (*Elbrecht & Steinke, 2019*). Additionally, all read counts with an abundance below 0.01% per sample were discarded as well as counts not present in both extraction replicates. Filtering all samples at 0.01% relative abundance ensures comparisons are done at the same sequencing depth, while reducing stochastic effects introduced by repeated rarefaction of the data. As downstream data analysis was performed with proportional read counts, filtering at 0.01% abundance is comparable to a sequencing depth of 10,000 sequences per sample. Additionally, since all reads were discarded that were not present in both extraction replicates with at least 0.01% abundance, the two replicates were combined for analysis. Since the GBIF database is clustered at 99% similarity, BOLDigger v1.1.3

(https://github.com/DominikBuchner/BOLDigger) was used to assign taxonomy to the OTU subset to improve taxonomic matching (setting "JAMP pipeline").

In silico analysis

To test if the partitioning of a sample into two size fractions would be sufficient to achieve optimal taxon recovery, the individually sequenced size fractions S+M (S < 2 mm, M > 2 mm) and L+XL (L > 4 mm, XL > 8 mm) were combined and pooled *in silico* in different ratios (Fig. 2, Fig. S3). First, based on data obtained from the separately sequenced size fractions, sequences of class S and M were pooled *in silico* for each sample after quality filtering (sequences present in both replicates, representing at least 0.01% of total observed reads per sample). Pooling was conducted to simulate unsorted samples, according to the dry weight of each size fraction. An identical process was carried out for the sequences from size fractions L and XL, in which they were pooled to simulate unsorted samples (Fig. S3). In the next step, the two reconstructed fractions (S+M and L+XL) were pooled *in silico* in different proportions, based on a sequencing depth of 10,000 reads. Starting with 100% S+M and 0% L+XL, the ratio was altered in 5% steps to 0% S+M and 100% L+XL, taking into account read numbers (*e.g.* 10 × read abundance S+M, 0 × read abundance L+XL). Relative read numbers were calculated for merged samples and a filtering threshold of 0.01% was applied. Thereby OTUs with a read abundance <0.01% per sample were discarded. Detailed information about *in silico* pooling is available in Supplementary Material Script S2 ("silico_pool_v1.R").

Rarefaction was applied to the unfiltered dataset (sequences present in both replicates, representing at least 0.01% of total observed reads per sample). The following examples detail a simulated sequencing depth of 500 reads (lowest sequencing depth) and 50,000 reads (highest sequencing depth (Fig. 3)). First, the sequencing depth was calculated using a filtering threshold defining the minimum abundance an OTU should have to be included in the analysis (for a sequencing depth of 500 reads: 100/500 = 0.2%; for a sequencing depth of 50,000 reads: 100/50,000 = 0.002%). In the next step, OTUs were excluded from the samples (g, equ, invg, fibo, log, 4×) if the minimum abundance of 0.2% (for 500 reads) or 0.002% (for 50,000 reads) was not reached respectively. Detailed information about *in silico* pooling is available in Supplementary Material Script S3 ("rarefaction.R").

RESULTS

Sorting dried specimens into four size fractions using the "sieving tower" (Fig. S1A) took 2–3 h per sample in total as statically charged specimens and broken off tissue pieces had to be collected from the sieve surface. The alternative "shaker" design was also affected by electrostatic charging. This problem could be overcome by moistening the dried samples before transferring them, which however again results in a higher total processing time. Wet sieving was not affected by electrostatic charging of specimens and enabled size sorting in approximately 15 min including sample transfer. Alternating between gentle shaking and letting small specimens sink through the four mm mesh resulted in clear
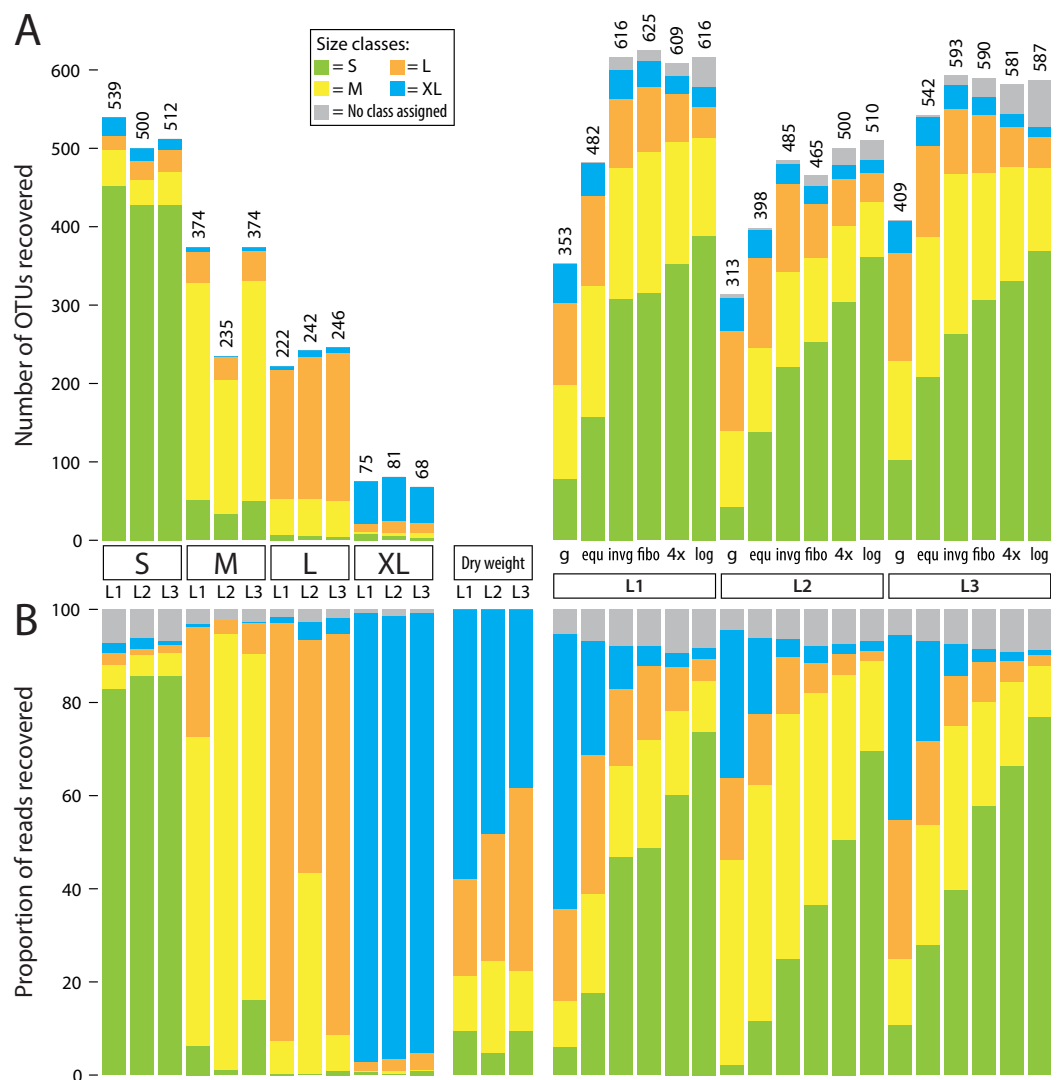
Figure 2 Detected OTUs across different samples (L1, L2 and L3). OTUs are assigned to size classes using the individually sequenced size fractions (S, M, L, XL). Some low-abundance OTUs were not recovered in individual size fractions, but in pooled samples. Thus, a size class was not assigned (grey bars). The four individual fractions were also pooled *in silico* at 1/4 sequencing depth (each filtered with a 0.04% threshold) to simulate equally pooling the samples. (A) barplots showing the number of OTUs recovered from individual size fractions as well as pooled samples (L1–L3). (B) barplots showing relative read abundance across samples (L1, L2 and L3). The "Dry weight" bars show the expected distribution of size classes based on relative abundance of dry weight of each size fraction, allowing a comparison to the proportionally pooled samples (g) that simulate a complete unsorted Malaise trap sample.

size fractionation with one order of magnitude difference in average specimen weight between the small and large size fractions (Fig. S1B).

The MiSeq run generated 21,911,822 PE reads which are available as raw data as well as demultiplexed samples under the NCBI SRA accession PRJNA625408. No-call events were observed on several tiles in read 2, leading to N being inserted at positions 16, 17 and 25. A total of 20 out of 38 tiles were affected by one or more No-call events, but as
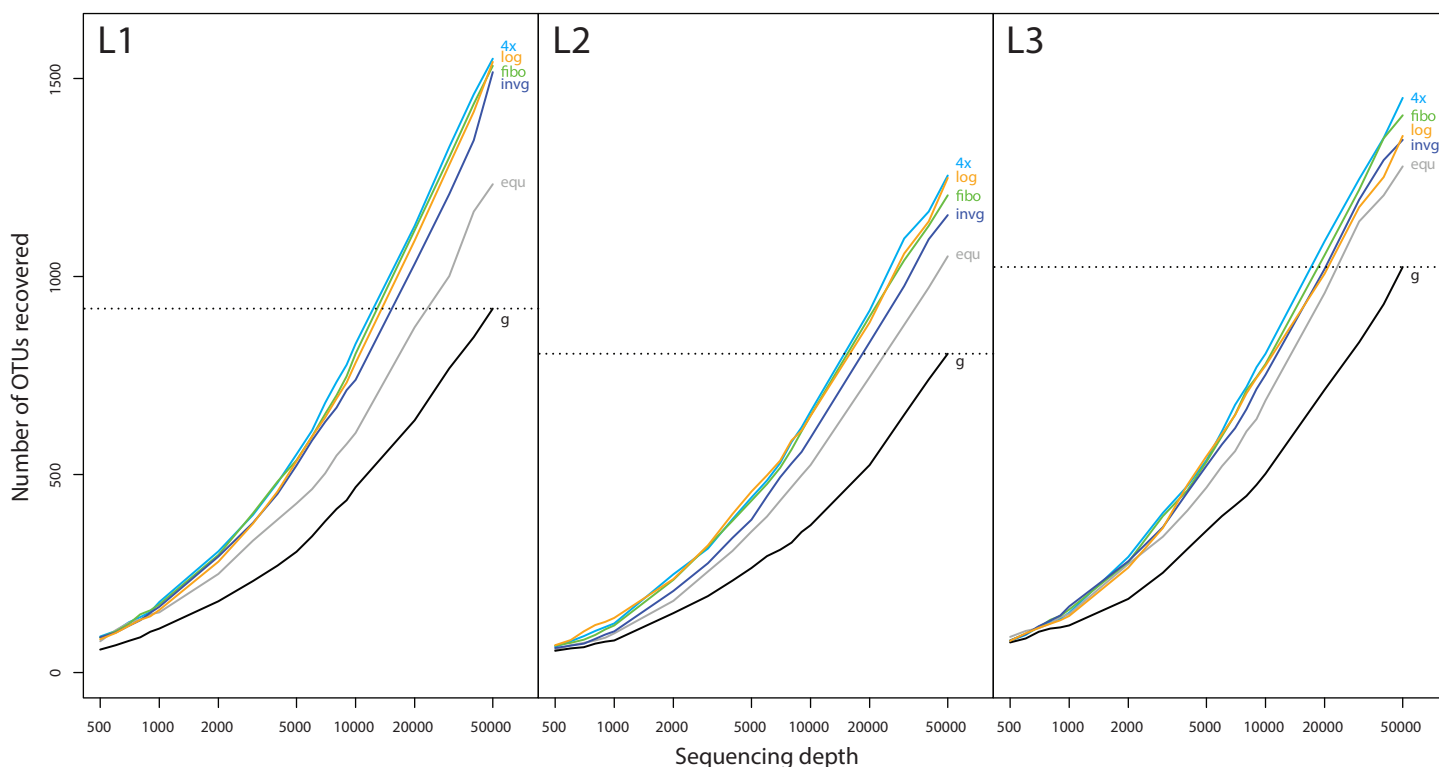
Figure 3 Rarefaction curves at different sequencing depths for all three samples (L1, L2 and L3) and lysates pooled in different proportions.
The taxon recovery of the unsorted Malaise sample (g) at 50,000 reads sequencing depth is used to estimate the increase in sequencing depth needed to recover a similar number of OTUs to what is recovered from samples that are size sorted and pooled in favour of smaller size fractions. The OTU counts in this rarefaction analysis are slightly higher than in our other analyses since replicates were pooled directly, without removing reads that are not in both replicates and without discarding OTUs or reads below 0.01% abundance.

respective areas were mostly removed by primer trimming, analysis was not impaired. Average sequencing depth per sample was 249,423 reads (SD = 92.075) with the lowest sample containing 146,344 reads. In bioinformatic processing on average 23.40% (SD = 4.16%) of reads were discarded. After complete quality filtering, 99% of the 2,197 OTUs were assigned to the phylum Arthropoda (Table S2).

To assign a specimen size to each OTU, relative read abundance of all three samples was combined for each size class. Then, the size class with the highest abundance was selected as a specimen size estimate. Three quarters of OTUs were present in only one size fraction (Figs. S3, S4, S5). Thus, most OTUs could be clearly assigned to a size class using this strategy, with a few exceptions in the large and medium size category (*e.g.* OTU1 *Myrmica ruginodis* and OTU11 *Sphaerophoria*). Here, the size category with the highest read number was assigned, even when an OTU was found in high abundance in two fractions. For some low-abundance OTUs, no reads were assigned in any of the separately sequenced size classes, but only in the pooled samples (this was the case for 9.38% of OTUs). Thus, the respective OTUs had no size class assigned. Most OTUs were assigned to smaller size classes, with 41.51% in S, 24.53% in M, 19.21% in L and 5.37% in XL (Table S1). The individual size fractions (S, M, L and XL) recovered mostly reads from their respective size class (Fig. 2), with only a few exceptions (*e.g.* OUT1, *Myrmica*

ruginodis). Here a high number of reads was assigned to size fraction L and M (43.2% and 55.7%). Thus, for OTU1 the size class M was assigned, despite the large size fraction also containing a substantial number of reads. The total number of taxa recovered was the highest in the smaller size fractions, and no more than 100 OTUs were recovered in the extra-large size fraction (Fig. 2).

When lysates were pooled based on dry size fraction weight (g) (simulating an unsorted Malaise trap sample), results were dominated by large specimens. Compared to the other pooling strategies, the simulated unsorted samples showed the weakest OTU recovery, followed by the pooling of all four lysates in equal proportions (equ). The other four pooling methods (invg, fibo, 4× and log) recovered more OTUs, with equally high numbers detected across the three Malaise trap samples. The logarithmic pooling (log) recovered the highest proportion of small specimens, while showing reduced recovery of large ones compared to the other methods. This was also supported by Jaccard and Bray–Curtis dissimilarity analysis (Fig. S6). Here, especially samples pooled using the log and 4× strategies cluster closely with samples where size class S had been sequenced individually. Samples pooled using strategies equ or g showed high similarity to size fraction L or even XL when read numbers were taken into account.

In silico pooling of the four individually sequenced size fractions at 1/4 sequencing depth of the lysates pooled in equal proportions (equ), showed that the individually sequenced size fractions recovered an average of 26.56% more OTUs than the equally pooled lysates (Fig. 2). However, the numbers of OTUs recovered from the *in silico* pool of individually sequenced size fractions were comparable to OTU counts recovered by lysates pooled in favour of small size fractions.

Rarefaction analysis showed that samples pooled in favour of small size fractions (invg, fibo, 4×, log) had a similar number of OTUs at around 15,000 reads sequencing depth compared to what was recovered from the unsorted sample (g) at 50,000 reads sequencing depth (Fig. 3). This represents a factor of approximately 3–4× reduction of sequencing depth compared to unsorted samples. In this comparison, unsorted samples (g) were more biased towards large specimens, while the samples pooled in favour of small specimens showed a higher detection rate of taxa assigned to the small size class (Fig. 2). Thus, while the number of OTUs recovered from the unsorted samples at 3–4× sequencing depth might be similar to that of size sorted samples, there are still differences in OTU composition between different pooling strategies (Fig. 2, Fig. S6).

To test if separation of a sample in only two size fractions would be sufficient to obtain optimal taxon recovery, individual size fractions S+M and L+XL were combined and pooled *in silico* in different ratios (Fig. S3). The recovery of OTUs was highest when adding 10–20% of the large fraction to the small fraction. The overall OTU recovery in the *in silico* pooled samples did in fact exceed the number of OTUs recovered by the processed lysate pools (invg, fibo, 4×, log) by up to 10–20%.

DISCUSSION

Previous studies have demonstrated that size sorting before metabarcoding can improve taxon recovery (*Elbrecht, Peinert & Leese, 2017*; *Creedy, Ng & Vogler, 2019*). This is

particularly relevant when dealing with bulk samples containing specimens with a wide variation in biomass. We were able to confirm these trends for Malaise trap samples. Experiments were done on three Malaise samples only, which represents low replication effort and thus does not allow for statistical testing of results. However, the results presented here clearly support expectations and findings of previous studies (*Elbrecht, Peinert & Leese, 2017*; *Creedy, Ng & Vogler, 2019*). As illustrated in Figs. 2 and 3, our results present distinct patterns clearly showing differences between the applied strategies and therefore justifying the limited number of replicates and the absence of statistical testing.

Small and medium sized fractions contained the majority of OTUs in each of the three samples, consistent with studies from marine benthos (*Wangensteen et al., 2018*). When the four size fractions were pooled in different proportions, the fractions simulating a completely unsorted Malaise trap were dominated by large specimens, showing consistently the lowest recovery of OTUs. Interestingly, all four ways of pooling in favour of the smaller fractions (inverse weight, logarithmically, fibonacci based and 4-fold increase) recovered a similar number of OTUs. This result was unexpected, as for example logarithmic pooling strongly favours small fractions while severely underrepresenting larger ones. A reason for this effect could be OTU sharing across the individually sequenced size classes due to small specimens attached to larger ones, consequently also detected in larger fractions. This would also indicate that sieving does not perfectly separate individuals of different size classes. However, for the dry sieving approach, the aggregation of specimens of different size classes was rarely observed and counteracted through the shaking amplitude. The phenomenon occurred more often with wet sieving but was counteracted with an increased mesh surface, the use of higher ethanol volumes during sieving and a higher sieving amplitude. Furthermore, smaller specimens may be detected as secondary prey items in the guts of larger specimens and would therefore also be detected in larger size fractions. Additionally, legs and other body parts can easily break off the brittle specimens during dry sieving, and small size fractions therefore contain tissue from large specimens. Because small pieces of tissue will not break off from all specimens equally and large taxa found in the small fractions might not be representative of the real diversity in the samples, specimen damage during sieving and the subsequent detection of larger specimens in smaller size classes should be avoided. As shown here, wet sieving is a promising approach which is more gentle than dry sieving. In addition, different size fractions from different samples should be pooled in the same proportions for each sample to ensure comparability. Choice of pooling strategy might depend on study goals, and on size class and taxonomic composition of the samples, which can vary greatly depending on location. If taxa of interest fall within a particular size range, pooling should be adjusted accordingly, given that the number of reads and OTUs of differently sized specimens depends on the pooling process.

While we have shown that size sorting can increase taxon recovery in Malaise trap samples, it remains to be discussed how this can be done most efficiently. This is of particular importance for upscaled metabarcoding as required for biodiversity monitoring. Size sorting can only be used in larger scale metabarcoding projects if the process is time and cost efficient and does not substantially increase laboratory work. When the

number of samples is limited and specimen biomass does not vary substantially, biomass biases can be mitigated (*Creedy, Ng & Vogler, 2019*) or minimized by only including a small part of the tissue (*e.g.* one leg) from large specimens (*Ji et al., 2013*; *Carew, Coleman & Hoffmann, 2018*). While manual sorting and processing of specimens can be a good strategy to reduce biomass bias, it is rather work intensive and difficult to scale or standardize. An automated approach would be desirable, but automated sieving techniques come with their own challenges. With the dry sieving approach used in this study, major methodological limitations arose, as it took several hours per sample to pick electrostatically charged individual specimens from the equipment. Consequently, we recommend using the alternative approach of wet sieving, completely immersing the specimens and sieve in ethanol. This approach can be automated more easily by stacking the sieve containers on shakers and thereby reducing hands-on time and increasing scalability. However, it needs to be considered that intensive cleaning of the equipment is required between sample processing steps to prevent cross contamination. Here, at least 15 min per sample should be calculated. Furthermore, wet sieving is conducted in pure 96% ethanol and the costs for additional amounts of ethanol need to be considered. Wet sieved specimens are less prone to damage and might be usable for morphological identification if non-destructive metabarcoding methods are applied, using for example preservative ethanol (*Hajibabaei et al., 2012*) or lysis buffer protocols (*Nielsen et al., 2019*).

Our *in silico* evaluations suggest that size sorting with a single sieve of four mm mesh size represents a good trade-off between taxon recovery and laboratory workload. Pooling the large and small size fractions in a proportion of 1:10 to 2:10 showed the highest taxon recovery in our *in silico* test, which however should be confirmed experimentally, since *in silico* comparisons can have limited reliability and due to the low replicate number processed in this study. For example, when *in silico* pooling 1/4 of each of the individually sequenced size fractions, more taxa were recovered than sequencing equally pooled lysate proportions. Thus, taxon recovery for the *in silico* estimates is slightly overrepresented, probably due to primer bias effects (*Elbrecht & Leese, 2015*; *Piñol et al., 2015*). For example, if a specific taxon in the large fraction is very well amplified, it will make up a large proportion of the reads in the pooled sample, decreasing the sequencing depth of all other specimens. When individual size fractions are sequenced, primer biases are limited to each individual fraction, increasing overall taxon recovery which is why some authors recommend the taxonomic sorting of bulk samples and sequencing each taxonomic group individually (*Morinière et al., 2016*; *Beentjes et al., 2019*). A taxonomic sorting step is however impractical for metabarcoding, as it substantially prolongs processing time per sample and does not necessarily increase taxon recovery as most arthropod primer binding sites show the same variability pattern (*Elbrecht et al., 2019*).

Other studies implement individual sequencing of each size fraction (*Cowart et al., 2015*; *Wangensteen et al., 2018*) and provided information about community composition for each size class to increase the detection of small and rare taxa. Sequencing of separate size classes doubles the costs for sample preparation and sequencing per sample. As an alternative, lysates or extracted DNA can be pooled proportionally to enrich the small size fractions sufficiently. Thereby, size sorting and subsequent pooling of lysates of each

fraction before the DNA extraction is particularly advantageous as it reduces costs and time required per sample compared to pooling strategies after DNA extraction or PCR.

In our study, we confirm that size sorting can be applied to increase taxon recovery in metabarcoding of Malaise traps samples. However, the implementation of techniques, *e.g.* number of size fractions or pooling proportions, needs to be adjusted to the aim of the study. In some cases, size sorting may be skipped, as a sufficient signal can be achieved with a high sequencing depth for samples with limited specimen size variability (*Creedy, Ng & Vogler, 2019*). While *Creedy, Ng & Vogler (2019)* do not give exact numbers on sequencing depth effects, we simulate that a 3- to 4-fold increase in sequencing depth recovers similar numbers of OTUs from an unsorted sample compared to size sorted and proportionally pooled samples. Calculations are based on rarefaction analysis, which is explained in the materials and methods section ("*In silico* analysis") and illustrated in Fig. 3. However, by increasing the sequencing depth, metabarcoding results will still be dominated by biomass rich taxa and it should be further experimentally tested how a higher sequencing depth can adequately replace size sorting.

Given that size sorting can be done in around 15 min using the wet sieving method demonstrated here, it seems like a worthwhile step to prepare Malaise trap samples for metabarcoding. Further Malaise trap sampling and additional size sorting experiments could be useful to quantify effects of large and abundant species on the detection of small or rare ones. Examination of Malaise trap samples or mock communities with known taxonomic composition and assigned size classes would be ideal to evaluate the results presented here. In addition to *in silico* assays, experimental approaches testing how a higher sequencing depth can compensate for size sorting and how different pooling proportions affect detected diversity patterns would be helpful to develop standardized monitoring protocols for flying arthropods.

CONCLUSIONS

We demonstrate that size sorting of Malaise trap samples can increase recovery of small specimens. While our analyses show distinct patterns clearly supporting previous studies, the limited number of samples processed and the resulting lack of statistical support needs to be carefully considered. Pooling lysates in the extraction process can be an efficient way of improving the representation of small taxa. The methods presented here reduce laboratory work and costs, in comparison to sequencing each size fraction individually. *In silico* tests suggest that sieving into two size fractions (four mm mesh size) and subsequent pooling in a proportion of 1:10 (one part size class L and nine parts size class S) or 2:10 can substantially increase taxon recovery. Increasing sequencing depth by 3 to 4-fold can also be an alternative to size sorting, but might still underrepresent small and rare taxa due to biomass biases. Wet sieving of specimens in ethanol may be the best compromise for optimizing taxon recovery with low effort, compared to other approaches. However, this method remains to be thoroughly tested. Ultimately, the decision of whether or not samples should be size sorted depends on the research goals, budget and timeline as well as on the sample composition and size heterogeneity of specimens in the samples.

ACKNOWLEDGEMENTS

We want to thank Claudia Etzbauer for laboratory infrastructure and supplying us with additional ATL buffer. Thomas Creedy, for helpful discussions and/or feedback to the manuscript. Sandra Kukowka for running the library on the Fragment Analyzer.

Funding

This work was supported by the Ministry for Environment, Agriculture, Conservation and Consumer Protection of the German State of North Rhine-Westphalia (No. III-1-620.08). There was no additional external funding received for this study. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
Ministry for Environment, Agriculture, Conservation and Consumer Protection of the German State of North Rhine-Westphalia: III-1-620.08.

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Vasco Elbrecht conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Sarah J. Bourlat conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Thomas Hörren performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Angie Lindner conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Adriana Mordente conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Niklas W. Noll analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Livia Schäffler conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Martin Sorg conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Vera M.A. Zizka analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Field Study Permissions

The following information was supplied relating to field study approvals (*i.e.*, approving body and any reference numbers):

Field experiments were approved by the nature conservation authority Euskirchen (reference number: 67.64.300/Pa).

Data Availability

The raw sequence used in this study is available at NCBI SRA: PRJNA625408.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.12177#supplemental-information.

REFERENCES

Aylagas E, Mendibil I, Borja Á, Rodríguez-Ezpeleta N. 2016. Marine sediment sample pre-processing for macroinvertebrates metabarcoding: mechanical enrichment and homogenization. *Frontiers in Marine Science* **3(96)**:203 DOI 10.3389/fmars.2016.00203.

Beentjes KK, Speksnijder AGCL, Schilthuizen M, Hoogeveen M, Pastoor R, van der Hoorn BB. 2019. Increased performance of DNA metabarcoding of macroinvertebrates by taxonomic sorting. *PLOS ONE* **14(12)**:e0226527 DOI 10.1371/journal.pone.0226527.

Braukmann TWA, Ivanova NV, Prosser SWJ, Elbrecht V, Steinke D, Ratnasingham S, de Waard JR, Sones JE, Zakharov EV, Hebert PDN. 2019. Metabarcoding a diverse arthropod mock community. *Molecular Ecology Resources* **19(3)**:711–727 DOI 10.1111/1755-0998.13008.

Bush A, Monk WA, Compson ZG, Peters DL, Porter TM, Shokralla S, Wright MTG, Hajibabaei M, Baird DJ. 2020. DNA metabarcoding reveals metacommunity dynamics in a threatened boreal wetland wilderness. *Proceedings of the National Academy of Sciences of the United States of America* **117(15)**:8539–8545 DOI 10.1073/pnas.1918741117.

Carew ME, Coleman RA, Hoffmann Ary A. 2018. Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ* **6**:e4980 DOI 10.7717/peerj.4980.

Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, Arnaud-Haond S. 2015. Metabarcoding is powerful yet still blind: a comparative analysis of morphological and molecular surveys of seagrass communities. *PLOS ONE* **10(2)**:e0117562 DOI 10.1371/journal.pone.0117562.

Creedy TJ, Ng WS, Vogler AP. 2019. Toward accurate species-level metabarcoding of arthropod communities from the Tropical Forest Canopy. *Ecology and Evolution* **9(6)**:3105–3116 DOI 10.1002/ece3.4839.

Deagle BE, Clarke LJ, Kitchener JA, Polanowski AM, Davidson AT. 2018. Genetic monitoring of open ocean biodiversity: an evaluation of dna metabarcoding for processing continuous plankton recorder samples. *Molecular Ecology Resources* **18(3)**:391–406 DOI 10.1111/1755-0998.12740.

Deagle BE, Thomas AC, McInnes JC, Clarke LJ, Vesterinen EJ, Clare EL, Kartzinel TR, Eveson JP. 2019. Counting with DNA in metabarcoding studies: how should we convert sequence reads to dietary data? *Molecular Ecology* **28(2)**:391–406 DOI 10.1111/mec.14734.

Edgar RC, Flyvbjerg H. 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. *Bioinformatics* **31(21)**:3476–3482 DOI 10.1093/bioinformatics/btv401.

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(1)**:e7745 DOI 10.7717/peerj.7745.

Elbrecht V, Leese F. 2015. Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass—sequence relationships with an innovative metabarcoding protocol. *PLOS ONE* **10(7)**:e0130324 DOI 10.1371/journal.pone.0130324.

Elbrecht V, Peinert B, Leese F. 2017. Sorting things out: assessing effects of unequal specimen biomass on DNA metabarcoding. *Ecology and Evolution* **7(17)**:6918–6926 DOI 10.1002/ece3.3192.

Elbrecht V, Steinke D. 2019. Scaling up DNA metabarcoding for freshwater macrozoobenthos monitoring. *Freshwater Biology* **64(2)**:380–387 DOI 10.1111/fwb.13220.

Elbrecht V, Vamos EE, Meissner K, Aroviita J, Leese F. 2017. Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution* **8(10)**:1265–1275 DOI 10.1111/2041-210X.12789.

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ. 2017. Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications* **8(1)**:1188 DOI 10.1038/s41467-017-01312-x.

Hajibabaei M, Spall JL, Shokralla S, van Konynenburg S. 2012. Assessing biodiversity of a freshwater benthic macroinvertebrate community through non-destructive environmental barcoding of DNA from preservative ethanol. *BMC Ecology* **12(1)**:28 DOI 10.1186/1472-6785-12-28.

Hardulak LA, Jerome M, Axel H, Lars H, Stefan S, Dieter D, Jörg M, Paul DNH, Gerhard H. 2020. DNA metabarcoding for biodiversity monitoring in a national park: screening for invasive and pest species. *Molecular Ecology Resources* **20(6)**:1–16 DOI 10.1111/1755-0998.13212.

Harris JK, Sahl JW, Castoe TA, Wagner BD, Pollock DD, Spear JR. 2010. Comparison of normalization methods for construction of large, multiplex amplicon pools for next-generation sequencing. *Applied and Environmental Microbiology* **76(12)**:3863–3868 DOI 10.1128/AEM.02585-09.

Hausmann A, Segerer AH, Greifenstein T, Knubben J, Morinière J, Bozicevic V, Doczkal D, Günter A, Ulrich W, Habel JC. 2020. Toward a standardized quantitative and qualitative insect monitoring scheme. *Ecology and Evolution* **10(9)**:4009–4020 DOI 10.1002/ece3.6166.

Hebert PDN, Cywinska A, Ball SL, deWaard JR. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270(1512)**:313–321 DOI 10.1098/rspb.2002.2218.

Ji Y, Ashton L, Pedley SM, Edwards DP, Tang Y, Nakamura A, Kitching R, Dolman PM, Woodcock P, Edwards FA, Larson TH, Hsu WW, Benedick S, Hamer KC, Wilcove DS, Bruce C, Wang X, Taal L, Lott M, Emerson BC, Yu DW. 2013. Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters* **16(10)**:1245–1257 DOI 10.1111/ele.12162.

Kelly RP, Shelton AO, Gallego R. 2019. Understanding PCR processes to draw meaningful conclusions from environmental DNA studies. *Scientific Reports* **9(1)**:12133 DOI 10.1038/s41598-019-48546-x.

Krehenwinkel H, Wolf M, Lim JY, Rominger AJ, Simison WB, Gillespie RG. 2017. Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports* **7(1)**:17668 DOI 10.1038/s41598-017-17333-x.

Lamb PD, Hunter E, Pinnegar JK, Creer S, Davies RG, Taylor MI. 2019. How quantitative is metabarcoding: a meta-analytical approach. *Molecular Ecology* **28(2)**:420–430 DOI 10.1111/mec.14920.

Leray M, Knowlton N. 2015. DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. *Proceedings of the National Academy of Sciences of the United States of America* **112(7)**:2076.

Leray M, Knowlton N, Ho S-L, Nguyen BN, Machida RJ. 2019. GenBank is a reliable resource for 21st century biodiversity research. *Proceedings of the National Academy of Sciences of the United States of America* **116(45)**:22651–22656 DOI 10.1073/pnas.1911714116.

Liu M, Clarke LJ, Baker SC, Jordan GJ, Burridge CP. 2020. A practical guide to DNA metabarcoding for entomological ecologists. *Ecological Entomology* **45(3)**:373–385 DOI 10.1111/een.12831.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *Next Generation Sequencing Data Analysis* **17(1)**:10 DOI 10.14806/ej.17.1.200.

Mata VA, Ferreira S, Campos RM, Silva LP, Veríssimo J, Corley MFV, Beja P. 2020. Efficient assessment of nocturnal flying insect communities by combining automatic light traps and DNA metabarcoding. *Environmental DNA* **3(2)**:398–408 DOI 10.1002/edn3.125.

Morinière J, de Araujo BC, Lam AW, Hausmann A, Balke M, Schmidt S, Hendrich L, Doczkal D, Fartmann B, Arvidsson S, Haszprunar G. 2016. Species identification in Malaise trap samples by DNA barcoding based on NGS technologies and a scoring matrix. *PLOS ONE* **11**:e0155497 DOI 10.1371/journal.pone.0155497.

Nielsen M, Gilbert MTP, Pape T, Bohmann K. 2019. A simplified DNA extraction protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity. *Environmental DNA* **1(2)**:144–154 DOI 10.1002/edn3.16.

Piñol J, Mir G, Gomez-Polo P, Agustí N. 2015. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources* **15(4)**:819–830 DOI 10.1111/1755-0998.12355.

Piper AM, Batovska J, Cogan NOI, Weiss J, Cunningham JP, Rodoni BC, Blacket MJ. 2019. Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *GigaScience* **8(8)**:10 DOI 10.1093/gigascience/giz092.

Renaud G, Stenzel U, Maricic T, Wiebe V, Kelso J. 2015. deML: robust demultiplexing of Illumina sequences using a likelihood-based approach. *Bioinformatics* **31(5)**:770–772 DOI 10.1093/bioinformatics/btu719.

Rognes T, Flouri T, Nichols B, Quince C, Mahé F. 2016. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4(17)**:e2584 DOI 10.7717/peerj.2584.

Siemann E, Tilman D, Haarstad J. 1996. Insect species diversity, abundance and body size relationships. *Nature* **380(6576)**:704–706 DOI 10.1038/380704a0.

Singer GAC, Fahner NA, Barnes JG, McCarthy A, Hajibabaei M. 2019. Comprehensive biodiversity analysis via ultra-deep patterned flow cell technology: a case study of eDNA metabarcoding seawater. *Scientific Reports* **9(1)**:5991 DOI 10.1038/s41598-019-42455-9.

Ssymank A, Sorg M, Doczkal D, Rulik B, Merkel-Wallner G, Vischer-Leopold M. 2018. Praktische Hinweise Und Empfehlungen Zur Anwendung von Malaisefallen Für Insekten in Der Biodiversitätserfassung Und Im Monitoring. *Series Naturalis* **1**:1–12.

Steinke D, Braukmann TWA, Manerus L, Woodhouse A, Elbrecht V. 2020. Effects of Malaise trap spacing on species richness and composition of terrestrial arthropod bulk samples. *BioRxiv* **19(1)**:711 DOI 10.1101/2020.09.30.321430.

Stork NE, Blackburn TM. 1993. Abundance, body size and biomass of arthropods in Tropical Forest. *Oikos* **67(3)**:483–489 DOI 10.2307/3545360.

Turunen J, Elbrecht V, Steinke D, Aroviita J. 2021. Riparian forests can mitigate warming and ecological degradation of agricultural headwater streams. *Freshwater Biology* **66(4)**:785–798 DOI 10.1111/fwb.13678.

Vamos E, Elbrecht V, Leese F. 2017. Short COI markers for freshwater macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics* **1(5)**:e14625 DOI 10.3897/mbmg.1.14625.

Wangensteen OS, Turon X. 2017. Metabarcoding techniques for assessing biodiversity of marine animal forests. In: *Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots.* Berlin: Springer International Publishing, 445–473.

Wangensteen OS, Palacín C, Guardiola M, Turon X. 2018. DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. *PeerJ* **6(Suppl S9)**:e4705 DOI 10.7717/peerj.4705.

Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MF, Grabowski Mł, Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaťovičová Z, Ferreira S, Dijkstra K-DB, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, van der Hoorn BB, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher JN, Mamos T, Paz G, Pešić V, Pfannkuchen DM, Pfannkuchen MA, Price BW, Rinkevich B, Teixeira MAL, Várbíró G, Ekrem T. 2019. DNA barcode reference libraries for the monitoring of aquatic biota in Europe: gap-analysis and recommendations for future work. *Science of the Total Environment* **678**:499–524 DOI 10.1016/j.scitotenv.2019.04.247.

Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z. 2012. Biodiversity soup: metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution* **3(4)**:613–623 DOI 10.1111/j.2041-210X.2012.00198.x.

Zizka VMA, Geiger MF, Leese F. 2020. DNA metabarcoding of stream invertebrates reveals spatio-temporal variation but consistent status class assessments in a natural and urban river. *Ecological Indicators* **115(4)**:106383 DOI 10.1016/j.ecolind.2020.106383.

Zizka VMA, Leese F, Peinert B, Geiger MF. 2019. DNA metabarcoding from sample fixative as a quick and voucher-preserving biodiversity assessment method. *Genome* **62(3)**:122–136 DOI 10.1139/gen-2018-0048.